

Title	Study on anomalous sound detection using instantaneous phase features
Author(s)	VO, TRAN QUANG TUAN
Citation	
Issue Date	2025-09
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/20028
Rights	
Description	Supervisor: 鵜木 祐史, 先端科学技術研究科, 修士 (知識科学)

Anomalous sound detection (ASD) is the task of identifying whether the sound produced by a specific machine is normal or anomalous. Because anomalous sounds exhibit signs of malfunction, early detection and prevention can enhance predictive maintenance efforts, ultimately improving machinery reliability and reducing downtime. Distinguishing between abnormal and normal sounds, a task that typically requires skilled and experienced machine engineers, faces significant challenges due to a shortage of human resources. The invention of ASD systems that leverage acoustical features related to human auditory perception is a promising solution for incorporating the strengths of both machine capabilities and human abilities to improve performance.

Most of the approaches in ASD concentrate on leveraging the superiority of deep-learning-based techniques, such as Autoencoder (AE) and acoustic features in the Mel scale, such as Mel spectrogram, etc., to model normal sound in an unsupervised manner in its latent space. The anomalous sound can be detected based on the large anomaly score after reconstructing the input spectrogram of AE-based models. Besides, the other approach, based on auditory perception analysis of Ota et al., attempts to tackle ASD by researching the primary differences between normal and anomalous sounds in hearing to develop timbral attributes. Despite obtaining attractive results in this approach, there exists a gap in ASD performance in detecting anomalous sound from some machine types in the MIMII dataset, which include Slider (ID 06) and Valve (ID 06). Based on the noticeable indicator of anomalous sounds emitted from these machine types, as argued by Ota et al., this study hypothesizes that the bearing faults of sliders or the beating sound of valves during malfunctions can cause sudden changes in the instantaneous frequency of these sounds. This hypothesis motivated this study to investigate the instantaneous phase and its derivative to detect phase interruptions, which represent the instantaneous changes in frequency better than amplitude caused by anomalous sound.

This study aims to propose a novel approach for ASD by utilizing instantaneous phase features. These features are derived from the outputs of an auditory filterbank, and then the derivative of phase is calculated along time, frequency, and time-frequency axes to capture interruptions holistically. The proposed phase-based features are presented in both the concepts of the derivation steps and the implementation. Later, the simulation with a frequency modulation signal is performed to validate the correctness of

them. Moreover, a supervised experiment employing a support vector machine (SVM) and phase-based features is conducted on the MIMII dataset to verify the effectiveness in ASD.

Secondly, the unsupervised ASD system utilizing phase-based features is investigated in this study to handle the lack of anomalous data scenario. By leveraging the AE-based Interpolation Deep Neural Network (IDNN) model as the backbone and the Area Under the Receiver Operating Characteristic curve (AUC-ROC) as evaluation criteria, the experimental results demonstrate that the proposed phase-based features work well in detecting anomalous sound from most of the machine types in the MIMII dataset unsupervisedly, including Slider, Fan, and Pump, outperforming other unsupervised methods using amplitude-based features. Additionally, the study acknowledges the poor performance in detecting anomalous sounds from Valve. Therefore, further investigation of phase-based features in detecting anomalous valve sound is necessary.

In conclusion, this study achieved two research goals as presented, including proposing instantaneous phase features for ASD, verifying the correctness of the concepts, and establishing an unsupervised ASD system utilizing those. Through the experiment in an unsupervised manner, the proposed method demonstrates superior performance compared to other unsupervised methods using amplitude-based information as discriminated features. Future work should address the remaining drawbacks in this study, such as detecting anomalous sound under low SNR conditions, and improve the performance in detecting anomalous sound from Valve while using instantaneous phase features.