Event Monitoring of Transformer Discharge Sounds based on Voiceprint

Rui Cai¹, Qian Wang¹, Yucheng Hou¹ and Haorui Liu²*

¹ State Grid Beijing Electric Power Research Institute, Fengtai, Beijing, 100075, China
² Tsinghua AI Plus, Beijing, 100084, China
* Corresponding author’s e-mail: liuhaorui@aithu.com

Abstract. This paper investigates the operation inspection and anomaly diagnosis of transformers in substations, and carries out an application study of artificial intelligence-based sound recognition technology in transformer discharge diagnosis to improve the timeliness and diagnostic capability of intelligent monitoring of substation equipment operation. In this study, a sound parameterization technology in the field of sound recognition is used to implement automatic discharge sound detections. The sound samples are pre-processed and then Mel-frequency cepstrum coefficients (MFCCs) are extracted as features, which are used to train Gaussian mixture models (GMMs). Finally, the trained GMMs are used to detect discharge sounds in the place of transformers in substations. The test results demonstrate that the audio anomaly detection based on MFCCs and GMMs can be used to effectively recognize anomalous discharge in the high scenario of transformers.

1. Introduction

Technicians who monitor the operation status of transformers rely on sound to diagnose any anomalies, which provides an idea for this project to propose a fault detection program for electrical equipment that simulates human auditory system. It is a safer and more convenient assistance for technicians in the daily maintenance and operation of substations, and reduces accidents by delivering early warning information through intelligent recognition of anomalous sounds. The substation audible sound diagnosis technology is a new diagnostic technology developed in recent years, which determines operation status of transformers by analyzing the audible sound signals in the noise band from 20 Hz to 20,000 Hz. It has a set of advantages, such as easy installation of acoustic probes, easy measurement of acoustic signals, non-contact measurement, high speed, no need to stick sensors in advance, online monitoring of moving targets, no electromagnetic signal generated in the acquisition and transmission processes of acoustic signals, no electrical connection with equipment, and no interference with equipment. Tremendous efforts have been made to examine the application of audible sound analysis for equipment fault diagnosis[1-3]. At present, researchers at home and abroad developed three aspects concerning the research interests and key topics: (1) transformer audible sound generation mechanism, such as fault causes, development trend; (2) transformer audible sound signal processing, such as noise cancellation technology, spectrum analysis; (3) development of expert diagnosis system. However, there are few studies focusing on the four parts of audio sound diagnosis technology: signal acquisition, signal data processing, signal noise cancellation, and signal feature quantity. Given all that, automatic detection and diagnosis of anomalous operation of substation transformers can be performed by collecting substation operation sound signals based on sound sampling.
acquisition sensor networks, mapping substation transformer anomaly sound features, and incorporating MFCCs, acting as an effective approach for monitoring conditions of substations. This paper addresses the detection of anomalous discharge sounds in the operation environment of transformers, and builds an anomalous discharge sound detection model based on MFCCs, in which real-time online monitoring sound data is scored and compared with preset scoring threshold to determine discharge anomalies at site.

2. Transformer sound feature extraction

2.1. Principle of MFCC for sound data acquisition

MFCCs are commonly used feature parameters in speech recognition, which combine the auditory perception of human ears and the principle of speech signals, with good recognition rate and robustness. Human can distinguish various voices, and the key is the cochlea. The cochlea acts as a filter bank that filters on a logarithmic frequency scale, where frequency below 1,000 Hz is linear, and frequency above 1,000 Hz is logarithmic. A frequency unit, Mel frequency, is redefined to transform speech signals to the perceptual frequency domain for analysis. Its relationship to the actual frequency is:

\[ F_{\text{mel}} = 2595 \log(1 + f/700) \]  

Where, \( F_{\text{mel}} \) is perceptual frequency, and \( f \) is actual frequency in Hz.

2.2. MFCCs for sound feature extraction

Mel-frequency cepstrum coefficients (MFCCs) are a representation of short-time energy spectrum of a sound, which are obtained by linear cosine transformation of logarithmic power spectrum in the non-linear Mel frequency scale. MFCCs are the plenary coefficients that form MFC. They stemmed from a cepstrum of audio fragments. The difference between cepstrum and Mel-frequency cepstrum is that the frequency bands uniformly distributed on the Mel-frequency scale are closer to the response of human auditory system than those linearly distributed in the conventional cepstrum. MFCCs have been widely used in the speech-processed speaker recognition).

2.3. Gaussian mixture models

GMMs use several Gaussian probability density functions by weighting to infinitely approximate and accurately quantify the mathematical models. For the extracted MFCCs of discharge signals, multidimensional Gaussian functions are used to establish GMMs to fit this dataset, and MCFF models of discharge sound are established. A GMM consists of several Gaussian functions, and the multidimensional Gaussian probability density function is:

\[
P(x \mid \theta) = \frac{1}{(2\pi)^{D/2} |\Sigma|^{1/2}} e^{-\frac{(x-\mu)^T \Sigma^{-1}(x-\mu)}{2}}
\]

Where, \( \theta \) is the expectation, covariance and weight coefficient of the sub-model, \( \Sigma \) is the covariance, \( \mu \) is the data mean, and \( D \) is the data dimension.

When establishing GMMs for MFCCs of discharge signals, let the mixture models with feature parameters \( x = \{x_1, x_2, \cdots, x_K\} \) is composed of \( K \) Gaussian distributions, then the probability distribution of GMMs \( p(x \mid \theta) \) is:

\[
p(x \mid \theta) = \sum_{k=1}^{K} a_k \phi(x \mid \theta_k)
\]

Where, in \( \theta_k = (\mu_k, \Sigma_k) \), \( K \) is the number of sub-Gaussian models, \( a_k \) is the probability of observed data of \( k \)-th sub-model \( (a_k \geq 0) \), and \( \phi(x \mid \theta_k) \) is the Gaussian distribution density function of \( k \)-th sub-model.
2.4. GMM parameter estimation

2.4.1. Maximum likelihood function. Maximum likelihood estimation is a common method to obtain estimated values. When estimating GMMs, maximum likelihood estimated parameters can be used to estimate parameters, which usually takes logarithm in the calculation:

$$\log L(\theta) = \sum_{j=1}^{N} \log P(x_j | \theta) = \sum_{j=1}^{N} \log \left( \sum_{k=1}^{K} a_k \phi(x_j | \theta_k) \right) \quad j = 1, 2, \cdots, N$$

Where, $L(\theta)$ is the logarithm estimated from the likelihood function, and $x_j$ is the $j$-th observed data. In the calculation of maximum likelihood function, because Equation (11) contains logarithmic and summation process, and each sub-model has unknown parameters, the model parameters cannot be obtained through simple derivative calculations, but only from iteration.

2.4.2. EM algorithm. EM algorithm is used for maximum likelihood estimation of probability model parameters with latent variables, which has two steps: expectation and maximization.

Step 1, expectation. Calculate the probability $r_{jk}$ of data $x_j$ from the model $k$ based on the current parameters:

$$r_{jk} = \frac{a_k \phi(x_j | \theta_k)}{\sum_{k=1}^{K} a_k \phi(x_j | \theta_k)} \quad j = 1, 2, \cdots, N; k = 1, 2, \cdots, K$$

Step 2, maximization. Calculate new iteration parameters, including updated mean $\mu_k'$, covariance matrix $\Sigma_k'$, and weight coefficient $a_k'$:

$$\mu_k' = \frac{\sum_{j=1}^{N} r_{jk} x_j}{\sum_{j=1}^{N} r_{jk}} \quad k = 1, 2, \cdots, K$$

$$\Sigma_k' = \frac{\sum_{k=1}^{N} r_{jk} (x_j - \mu_k)(x_j - \mu_k)^T}{\sum_{j=1}^{N} r_{jk}} \quad k = 1, 2, \cdots, K$$

$$a_k' = \frac{1}{N} \sum_{j=1}^{N} r_{jk} \quad k = 1, 2, \cdots, K$$

Repeat steps 1 and 2, and set a very small positive number as the termination condition of iteration, when $\| \theta' + 1 - \theta \| < \varepsilon$, iterated parameters hardly change, which meet the requirements, and the iteration is terminated.

3. Mathematical Equations of transformer sound feature extraction

In this experiment, 9 kinds of audio files of discharge sounds are collected. The length of each sound is about 30s, and the sampling rate is 32kHz. 9 kinds of discharge sounds are trained at the same time to obtain a discharge sound event detection model. The real environmental noise and discharge sound collected around the transformer of the substation are mixed to simulate the environmental discharge sound of the substation for testing.

In this experiment, the feature uses 40-dimensional MFCC and its first-order and second-order derivative splicing. After splicing, the feature dimension is 120 dimensions. The MFCC extraction parameters are: frame length 40ms, frame shift 10ms. The number of mixed GMM models is 5. After the model is scored, the scores of adjacent 5 frames (100ms) are averaged as the final abnormal score, and the event detection time difference is within 100ms.

Data with different signal-to-noise ratios are tested (here, the signal-to-noise ratio is defined as the ratio of the maximum amplitude of the spark to the maximum amplitude of the noise). Figure 2 and
Figure 3 shows the detection results when the signal-to-noise ratio is 0.5. Figure 2 shows that the signal is the original data waveform, and the discharge signal is 50% of the background noise amplitude. The blue line is the background noise waveform, and the yellow line is the discharge sound waveform. Figure 3 shows the spectrogram and the final model scoring results, where the left ordinate is the signal frequency scale, the right ordinate is the final estimated score scale, and the blue line is the final scoring value. It can be seen from the figure that the scoring result at this time can accurately determine the existence of the discharge signal, and the high score period and the discharge occurrence period have a high degree of consistency. Set the detection operating point threshold to -50, which can be seen from the figure. According to the calculation results in this article, the alarm of the discharge event can be well realized. The number of false alarms is 0. One event (about 29s) is missed, and further post-processing is passed. And optimized selection of detection operating points can better control false alarms and false alarms.

4. Conclusions
MFCC parameters of transformer discharge sound signals vary substantially and can be used to describe different discharge features. GMMs are established for training and testing, where five types of discharge faults are identified. Intelligent recognition of transformer faults is performed based on MFCCs and GMMs. It is suggested to extend sample libraries for other types of faults, providing reference for online monitoring of transformers and other substation equipment.
References

[1] Shatnawi, Y., Al-Khassaweneh, M. (2013). Fault diagnosis in internal combustion engines using extension neural network. IEEE Transactions on Industrial Electronics, 61(3), 1434-1443.

[2] Padmanabhan, J., Johnson Premkumar, M. J. (2015). Machine learning in automatic speech recognition: A survey. IETE Technical Review, 32(4), 240–251.

[3] Chauhan, V., Dwivedi, S., Karale, P., Potdar, S. M. (2016). Speech to text converter using Gaussian Mixture Model (GMM). International Research Journal of Engineering and Technology (IRJET), 3(5), 160–164.

[4] Gupta, A., Sarkar, K. (2018). Recognition of spoken bengali numerals using MLP, SVM, RF based models with PCA based feature summarization. International Arab Journal of Information Technology, 15(2), 263–269.

[5] Gamit, M. R., Dhameliya, K. (2015). Isolated words recognition using MFCC, LPC and neural network. International Journal of Research in Engineering and Technology, 4(6), 146–149.