






Acoustic-Based Detection Technique for Identifying Worn-Out Components in Large-Scale Industrial Machinery

Christof Pichler^{1*} , Markus Neumayer^{1**} , Bernhard Schweighofer¹ , Christoph Feilmayr², Stefan Schuster^{2**} , and Hannes Wegleiter^{1**} 

¹Christian Doppler Laboratory for Measurement Systems for Harsh Operating Conditions, Institute of Electrical Measurement and Sensor Systems, Graz University of Technology, 8010 Graz, Austria

²voestalpine Stahl GmbH, 4020 Linz, Austria

*Graduate Student Member, IEEE

**Member, IEEE

Manuscript received 31 July 2023; accepted 8 August 2023. Date of publication 22 August 2023; date of current version 4 September 2023.

Abstract—This letter addresses the challenge of monitoring large-scale machine halls, particularly in the context of iron making processes. We propose an acoustic sound-based condition monitoring (ASCM) system to detect potential faults and damages in machinery. The letter focuses on selecting suitable audio features, integrating physical insights regarding the fault, and determining optimal window lengths for feature extraction. Our fault detection method utilizes outlier detection with a Gaussian mixture model trained on features extracted only from normal operating conditions. We compare conventional audio features with physically motivated features and conduct a window length analysis. The results demonstrate the effectiveness of our approach and the impact of incorporating physically motivated features for fault detection performance.

Index Terms—Sensor applications, audio features, fault detection, Gaussian mixture model (GMM).

I. INTRODUCTION

In various industries, ensuring the optimal functionality of individual machines is critical for uninterrupted production processes and minimizing financial losses. Particularly in industries, such as blast furnace iron production, where a continuous production chain is essential. The preprocessing of raw materials plays a crucial role in blast furnace iron production and any shutdown should be avoided due to the significant energy and time required for the start-up process of the blast furnace [1].

Manual monitoring alone is insufficient in industries where a continuous production chain is vital. Hence, condition monitoring systems are necessary to detect potential faults and damages in machinery [2]. Traditional condition monitoring systems, which often utilize vibration sensors, may not be suitable for large-scale machinery due to the requirement for a significant number of sensors. As an alternative, acoustic sound-based condition monitoring (ASCM) systems have been proposed [3]. These systems take advantage of the fact that experienced maintenance personnel can identify indications of damage events based on the acoustic signature of machinery. The broad applicability of ASCM to different types of machines has been explored in [2], and general research on sound scene analysis and the applicability of ASCM in large machine halls has been conducted in [4].

While acoustic condition monitoring has been widely used for single machines or small machine parts, such as rotating machines, roller bearings, and tool conditions [5], [6], [7], [8], [9], [10], [11], [12], [13], there is a need to address the challenges of monitoring large-scale machine halls. These halls, such as burden preparation halls for blast furnaces, are vast and highly reverberant environments [14]. Thus, the acoustic signals consist of machine sounds and reverberations,

where the activation and deactivation of specific machines for material mixing introduce variations in the operational conditions. As a result, the normal acoustic signal exhibits a wide range of variations across the signal space, indicating its nonuniqueness. Analyzing these diverse audio signals requires specialized techniques.

Most methods proposed in the aforementioned works [5], [6], [7], [8], [9], [10], [11], [12], [13] leverage acoustic knowledge of the fault case, particularly machine learning-based approaches that rely on a large amount of data in case of faulty machinery. However, our particular scenario presents a challenge as we have limited access to a substantial amount of audio data for fault cases. In addition, the fault signal is substantially overwhelmed by the background sound, indicating the challenging nature of dealing with a weak signal problem. Anomalous sound detection (ASD) methods have been proposed as effective solutions for detecting abnormal sounds in various applications, including small-scale industrial applications [15], [16], [17]. One key advantage of these methods is their ability to detect anomalies using only data from the normal state, employing outlier detection techniques.

In this research, we propose a fault detection system for huge industrial machinery based solely on normal operation state data. We place emphasis on the selection of appropriate audio features, utilizing both standard audio features and physical motivated audio features. A comprehensive analysis was conducted to assess the suitability of different combinations of audio features for such a detection system. In addition, we determined the optimal window length for data framing during the preprocessing stage and our findings indicate that incorporating physical knowledge can provide insights into determining the optimal window length.

II. MEASUREMENT SYSTEM AND DATA ACQUISITION

To perform audio measurements in a burden preparation hall, a microphone array was developed taking the challenging environmental

Corresponding author: Christof Pichler (email: christof.pichler@tugraz.at).

Associate Editor: B. Jakoby.

Digital Object Identifier 10.1109/LENS.2023.3307103

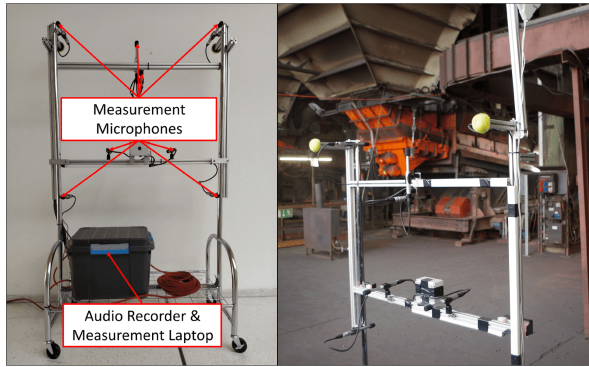


Fig. 1. Left: Picture of the measurement setup. The measurement setup is a microphone array with eight measurement microphones and an audio recorder and a measurement laptop. Right: Picture of the installed system in a burden preparation hall.

conditions into consideration. The volume level in the hall was found to be consistently high, averaging around 95 dB SPL. In addition, the hall is subjected to high dust loads, which necessitated the use of a dust-protected setup and microphones with a high dynamic range.

A. Microphone Array

To conduct audio measurements in the burden preparation hall, we set up an eight-microphone array using Behringer ECM8000 measurement condenser microphones. These microphones were chosen due to their high dynamic range and flat frequency response, which range from 20 to 20 000 Hz, making them well suited for our application. The audio signals were recorded using an 8-channel Zoom F8n MultiTrack Field Audio Recorder with a sampling rate of 48 kHz. Over multiple weeks, each hour a 20 s sample was stored as a multichannel WAV-file. To ensure an autonomous operation, a measurement laptop with remote access was used to operate the system. The arrangement and orientation of the microphones can be seen in Fig. 1, while a foam cover was used to protect the microphones from the high levels of dust present in the hall. To safeguard the sensitive components of the system, such as the audio recorder and measurement laptop, they were housed in a dust-protected box.

B. Measurements and Signal Analysis

The acoustic signal in the hall is a combination of multiple sound sources, including 13 sieves processing different materials. The operating states of these sieves, which are turned-ON or OFF based on blast furnace requirements, result in various combinations with different acoustic signatures. A previous study [4] demonstrated that the acquired acoustic signals are short-time stationary signals, making power spectral density (PSD) analysis suitable for investigating the frequency characteristics. The study in [4] extensively showed that normal operating states exhibit different PSDs, indicating a wide coverage of the signal space and diverse acoustic characteristics within normal operation.

Obtaining audio files of actual fault cases is crucial for testing the proposed ASD system. However, such data are often not readily available and despite our long term automated measurement routine, we could not capture a fault case. To overcome this limitation, we conducted fault tests on the machines and recorded the resulting audio in an emulated setting. Common fault scenarios, such as spring ruptures, involve the collision of loose metal parts with the machine during operation, generating a distinctive knocking sound. To replicate

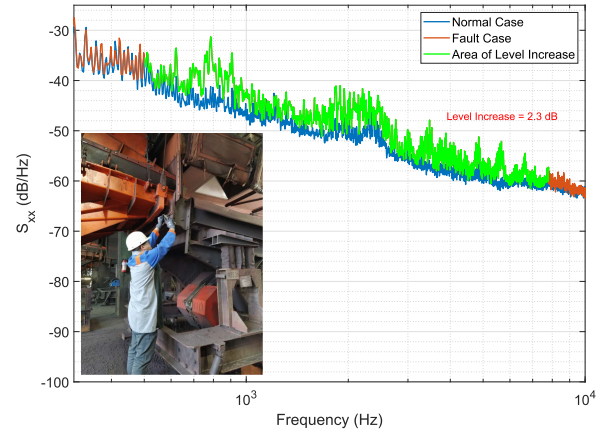


Fig. 2. PSDs for the normal condition and the emulated fault case. The fault case manifests itself in a level increase in a specific frequency band. The picture in the left lower corner of the plot shows the emulation of a fault case.

this sound, fault case experiments were performed with the assistance of experienced personnel, where a $100 \times 200 \times 10$ mm steel plate was placed against the machine during operation. Care was taken to ensure that the resulting sound closely resembled that of a real fault case. Fig. 2 illustrates the experimental setup and displays the PSDs of the normal operating condition just before the fault experiment and the simulated fault. This approach facilitated the collection of multiple datasets containing simulated yet authentic-sounding fault cases.

As depicted in Fig. 2, the level increase due to the fault mainly appears in the frequency range between 300 and 10 000 Hz. Consequently, for all subsequent analyses and assessments, we applied bandpass filtering to the recorded data, exclusively focusing on this specific frequency range.

C. Characteristics of Fault Case

The objective of analyzing audio data from fault cases is to extract relevant physical features and assess whether their incorporation enhances the effectiveness of fault detection of an ASD system. By identifying fundamental underlying physical properties, our aim is to integrate them into the fault detection process for improved performance.

Since the knocking in the fault data is expected to have periodic characteristics due to the periodic oscillation of the sieve, we conducted further analysis using the pitch detector [18]

$$T(\mathbf{x}) = \sum_{i=1}^{M/2-1} I(f_i) \quad (1)$$

where $I(f_i)$ is the periodogram of the data vector \mathbf{x} , M being the discrete period of the signal in samples, and $f_i = i/M$ are the frequencies in the interval $[0, 1/2]$. Evaluating $T(\mathbf{x})$ for different periods M and determining the maximum value gives access to the fundamental frequency if a periodic signal is present.

The underlying principle of this approach involves summing the higher harmonics and the fundamental frequency in the spectra to improve the SNR for accurate fundamental frequency estimation and detection. This test statistic can be derived by hypothesis testing, taking into account the structure of the PSD of a periodic signal. Now, to detect the presence of a periodic signal, the test statistic is compared with a threshold γ , and it is decided that a periodic signal is present if $T(\mathbf{x}) > \gamma$. In such scenarios, where background noise exhibits

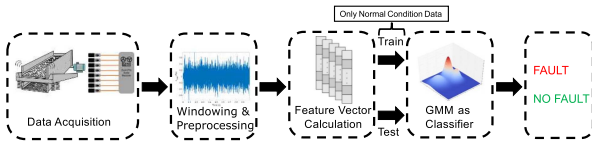


Fig. 3. ASD system is illustrated schematically, highlighting the key steps involved.

significant variations in an unknown manner, we suggest employing a constant false alarm rate detector to determine the threshold parameter γ [19]. For a more comprehensive understanding of pitch estimation and detection, see [18] and [20]. In comparing the estimated pitches between normal and fault cases, we observed a fundamental frequency corresponding to the sieve frequency in both scenarios. However, it is important to note that fault cases exhibited higher amplitudes to some extent, making them potential indicators of damage. Nonetheless, relying solely on pitch for fault detection is insufficient due to the overlap inside the feature space. Analyzing the PSDs revealed that fault cases present as bandpass signals. Based on these analysis, the utilization of bandpass energies, that can be extracted by employing a filterbank, as physically motivated features can be inferred. Combining these features with pitch estimates offers a promising approach for fault detection. Furthermore, our investigations into pitch detection and estimation also provided valuable insights into determining the optimal window size for feature extraction.

III. FAULT DETECTION METHOD

This section presents a comparative analysis in the context of an ASD-based approach, examining the effectiveness of different audio features and window lengths. A Gaussian mixture model (GMM) is trained on feature vectors extracted from a training set, exclusively composed of data from normal cases. Thus, the GMM serves as a statistical representation of the normal operating conditions in the feature space. GMMs are known for their ability to capture complex decision boundaries [15], making them well suited for the ASD application. Such systems use outlier detection methods based on probability values and a threshold associated with feature vectors and the model. Instead of employing a fixed threshold for outlier detection, we employ a receiver operating curve (ROC) analysis to evaluate the system’s overall performance across various threshold values, providing a comprehensive assessment of its fault detection capabilities. Such a system is schematically depicted in Fig. 3. Consequently, we conducted a comparison between systems utilizing conventional audio features and those employing the aforementioned physically based features. One limitation of using common features (CF) lies in the challenge of determining which specific features should be utilized. In addition, there is the issue of selecting appropriate parameters, such as windowing length, among others. Hence, utilizing physically motivated features based on the properties of fault signals provides a more intuitive and straightforward approach.

A. Window Length Analysis

The initial step in feature extraction involves framing the audio data using a suitable window function and length. According to [21], an antitruncation window is recommended. Therefore, we opted to utilize a Hamming window. To determine the optimal window length, we conducted a comprehensive analysis with different window sizes that encompassed the entire ASD system, utilizing the areas under the receiver operating curve (AUC) as the criterion for optimization.

TABLE 1. Analysis of Window Length

Window Length / ms	20	30	50	70	80	83	90	100
AUC (System with MFCC)	0.87	0.90	0.91	0.89	0.92	0.88	0.89	0.89

TABLE 2. List of Used Audio Features

Name	Category	Domain
Zero Crossing Rate	CF	Time-Domain
Short Time Energy	CF	Time-Domain
Crest Factor	CF	Time-Domain
Spectral Roll Off Frequency	CF	Frequency-Domain
Spectral Centroid	CF	Frequency-Domain
Audio Spectral Centroid	CF	Frequency-Domain
Audio Spectrum Spread	CF	Frequency-Domain
Mel-Frequency Cepstral Coefficients	MFCC	Frequency-Domain
Pitch	PF	Frequency-Domain
Maximum of Pitch Teststatistics	PF	Frequency-Domain
Bandenergies	BE	Frequency-Domain

CF, . . Common Features. MFCC, . . Mel Frequency Cepstral Coefficients. PF, . . Pitch Features. BE, . . Band Energies.

TABLE 3. Results of ROC-Analysis of the Different GMMs

Audio Features	CF	MFCC	MFCC and CF	PF + BE	MFCC and PF
AUC	0.84	0.92	0.92	0.90	0.91

The length analysis can be observed in Table 1, which shows that the performance exhibits a high degree of similarity, yet it also illustrates that the optimal window length for our specific purpose corresponds to the period of the sieve vibration (80 ms), which can be estimated through pitch estimation.

For all subsequent analysis, we employ a fixed window length of (80 ms).

B. Performance Comparison of the ASD System With Different Audio Features

For the ASD model, we have selected a set of features commonly employed in audio signal processing [21]. In addition, we leverage the estimated pitch, its corresponding amplitude and bandpass energies as audio features in our analysis. A comprehensive list of the features utilized in our analysis can be found in Table 2.

The audio features mentioned above were extracted from the dataset, specifically capturing sounds under normal operating conditions, and then divided into a training set (3000 windows of data) and a test set (3700 windows of data). In addition, extracted audio features from the fault case were added to the test set (1400 windows of data). The extracted features were utilized to construct feature vectors. Different feature combinations were used to construct distinct feature vectors, and for each combination, a GMM was fitted to the training set. This exploration aimed to assess the effectiveness of different feature types in the ASD approach. The testing dataset was then utilized to evaluate the fault detection performance of the trained GMMs. The performance evaluation included a detailed analysis using ROC, and the results are presented in Fig. 4. Furthermore, Table 3 gives the calculated AUC.

To interpret the AUC value, a rule of thumb according to [22] is that a value around 0.9 indicates excellent discrimination of the data. While the physically motivated features (PF + BE) showed good performance, the most effective results were obtained by using mel frequency cepstral coefficient (MFCCs) as features for our specific task. This finding highlights the practical value of MFCCs in capturing relevant information necessary for accurate fault detection within our context. Our investigations revealed that incorporating additional types of features alongside MFCCs did not yield any notable enhancements in the overall performance of the system. Nevertheless, the system

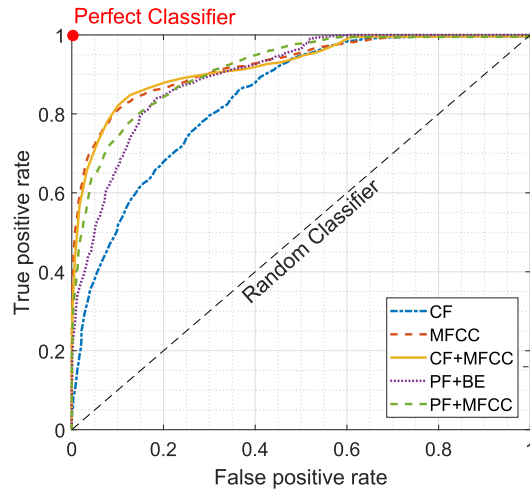


Fig. 4. ROC curve analysis for different sets of used features.

utilizing physically motivated features exhibits a high level of detectability and offers the advantage of providing a more interpretable representation of the data.

IV. CONCLUSION

In conclusion, our study reveals several findings. First, we found that the proposed ASD system is well suited for fault case detection in our specific application. Second, we observe that physical motivated features exhibit a better performance compared with CF, confirming their effectiveness in capturing relevant information for fault detection. In addition, while the MFCCs demonstrate only a slight improvement over the physical motivated features, they still emerge as the top-performing features for our specific task.

The physically motivated features have better interpretability and are easier to calculate, making them practical for implementation. The pitch information helps determine the optimal window length for both CF and MFCCs. These findings improve feature selection and contribute to more understandable and efficient fault detection systems in industries. The proposed system is the starting point for a complete detection system. It has good detection rates but a notable false positive rate as well. To improve the system, we are working on using the continuous sound produced during failures and considering time-related information. This is an ongoing research focus.

ACKNOWLEDGMENT

This work was supported in part by the Austrian Federal Ministry of Digital and Economic Affairs, in part by the National Foundation for Research, Technology and Development, Austria, in part by the Christian Doppler Research Association, and in part by the TU Graz Open Access Publishing Fund, Austria.

REFERENCES

- [1] T. Yamamoto, K. Wada, S. Kanoko, and A. Higuchi, "Present status and future of equipment diagnosis technology for iron- and steelmaking plants," *Nippon Steel Sumitomo Metal Tech. Rep.*, no. 112, 2016.
- [2] H. M. Hashemian and W. C. Bean, "State-of-the-art predictive maintenance techniques," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 10, pp. 3480–3492, Jan. 2011.
- [3] J. Yenus, G. Brooks, M. Dunn, and R. Kadam, "Application of vibration and sound signals in monitoring iron and steelmaking processes," *Ironmaking Steelmaking*, vol. 47, no. 2, pp. 178–187, 2020.
- [4] C. Pichler et al., "Acoustic condition monitoring: Signal analysis for large machinery halls," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf.*, 2022, pp. 1–6.
- [5] N. Rajapaksha, S. Jayasinghe, H. Enshaei, and N. Jayarathne, "Acoustic analysis based condition monitoring of induction motors: A review," in *Proc. IEEE Southern Power Electron. Conf.*, 2021, pp. 1–10.
- [6] C. S. Gong et al., "Design and implementation of acoustic sensing system for online early fault detection in industrial fans," *J. Sensors*, vol. 2018, pp. 1–15, 2018.
- [7] B. Bagheri, H. Ahmadi, and R. Labbafi, "Implementing discrete wavelet transform and artificial neural networks for acoustic condition monitoring of gearbox," *Elixir Mech. Eng.*, vol. 35, pp. 2909–2911, 2011.
- [8] M. Entezami et al., "Acoustic analysis techniques for condition monitoring of roller bearings," in *Proc. 6th IET Conf. on Railway Condition Monit.*, 2014, pp. 1–8.
- [9] R. B. Randall and J. Antoni, "Rolling element bearing diagnostics—A tutorial," *Mech. Syst. Signal Process.*, vol. 25, no. 2, pp. 485–520, 2011.
- [10] N. Nikolaou and I. Antoniadis, "Rolling element bearing fault diagnosis using wavelet packets," *NDT E. Int.*, vol. 35, no. 3, pp. 197–205, 2002.
- [11] A. J. Torabi, M. J. Er, X. Li, B. S. Lim, and G. O. Peen, "Application of clustering methods for online tool condition monitoring and fault diagnosis in high-speed milling processes," *IEEE Syst. J.*, vol. 10, no. 2, pp. 721–732, Jun. 2016.
- [12] C. Cooper et al., "Convolutional neural network-based tool condition monitoring in vertical milling operations using acoustic signals," *Procedia Manuf.*, vol. 49, pp. 105–111, 2020.
- [13] M.-K. Liu, Y.-H. Tseng, and M.-Q. Tran, "Tool wear monitoring and prediction based on sound signal," *Int. J. Adv. Manuf. Technol.*, vol. 103, pp. 3361–3373, 2019.
- [14] M. Crocco, M. Cristani, A. Trucco, and V. Murino, "Audio surveillance: A systematic review," *ACM Comput. Surv.*, vol. 48, no. 4, pp. 1–46, 2016.
- [15] J. Wu, F. Yang, and W. Hu, "Unsupervised anomalous sound detection for industrial monitoring based on arface classifier and Gaussian mixture model," *Appl. Acoust.*, vol. 203, 2023, Art. no. 109188.
- [16] E. Di Fiore, A. Ferraro, A. Galli, V. Moscato, and G. Sperli, "An anomalous sound detection methodology for predictive maintenance," *Expert Syst. Appl.*, vol. 209, 2022, Art. no. 118324.
- [17] K. Suefusa, T. Nishida, H. Purohit, R. Tanabe, T. Endo, and Y. Kawaguchi, "Anomalous sound detection based on interpolation deep neural network," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2020, pp. 271–275.
- [18] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume II: Detection Theory*, 1st ed. Upper Saddle River, NJ, USA: Prentice Hall, 1998.
- [19] L. L. Scharf and C. Demeure, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Reading, MA, USA: Addison-Wesley Publishing Company, 1991.
- [20] M. Christensen and A. Jakobsson, *Multi-Pitch Estimation, Ser. Synthesis Lectures on Speech and Audio Processing*. San Rafael, CA, USA: Morgan and Claypool, 2009.
- [21] F. F. Li and T. J. Cox, *Digital Signal Processing in Audio and Acoustical Engineering*. Boca Raton, FL, USA: CRC, 2019.
- [22] D. W. Hosmer Jr, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, Hoboken, NJ, USA: Wiley, 2013.