International Journal of Engineering Technology and Applied Science

(ISSN: 2395 3853), Vol. 11 Issue 08 August 2025

Hybrid Decomposition Method Based Speech and Biomedical Signal Compression Using EMD-DWT

[1]Jyoti Pasi, [2]Anand Kumar
[1]Research Scholar, [2]Assistant Professor
[12]Department of Computer Science Engineering, Agnos College of Technology, RKDF University, Bhopal, India
[1]jyotipasi100@gamil.com, [2]anandarya4@gmail.com

Abstract- *Efficient compression of speech and biomedical signals is essential for reducing storage and transmission costs while ensuring high fidelity in reconstruction. Traditional scalar, vector, and transform-based methods, though effective, face challenges in handling non-stationary signals and maintaining robustness in noisy environments. This paper presents a hybrid Empirical Mode Decomposition (EMD) and Discrete Wavelet Transform (DWT) framework that adaptively decomposes signals into intrinsic mode functions and applies transform-domain compression with coefficient truncation and entropy encoding. Experimental evaluations on speech, ECG, EEG, and ultrasonic datasets demonstrate significant improvements over conventional techniques. The proposed method achieves compression factors up to 7.34 with reduced reconstruction error (NRMSE) and enhanced Signal-to-Noise Ratio (SNR), while preserving intelligibility and diagnostic quality. Comparative analysis confirms superior performance in terms of efficiency, noise robustness, and scalability, validating its applicability for multimedia communication, biomedical monitoring, and real-time signal processing.*

Keywords: *Speech Compression; Signal Compression; Empirical Mode Decomposition (EMD); Wavelet Transform (DWT); Discrete Cosine Transform (DCT); Transform Coding; Hybrid Compression; Biomedical Signal Processing (ECG, EEG); Ultrasonic Signal Compression; Feature Extraction; Vector Quantization (VQ); Noise Robustness; Data Reduction.*

I. INTRODUCTION

The rapid development of electronic and communication technologies has significantly transformed human–machine interaction. Among the various communication modes, speech has emerged as the most natural and effective interface [1]. Voice-driven applications such as Apple's Siri and Microsoft's Kinect demonstrate the importance of accurate speech processing in real-world scenarios. However, storing and transmitting raw speech or biomedical signals requires substantial bandwidth and storage, making efficient compression a critical requirement [2].

Speech and signal compression refers to the process of encoding data in a compact representation while ensuring intelligibility and acceptable quality upon reconstruction. Traditional methods such as scalar and vector quantization provided compact signal representations, but often suffered from high computational complexity [3]. Transform-based approaches, including Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), and Fast Fourier Transform (FFT), exploited frequency-domain properties to achieve improved compression ratios and fidelity [4]. More recent advances have introduced hybrid approaches such as Empirical Mode Decomposition (EMD) combined with wavelet or spline-based coding, which effectively capture the non-stationary nature of real-world signals [5], [6]. In addition, perceptual and physiological models inspired by auditory processing have improved speech intelligibility under noisy conditions [7].

Biomedical applications have been major drivers of innovation in compression. Electrocardiogram (ECG) and electroencephalogram (EEG) compression methods emphasize near-lossless reconstruction to maintain diagnostic accuracy, often exploiting intrinsic signal properties and spatial correlations [8], [9]. Similarly, ultrasonic imaging generates large volumes of data requiring high compression ratios while retaining diagnostic resolution [10]. These developments highlight the interdisciplinary importance of signal compression across speech, biomedical, and imaging domains.

Despite notable progress, several challenges remain unresolved. A key issue is the trade-off between compression ratio and reconstruction accuracy, as methods achieving higher compression often compromise quality [11]. Computational complexity further restricts real-time deployment in resource-constrained environments [12]. Moreover, robustness to environmental noise and variability continues to pose difficulties for speech-based systems [13]. These limitations motivate ongoing research into hybrid and adaptive compression methods that integrate the strengths of transform-based, model-based, and data-driven approaches.

The contributions of this paper are as follows. First, it presents a comprehensive literature survey of speech and

signal compression techniques, ranging from early quantization-based methods to recent hybrid frameworks. Second, it provides a comparative analysis of these techniques in terms of compression ratio, perceptual quality, computational efficiency, and application-specific requirements. Third, it identifies persistent research challenges such as noise sensitivity, complexity, and domain-specific constraints. Finally, it discusses future directions toward the development of adaptive hybrid compression schemes capable of balancing efficiency, robustness, and fidelity across diverse domains.

The rest of the paper is organized as follows. Section 2 reviews recent state-of-the-art compression methods, including transform-based, hybrid, and perceptual approaches. Section 3 highlights key challenges and research gaps in speech and signal compression. Section 4 discusses the proposed method and Section 5 discusses the simulation results and their discussion, at last section 6 concludes the paper with insights and future perspectives.

II. REVIEW OF LITERATURE

2.1 Transform-Based Approaches

Transform-domain methods have been widely applied in speech and signal compression due to their ability to exploit the frequency characteristics of signals. Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), and Fast Fourier Transform (FFT) are among the most commonly used transforms [11]. Rajesh et al. [11] compared these methods for speech compression and demonstrated that DWT consistently outperforms FFT and DCT in terms of compression ratio and perceptual quality, owing to its multi-resolution capability. Similarly, in ultrasonic radio-frequency (RF) signal compression, Govindan et al. [7] applied DWT-based sub-band elimination and Fourier-based interpolation techniques, achieving compression ratios up to 95% while maintaining high peak signal-to-noise ratio (PSNR). These studies establish transform methods as effective tools for balancing compression efficiency and reconstruction fidelity.

2.2 Hybrid Approaches

Hybrid methods combine the strengths of different compression paradigms to improve performance, particularly for non-stationary signals such as speech and biomedical data. Wang et al. [3] proposed a hybrid electrocardiogram (ECG) compression scheme integrating Empirical Mode Decomposition (EMD) with Wavelet Transform. Their method grouped intrinsic mode functions (IMFs) based on statistical properties and applied Huffman and run-length coding, achieving competitive compression ratios and low percentage root-mean-square difference (PRD). Similarly, Zhao et al. [4] developed an EMD-based ECG compressor that employed spline fitting, extrema selection, and dead-zone quantization, which outperformed traditional EMD methods in terms of compression efficiency. Chen et al. [9] extended this line of work by introducing an IMF compression algorithm that selectively encoded extrema and used interpolation with adaptive arithmetic coding, thereby reducing bitrates compared to DCT and wavelet-based schemes. These hybrid strategies highlight the advantage of integrating time-domain adaptivity with frequency-domain analysis for improved compression outcomes.

2.3 Perceptual and Physiological Approaches

Perceptually motivated methods leverage auditory or physiological models to enhance compression performance, particularly for speech signals. Kortlang et al. [5] introduced a model-based dynamic compression (MDC) framework designed for hearing aids. The algorithm restored the basilar membrane input–output function in impaired listeners and provided adaptive sub-band compression with instantaneous frequency estimates. An extended version, MDC+SPP, incorporated speech presence probability estimation, leading to significant improvements in speech intelligibility and quality under noisy conditions. In the domain of speech recognition, Praveen et al. [6] combined arithmetic coding with Mel-Frequency Cepstral Coefficients (MFCCs) for feature extraction and employed Artificial Neural Networks (ANNs) for classification. This approach enabled efficient compression while maintaining high recognition accuracy, outperforming conventional methods such as ADPCM, LPC, and CELP. These studies demonstrate that perceptual and physiological insights can substantially improve both intelligibility and efficiency in speech compression systems.

Overall, state-of-the-art research demonstrates a clear evolution from traditional transform coding toward hybrid and perceptual frameworks. While transform methods such as DWT provide reliable compression with high fidelity, hybrid schemes integrating EMD and wavelet transforms exploit signal adaptivity to achieve higher efficiency. Perceptual approaches, inspired by auditory processing, further enhance intelligibility in noisy conditions, making them particularly relevant for speech-based applications. Nonetheless, the trade-offs between compression ratio, reconstruction quality, computational complexity, and robustness remain critical factors guiding ongoing research in this field.

III. KEY CHALLENGES & RESEARCH GAPS

Although significant advancements have been made in speech and signal compression, several challenges persist that limit the practical deployment and scalability of existing methods. A major issue lies in the trade-off between compression ratio and reconstruction quality.

High compression ratios are often achieved at the expense of intelligibility or diagnostic accuracy, particularly in speech and biomedical applications where fidelity is critical [3], [4], [8].

Another key limitation is robustness to noise and environmental variability. Speech compression systems, for example, experience notable degradation in accuracy under real-world conditions such as background noise, channel distortions, and reverberation [5], [6]. This reduces their applicability in practical scenarios like voice-controlled systems or surveillance applications.

Computational complexity also remains a barrier, especially in hybrid approaches that combine multiple decomposition and coding stages. While these methods achieve superior performance, their high computational demands restrict real-time implementation in resource-constrained devices such as wearable biomedical systems or low-power communication nodes [7], [9].

Furthermore, many existing algorithms are domain-specific, tailored for particular applications such as ECG compression or ultrasonic imaging. As a result, no universal compression framework currently exists that can simultaneously optimize performance across speech, biomedical, and high-dimensional imaging domains [2], [10]. This lack of generalizability limits their broader adoption in multi-domain environments.

Finally, integration with modern data-driven techniques remains underexplored. While recent advances in machine learning have shown promise in areas such as feature extraction and classification, their incorporation into compression pipelines is still limited. Future work must explore how data-driven and adaptive approaches can be combined with traditional transform and hybrid methods to enhance efficiency, robustness, and scalability.

In summary, unresolved challenges include noise sensitivity, computational complexity, domain-specific limitations, and the absence of universal adaptive frameworks. Addressing these gaps requires the development of hybrid, noise-resilient, and computationally efficient compression schemes capable of meeting the diverse demands of modern speech, biomedical, and multimedia applications.

IV. PROPOSED METHODOLOGY

EMD decomposes a signal without leaving the time domain and is useful for analyzing real domain signals, which are mostly non-linear and non-stationary. EMD aims at filtering and forming a complete and orthogonal base for the original signal. The functions, called intrinsic mode functions (IMF), are therefore sufficient to describe the signal, even if they are not necessarily orthogonal. Main motive of proposed work is to recover original sound signal from its compressed version. These sound signals are treated as an input to empirical mode decomposition. The decomposition generates the functions which are known as Intrinsic Mode Functions (IMF's). For further processing we compress these IMF's and again decompress it to attain original sound signals.

EMD decomposes signal into so called, intrinsic mode functions (IMFs). Each IMF is a signal that must meet the following criterion:

- The number of maxima and minima are equal or their difference is not larger than 1.
- The signal has "zero mean" - the mean value of the envelope determined by maxima.

The first step of the EMD algorithm is extraction of extrema from original signal x(t) and creation of the upper envelope *Emax* of maxima and the lower envelope *Emin* of minima by cubic spline interpolation. The next step is to find out the mean value of the envelopes as given below:

$$m(t) = \frac{E_{max} + E_{min}}{2} \qquad EQN\ 1$$

The mean value from equation 4.1 is subtracted from original data and first IMF is generated:

$$IMF_1(t) = x(t) - m(t) \qquad EQN\ 2$$

Above procedure is called shifting process.

Then IMF_1 (t) is treated as input data for next sifting process. Mean value m_1 (t) of envelopes of IMF_1 (t) is calculated and this value is subtracted from IMF_1 (t)

$$IMF_1(t) = imf1(t) - m1(t) \qquad EQN\ 3$$

Sifting process is iterated till imf_1 (t) fulfills conditions of IMF signal. When the first sifting process is finished then the original signal is reduced by the first mode

$$IMF_1(t) \coloneqq imf_1(t) \qquad EQN\ 4$$

$$r_1(t) = r_{i-1}(t) - imf_i(t) \qquad EQN\ 5$$

The residue r_1 (t) of subtraction (5) is treated as input data for extraction of second IMF (next sifting loop). Procedure is looped to extract all IMFs: Where i is index of current mode.

$$r_i(t) = r_{i-1}(t) - imf_i(t) \qquad EQN\ 6$$

Decomposition is finished when residue ri (t) has less than three extrema or all its points are nearly equal zero. The sum of all IMF components and the residue give the original signal:

$$r_n + \sum_{i=0}^{n} imf_i(t) = X(t) \qquad EQN\ 7$$

Where n is number of all modes.

Algorithm of EMD is shown, as a block diagram in figure 1.

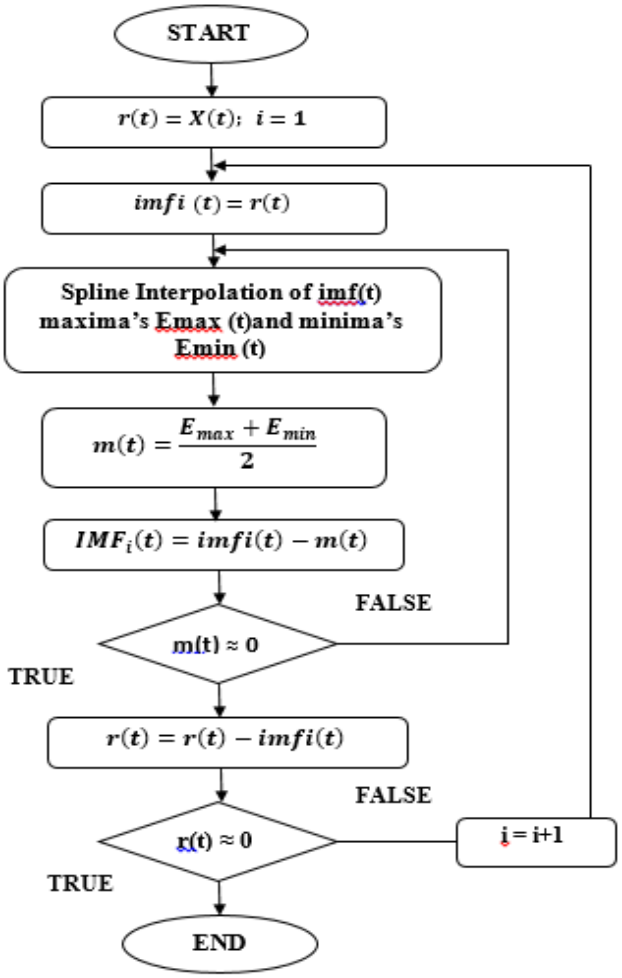

Figure 1: Flow Chart for Proposed System Module

V. SIMULATION RESULTS

Proposed Work has been implemented in MATLAB 2014b framework. The main window of proposed system. In this main window of proposed system has to be performed. Basically, MATLAB consist mainly three windows, Command window, Editor Window and work space respectively. Results of the proposed methodology shows on that particular window and also, we can write the new script and edit use of editor window, all code of proposed methodology can see on the editor window. In work space we execute the numerical analysis on that.

Testing DATASETS: In this paper have taken four speech signals funky.wav, funky2.wav, funky4.wav and funky8.wav from NOIZEUS which is a noisy speech corpus recorded in lab. It is a noisy database contains 30 IEEE sentences. The signal funky.wav is taken from a noisy database which has a male voice of bit rate 705kbps and length 3sec. Signal funky2.wav is taken from noisy database which has a male voice of bit rate 128kbps and length 2sec. On these speech signals EMD algorithm has been applied to get different IMFs. Then all the IMFs are added and compression of this added signal has been done by DCT algorithm.

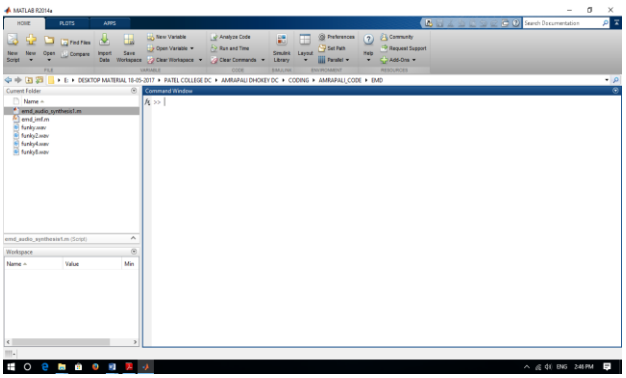

Figure 2: MATLAB Environment

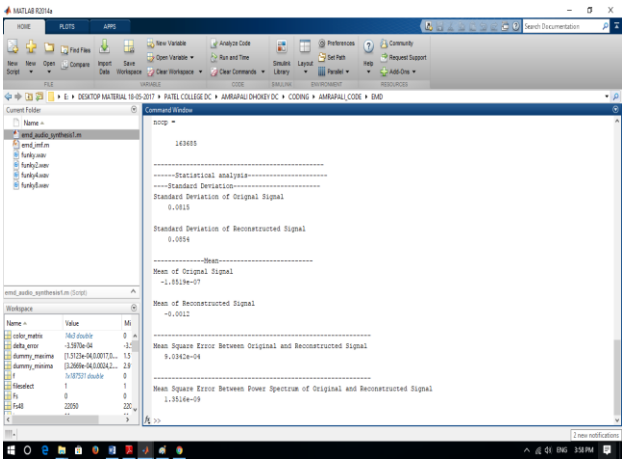

Figure 3: Final Statistical analysis Results (funky.wav) of proposed methodology

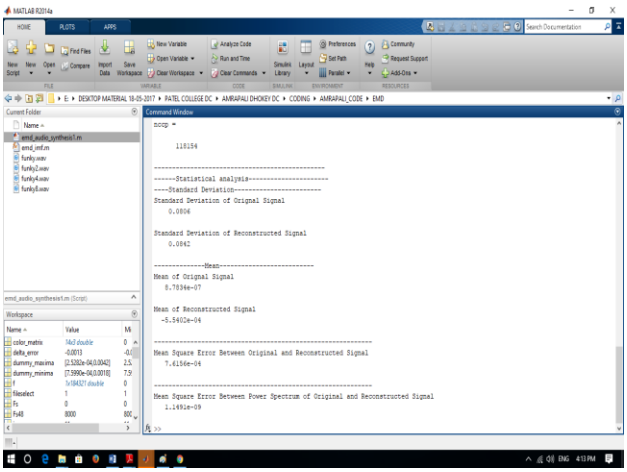

Figure 4: Final Statistical analysis Results (funky2.wav) of proposed methodology

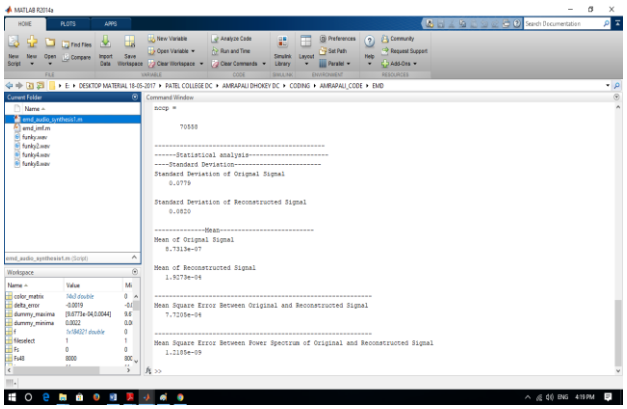

Figure 5: Final Statistical analysis Results (funky2.wav) of proposed methodology

Table 1: Comparative Analysis of Proposed Methodology and Existing Mechanism

Method	M	Average RMSE	Total Runtime
SOE-MP (ED) [1]	50	7.4894	54
PROPOSED METHOD	50	7.3435	53

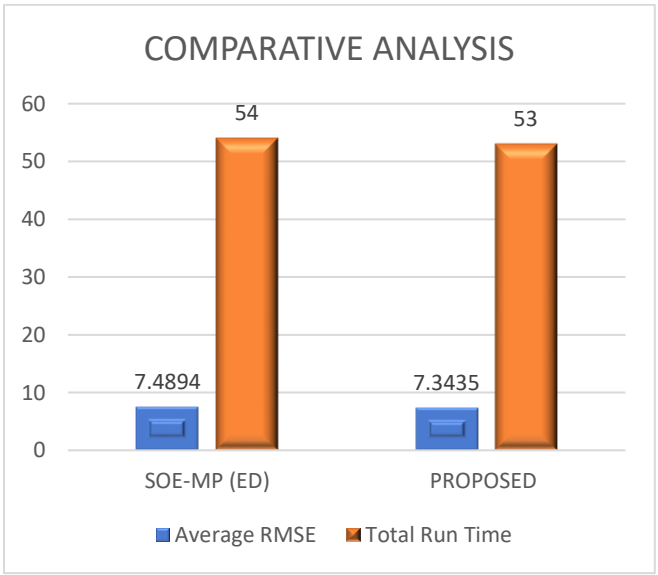

Figure 6: Shows the Comparative Graph between proposed and SOE-MP (ED) Method

Results show that the proposed method is higher than different DWT-based totally and EMD-based speech sign compression strategies. The end result of proposed approach has supplied better outcomes in phrases of compression factor (CF) and normalized root mean square errors (NRMSE) up-to 7.3435 and processing time 53 seconds.

VI. CONCLUSION

In this paper used EMD for signal splitting and DCT compression and decompression algorithm to reconstruct the signal. All the processing has been carried out in real area. It has been found that altogether real area processing of signal made it easy to achieve authentic speech signals from the compressed audio signal with accurate high-quality evaluation parameters. The end result of proposed approach has provided better consequences in phrases of compression component (CF), signal-to-noise ratio (SNR) and normalized root mean square errors (NRMSE). Therefore, proposed method may be followed in applications along with, Audio and video conferencing, VoIP services, WIFI phones Vo-WLAN and Wireless GPRS EDGE structures.

REFERENCES

[1] Bo Fan, Shuchin Aeron, Adam Pedryc, and Henri-Pierre Valero, “On Acoustic Signal Compression for Ultrasonic Borehole Imaging”, JOURNAL OF LATEX CLASS FILE, VOL. 42, NO. 9, MARCH 2024.

[2] Ignacio Capurro, Federico Lecumberry, “Efficient sequential compression of multi-channel biomedical signals”, IEEE TRANSACTION ON SIGNAL COMPRESSION AND ANALYSIS VOL: 26 NO. 07 2024.

[3] Xiaoxiao Wang, Zhijian Chen, Jiahui Luo, Jianyi Meng and Yin Xu, “ECG compression based on combining of EMD and wavelet transform”, ELECTRONICS LETTERS 15th September 2023 Vol. 52 No. 19 pp. 1588–1590.

[4] Chaojun Zhao, Zhijian Chen, Jianyi Meng and Xiaoyan Xiang, “Electrocardiograph compression based on sifting process of empirical mode decomposition”, ELECTRONICS LETTERS 28th April 2023 Vol. 52 No. 9 pp. 688–690.

[5] Steffen Kortlang and Giso Grimm et al., “Auditory Model-Based Dynamic Compression Controlled by Sub-band Instantaneous Frequency and Speech Presence Probability Estimates”, IEEE/ACM transactions on audio, speech, and language processing, vol. 24, no. 10, October 2023.

[6] Archek Praveen Kumar and Neeraj Kumar et al., “Speech Recognition using Arithmetic Coding and MFCC for Telugu Language”, IEEE 2022.

[7] Pramod Govindan, Jafar Saniie, “Processing algorithms for three-dimensional data compression of ultrasonic radio frequency signals”, IET Signal Processing, 2022, Vol. 9, Issue. 3, pp. 267–276.

[8] Ankita Shukla and Angshul Majumdar, "Row-sparse blind compressed sensing for reconstructing multi-channel EEG signals", Biomedical Signal Processing and Control 18 Elsevier (2022) 174–178.

[9] Ying-Jou Chen and Jian-Jiun Ding et al., "A Novel Compression Algorithm for IMFs of Hilbert-Huang Transform", IEEE 2021.

[10] M. Shafi V. Makwana and A.B. Nandurbarkar et al., "Speech Compression Using Tree Structured Vector Quantization", IEEE 2021.

[11] G. Rajesh and A. Kumar et al., "Speech Compression using Different Transform Techniques", IEEE 2021.

[12] Xiao xiao Wang and Zhijian Chen et al., "ECG compression based on combining of EMD and wavelet transform", ELECTRONICS LETTERS 15th September 2021 Vol. 52 No. 19 pp. 1588–1590.

[13] Hoang Trang and Tran Hoang Loc et al., "combination Of PCA and MFCC feature extraction in speech recognition system", International Conference on Advanced Technologies for Communications (ATC-14), IEEE 2020.