JSS MAHAVIDYAPEETHA

JSS SCIENCE AND TECHNOLOGY UNIVERSITY
SRI JAYACHAMARAJENDRA COLLEGE OF ENGINEERING
JSS TECHNICAL INSTITUTIONS CAMPUS, MYSURU - 570 006

Subject: Digital Compression Techniques(20EC813) Event – 4 Report

"Speech compression using Huffman Coding and Wavelet Transform"

Course outcome covered in this event:

CO4: Simulate real time computing problems using modern software tools through group projects and give oral presentation with documentation.

Submitted by

Sl. No.	USN	NAME
1.	01JST20EC033	GURURAJ DEEPAK TODURKAR
2.	01JST20EC045	KRISHNAKUMAR BALACHANDRA BHAT
3.	01JST20EC111	VADIRAJ B HAYAGREEVA
4.	01JST20EC015	B R MANOJ

Submitted to

Prof .Pavithra D R
Assistant Professor
Department of Electronics & Communication SJCE, JSS Science and Technology University Mysuru

DEPARTMENT OF ELECTRONICS AND COMMUNICATION
JSS SCIENCE AND TECHNOLOGY UNIVERSITY
JSS TECHNICAL INSTITUTIONS CAMPUS
MYSORE-570006

2023-2024

JSS MAHAVIDYAPEETHA

JSS SCIENCE AND TECHNOLOGY UNIVERSITY

SRI JAYACHAMARAJENDRA COLLEGE OF ENGINEERING

JSS TECHNICAL INSTITUTIONS CAMPUS MYSURU - 570 006

Subject: Digital Compression Techniques (20EC813)

Event – 4

Report Evaluation sheet

Course outcome covered in this event:

CO4: Simulate real time computing problems using modern software tools through group projects and give oral presentation with documentation.

Name	GURURAJ DEEPAK TODURKAR
Roll Number	13
Section	B
USN	01JST20EC033
Submission Date	

Evaluation Component	**Max. Marks**	**Marks Scored**
Aim formulation and background work	**4**	
Identification of application with background study, Detailed analysis of Feasibility, Objectives and Methodology	**10**	
Quality of report writing	**4**	
Interaction/Viva	**2**	
Total	**20**	

Name of the faculty	**Prof. Pavithra D R**
Signature of the faculty	
Signature of the student	

2023 – 2024

JSS MAHAVIDYAPEETHA

JSS SCIENCE AND TECHNOLOGY UNIVERSITY

SRI JAYACHAMARAJENDRA COLLEGE OF ENGINEERING

JSS TECHNICAL INSTITUTIONS CAMPUS MYSURU - 570 006

Subject: Digital Compression Techniques (20EC813)

Event – 4

Report Evaluation sheet

Course outcome covered in this event:

CO4: Simulate real time computing problems using modern software tools through group projects and give oral presentation with documentation.

Name	B R MANOJ
Roll Number	05
Section	B
USN	01JST20EC015
Submission Date	

Evaluation Component	**Max. Marks**	**Marks Scored**
Aim formulation and background work	**4**	
Identification of application with background study, Detailed analysis of Feasibility, Objectives and Methodology	**10**	
Quality of report writing	**4**	
Interaction/Viva	**2**	
Total	**20**	

Name of the faculty	**Prof. Pavithra D R**
Signature of the faculty	
Signature of the student	

2023 – 2024

JSS MAHAVIDYAPEETHA

JSS SCIENCE AND TECHNOLOGY UNIVERSITY
SRI JAYACHAMARAJENDRA COLLEGE OF ENGINEERING
JSS TECHNICAL INSTITUTIONS CAMPUS MYSURU - 570 006

Subject: Digital Compression Techniques (20EC813)
Event – 4
Report Evaluation sheet

Course outcome covered in this event:
CO4: Simulate real time computing problems using modern software tools through group projects and give oral presentation with documentation.

Name	KRISHNAKUMAR BALACHANDRA BHAT
Roll Number	18
Section	B
USN	01JST20EC045
Submission Date	

Evaluation Component	Max. Marks	Marks Scored
Aim formulation and background work	**4**	
Identification of application with background study, Detailed analysis of Feasibility, Objectives and Methodology	**10**	
Quality of report writing	**4**	
Interaction/Viva	**2**	
Total	**20**	

Name of the faculty	**Prof. Pavithra D R**
Signature of the faculty	
Signature of the student	

2023 – 2024

JSS MAHAVIDYAPEETHA

JSS SCIENCE AND TECHNOLOGY UNIVERSITY
SRI JAYACHAMARAJENDRA COLLEGE OF ENGINEERING
JSS TECHNICAL INSTITUTIONS CAMPUS MYSURU - 570 006

Subject: Digital Compression Techniques (20EC813)
Event – 4
Report Evaluation sheet

Course outcome covered in this event:

CO4: Simulate real time computing problems using modern software tools through group projects and give oral presentation with documentation.

Name	VADIRAJ B HAYAGREEVA
Roll Number	40
Section	B
USN	01JST20EC111
Submission Date	

Evaluation Component	Max. Marks	Marks Scored
Aim formulation and background work	**4**	
Identification of application with background study, Detailed analysis of Feasibility, Objectives and Methodology	**10**	
Quality of report writing	**4**	
Interaction/Viva	**2**	
Total	**20**	

Name of the faculty	**Prof. Pavithra D R**
Signature of the faculty	
Signature of the student	

2023 – 2024

2023 – 2024

Abstract

In this paper, the usage of Huffman coding and the Wavelet transform for file compression is examined. The Wavelet transform approach successfully removes the redundant information included in the signals, with a focus on voice signals. Huffman coding is then used to obtain even more compression. The combined technique provides effective file compression, resulting in smaller files while keeping the signal integrity and perceptual quality. The suggested method offers a promising approach to speech signal compression, with potential uses in transmission and storage systems where resource efficiency is crucial.

Introduction

The removal of unnecessary material while maintaining the perceptual properties of speech requires the employment of compression techniques for effective speech signal storage and transmission. Speech compression decreases the amount of memory space needed for storing and allows for quicker retrieval by eliminating repetitious information. Additionally, since compressed speech requires fewer bits per second during transmission, transmission bandwidth is lowered, and inter-symbol interference is avoided. To achieve speech compression, the signal is split up into precise and approximate levels using the Wavelet transform, resulting in high-pass and low-pass subbands, respectively. The length of the coefficients is halved at each stage of decomposition, further dividing the low-pass subband into high-pass and low-pass subbands.

The Discrete Wavelet Transform (DWT) generates larger coefficients for representing the important components of the signal, while noise is represented by lower coefficients. By selecting a suitable threshold, coefficients with extremely small values can be eliminated. The remaining coefficients above the threshold are then compressed using the Huffman source coding method after the Wavelet transform. The Huffman coding technique employs variable-length coding, using longer code words for messages with lower odds of occurrence. This approach maximizes the average amount of information contained in each binary digit, effectively compressing the speech signal. By utilizing these compression methods, speech compression can preserve crucial speech qualities while making optimal use of transmission and storage resources. This enables efficient speech signal storage and transmission, reducing memory space requirements, lowering transmission bandwidth, and ensuring the preservation of important speech characteristics.

Literature Survey:

1. Speech Compression with Wavelet Transform and Huffman Coding Ranjushree Pal, Department of Electronics and Telecommunication Engineering, Dwarkadas J Sanghvi College of Engineering, Mumbai: 400056, India.

   The literature survey highlights the need for efficient speech compression techniques to reduce redundancy in speech signals. The research paper focuses on using Wavelet Transform and Huffman coding to achieve higher compression ratios compared to Wavelet Transform alone. The combined approach retains perceptual qualities and maximum energy while slightly increasing distortion. Overall, it demonstrates the effectiveness of integrating Wavelet Transform and Huffman coding for efficient speech compression.

2. Speech Compression Exploiting Hamming Correction Code Compressor Islam Amro Faculty of Information Technology Al-Quds Open University Ramallah.

   The literature survey examines the utilization of Hamming Correction Code Compressor (HCDC) for compressing speech signals directly in Pulse Code Modulation (PCM) form. The average compression rate achieved with HCDC is approximately 2.3 for the tested datasets. A comparison with the Free Lossless Audio Compression (FLAC) package reveals that HCDC performs slightly better in compressing 8-bit resolution samples, while the compression rate for 16-bit resolution is lower, averaging around 1.4. HCDC's main advantages lie in its simplicity, lower time consumption, and reduced space requirements compared to other lossless speech compression techniques.

3. Wavelet transform for speech compression and denoising Fatma zohra Chelali* and Noureddine. Cherabit, University of Science and Technology Houari Boumediene of Algiers, ALGERIA Amar. Djeradi and Leila.Falek

   The literature survey focuses on speech compression techniques using Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT), as well as speech denoising with packet wavelets. The efficiency of packet wavelets in reducing noise in acoustic signals is highlighted. Various audio compression standards are discussed, emphasizing adherence to these standards during signal compression. The quality of acoustic compression is evaluated using Peak Signal to Noise Ratio (PSNR) and mean square error metrics. The results demonstrate superior performance of the Wavelet transform compared to DCT in speech compression.

4. Continuous Speech Coding Using Coiflets Wavelet. Shijo M Joseph1 Babu Anto

   The literature survey highlights the effectiveness of wavelet transforms for signal compression, particularly in retaining both time and frequency aspects of the signal. The challenge lies in selecting the optimal wavelet and decomposition level for speech coding. The paper focuses on analyzing the performance of Coiflets wavelet in continuous Malayalam speech coding, a first-of-its-kind attempt in Indian languages. The experiment involves decomposing the signal using Coiflets wavelet, applying thresholding based on the Birge-Massart strategy to eliminate insignificant coefficients, and employing a lossless encoding algorithm for the remaining coefficients. Performance evaluation metrics include SNR, PSNR, NRMSE, RSE, MOS, and compression ratio. The study successfully compresses and reconstructs continuous Malayalam spoken sentences while maintaining perfect audibility.

5. A Comparative Study between Discrete Wavelet Transform and Linear Predictive Coding D.Ambika, V.Radha

   The literature survey discusses the analysis of the compression process by comparing the compressed signal with the original signal. The study implements powerful speech analysis and compression techniques, namely Linear Predictive Coding (LPC) and Discrete Wavelet Transform (DWT), using MATLAB. Nine samples of spoken words from different speakers are collected for implementation. The results obtained from LPC are compared with the Discrete Wavelet Transform method. Evaluation of the results includes compressed ratio (CR), Peak Signal-to-Noise Ratio (PSNR), and Normalized Root-Mean-Square Error (NRMSE). The findings indicate that DWT outperforms LPC in terms of performance for the given samples.

6. An Audio Compression Method Based on Wavelet Packet Decomposition, Ordering, and Polynomial Approximation of Expressive Coefficients Lucas L. Motta∗ , Rafael S. Mendes∗ , Max H. M. Costa.

   The paper presents an audio compression method using wavelet packet decomposition, coefficient ordering, and polynomial approximation. The encoder achieves competitive SNR performance but shows lower results in perceptual quality compared to other codecs.

Further exploration of the method's variance and evaluation with different psycho-acoustic models is recommended for future work.

7. Achieving Lossless Audio Encoder through integrated approaches of Wavelet Transform, Quantization and Huffman encoding (LAEIWQH) Asish Debnath ,Uttam Kr. Mondal , Niranjan Panja , Bibek Bikash.

   The paper introduces a lossless audio compression technique using discrete 2-D wavelet transformation and quantization followed by Huffman encoding. The approach achieves high compression rates while preserving audio information and quality. Experimental results demonstrate its effectiveness compared to existing lossless techniques. Further improvements can be made by refining the encoding rules.

8. M. S. Serina and S. G. Mosin, "Digital Audio Information Compression Using Wavelets," Proceedings of the International Conference Mixed Design of Integrated Circuits and System.

   The research explores the application of Wavelet Transform for audio compression and investigates the influence of different factors on the compression ratio. The developed software model demonstrates that the initial audio file can be compressed by 2-4 times with less than 10% deformation in the restored file. The study finds that Daubechies wavelets provide the best results in terms of quality and compression ratio, while Haar wavelets offer easier hardware implementation with slightly lower quality. Additionally, using fewer decomposition levels can be sufficient if high-quality restoration is not a requirement. Future work aims to further improve compression ratio and restoration quality using Wavelet Transform for speech and sound processing.

9. M. H. Martínez, N. F. Rosas and C. Camargo, "Compressing audio signals with the use of the wavelet transform and Embedded Systems"

   The paper presents the development of a perceptual audio compressor using Wavelet Transform in an Embedded System. The proposed compressor achieves remarkable compression ratios while preserving the original format (.wav) and maintaining appropriate time-frequency tracking without perceptual losses. By using the Daubechies-Wavelet Type 4, the transient representation and highly variable spectral components in audio signals are effectively addressed. The implemented perceptual audio coder based on Wavelets-Daubechies 4 Transform provides a suitable alternative to the classical Fourier representation in psychoacoustic modeling.

10. G. Kemper and Y. Iano, "An Audio Compression Method Based on Wavelets Subband Coding,"

    This paper presents an audio compression method based on wavelets sub-band quantization and coding. The proposed coder utilizes wavelet packets transform to capture the critical bands of the human auditory system. Results from the MPEG layer 2 psychoacoustic model are incorporated in the wavelet coefficients coding. The coder employs scalar and vector quantization methods tailored to the sensitivity of each wavelet sub-band. Entropy coding is also employed to enhance performance. Subjective evaluations demonstrate that the proposed coder achieves transparent coding of monophonic CD signals at bit rates of 80-96 Kbit/sec. Optimizing the involved algorithms can result in a high-performance coder for audio compression.

11. A. K. Sinha, S. K. Dutta, B. Dev, Jaydipta, R. Ranjan and R. Kumari, "Wavelet based Speech Coding technique using median function thresholding"

    The paper introduces a speech coding technique based on wavelet theory and incorporates the median function thresholding into the Discrete Wavelet Transform (DWT). The proposed method achieves high compression scores of 2.8720 compared to 1.7364 in another method. The reconstructed signal maintains good quality, retaining 99.8184% of the signal energy with a small error of 0.0423%. The technique also effectively reduces noise content in the speech signal. The proposed coding method shows promise for real-time systems.

12. F. z. Chelali, N. Cherabit, A. Djeradi and L. Falek, "Wavelet transform for speech compression and denoising,"

    This paper explores speech compression using DCT and DWT, along with speech denoising using packet wavelet. Wavelet transform yields better results compared to DCT. The authors implemented a system based on DCT and DWT, achieving optimal compression. Packet wavelet and softthresholding effectively reduce noise. DWT outperforms DCT with a PSNR of approximately 50 dB. Subjective evaluation and Simulink diagrams were used for analysis.

13. J. Pang, S. Chauhan and J. M. Bhlodia, "Speech Compression FPGA Design by Using Different Discrete Wavelet Transform Schemes,"
This paper presents FPGA designs for speech compression using different discrete wavelet transform (DWT) schemes. The Daubechies DWT and Daubechies Lifting Scheme DWT are utilized. The speech signals are converted from analog to digital format using an audio CODEC chip. The compressed speech is represented by the low pass filtered components of the DWT result. The FPGA implementation allows for real-time compression and storage in SDRAM. The compressed speech can be read back, upsampled, and played clearly through external speakers. The results demonstrate successful speech compression with minimal background noise using both post-processing and real-time processing approaches.

Literature Review

Paper Title	Authors	Key Techniques	Findings
Speech Compression with Wavelet Transform and Huffman Coding	Ranjushree Pal	Wavelet Transform, Huffman coding	Wavelet Transform and Huffman coding together achieve higher compression ratios with slightly increased distortion, demonstrating effective speech compression.
Speech Compression Exploiting Hamming Correction Code Compressor	Islam Amro	Hamming Correction Code Compressor (HCDC)	HCDC achieves an average compression rate of approximately 2.3. Performs better than FLAC for 8-bit resolution samples, but lower for 16-bit resolution. Simplicity, lower time consumption, and reduced space requirements are advantages of HCDC.

Paper Title	Authors	Key Techniques	Findings
Wavelet transform for speech compression and denoising	Fatma zohra Chelali, Noureddine Cherabit, Amar Djeradi, Leila Falek	Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), packet wavelets	Packet wavelets effectively reduce noise in acoustic signals. Wavelet transform performs better than DCT in speech compression. Quality evaluation using PSNR and mean square error metrics favor Wavelet transform.
Continuous Speech Coding Using Coiflets Wavelet	Shijo M Joseph, Babu Anto	Coiflets wavelet	Coiflets wavelet performs well in continuous Malayalam speech coding, achieving compression while maintaining audibility. Performance evaluation metrics include SNR, PSNR, NRMSE, RSE, MOS, and compression ratio.
A Comparative Study between Discrete Wavelet Transform and Linear Predictive Coding	D.Ambika, V.Radha	Linear Predictive Coding (LPC), Discrete Wavelet Transform (DWT)	DWT outperforms LPC in terms of performance for the given samples, based on compressed ratio (CR), PSNR, and NRMSE evaluation.
An Audio Compression Method Based on Wavelet Packet Decomposition, Ordering, and Polynomial Approximation of	Lucas L. Motta, Rafael S. Mendes, Max H. M. Costa	Wavelet packet decomposition, coefficient ordering, polynomial approximation	The encoder achieves competitive SNR performance but lower perceptual quality compared to other codecs.

Paper Title	Authors	Key Techniques	Findings
Expressive Coefficients			
Achieving Lossless Audio Encoder through integrated approaches of Wavelet Transform, Quantization and Huffman encoding (LAEIWQH)	Asish Debnath, Uttam Kr. Mondal, Niranjan Panja, Bibek Bikash	2-D wavelet transformation, quantization, Huffman encoding	High compression rates are achieved while preserving audio information and quality. Experimental results show effectiveness compared to existing lossless techniques.
Digital Audio Information Compression Using Wavelets	M. S. Serina, S. G. Mosin	Wavelet Transform (Daubechies wavelets, Haar wavelets)	Daubechies wavelets provide the best results in quality and compression ratio, while Haar wavelets offer easier hardware implementation with slightly lower quality. Fewer decomposition levels can suffice if high-quality restoration is not necessary.
Compressing audio signals with the use of the wavelet transform and Embedded Systems	M. H. Martínez, N. F. Rosas, C. Camargo	Wavelet Transform (Daubechies-Wavelet Type 4)	The proposed compressor achieves remarkable compression ratios while preserving the original format and maintaining appropriate time-frequency tracking without perceptual losses.
An Audio Compression Method	G. Kemper, Y. Iano	Wavelet packets transform, scalar and vector	The proposed coder achieves transparent coding of monophonic CD signals at bit rates of 80-

Paper Title	Authors	Key Techniques	Findings
Based on Wavelets Subband Coding		quantization, entropy coding	96 Kbit/sec. Optimization of algorithms can result in a high-performance coder for audio compression.
Wavelet based Speech Coding technique using median function thresholding	A. K. Sinha, S. K. Dutta, B. Dev, Jaydipta, R. Ranjan, R. Kumari	Discrete Wavelet Transform (DWT), median function thresholding	The proposed method achieves higher compression scores and maintains good quality with minimal error. It effectively reduces noise in the speech signal, showing promise for real-time systems.
Wavelet transform for speech compression and denoising	F. z. Chelali, N. Cherabit, A. Djeradi, L. Falek	DCT, DWT, packet wavelet, soft thresholding	Wavelet transform outperforms DCT with higher PSNR. Packet wavelet and soft thresholding effectively reduce noise. Subjective evaluation and Simulink diagrams were used for analysis.

In conclusion, the various methods and approaches has used of various wavelet types (e.g., Daubechies, Coiflets), different transform techniques (e.g., DCT, DWT), incorporation of coding schemes (e.g., Huffman encoding, quantization), and exploration of additional techniques (e.g., packet wavelets, polynomial approximation, and median function thresholding) are examples of differences in methods used across the papers.

In order to eliminate redundancy in speech signals, the evaluated studies emphasise the significance of effective speech compression techniques. Wavelet Transform is useful for increasing compression ratios while preserving perceptual features, especially when used with Huffman coding. The advantage of Wavelet Transform over DCT in voice compression is highlighted by the exploration of several wavelet types and transform methodologies. The effectiveness and benefits of additional techniques including the Hamming Correction Code Compressor, linear predictive coding, and sub band coding are demonstrated.

Problem Statement:

Efficient storage and transmission of speech signals are vital in various applications. However, the inherent redundancy in speech signals leads to wastage of resources. Therefore, there is a need for an effective speech compression technique that can remove redundant information while retaining the perceptual qualities of speech. ***The challenge lies in developing a compression approach that can achieve a high compression ratio, minimize distortion in the reconstructed waveform, and preserve the maximum percentage of signal energy.*** Additionally, the technique should be applicable specifically to female speech signals recorded using Praat software. Addressing these challenges will contribute to more efficient storage and transmission of speech signals, optimizing resource utilization and enhancing overall system performance.

Objectives:

1. Develop a speech compression technique that efficiently removes redundant information from speech signals.
2. Investigate the application of Wavelet Transform to decompose the speech signals into high pass and lowpass sub bands, thereby eliminating noise and retaining essential information.
3. Implement Huffman coding to further compress the speech signals, optimizing the average information per binary digit and achieving a higher compression ratio.

4. Evaluate the performance of the proposed technique by assessing the distortion in the reconstructed waveform, measuring the percentage of preserved signal energy, and analysing the perceptual qualities of the compressed speech.

Block Diagram:

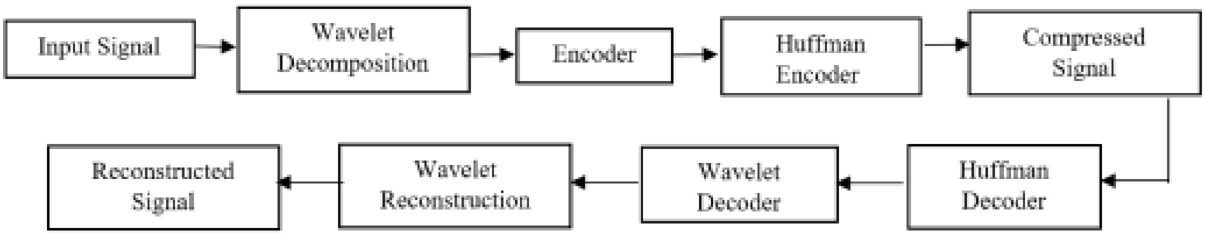

Figure 1. Block Diagram of the Flow of Implementation

Methodology:

The method of voice compression and reconstruction utilising wavelets and Huffman coding is shown in a flowchart in Figure 1. The Python interpreter in the VS Code IDE was used to implement the entire article. The Wavelet transform is initially used to decompose the input signals. The calculation of a threshold value results in the disregard of DWT coefficients below the threshold. As a result, compression is achieved by concentrating the information into fewer DWT coefficients. The compression works well despite the slight distortion present. The Huffman source coding technique is used by the encoder to compress the signal after encoding the coefficients. The Huffman decoder and then the wavelet decoder are used to decode the compressed signal during the decoding stage. The signal that was rebuilt is a WAV file.

Compression with Wavelets

The db6 wavelet was chosen for this experiment because it had the most energy after the first level of decimation and produced a sizable proportion of zero counts after the fifth level. Using the db6 wavelet, the recorded signals were wavelet decomposed up to the sixth level. There were two parts to each degree of decomposition: an approximate component and a detailed part. The Approximation component was further divided into an Approximation part and a Detailed part, with the Detailed part remaining unaltered. The majority of coefficient values became negligible and could be deleted as a result of this decomposition procedure, which was carried out up to the sixth level. The Python interpreter in the VS Code IDE was used to determine a threshold that would compress the signal.

The coefficient values that above the threshold were kept after the thresholding procedure and formed a row matrix to represent the compressed signal. The non-zero elements' indexes, which were crucial for signal processing, were also saved by the encoder.

Compression with Huffman Coding

The research investigates the effectiveness of wavelet transform alone versus wavelet transform combined with Huffman coding for speech compression. The wavelet encoder rounds up non-zero coefficients to the nearest integer, which are then used as symbols in the subsequent Huffman encoding process. The implementation of the Huffman dictionary and encoder is carried out using MATLAB, resulting in binary string outputs. The length of these output signals represents the number of bits required for transmission. By calculating the compression ratio, which compares the original signal size to the compressed signal size, the efficiency of the compression techniques can be assessed. Moreover, distortion analysis is conducted to evaluate the accuracy of the reconstructed waveform in comparison to the original signal.

$$Compression\ Ratio = \frac{Length\ of\ orignal\ signal\ in\ bits}{Length\ of\ compressed\ signal\ in\ bits}$$

Introducing Machine Learning for threshold calculations:

Training a Decision Tree Regressor to forecast thresholds for adaptive thresholding is the machine learning component. Let's get into greater detail about this machine learning element:

Decision Tree Regressor Training: A Decision Tree Regressor is trained to predict the best threshold for adaptive thresholding for each level of coefficients generated from the wavelet transform. The sklearn.tree module's DecisionTreeRegressor() method is used to create the regressor.

Getting ready the training data By flattening the coefficients and generating the input feature x_train and the target variable y_train, the training data for the regressor is created. Using the flatten() method on the coefficient array, the coefficients are made flat.

The target variable (y_train) is the set of flattened coefficients.

How to fit the regressor: By executing the fit() method on the regressor object and providing the inputs x_train and y_train, the Decision Tree Regressor is trained. In this step, the model is trained to understand how the input feature (indices) and the target variable (flattened coefficients) transfer to one another.

```
# Use a machine learning model (Decision Tree Regressor) to predict the threshold
# Train the model on the wavelet coefficients to predict optimal thresholds
regressors = []
thresholds = []

for i, coeff in enumerate(coeffs):
    regressor = DecisionTreeRegressor()
    flatten_coeff = coeff.flatten()
    x_train = np.arange(len(flatten_coeff)).reshape(-1, 1)
    y_train = flatten_coeff
    regressor.fit(x_train, y_train)
    regressors.append(regressor)
    thresholds.append(np.mean(np.abs(regressor.predict(x_train) - y_train)))
```

The thresholds' prediction: np.mean(np.abs(regressor.predict(x_train) - y_train)) is used to find the mean absolute difference between the predicted values and the actual values of the training data after the regressor has been trained. The projected threshold for each level of coefficients is represented by this value.

Regressors and thresholds are stored in separate lists (regressors and thresholds) for use in the code at a later time. This includes the trained regressor and the corresponding threshold for each level of coefficients.

The code makes use of machine learning to forecast the ideal thresholds for adaptive thresholding, which is a critical step in the audio compression process, utilising the Decision Tree Regressor. The algorithm can decide which coefficients to preserve or discard during compression thanks to the trained regressors, which capture the patterns and relationships within the coefficients.

Reconstruction of original signal

The Python interpreter in the VS Code IDE is used to implement the Huffman and wavelet decoders. The waveform reconstruction is stored as a WAV file so that it can be played back by various audio software programmes. The Python matplotlib module is used to plot the original input signal as well as the reconstructed output signal. Metrics like distortion, quantified as the Mean Square Error (MSE), are calculated to judge the integrity of the reconstructed signal. To gauge the degree of information preservation, the retained signal energy of the rebuilt signal is assessed. The research ensures effective application and evaluation of the compression and reconstruction strategies by utilising the Python interpreter in VS Code.

$$Retained\ Energy = \frac{Energy\ in\ reconstructed\ signal}{Energy\ in\ orignal\ signal}$$

$$Mean\ Square\ Error = \frac{1}{N}\sum_{k=0}^{N} \quad (Rsignal(k) - Osingnal(k))^2$$

N = Total number of samples

Rsignal = Reconstructed Signal

Osignal = Original Signal

Results and Discussion:

Original Audio:

This plot shows the waveform of the original audio signal. It is the raw audio data captured from the .wav file before any compression is applied. The x-axis represents the sample index, and the y-axis represents the amplitude of the audio signal.

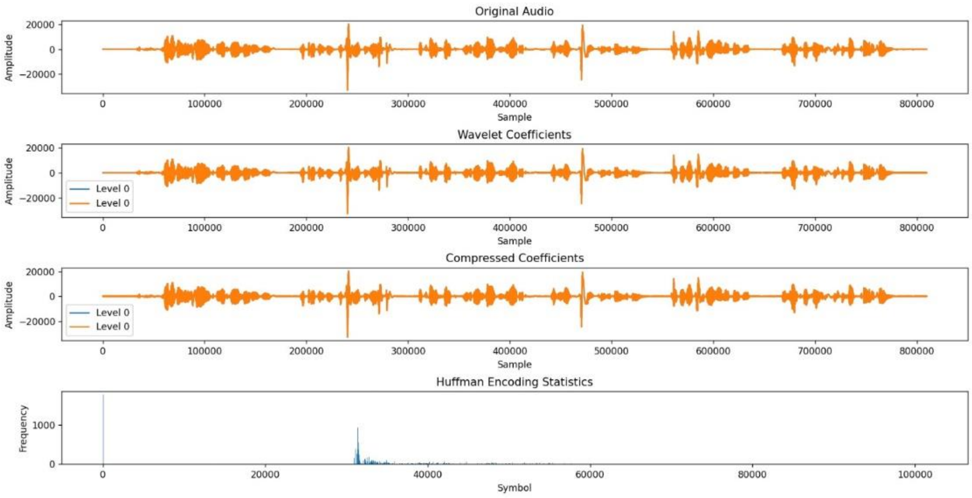

Figure 2 Compressed Speech Waveforms

Wavelet Coefficients:

This plot displays the wavelet coefficients obtained from the wavelet transform of the audio signal. The wavelet transform decomposes the original audio signal into different frequency components across multiple scales. The plot consists of multiple subplots, each representing a level of the wavelet decomposition. The x-axis represents the sample index, and the y-axis represents the amplitude of the wavelet coefficients.

```
Compressed Audio Data:
[[ nan  nan]
 [ nan   1.]
 [ -1.  -1.]
 ...
 [-57. -57.]
 [-57. -58.]
 [-61. -60.]]
```

Figure 3 Compressed audio details

Compressed Coefficients:

This plot shows the wavelet coefficients after applying adaptive thresholding. The adaptive thresholding process reduces the noise-like or less significant components of the wavelet coefficients while retaining the important features of the audio signal. Similar to the "Wavelet Coefficients" plot, this plot also consists of multiple subplots, each representing a level of the wavelet decomposition. The x-axis represents the sample index, and the y-axis represents the amplitude of the compressed wavelet coefficients.

```
Number of Bits before Compression: 12952128
Number of Bits after Compression: 8
Compression Ratio: 6.176591213428404e-07
```

Figure 4 Compression result

Huffman Encoding Statistics:

This plot displays the statistics of the Huffman encoding process used to encode the compressed wavelet coefficients. The plot shows the frequency of occurrence of each unique tuple coefficient after applying the adaptive thresholding. The x-axis represents the symbols (unique tuple coefficients), and the y-axis represents their corresponding frequencies.

Future Scope:

The research also considers additional aspects such as computational complexity, perceptual quality, and subjective evaluations. Computational complexity provides insights into the feasibility and efficiency of implementing the compression technique in real-time systems or resource-constrained environments. Perceptual quality assessments, including listening tests, measure the perceived similarity between the original and reconstructed speech signals, considering factors like speech intelligibility, naturalness, and overall quality. By considering these factors, the research aims to provide a comprehensive evaluation of the proposed compression technique, optimizing the trade-offs between compression efficiency and signal fidelity for practical applications.

Conclusion

In conclusion, this research explores the application of wavelet transform and Huffman coding for speech compression. Voice signals were considered, and the compression performance was evaluated based on metrics such as compression ratio, distortion, and retained signal energy. The results indicate that wavelet transform with Huffman coding achieves a higher compression ratio compared to wavelet transform alone, while maintaining low distortion levels. The reconstructed waveforms closely resemble the original signals, and the retained signal energy demonstrates the effectiveness of the compression techniques. The study provides valuable insights into efficient speech compression for storage and transmission, paving the way for improved utilization of resources and enhanced perceptual quality in speech communication systems.

References

[1] Speech Compression with Wavelet Transform and Huffman Coding Ranjushree Pal, Department of Electronics and Telecommunication Engineering, Dwarkadas J Sanghvi College of Engineering, Mumbai: 400056, India.

[2] L. L. Motta, R. S. Mendes and M. H. M. Costa, "An Audio Compression Method Based on Wavelet Packet Decomposition, Ordering, and Polynomial Approximation of Expressive Coefficients," 2019 IEEE Latin-American Conference on Communications (LATINCOM), Salvador, Brazil, 2019, pp. 1-6, doi: 10.1109/LATINCOM48065.2019.8937917.

[3] Speech Compression Exploiting Hamming Correction Code Compressor Islam Amro Faculty of Information Technology Al-Quds Open University Ramallah.

[4] A. Debnath, U. K. Mondal, B. Bikash Roy and N. Panja, "Achieving Lossless Audio Encoder through integrated approaches of Wavelet Transform, Quantization and Huffman encoding (LAEIWQH)," 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA), Gunupur, India, 2020, pp. 1-5, doi: 10.1109/ICCSEA49143.2020.9132865.

[5] M. S. Serina and S. G. Mosin, "Digital Audio Information Compression Using Wavelets," Proceedings of the International Conference Mixed Design of Integrated Circuits and System, 2006. pp. 660-664, doi: 10.1109/MIXDES.2006.1706666.
[6] M. H. Martínez, N. F. Rosas and C. Camargo, "Compressing audio signals with the use of the wavelet transform and Embedded Systems," Symposium of Signals, Images and Artificial, Bogota, Colombia, 2013, pp. 1-5, doi: 10.1109/STSIVA.2013.6644947.
[7] Wavelet transform for speech compression and denoising Fatma zohra Chelali* and Noureddine. Cherabit, University of Science and Technology Houari Boumediene of Algiers, ALGERIA Amar. Djeradi and Leila.Falek
[8] G. Kemper and Y. Iano, "An Audio Compression Method Based on Wavelets Subband Coding,"
[9] A. K. Sinha, S. K. Dutta, B. Dev, Jaydipta, R. Ranjan and R. Kumari, "Wavelet based Speech Coding technique using median function thresholding," 2014 International Conference on Electronics and Communication Systems (ICECS), Coimbatore, India, 2014, pp. 1-4, doi: 10.1109/ECS.2014.6892538.
[10] A Comparative Study between Discrete Wavelet Transform and Linear Predictive Coding D.Ambika, V.Radha
[11] F. z. Chelali, N. Cherabit, A. Djeradi and L. Falek, "Wavelet transform for speech compression and denoising," 2018 6th International Conference on Multimedia Computing and Systems (ICMCS), Rabat, Morocco, 2018, pp. 1-7, doi: 10.1109/ICMCS.2018.8525996.
[12] Continuous Speech Coding Using Coiflets Wavelet. Shijo M Joseph1 Babu Anto
[13] J. Pang, S. Chauhan and J. M. Bhlodia, "Speech Compression FPGA Design by Using Different Discrete Wavelet Transform Schemes," Advances in Electrical and Electronics Engineering - IAENG Special Edition of the World Congress on Engineering and Computer Science 2008, San Francisco, CA, USA, 2008, pp. 21-29, doi: 10.1109/WCECS.2008.11.