INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2024, VOL. 70, NO. 2, PP. 323-329
Manuscript received February 10, 2024; revised June, 2024. doi: 10.24425/ijet.2024.149548

EMD-based time-frequency analysis methods of audio signals

Marcin Lewandowski, and Qizhang Deng

Abstract—**Using appropriate signal processing tools to analyze time series data accurately is essential for correctly interpreting the underlying processes. Commonly employed methods include kernel-based transforms that utilize base functions and modifications to depict time series data. This paper refers to the analysis of audio data using two such transforms: the Fourier transform and the wavelet transform, both based on assumptions regarding the signal's linearity and stationarity. However, in audio engineering, these assumptions often do not hold as the statistical characteristics of most audio signals vary over time, making them unsuitable for treatment as outputs from a Linear Time-Invariant (LTI) system. Consequently, more recent methods have shifted towards breaking down signals into various modes in an adaptive, data-specific manner, potentially offering benefits over traditional kernel-based methods. Techniques like empirical mode decomposition and Holo-Hilbert Spectral Analysis are examples of this. The effectiveness of these methods was tested through simulations using speech signals for both kernel-based and adaptive decomposition methods, demonstrating that these adaptive methods are effective for analyzing audio data that is both nonstationary and an output of the nonlinear system.**

Keywords—**empirical mode decomposition; nonstationary audio data; time-frequency analysis**

I. Introduction

TIME-FREQUENCY analysis is a crucial tool for data processing, beneficial for analyzing signals and extracting significant information. This technique breaks down a signal into vital time-varying features that reveal the system's inherent behavior, help set parameters for system modeling, or verify the accuracy of an existing model. Fourier spectral analysis has traditionally been a staple for mapping overall energy-frequency distributions. The spectrogram, which applies Fourier spectral analysis within a constrained time window, is fundamental to achieving time-frequency distribution. The Fourier transform applies under general conditions [1]; however, Fourier spectral analysis yields meaningful results only when the system is linear and the data is strictly periodic or stationary [2]. Typically, the signals used for analysis, whether derived from physical measurements or computational models, tend to be nonstationary, represent nonlinear processes, and the total signal span is too short for proper analysis [2]. The stationarity requirement is, for all t:

$$\begin{gathered} E(|X(t)^2|) < \infty, \\ E(|X(t)|) = m, \\ C\big(X(t_1), X(t_2)\big) = C\big(X(t_1+\tau), X(t_2+\tau)\big) = C(t_1 - t_2) \end{gathered} \tag{1}$$

in which $E(\cdot)$ is the expected value defined as an ensemble average and $C(\cdot)$ is a covariance function. A less rigorous definition is for piecewise stationarity when a signal is stationary within a limited time span [2]. In practical scenarios, real-world data, such as audio signals, is continuously collected and analyzed over a finite and limited duration. As a result, it is inherently nonstationary [3]. Nonlinearity is a fundamental characteristic of many natural phenomena, often further complicated by numerical models, errors, and imperfections in measurement tools, detectors, and data acquisition systems. While linear systems can approximate real-world processes, these complexities may render the resulting data nonlinear.

Consequently, Fourier-based spectral analysis methods have limited applicability and can produce misleading results when they approximate nonstationary and nonlinear data with simpler models. This limitation stems from two main issues. First, the Fourier spectrum requires numerous additional harmonic components to simulate nonstationary data that varies globally. While these may be mathematically valid, they often lack physical relevance. Second, Fourier spectral analysis relies on a linear combination of trigonometric functions or kernels [2]. Thus, if the data's form deviates from simple sine or cosine waves, the resulting spectrum will include harmonics that do not necessarily correspond to the signal's energy-frequency distribution.

Audio data, such as music or speech, changes over time and thus cannot be described by a mathematical expression or accurately represented by a linear system. As a result, modern methods for analyzing audio data often involve breaking down a signal into various modes that adapt to the data, offering potential improvements over traditional kernel-based transforms. One such technique is empirical mode decomposition (EMD) [2], which has proven effective across numerous signal processing applications. EMD-based approaches decompose a signal into intrinsic mode functions (IMFs), derived by identifying local extremes and creating upper and lower envelopes. These IMFs, which can have well-defined Hilbert transforms, allow for the computation of instantaneous frequencies, facilitating the precise localization of events within the signal in both time and frequency domains [2].

First Author is with Warsaw University of Technology (e-mail: marcin.lewandowski@pw.edu.pl).

Second Author is with University of New South Wales Sydney (e-mail: z5491364@ad.unsw.edu.au).

© The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0, https://creativecommons.org/licenses/by/4.0/), which permits use, distribution, and reproduction in any medium, provided that the Article is properly cited.

Despite its effectiveness, the standard EMD method has some drawbacks, such as mode-mixing, noise sensitivity, and data sampling issues. However, modifications to the EMD method have been implemented to address these issues.

II. Review of Nonstationary Data Processing Methods

Methods for processing nonstationary data generally fall into kernel-based and data-dependent (adaptive) categories. Kernel-based methods utilize a predetermined basis, such as trigonometric functions in Fourier analysis or specific wavelet functions in wavelet analysis, and operate under the assumption that the signal being analyzed comprises these basis functions. On the other hand, data-dependent methods break down a signal into a collection of intrinsic mode functions. This decomposition is achieved by directly extracting the energy linked to different intrinsic time scales without presupposing the signal's characteristics.

A. The spectrogram

The classical short-time Fourier transform (STFT) captures a time-varying signal representation by applying a finite window length to localize the signal in time, followed by performing a Fourier transform on each time segment [4]. STFT can be expressed as:

$$\mathrm{STFT}\{x(t)\}(t,\omega) = \int_{-\infty}^{\infty} x(t)\, h^*(t-\tau) e^{-i\omega t}, \tag{2}$$

where h^* is a specific window function. Because it is based on Fourier analysis with a fixed window size, the data is assumed to be piecewise stationary, leading to a consistent time-frequency resolution throughout the analysis. Accurately localizing an event in time requires a narrow window width, which leads to diminished frequency resolution and vice-versa.

B. The wavelet analysis

The wavelet transform [5] creates a signal representation using a variable time window that adjusts the scale of the basic wavelet function, thereby providing a multiscale representation of the signal. The wavelet transform is generally defined as follows:

$$\mathrm{W}(s,u) = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{s}} \psi^*\left(\frac{t-u}{s}\right), \tag{3}$$

in which $\psi^*(\cdot)$ is the basic wavelet function that satisfies certain general conditions, namely, it is orthonormal, and its mean value is zero, u is a translation of the origin (which gives a temporal time location of an event), and $1/\sqrt{s}$ gives a frequency scale [2]. Wavelet analysis offers uniform resolution across all scales, which is a significant benefit compared to the STFT. The time-frequency resolution grids of both the wavelet and STFT transforms are illustrated in Fig. 1 below.

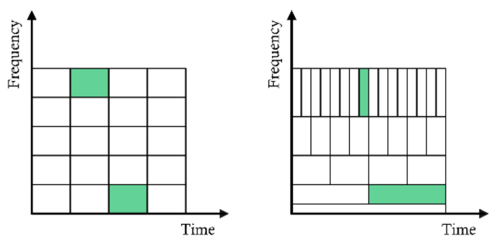

Fig. 1. Time-frequency resolution grid for STFT (left panel) and wavelet transform (right panel).

C. Empirical mode decomposition

Empirical mode decomposition is an adaptive method for time-frequency analysis that breaks down signals into intrinsic mode functions (IMFs). Each IMF adheres to two criteria: the number of extrema and zero-crossings must either be equal or differ by no more than one, and the average value between the envelope formed by local maxima and the envelope formed by local minima must be zero [2]. The determination of each IMF component involves subtracting the mean of its envelope (calculated as the average of the upper and lower envelopes, illustrated in Fig. 2) from the signal itself:

$$u_i = s(t) - m_i, \tag{4}$$

where m_i is a signal's mean of the envelope.

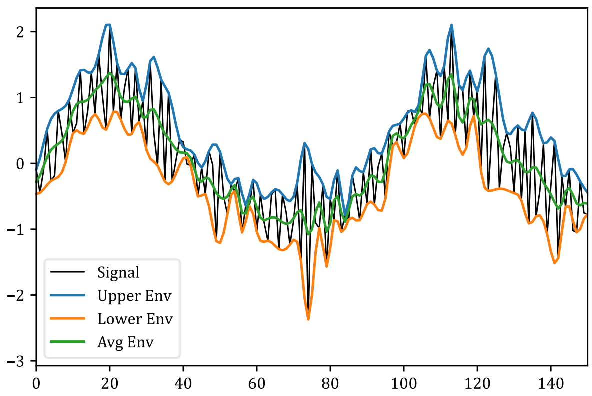

Fig. 2. Mean values of the upper and lower envelope of the input signal.

After the first IMF is calculated and it fulfills the previously mentioned requirements, it is subtracted from the input signal $s(t)$ (see Fig. 3) and the whole process is repeated with a residual signal r_1 in Eq. (5):

$$r_1 = s(t) - u_1 \tag{5}$$

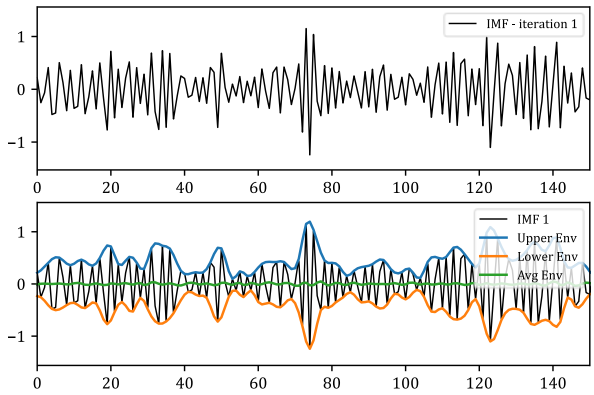

Fig. 3. First intrinsic mode function derived during the empirical mode decomposition analysis sifting phase.

Thus, the analyzed signal is decomposed into a set of n components (see Fig. 4 with six IMFs), and the last residual r_n:

$$s(t) = \sum_{j=1}^{n} u_j + r_n, \tag{6}$$

where r_n is dc offset or general trend in the analyzed signal.

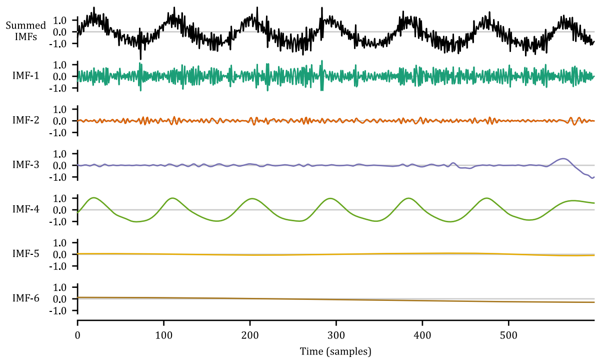

Fig. 4. Sample input signal decomposed into 6 IMFs.

D. Ensemble empirical mode decomposition

Traditional empirical mode decomposition (EMD) encounters several issues, notably mode-mixing, where similar frequency information appears across multiple intrinsic mode functions (IMFs). This is especially prominent when bursts of high-frequency content occur in the signal. An example of such a decomposition is illustrated in Fig. 5, where bursts of high frequencies are distributed from IMF1 to IMF3.

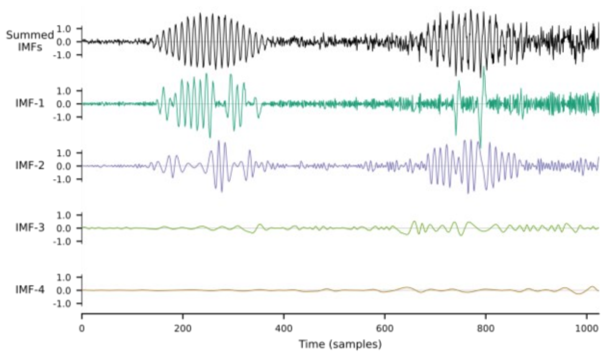

Fig. 5. Mode-mixing in EMD signal's decomposition.

Ensemble Empirical Mode Decomposition (EEMD) addresses the mode-mixing issue by averaging the decomposition outcomes across multiple noisy iterations of the original signal [6]. As depicted in Fig. 6, this approach significantly reduces mode-mixing, though noise persists in the intrinsic mode functions (IMFs), resulting in different IMF realizations for the same input signal.

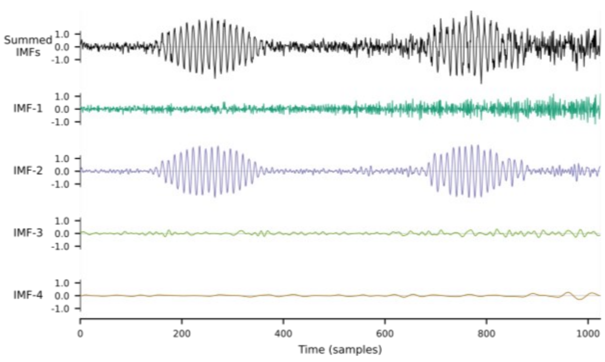

Fig. 6. Mode-mixing in EMD corrected with EEMD analysis.

To address these issues, the Complete Ensemble EMD with Adaptive Noise (CEEMDAN) method was introduced in [7] and later refined in [8]. This approach involves adding specific noise at each stage of the decomposition process and calculating a distinct residue to extract each IMF mode. However, despite these improvements, CEEMDAN still encounters challenges such as some degree of mode-mixing, unwanted residual noise, and the potential splitting of the signal across different modes [9].

E. Masked and iterated masked EMD

A novel method to enhance EMD decomposition, Masked EMD, was recently suggested in [10]. This technique involves introducing a masking signal into the input before sifting. Doing so reduces mode-mixing since the sift overlooks signal components that change slower than the frequency of the masking signal [9]. In this method, any signal content with frequencies significantly lower than the masking frequency is disregarded during that iteration and replaced by the mask. Eventually, the mask is removed, enabling accurate retrieval of intermittent signal activity [9]. While selecting the appropriate masking signals is an active area of research, it often remains a manual task, requiring expertise and potentially introducing subjective bias [10] - [14].

The problem of selecting an appropriate masking signal has been recently addressed with the introduction of iterated masked EMD (itEMD) [9]. This sifting technique automatically selects mask signal frequencies directly from the data, enabling it to detect oscillations and reduce mode-mixing without manually specifying masking signals. Figure 7 below illustrates the itEMD decomposition of the same signal depicted in Fig. 5 and Fig. 6.

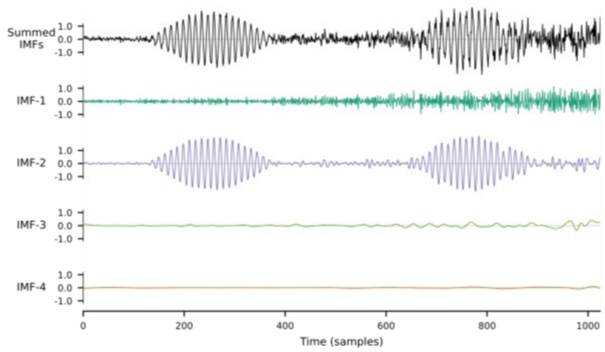

Fig. 7. EMD's mode-mixing problem resolved with itEMD method.

F. Hilbert-Huang transform

The Hilbert-Huang transform analyzes a signal's energy or power distribution across frequency and time. It employs Hilbert spectral analysis (HSA) [2], [15] to study the instantaneous frequency of the input signal as a function of time, making it particularly effective for analyzing nonlinear and nonstationary data. Instantaneous frequency is determined using the Hilbert transform, allowing the signal to be represented as follows:

$$s(t) = Real\left\{\sum_{j=1}^{n} a_j(t)\exp\left(i\int \omega_j(t)\,dt\right)\right\}, \tag{7}$$

where $a_j(t)$ and $\omega_j(t)$ are instantaneous amplitude and frequency, respectively. For this analysis to be valid, the input signal must fulfill specific prerequisites met by each of the intrinsic mode functions (IMFs) derived via EMD-based methods. The outcome is a frequency-time distribution of the signal's amplitude (or energy), enabling the detection of localized features. For illustration, Fig. 8 and Fig. 9 below display the Hilbert-Huang transform applied to a nonlinear and nonstationary signal.

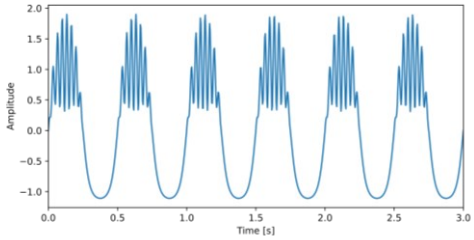

Fig. 8. Nonlinear and nonstationary synthetic signal example.

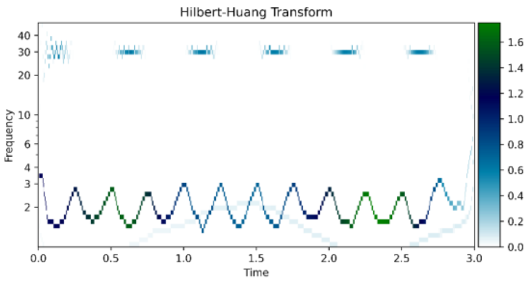

Fig. 9. Hilbert-Huang transform of signal shown in Fig. 8.

G. Other methods

The synchrosqueezed transform (SST) [16] enhances the continuous wavelet transform (CWT) by calculating instantaneous frequencies and using a reassignment algorithm to concentrate them at the core of the time-frequency area [17]. This process results in more precise time-frequency representations than those produced by the STFT and CWT, which can suffer from spectral smearing due to finite sampling lengths. Additionally, this "squeezing" effectively reduces the number of unnecessary wavelet coefficients, yielding a more sparse representation.

Singular Spectrum Analysis (SSA) [18] is another technique that decomposes a time series into a limited number of distinct and interpretable components, such as a slowly varying trend, oscillatory components, and unstructured noise. SSA utilizes a specifically constructed matrix's singular value decomposition (SVD) based on the time series data. It does not require the assumption of a parametric model or stationarity conditions for the time series, making SSA a model-free approach with broad applicability.

Other methods are Intrinsic-Time Scale Decomposition (ITD) [19] and Fourier Decomposition Method (FDM) [20].

H. Holo-Hilbert Spectral Analysis

The Holo-Hilbert Spectral Analysis (HHSA) proposed in [21] can be applied to overcome the limitations of the previous methods. HHSA is a nonlinear analysis tool that uses empirical mode decomposition to detect intrinsic amplitude modulations by representing data across multiple dimensions, including amplitude modulation, carrier, and time. The carrier frequency in HHSA aligns with the frequency dimension in traditional kernel-based spectral analyses. Unlike simple measurements of pairwise couplings, HHSA utilizes instantaneous frequency information to comprehensively capture the energy and content of all potential modulating and carrier frequencies in data derived from nonstationary and nonlinear processes. In essence, the HHSA method extends HHT with the second layer of sifting. An example of a HHSA analysis is presented with the synthetic signal shown in Fig. 10. The signal consists of slow 5Hz data with 0.5Hz amplitude modulation (red signal in Fig. 10) and fast 37Hz data modulated by the slow signal (blue signal in Fig. 10). Iterated masked EMD was used to decompose this nonstationary signal, and sifting process is showed in Fig. 11.

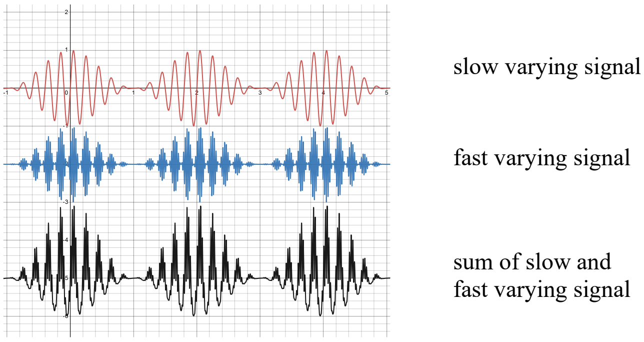

Fig. 10. Nonstationary signal with slow (red) and fast (blue) components with amplitude modulations (black plot is the sum).

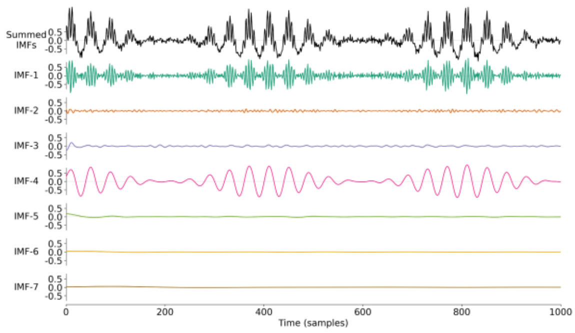

Fig. 11. itEMD first layer sifting process. IMF-1 and IMF-4 contain frequencies of 37Hz and 5Hz.

It turns out that each IMF has its modulations, which could be further investigated. In Fig. 12, the amplitude modulations in IMF-1 and IMF-4 are visible.

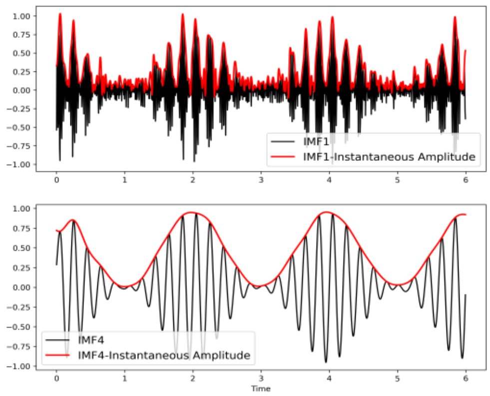

Fig. 12. Amplitude modulations in IMF-1 and IMF-4.

The second level of sifting on each IMF reveals all amplitude modulations and intrinsic couplings of components to understand the analyzed signal better. Figure 13 shows the result where the distribution of power versus frequency is on the left panel (carrier frequency spectrum), time-frequency analysis is in the middle, and holospectrum (amplitude modulation spectrum) is on the right panel. Two main input signal components (at 5Hz and 37Hz) are modulated with 0.5Hz and 5Hz, respectively, corresponding to the parameters of a synthetic signal in Fig. 10.

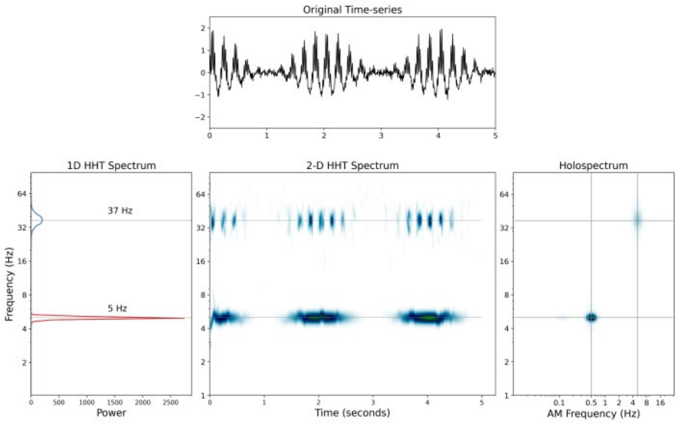

Fig. 13. Holo-Hilbert spectral analysis of nonstationary signal from Fig. 10.

A step further would be to show the time-varying HHSA analysis already done in [22]. Figure 14 illustrates the outcome of HHSA on synthesized data with time-varying coupling strength. A synthesized signal $x(t)$ with time-varying coupling strength from 0 to 1 is shown in panel A, where the modulated signal $x_{fA}(t)$ shows a power increase corresponding to the coupling strength. In panel B, the time-resolved power spectrum derived from the Hilbert-Huang Transform (HHT) is displayed, while panel C shows the amplitude spectrum of the envelope at four different time points obtained via Holo-Hilbert Spectral Analysis (HHSA). These demonstrate how HHSA tracks four distinct levels of coupling strength over time [22].

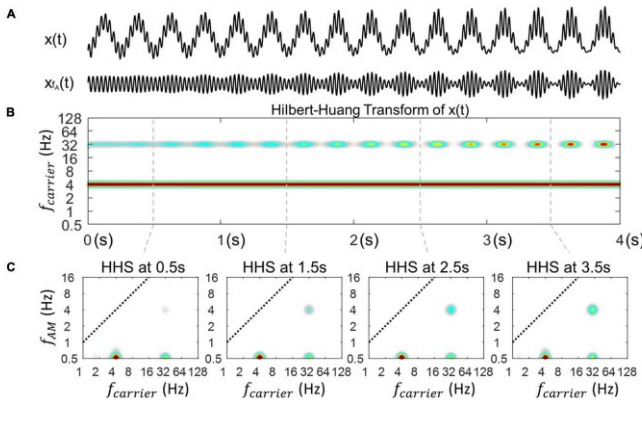

Fig. 14. Time-varying HHSA example analysis [22].

These methods have been successfully adopted in biomedical and neuroscience data analysis [27] - [29], and they can also be applied in the analysis of audio data.

III. Application of EMD-based Methods to Speech Data

The speech signal, shaped by a complex psycho-acoustic process developed over thousands of years of human evolution, contains information such as the speaker's age, height, emotion, accent, health, physiological disorders, and identity, fueling various subfields of Speech Processing [23], [24]. However, speech is inherently nonlinear and nonstationary [25], making the extraction of such detailed information a complex task [24], [26]. Methods based on Empirical Mode Decomposition (EMD) might uncover more insights into speech signals than traditional Fourier and wavelet-based methods.

The specific signal analyzed was the vowel 'A' spoken by a female and recorded at 48kHz sampling frequency. This speech signal's waveform and Fourier spectrum are displayed in Fig. 15. STFT and wavelet analysis are illustrated in Fig. 16 (left and right panels, respectively). Although these transforms clearly show the speech signal's fundamental frequency and harmonics at approximately 200, 400, 600, 800, 950, 1050, and 1100 Hz, they fail to reveal how these components interact or how their parameters evolve over time.

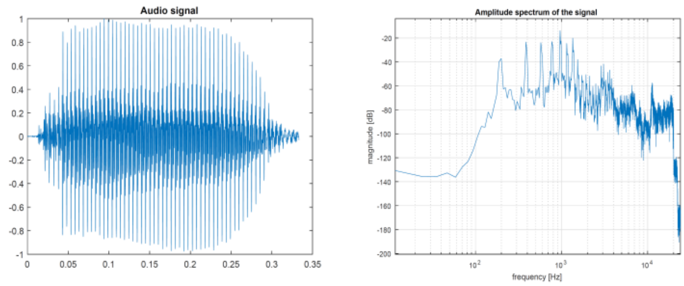

Fig. 15. Female voice letter ‘A’ - its waveform on the left and FFT on the right.

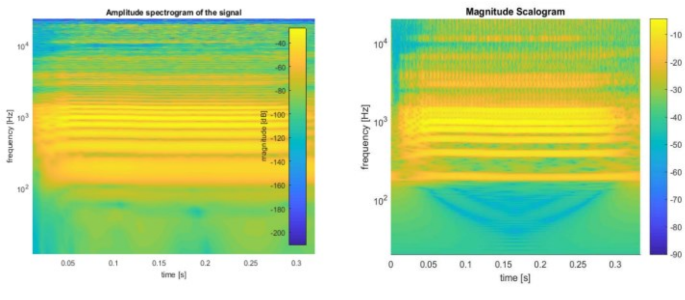

Fig. 16. STFT (left panel) and wavelet transform (right panel) of the speech signal (letter 'A').

Iterated masked EMD was used to decompose speech into 6 IMFs, the first 3 providing meaningful signal information. These 3 IMFs were analyzed with another level of decomposition to get the amplitude modulations hidden in the original data. Figures 18, 19, and 20 show the HHT spectrum of carrier signals in IMF1 - IMF3, respectively.

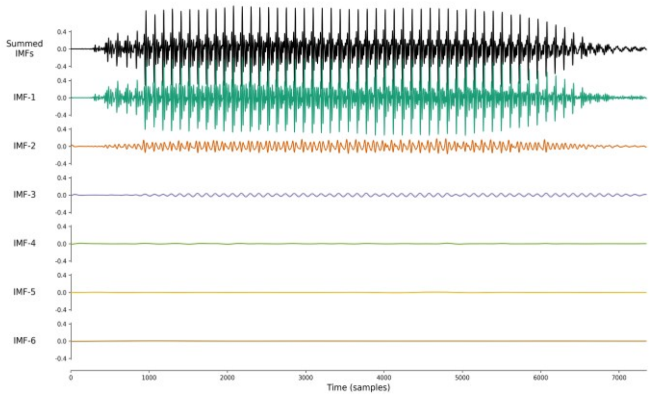

Fig. 17. Iterated Masked EMD sifting process where IMF-1 to IMF-3 contains meaningful speech signal information.

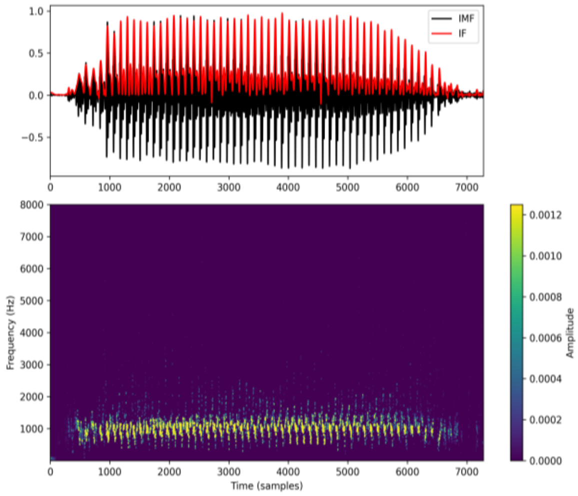

Fig. 18. HHT spectrum of carrier signal IMF-1 changing around 1kHz.

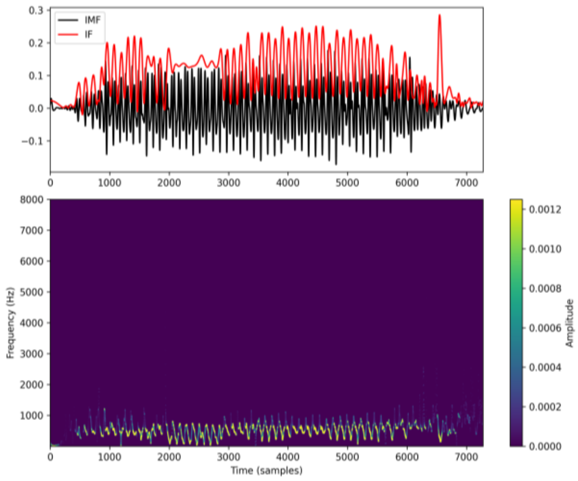

Fig. 19. HHT spectrum of carrier signal IMF-2 changing around 800Hz.

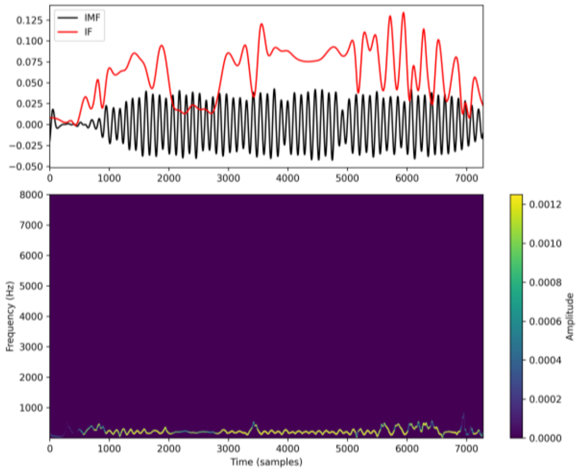

Fig. 20. HHT spectrum of carrier signal IMF-3 changing around 400Hz.

The holospectrum of the speech signal is shown in Fig. 21, where it is evident that the component changing around 1kHz with about 380Hz amplitude modulation is the essential part of letter A spoken by this particular person.

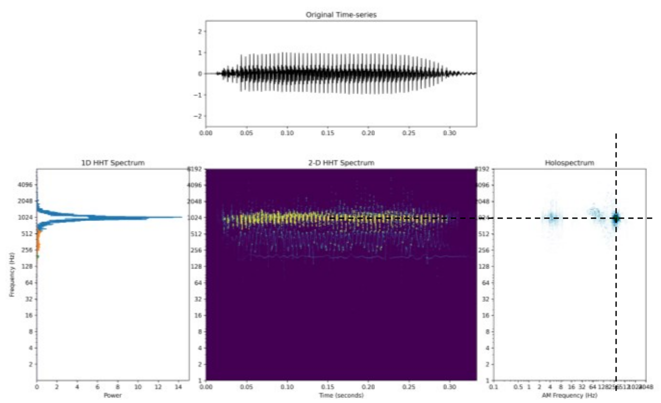

Fig. 21. Holo-Hilbert spectral analysis of speech signal.

IV. Conclusions

This paper provides an overview of data analysis and processing from a nonlinear and nonstationary perspective, contrasting with the traditional short-time linear and stationary analysis approach. Audio data, such as music or speech, is inherently nonstationary and cannot be fully captured by mathematical expressions or approximated through linear systems. This has led to the adoption AM-FM representations for audio data, which conceptualize audio as a collection of AM-FM signals. A speech signal is just one example of audio data that defies linear modeling. For instance, quantization noise is typically modeled as white noise added to a system, but in reality, it is a nonstationary, input-dependent, and nonlinear error signal.

Additionally, the time series of state variables in systems with feedback mechanisms - such as modulators, adaptive filters, prediction algorithms, and IIR filters - tend to alter their parameters over time. Methods like Empirical Mode Decomposition (EMD) are inherently empirical, and each technique discussed in the literature has limitations. As a result, time-frequency analysis techniques, such as the Holo-Hilbert Spectrum analysis method, remain a dynamic field of study.

REFERENCES

[1] S. Bochner, "Fourier Integrals: Introduction to the Theory of Fourier Integrals. By EC Titchmarsh. Oxford, Clarendon Press, 1937.", Science, t. 87, nr2260, s. 370–370, 1938. https://www.science.org/doi/abs/10.1126/science.87.2260.370.a

[2] N. E. Huang et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis", Proc. R. Soc. Lond. Ser. Math. Phys. Eng. Sci., t. 454, nr 1971, s. 903–995, mar. 1998, doi:10.1098/rspa.1998.0193. https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1998.0193

[3] P. J. Brockwell and R. A. Davis, Time series: theory and methods. Springer science & business media, 2009.

[4] J. Allen, "Short term spectral analysis, synthesis, and modification by discrete Fourier transform", IEEE Trans. Acoust. Speech Signal Process., t. 25, nr 3, s. 235–238, 1977. https://ieeexplore.ieee.org/abstract/document/1162950

[5] Y. Meyer, "Wavelets and Operators: Volume 1". Cambridge university press, 1992. https://doi.org/10.1017/CBO9780511623820

[6] Z. Wu i N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method", Adv. Adapt. Data Anal., t. 1, nr 01, s. 1–41, 2009. https://www.worldscientific.com/doi/abs/10.1142/S1793536909000047

[7] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise", in 2011 IEEE international conference on acoustics, speech and signal processing (ICASSP), 2011, s. 4144–4147. https://ieeexplore.ieee.org/abstract/document/5947265

[8] M. A. Colominas, G. Schlotthauer, and M. E. Torres, "Improved complete ensemble EMD: A suitable tool for biomedical signal processing", Biomed. Signal Process. Control, t. 14, s. 19–29, 2014. https://www.sciencedirect.com/science/article/abs/pii/S1746809414000962

[9] M. S. Fabus, A. J. Quinn, C. E. Warnaby, and M. W. Woolrich, "Automatic decomposition of electrophysiological data into distinct nonsinusoidal oscillatory modes", J. Neurophysiol., t. 126, nr 5, s. 1670–1684, 2021. https://journals.physiology.org/doi/full/10.1152/jn.00315.2021

[10] R. Deering and J. F. Kaiser, "The use of a masking signal to improve empirical mode decomposition", in Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005., 2005, t. 4, s. iv–485. https://ieeexplore.ieee.org/abstract/document/1416051

[11] A. J. Quinn et al., "Within-cycle instantaneous frequency profiles report oscillatory waveform dynamics", J. Neurophysiol., t. 126, nr 4, s. 1190–1208, 2021. https://journals.physiology.org/doi/full/10.1152/jn.00201.2021

[12] Y. Yang, J. Deng, and D. Kang, "An improved empirical mode decomposition by using dyadic masking signals", Signal Image Video Process., t. 9, nr 6, s. 1259–1263, 2015. https://link.springer.com/article/10.1007/s11760-013-0566-7

[13] O. B. Fosso and M. Molinas, "EMD mode mixing separation of signals with close spectral proximity in smart grids", w 2018 IEEE PES innovative smart grid technologies conference Europe (ISGT-Europe), 2018, s. 1–6. https://ieeexplore.ieee.org/abstract/document/8571816

[14] S. Cole and B. Voytek, "Cycle-by-cycle analysis of neural oscillations", J. Neurophysiol., t. 122, nr 2, s. 849–861, 2019. https://journals.physiology.org/doi/full/10.1152/jn.00273.2019

[15] A. V. Oppenheim, "Discrete-time signal processing". Pearson Education India, 1999. https://dl.acm.org/doi/10.5555/294797

[16] I. Daubechies, J. Lu, and H.-T. Wu, "Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool", Appl. Comput. Harmon. Anal., t. 30, nr 2, s. 243–261, 2011. https://www.sciencedirect.com/science/article/pii/S1063520310001016

[17] F. Auger et al., "Time-frequency reassignment and synchrosqueezing: An overview", IEEE Signal Process. Mag., t. 30, nr 6, s. 32–41, 2013. https://ieeexplore.ieee.org/abstract/document/6633061

[18] J. B. Elsner and A. A. Tsonis, "Singular spectrum analysis: a new tool in time series analysis". Springer Science & Business Media, 1996. https://doi.org/10.1007/978-1-4757-2514-8

[19] M. G. Frei and I. Osorio, "Intrinsic time-scale decomposition: time–frequency–energy analysis and real-time filtering of nonstationary signals", Proc. R. Soc. Math. Phys. Eng. Sci., t. 463, nr 2078, s. 321–342, luty 2007, doi: 10.1098/rspa.2006.1761. https://royalsocietypublishing.org/doi/abs/10.1098/rspa.2006.1761

[20] P. Singh, S. D. Joshi, R. K. Patney, and K. Saha, "The Fourier decomposition method for nonlinear and nonstationary time series analysis", Proc. R. Soc. Math. Phys. Eng. Sci., t. 473, nr 2199, s. 20160871, 2017. https://royalsocietypublishing.org/doi/full/10.1098/rspa.2016.0871

[21] N. E. Huang and in., "On Holo-Hilbert spectral analysis: a full informational spectral representation for nonlinear and nonstationary data", Philos. Trans. R. Soc. Math. Phys. Eng. Sci., t. 374, nr 2065, s. 20150206, kwi. 2016, doi: 10.1098/rsta.2015.0206. https://scholarlypublications.universiteitleiden.nl/handle/1887/112702

[22] C.-H. Juan et al., "Revealing the dynamic nature of amplitude modulated neural entrainment with Holo-Hilbert spectral analysis", Front. Neurosci., t. 15, s. 673369, 2021. https://www.frontiersin.org/articles/10.3389/fnins.2021.673369/full

[23] L. R. Rabiner, R. W. Schafer, and others, "Introduction to digital speech processing", Found. Trends® Signal Process., t. 1, nr 1–2, s. 1–194, 2007 https://doi.org/10.1561/2000000001

[24] J. Benesty, M. M. Sondhi, Y. Huang, and others, Springer handbook of speech processing, t. 1. Springer, 2008. https://link.springer.com/book/10.1007/978-3-540-49127-9

[25] D. Kapilow, Y. Stylianou, and J. Schroeter, "Detection of non-stationarity in speech signals and its application to time-scaling", 1999. https://doi.org/10.21437/Eurospeech.1999-503

[26] R. S. Holambe and M. S. Deshpande, Advances in nonlinear modeling for speech processing. Springer Science & Business Media, 2012. https://link.springer.com/book/10.1007/978-1-4614-1505-3

[27] P.-L. Lee et al., "The Full Informational Spectral Analysis for Auditory Steady-State Responses in Human Brain Using the Combination of Canonical Correlation Analysis and Holo-Hilbert Spectral Analysis", J. Clin. Med., t. 11, nr 13, s. 3868, 2022. https://www.mdpi.com/2077-0383/11/13/3868

[28] N. Moradi, P. LeVan, B. Akin, B. G. Goodyear, and R. C. Sotero, "Holo-Hilbert spectral-based noise removal method for EEG high-frequency bands", J. Neurosci. Methods, t. 368, s. 109470, 2022. https://www.sciencedirect.com/science/article/abs/pii/S0165027021004052

[29] W.-K. Liang, P. Tseng, J.-R. Yeh, N. E. Huang, and C.-H. Juan, "Frontoparietal beta amplitude modulation and its interareal cross-frequency coupling in visual working memory", Neuroscience, t. 460, s. 69–87, 2021. https://www.sciencedirect.com/science/article/abs/pii/S0306452221000865