Iranian Society of Acoustics and Vibration
The 14th ISAV2024 International Conference on Acoustics and Vibration
11-12 Dec 2024 Karaj - Iran
Kharazmi University

Acoustic Signal Processing Revisited: Exploring Hilbert-Huang Transform and the Challenge of Cross-Term Errors

Javad Isavand[a], Andrew Peplow[b], Jihong Yan [a*]

[a] *School of Mechatronics Engineering, Harbin Institute of Technology, Harbin, China.*

[b] *Principal Acoustic Consultant, Hawkins & Associates, Cambridge, England.*

* *Corresponding author e-mail: jyan@hit.edu.cn*

Abstract

Numerous comprehensive studies have investigated the effectiveness of joint time-frequency transformations for analyzing non-stationary time series data. The primary objective of these investigations is to improve accuracy in both time and frequency domains, which is crucial for a wide range of applications. These methods have demonstrated significant efficacy in various research fields, particularly in audio and acoustic signal processing. Despite their success, several challenges persist, such as the occurrence of cross-term errors. This paper presents a comparative analysis of two prominent time-frequency analysis methods: The Short-Time Fourier Transform (STFT) and the Hilbert-Huang Transform (HHT). We employ three acoustic signal types, drawn from industrial applications, music, and audio processing, to evaluate the performance of each method. Our findings reveal that the STFT outperforms the HHT, providing more accurate results across all tested signal types. Notably, the HHT introduces a higher risk of cross-term errors, which can compromise the clarity and usability of the analyzed data.

Keywords: Acoustic Signal Processing; Hilbert-Huang Transform; Cross-term Error.

---

1. Introduction

In diverse domains such as medicine, industry, and music, our environment is filled with a plethora of non-stationary time series and acoustic signals. These signals display a range of qualitative characteristics, ranging from pleasant to unpleasant, yet each can include valuable informational content. Notably, a significant proportion of acoustic signals manifest as non-stationary time series, inherently characterized by time-varying components. Hence, analyses that treat these signals solely in either the time or frequency domain may not fully capture all essential information contained within them. Consequently, considerable efforts have been dedicated to conducting research, with the objective of formulating a diverse range of time-frequency methods. These methodologies offer a more precise and comprehensive analysis of these signals. Categories of these methods are the Time-Frequency Distribution (TFD), Wavelet Transform (WT), Synchro squeezed approaches, and Data-driven mode decomposition methods,[1], [2].

Time-Frequency Distribution (TFD) methods can be classified into two principal categories: linear and nonlinear techniques. Among the linear approaches, the Short-Time Fourier Transform (STFT) stands out as the most straightforward method. This technique basically segments the time series into short intervals through a sliding window process, followed by the computation of the Fourier transform for each segment, [3].

$$X(t,\omega) = \int_{-\infty}^{\infty} w(t-\tau)x(\tau)e^{-j\omega t}d\tau \qquad (1)$$

$$S_{STFT}(t,\omega) = \frac{1}{2\pi}|X(t,\omega)|^2 \qquad (2)$$

In contrast, Cohen introduced a nonlinear TFD method known as Cohen's Bilinear Distribution (BD) which calculates the Fourier transform of a time-varying covariance function using a Kernel function [4]. Subsequently, diverse BD techniques have emerged, utilizing distinct kernel functions to achieve a high-resolution energy distribution. Notable instances of such techniques include the Wigner-Ville Distribution, Choi-William Kernel, and Cone Kernel, among others [1], [5]. Beyond BD methods, the Gabor Transform constitutes another approach dedicated to transmuting time series into the time-frequency domain. This method involves the application of a sliding Gaussian function to the time series, followed by the Fourier transform on each segmented portion [4]. Furthermore, an enhanced iteration of the Gabor Transform is the S-transform, also known as the Stockwell Transform. The S-transform, rooted in the STFT, employs a frequency-dependent Gaussian window, thereby augmenting its analytical performance [6].

Instead of applying a uniform window through each time intervals, the Wavelet Transform (WT) employs a diverse set of filters with varying bandwidths that span from the lowest to the highest frequencies. Extensive research has been conducted on wavelet transform techniques to enhance its accuracy. Amongst others, the techniques have been improved by utilizing a variety of “mother wavelet” forms and innovative approaches [7], [8], [9], [10].

Also, in recent years, there has been significant attention given to data-driven mode decomposition methods. These methods are designed to effectively break down a given signal into multiple zero-mean Intrinsic Mode Functions (IMFs). The Hilbert-Huang Transform (HHT)[11] is one such method that is commonly used in conjunction with Empirical Mode Decomposition (EMD) [12].

$$x[n] = \sum_{i=1}^{L} IMF[n] + Res[n] \qquad (3)$$

Where L is the number of IMFs and $Res[n]$ is the residue. EMD have successfully been used in the analysis of non-stationary signals.

Other related approaches, such as Bivariate Empirical Mode Decomposition (BEMD), Multivariate Empirical Mode Decomposition (MEMD), Ensemble and Empirical Mode Decomposition (EEMD), have also gained traction and found applications across a wide range of research areas [13], [14]. Besides EMD, the Variational Mode Decomposition (VMD)[15], Fourier Mode Decomposition (FMD)[16], and Dynamic Mode Decomposition (DMD) [17] have also been extensively investigated and implemented in various studies. These alternative methods offer additional opportunities for analysis and study in the field of mode decomposition.

In order to improve the time and frequency resolutions, authors have also explored the potential of Reassignment Methods (RM) and Synchrosqueezed Transform techniques (SST), as alternative approaches to overcome the limitations of existing techniques, [8]. A noteworthy example is the Fourier Synchro-squeezing Transform (FSST)[18], which involves performing the Short-Time Fourier Transform (STFT) of the signal using a windowing technique. Subsequently, the Instantaneous Frequency (IF) is estimated by taking the derivative of the transformed signal with respect to time. Finally, the FSST coefficients are obtained by selecting the STFT coefficients that correspond to the calculated IFs. By incorporating the concept of instantaneous frequency into the inherent capabilities of wavelet methods, the Wavelet Synchro-Squeezing Transform (WSST) achieves enhanced accuracy and robustness, [9]. Despite the successful implementation and improved accuracy demonstrated by various transforming techniques across different research domains, their usage encounters challenges such as mode mixing, cross-term errors, and high-frequency harmonic distortion, [19], [20], [21].

Joint time-frequency analysis techniques have been widely implemented in the field of audio and acoustic signal processing such as acoustic-based condition monitoring [22], signal identification [23] and reconstruction [24], music [25], and voice recognition [26]. Beside these areas, research on reducing the computational costs and data volume highlights remarkable challenges, especially when applying these methods in research areas such as Internet of Things (IoT), Cloud Computing, where vast amounts of data are encountered [27], [28]. To tackle this issue, research groups have turned to the Reduced-rank Spectral approach to reduce model complexity [29], the application of Sparse representations [30], [31], or reduced-order machine-learning-based techniques [32].

2. Case Study Example I: Acoustic Signal Processing for Condition Monitoring

In the context of Industry 4.0 and the increasing utilization of Internet of Things (IoT) and Industrial Cloud Computing (ICC) approaches, the use of acoustic signals in industrial condition monitoring has attracted significant attention from researchers, [22], [35]. Therefore, this section presents the concerns of the application of the Hilbert-Huang transforming technique to an industrial acoustic signal which is usually used for condition monitoring.

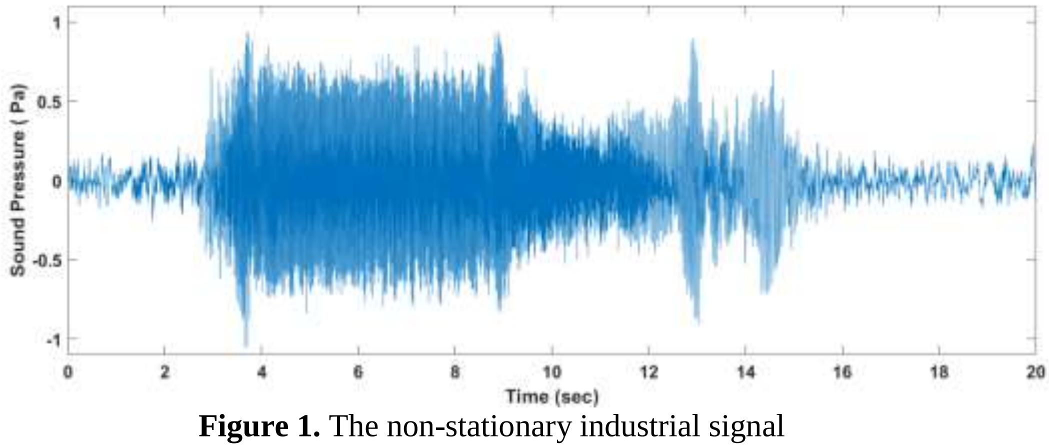
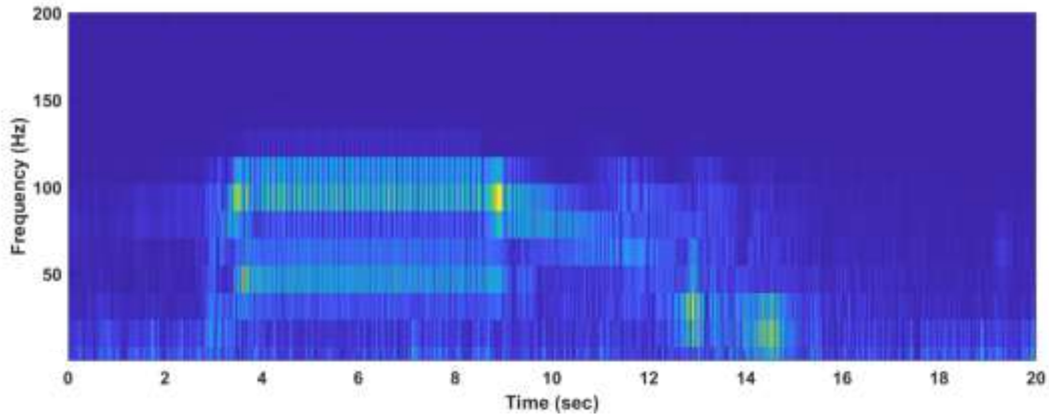


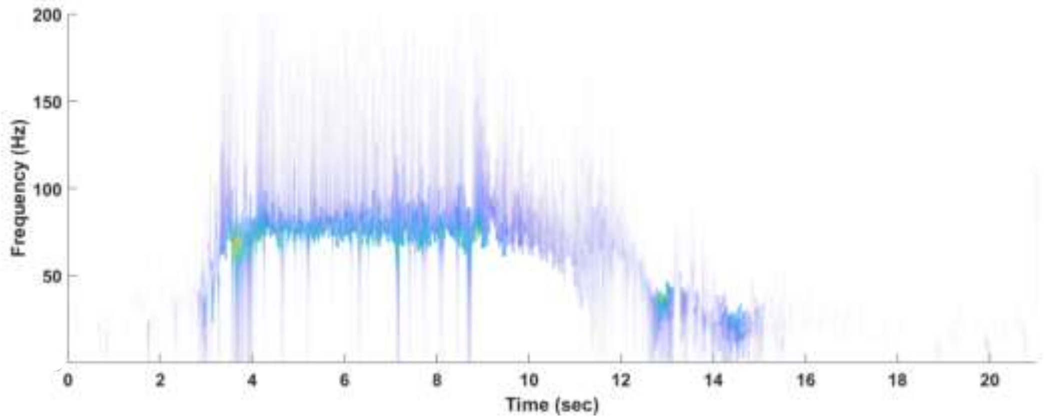
Figure 1. The non-stationary industrial signal

As rotary machines have become more complex, the balancing process has become crucial in condition monitoring to ensure reliable and safe machine operations. Imbalance is a common issue in flexible rotating machinery, which can lead to severe vibration and noise levels. However, an unbalanced object can induce significant unwanted deflection through resonant vibration at frequencies close to certain rotational speeds, known as critical speeds. This is particularly critical for flexible machines that typically operate at rotations above their critical speeds. To analyse such complicated time series during the run-up and shutdown stages, including critical speeds, researchers in condition monitoring can consider the vibration and/or acoustic responses of the rotary machine. In this research, the proposed method's effectiveness is evaluated by examining the signal displayed in Fig. 1 which was measured during the run-up and shutdown stages of a rotating machine with the sampling rate 2000 samples per second.

In contrast to the Short-Time Fourier Transform (STFT), the spectra generated by the implementation of the Hilbert-Huang Transform (HHT) are not accurate enough. Specifically, the diagrams reveal the presence of the second harmonic of the original signal when the rotational speed of the machine is 2970 rpm (equivalent to 49.5 Hz), with the second harmonic appearing at a frequency of 99 Hz. Notably, the HHT spectrum displays a cross-term error, wherein the amplitudes between 49.5 Hz and 99 Hz are larger than those associated with these frequencies.



a. The Short-Time Fourier Transform (STFT) Technique



b. The Hilbert-Huang Transform (HHT) Technique

Figure 2. The joint time-frequency domain using different techniques

3. Case Study Example II: Music Signal Processing

Musical signals, characterized by their non-stationary nature, exhibit complexity and temporal-frequency variations. Therefore, a comprehensive analysis of such signals in the time and frequency domains is essential, despite the inherent complexity associated with music. To compare the proposed technique with established methods, a non-stationary time series (a 15-second violin music piece recorded in a defused field environment at a sampling rate of 44100 samples per second) is selected for analysis, considering the potential presence of echo effects and cross-term errors, as shown in Fig.3. Figure 4 illustrates the results obtained by employing various methods. One of the main challenges lies in implementing the HHT and STFT methods, which demonstrates insufficient precision. It is anticipated that the spectrum may display a single frequency at a given time as a harmonic sound, commonly known as a “musical note” by musicians. Since a violin-generated sound can be treated as a pure harmonic signal, it is expected to able to clearly observe such a characteristic, [36].

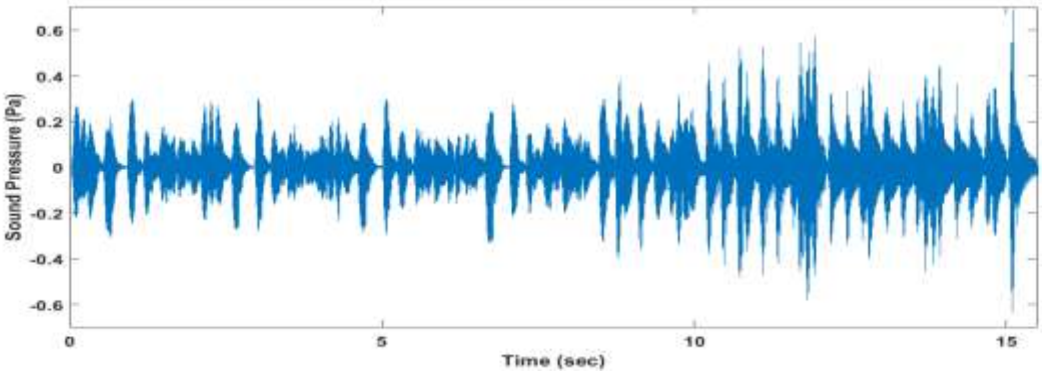
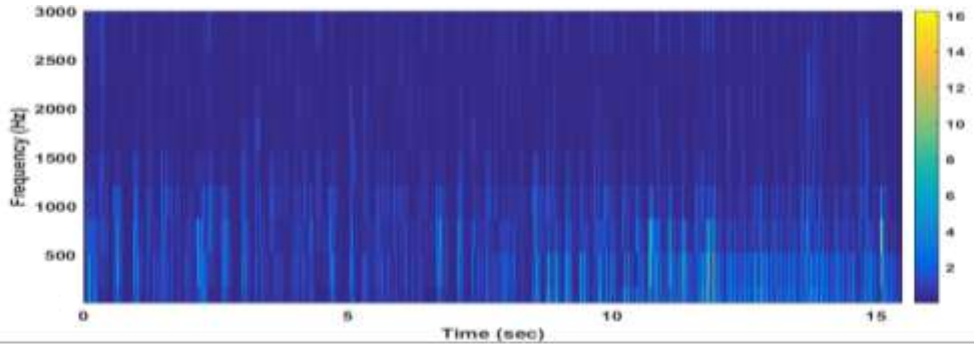
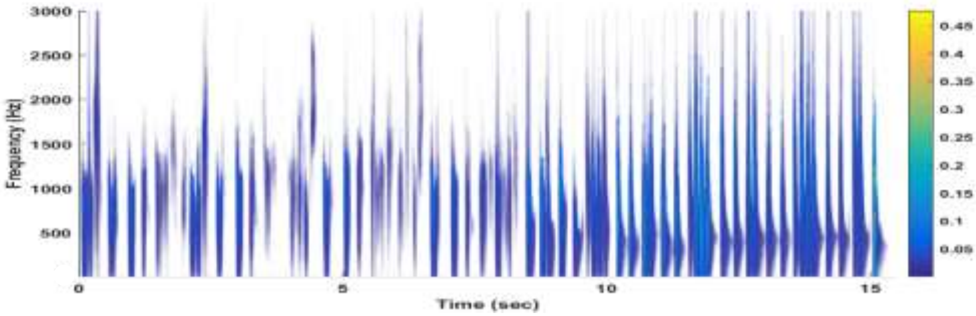


Figure 3. The non-stationary music signal



a. The Short-Time Fourier Transform (STFT) Technique



b. The Hilbert-Huang Transform (HHT) Technique

Figure 4. The joint time-frequency domain using different techniques

In contrast, utilizing the HHT and STFT methods, as depicted in Figure 4, analysis becomes more intricate and challenging due to the presence of harmonics and cross-term errors in the joint time-frequency spectrum. Certainly, analysing a piece of music using the joint time-frequency domain requires specialized knowledge in related research areas. Readers are encouraged to refer to the references for a more in-depth analysis of the joint time-frequency spectrum of music [36].

4. Case Study Example III: Audio Signal Processing

Despite extensive research on audio signal processing techniques, achieving high-quality audio signals remains a continuous challenge, particularly in terms of noise cancellation. A specific challenge arises when the original signal and the noise contain overlapping frequencies, known as Frequency Overlap [38]. To tackle this issue, a joint time-frequency analysis emerges as a potential solution. In this section, these methods are applied to a "noisy audio signal, as shown in Fig.5 which was recorded with sampling rate of 44100 samples per second. The main challenge with this signal is the presence of frequency overlap, which adds complexity to the analysis and subsequent noise cancellation process.

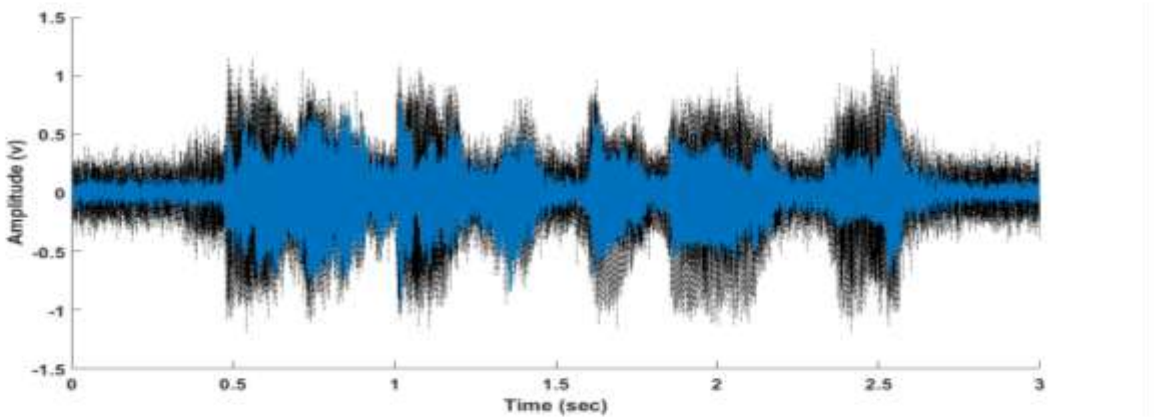
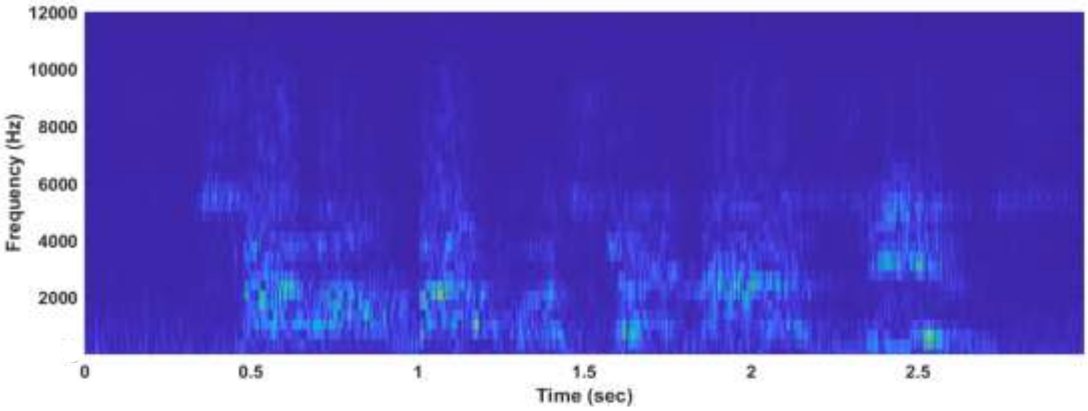
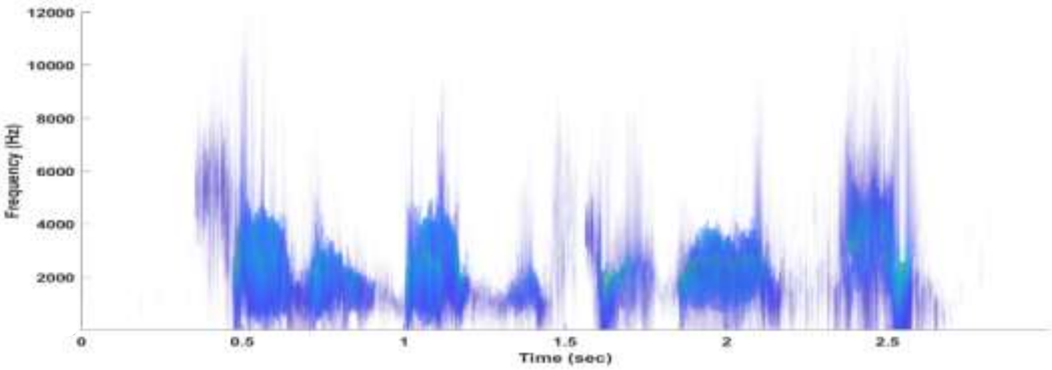


Figure 5. The non-stationary original and noisy audio signals

Figure 6 showcases the time-frequency analysis of the given audio signal using the Hilbert-Huang transform, Short-Time Fourier Transform, and the proposed method. The Hilbert-Huang transform exhibits ambiguous areas, whereas the proposed method accurately identifies frequencies. Using these calculated frequencies, a simple bandpass filter can be designed to demonstrate the effectiveness of this method.



a. The Short-Time Fourier Transform (STFT) Technique



b. The Hilbert-Huang Transform (HHT) Technique

Figure 6. The joint time-frequency domain using different techniques

When it comes to noise cancellation through filter design, the method's capacity to target specific frequencies, as illustrated in Figure 6, can offer a potential solution. This capability can also be considered an advantage of the proposed method. Readers interested in detailed designs for audio filtering and noise cancellation are strongly encouraged to consult the referenced paper [37].

5. Conclusion

While time-frequency transformation methods have been extensively researched and applied in various domains, including audio and acoustic signal processing, achieving higher accuracy remains a significant challenge. In response to this need, this study presents a comparative analysis of two well-established methods: the Short-Time Fourier Transform (STFT) and the Hilbert-Huang Transform (HHT). The analysis uses three different acoustic signals from the fields of industrial sound, music, and audio signal processing. The results demonstrate that the STFT technique demonstrates notably superior performance in this comparison. This study not only highlights the strengths of STFT in practical applications but also underscores the limitations of HHT, contributing valuable insights for researchers and practitioners in the field of acoustic signal processing.

REFERENCES

1. A. Akan and O. K. Cura, "Time–frequency signal processing: Today and future," Digital Signal Processing, vol. 119, p. 103216, 2021, doi: 10.1016/j.dsp.2021.103216.
2. X. Zhu, B. Li, K. Yang, Z. Zhang, and W. Li, "Parameter analysis of chirplet transform and high-resolution time-frequency representation via chirplets combination," Signal Processing, vol. 205, p. 108824, 2023, doi: 10.1016/j.sigpro.2022.108824.
3. L. Lu, W.-X. Ren, and S.-D. Wang, "Fractional Fourier transform: Time-frequency representation and structural instantaneous frequency identification," Mechanical Systems and Signal Processing, vol. 178, p. 109305, 2022, doi: 10.1016/j.ymssp.2022.109305.
4. L. Cohen, "Time-frequency Analysis, 778 Prentice Hall PTR," Englewood Cliffs, NJ, 1995.
5. R. Salles, K. Belloze, F. Porto, P. H. Gonzalez, and E. Ogasawara, "Nonstationary time series transformation methods: An experimental review," Knowledge-Based Systems, vol. 164, pp. 274–291, 2019, doi: 10.1016/j.knosys.2018.10.041.
6. P.-P. Yuan, J. Zhang, J.-Q. Feng, H.-H. Wang, W.-X. Ren, and C. Wang, "An improved time-frequency analysis method for structural instantaneous frequency identification based on generalized S-transform and synchroextracting transform," Engineering Structures, vol. 252, p. 113657, 2022, doi: 10.1016/j.engstruct.2021.113657.
7. H. Dong, G. Yu, T. Lin, and Y. Li, "An energy-concentrated wavelet transform for time-frequency analysis of transient signal," Signal Processing, vol. 206, p. 108934, 2023, doi: 10.1016/j.sigpro.2023.108934.
8. F. Auger et al., "Time-frequency reassignment and synchrosqueezing: An overview," IEEE Signal Processing Magazine, vol. 30, no. 6, pp. 32–41, 2013, doi: 10.1109/MSP.2013.2265316.
9. I. Daubechies, J. Lu, and H.-T. Wu, "Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool," Applied and computational harmonic analysis, vol. 30, no. 2, pp. 243–261, 2011, doi: 10.1016/j.acha.2010.08.002.
10. W. Li, F. Auger, Z. Zhang, and X. Zhu, "Newton time-extracting wavelet transform: An effective tool for characterizing frequency-varying signals with weakly-separated components and theoretical analysis," Signal Processing, vol. 209, p. 109017, 2023, doi: 10.1016/j.sigpro.2023.109017.
11. U. B. de Souza, J. P. L. Escola, and L. da Cunha Brito, "A survey on Hilbert-Huang transform: Evolution, challenges and solutions," Digital Signal Processing, vol. 120, p. 103292, 2022, doi: 10.1016/j.dsp.2021.103292.
12. G. Rilling, P. Flandrin, and P. Goncalves, "On empirical mode decomposition and its algorithms," presented at the IEEE-EURASIP workshop on nonlinear signal and image processing, Grado: IEEE, 2003, pp. 8–11.
13. W. Zhou, Z. Feng, Y. Xu, X. Wang, and H. Lv, "Empirical Fourier decomposition: An accurate signal decomposition method for nonlinear and non-stationary time series analysis," Mechanical Systems and Signal Processing, vol. 163, p. 108155, 2022, doi: 10.1016/j.ymssp.2021.108155.
14. J. Gu and Y. Peng, "An improved complementary ensemble empirical mode decomposition method and its application in rolling bearing fault diagnosis," Digital Signal Processing, vol. 113, p. 103050, 2021, doi: 10.1016/j.dsp.2021.103050.
15. K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," IEEE transactions on signal processing, vol. 62, no. 3, pp. 531–544, 2013, doi: 10.1109/TSP.2013.2288675.
16. P. Singh, S. D. Joshi, R. K. Patney, and K. Saha, "The Fourier decomposition method for nonlinear and non-stationary time series analysis," Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, vol. 473, no. 2199, p. 20160871, 2017, doi: 10.1098/rspa.2016.0871.

17. S. L. Brunton and J. N. Kutz, Data-driven science and engineering: Machine learning, dynamical systems, and control. Cambridge University Press, 2019.
18. T. Oberlin, S. Meignen, and V. Perrier, "The Fourier-based synchrosqueezing transform," presented at the 2014 IEEE international conference on acoustics, speech and signal processing (ICASSP), IEEE, 2014, pp. 315–319.
19. L. Tang, X.-Q. Shang, T.-L. Huang, N.-B. Wang, and W.-X. Ren, "An improved local maximum synchrosqueezing transform with adaptive window width for instantaneous frequency identification of time-varying structures," Engineering Structures, vol. 292, p. 116543, 2023, doi: 10.1016/j.engstruct.2023.116543.
20. J. Li, H. Wang, P. He, S. M. Abdullahi, and B. Li, "Long-term variable Q transform: A novel time-frequency transform algorithm for synthetic speech detection," Digital Signal Processing, vol. 120, p. 103256, 2022, doi: 10.1016/j.dsp.2021.103256.
21. A. K. Choudhary, P. Kumar, and S. K. Verma, "An appropriate discrete-transformation technique for order reduction methodology," Array, vol. 14, p. 100155, 2022, doi: 10.1016/j.array.2022.100155.
22. J. Isavand, A. Kasaei, A. Peplow, B. Afzali, and E. Shirzadi, "Comparison of vibration and acoustic responses in a rotary machine balancing process," Applied Acoustics, vol. 164, p. 107258, 2020, doi: 10.1016/j.apacoust.2020.107258.
23. D. Wu, Z. Yang, Y. Ruan, and X. Chen, "Blind single-channel lamb wave mode separation using independent component analysis on time-frequency signal representation," Applied Acoustics, vol. 216, p. 109810, 2024, doi: 10.1016/j.apacoust.2023.109810.
24. L. Geng, G. Zhang, L. Yu, J. Ma, C.-L. Shao, and F. Xie, "Reconstruction and analysis of a non-stationary sound field in a uniformly flow medium," Applied Acoustics, vol. 201, p. 109111, 2022, doi: 10.1016/j.apacoust.2022.109111.
25. Q. Meng, G. Liu, L. Tian, M. Zeng, X. Lu, and J. Yan, "An improved vocoder algorithm based on music harmonics and time sampling," Applied Acoustics, vol. 205, p. 109288, 2023, doi: 10.1016/j.apacoust.2023.109288.
26. X. Zhang, X. Zhu, C. Zhou, Z. Tao, and H. Zhao, "Pathological voice classification based on the features of an asymmetric fluid–structure interaction vocal cord model," Applied Acoustics, vol. 207, p. 109348, 2023, doi: 10.1016/j.apacoust.2023.109348.
27. A. Verma, A. Goyal, S. Kumara, and T. Kurfess, "Edge-cloud computing performance benchmarking for IoT based machinery vibration monitoring," Manufacturing Letters, vol. 27, pp. 39–41, 2021, doi: 10.1016/j.mfglet.2020.12.004.
28. R. Das and M. M. Inuwa, "A review on fog computing: issues, characteristics, challenges, and potential applications," Telematics and Informatics Reports, p. 100049, 2023, doi: 10.1016/j.teler.2023.100049.
29. A. Fradi and K. Daoudi, "Reduced-rank spectral mixtures Gaussian processes for probabilistic time–frequency representations," Signal Processing, vol. 218, p. 109355, 2024.
30. A. Pinkus, "Sparse representations and approximation theory," Journal of Approximation Theory, vol. 163, no. 3, pp. 388–412, 2011, doi: 10.1016/j.jat.2010.10.007.
31. X. Chen, H. Chen, Y. Hu, Y. Xie, and S. Wang, "A Sparse Time-Frequency Reconstruction Approach from the synchroextracting domain," Signal Processing, p. 109517, 2024.
32. J. Isavand, A. Kasaei, A. Peplow, and J. Yan, "Enhancing response estimation and system identification in structural health monitoring through data-driven approaches," Building Acoustics, p. 1351010X231219662, 2024, doi: 10.1177/1351010X231219662.
33. N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete cosine transform," IEEE transactions on Computers, vol. 100, no. 1, pp. 90–93, 1974, doi: 10.1109/T-C.1974.223784.
34. J. Isavand, A. Kasaei, A. Peplow, X. Wang, and J. Yan, "A reduced-order machine-learning-based method for fault recognition in tool condition monitoring," Measurement, vol. 224, p. 113906, 2024, doi: 10.1016/j.measurement.2023.113906.
35. A. Peplow, J. Isavand, A. Kasaei, B. Afzali, and D. Bard-Hagberg, "A speed-variant balancing method for flexible rotary machines based on acoustic responses," Sustainability, vol. 13, no. 13, p. 7237, 2021, doi: 10.3390/su13137237.
36. K. Chithra and M. Sinith, "A Comprehensive Study of Time-Frequency Analysis of Musical Signals," presented at the 2023 International Conference on Signal Processing, Computation, Electronics, Power and Telecommunication (IConSCEPT), IEEE, 2023, pp. 1–5.
37. Y. D. Mistry, G. K. Birajdar, and A. M. Khodke, "Time-frequency visual representation and texture features for audio applications: a comprehensive review, recent trends, and challenges," Multimedia Tools and Applications, pp. 1–35, 2023.