IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, July 2023

Audio Compression on Multimedia Compression Techniques

Hema P[1] and Kala L[2]
Student, NSS College of Engineering, Palakkad, India[1]
Associate professor, NSS College of Engineering, Palakkad, India[2]

Abstract: *Recently, the study of audio compression has become more prominent. Today's apps make use of advancements in audio signal processing, including advanced audio coding (AAC), perceptual audio coding techniques (MP3 encoding), internet radio, and other lossless audio coding systems. In this essay, we provide a summary and contrast the algorithms Huffman and Arithmetic. The method is used with a comparison of the two algorithms, Huffman and Arithmetic, on raw WAV files. The fundamental objective of the steganography technique is to increase the security of the transmitted data. Unauthorized users are unable to access or misuse the steganographic file. Audio steganography is applicable to non-technical areas as well in order to protect the privacy and confidentiality of the data. This paper also includes a review of recent work on audio steganography. For database applications like storage and transfer, audio compression is essential. This essay discusses how compression techniques are applied to audio using various transform coding and focuses on the advantages of transform coding in compared to current methods. By using the transform coding technique, numerous attempts have been made to completely eradicate or minimize audio noise, and the study has produced a number of fruitful results. Examining current fractal coding techniques for digital multimedia compression is the aim of the fractional compression study. It discusses strategies for reducing encoding time, which is considered to be the primary challenge in fractal compression, as well as suggested fractal coding techniques in the audio and image domains. In order to use the most widely used communication platforms, large audio file sizes must be transferred via digital audio; this poses substantial challenges for storage and preservation. This article makes good use of audio.It is advised to use a mixed transform coding scheme as a compressive method. The audio file size was greatly decreased while maintaining high quality, and the compression results are promising. The cascaded prediction approach was improved and reportedin Audio Compression utilizing OLS+ and CDCCR Method. Comparing three primary predictor block types with varying levels of complexity included two complicated prediction tech- niques with backward adaptation, namely extended Active Level Classification. Extended Ordinary Least Square (OLS+) and the Extended ALCM+ Model.*

Keywords: AAC, MP3 Encoding.

I. INTRODUCTION

Signal processing uses the method of compression. Becauseenormous volumes of data are frequently sent across a net-work's communication channel, it is of utmost significance. For many sorts of information, including music, video, pic-tures, and text, data compression is necessary. Compressing audio (speech) signals makes transmission and receiving con- siderably simpler, which enhances communication. Since the amount of the data grows larger and more complicated as it is extended, it is essential to apply compression techniques, which must be supported by advancements in continuous bandwidth technology. Humans by nature favour data of the highest quality and smallest possible size (quantity). When taking into account the aforementioned problems, the approachis to maximise the current compression by lowering the number of slots that the compressed data occupies. This paper gives a review of various audio steganographic techniques. Reducing the number of bits required to represent an audio signal with the purpose of decreasing memory storage costs and transmission bandwidth requirements is the main driving force behind the development of speech/audio compression systems. The fundamentals of audio compression rely on eliminating redundant parts of the stream while maintaining its clarity. Fractional Compression study study discusses the

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, July 2023

improvement on the fractal coding and shows the performance of applying it on image and audio files. Utilizing DCT and wavelet transform, several techniques have been proposed for the digital compression of audio signals. Following his study of the usage of DCT and DWT to compress voice signals, Kaur used DCT to the signal before using DWT to decode the encoded data. When existing bits are removed, digitized data may often be represented by a smaller number of bits.

II. PROBLEM DEFINITION AND ALGORITHM DEVELOPMENT

Waveform audio files are compressed in this study usinga lossless technique. This method makes use of a variety of algorithms, including Huffman and Arithmetic. In the Huffman approach, alphabetical symbols are initially listed in order of likelihood. This results in a tree with symbols on each leafthat grows from bottom to top. The two symbols with the lowest probability are selected at each level, added to the top of the incomplete tree, checked off the list, and replaced with auxiliary symbols that stand in for the two original symbols. This is performed through a succession of phases. Once there is just one auxiliary symbol left on the list, the tree is finished. The tree is then passed in order to determine the symbolcode[1].

2.1 Huffman Algorithm

For lossless data compression, the Huffman algorithm, an effective prefix code, is widely used. Finding or using the codecontinued thanks to Huffman coding. A variable-length code table created by the Huffman method can be used to encrypt source symbols. Similar to other entropy coding methods, the more often used symbols are generally represented using fewer bits than the less widely used symbols. The Huffman technique may successfully be used and find the code in linear time withthe number of input weights if the input weights are sorted. Huffman encoding is not necessarily the best when compared to other compression techniques, even though there are partial encoding methods. The first stage in the Huffman algorithm flow is to assign a portion of the code from each piece ofdata to create a variety of symbols. [1]The subsequent phase is written as the initial sign. This symbol must be included to the code before the following symbol. It will ascertain whetherthe last symbol is still there. The repeating process is initiated by adding extra code to the succeeding symbol if the last symbol is missing. The last symbol is verified once again. If this last symbol is not present, the procedure is finished. The vertex and base of the node must then be filled in with the integers 0 and 1, respectively.

2.2 Arithmetic Coding

Since the technique has successfully replaced the Huffman coding for 25 years, arithmetic coding has a significant history. Arithmetic coding is a type of entropy encoding, which changes data into a different form by seldom using more bits and more frequently using a small number of characters. With this method of coding, the input message is divided into symbols, and each symbol is exchanged for a floating- point number. [1]A fractional integer n is used in arithmetic coding to represent the complete message, where (0.0 n 1.0). A series of entered symbols are replaced with floating point via the arithmetic coding process. Given the length and complexity of the encoded messages, certain bits are required to meet this demand. All real numbers from 0 to 1, including
0 but excluding 1, are represented by a series of symbols into a single number with interval (0.1) and notation (0.1), alternatively it may be expressed as 0 = x 1. Since this numbercan only be deciphered once, it produces a series of symbols that are utilised to produce these numbers. The encoding is described as follows:

Set low = 0.0 (initial conditions)
Set high = 1.0 (initial conditions)
While (input symbol is still there) do
Take the symbol input.
CR = high – low.
High = low + CR*high_range(symbol)
Low = low + CR*low_range(symbol)
End while
Print low

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.301

Volume 3, Issue 1, July 2023

III. AUDIO COMPRESSION SYSTEM

Utilizing lossless audio compression techniques like FFT,DWT, and the DCT, audio signals may be compressed. It is still difficult to reconstruct the original audio signal from the compressed version using these approaches due to its low CR, simplicity, and PSNR. The DCT method is used and improved for better compression. The updated DCT algorithm was created using the C programming language and Octave. The signal was converted from the time domain to the frequency domain using DCT, and samples with an energy level of less than 1 were eliminated. Additionally, samples that had a lower frequency than the sample frequency before them were hidden. The Octave Signal Processing toolbox was used for all simulations, and the C programming language was used for all coding. Real-time data was gathered at the output while the Arduino Microcontrollers were employed to regulate signal transmission[2].

3.1 The System Requirement

The Octave Signal Processing toolbox or the MATLABSignal Processing (MSP) toolkit were both used throughout this work to do all Digital Signal Processing (DSP). Utilizing an Arduino microcontroller that had been C-programmed and external Octave scripts, the transmission simulations were run.For a more reliable outcome, several voice signals with various sampling rates were employed. For ease of processing, each sampled voice signal is saved in.wav files[2]. Software built in the C programming language and Octave were used to process the signal further. FFT, which uses floating point format, is done on some of the arrays. The C programming language made it simple to process the data included in the voice signals. The signal must be scaled down if it exceeds the specified threshold since the MSP values must all be kept between -1 and 1. (-1 and 1). The DSP is in charge of scaling. The technique described above is simply reversed in the compressed audio signal.

3.2 The System Architecture

There are three key components to the DCT audio compression technology as a whole. Together, these system components achieve the intended outcome. The signal processing device (SPD), the transmitting unit, and the receiving unit are the three primary components of the system[2].

3.3 Compressing the audio input signal with the SPD

With the aid of the signal processing toolbox, the audio file was compressed. The wav file was read back into memory as the initial action. At immediately, the audio (voice) stream was divided into windows (small parts), and each window underwent its own compression process. The coefficients in each window correspond to audible frequencies. We also take into account the data's binary format. The DCT findings might have been saved as floating point numbers, however it would need 32 bits per coefficient and produce incredibly large values. Typically, wav files are saved as 16-bit integers. The coefficients were linearly translated into 16-bit integers after being converted to the frequency domain. As a result, in addition to their coefficients, there indices were also kept. The indices of the coefficient are also recorded as 16-bit integers. The auditory masking filter is used to further eliminate inaudi- ble samples once the inaudible samples have been eliminated from the new audio stream. The idea of audio masking and how the human ear works served as the foundation for the development of the auditory filter. To verify if any of the three coefficients were hidden, the filter repeatedly evaluated the previous and following coefficients of the new audio signal in the frequency domain[2].

IV. AUDIO STEGANOGRAPHIC METHODS

Watermarking audio signals is difficult because we must pay close attention to certain factors, such how much the signal's quality should not be affected by the watermarkingof messages. In order to keep noise levels low and audio signal quality high, the human auditory system's sensitivityto additive white Gaussian noise, or AWGN, should be as lowas 70 dB below ambient level. The information underlyingan audio cover file may be concealed using a wide range of ways without degrading the signal quality. These methods give the sender the ability to conceal data so effectively that any changes made to the audio file are undetectable to outside parties.Audio steganographic approaches can be categorised using a variety of domains, including time-based, frequency- based, transformation-based, and others. These domains are further broken into subcategories, such as the transform do- main, which is further separated

DOI: 10.48175/IJARSCT-12020

ISSN 2581-9429 IJARSCT

into the wavelet and frequency domain. The primary related approaches for each of these domains are defined as follows[3]:

4.1 LSB Coding

The least significant bit (LSB) algorithm is the oldest technique for obscuring audio data. By changing the least significant bits in the cover file to include a series of bytes,the user may turn an image into audio and audio into a picture using this approach. By rearranging the information message before hiding it in the audio document, this is done to increasesecurity. To protect the data from an attacker, this is done. If the attacker discovers the secret data, he will not be able to decipher the data's originality because it is in a coded format. The criteria for file selection are used to generate random sequences in order to conceal data. When there is a highchannel bit rate and the LSB replacement doesn't result in quality loss, that process is typically successful. The LSB is thebit location in a binary integer that, in relation to the available unit value, determines whether a number is odd or even in terms of computation. As a result, its complexity is smallerand data concealment and extraction take place faster and withless delay. The optimal LSB coding data transmission rate is 1 kbps every 1 kHz. Two message bits are used in place of a sample's two least significant bits.

4.2 Parity Coding

A reliable audio steganographic method is parity coding. Instead of breaking the signal into individual samples, this approach splits the signal across several samples and embeds each bit of the secret message from a parity bit. Assume that if someone wishes to decode the signal and the parity bit of a chosen area does not match the secret bit to be encoded, they will invert the LSB of one of the samples and have one more option for encoding the secret bits information since it cannot be easily encoded in another manner. The biggest drawbackis that they are weak and worry that some data may be lost if the information is resampled[3].

4.3 Phase Coding

This kind of watermarking is also effective. The user then substitutes the phase of an initial audio signal with a referencephase carrying concealed data after we have first converted themessage into blocks and then embedded it in phase. Using the remaining phase, the relative phase between different segmentsis changed and organised. Regarding the signal to noise ratio, this coding process is regarded as a powerful and effective one. When there is a significant change in the phase and frequency components, this method exhibits a modest phase dispersion. However, unclear coding is still possible if the phase variationis minimal[3].

V. AUTO ENCODER (AE) MODEL

There are two components to the visual data compression module. Encoder AE was used to initially process the original lip pictures. The AEmodel, which comprises of three 2D convolutional layers as the encoder and three 2D transposed convolutional layers as the decoder, is trained in a self-learningway in advance. The hidden representation of the AE hasdimensions of 32*8*8, which, when flattened, result in 2048 dimensions. The visual characteristic is this latent representation, which is just 16.67 the size of the original lip picture (3*64*64). To further minimise the amount of bits required foreach feature piece, Qualatent processes the extracted visual features. The SE model then makes use of the compressed visual characteristics. The EncoderAE and Qualatent modules may be fitted in the visual sensor in real-world applications to improve the effectiveness of online computing. By blurring the pictures of lips and reducing transmission costs using a two- stage visual data compression technique, the issue of privacy can be somewhat addressed.

VI. SPEECH AUDIO COMPRESSION SYSTEM

The audio is initially in the spatial domain, which is difficult for audio processing and compression, and must be converted into the frequency domain, where the majority ofthe audio information is found. For this reason, the audio compression system employed the Wavelet Transform (DWT) and Discrete Cosine Transform (DCT) approaches. While em- ploying different encoding systems, compression of the signal is accomplished via transform algorithms, which only offer information about the signal. By disregarding tiny magnitude coefficients as insignificant data and then eliminating them, compression is accomplished[5]. The first unit in the audio compression system is the encoding unit, and the second isthe decoding unit. Each unit is completed in a variety ofphases. the discrete cosine transform, cosine packet trans- formations, discrete wavelet transform, and wavelet packettransform are some of the transform functions (techniques) used to breakdown the input speech/audio signal into differentresolution or frequency bands in the initial stage. Following decomposition, quantization is used to decrease the information obtained in the transform coefficients in a mannerthat produces perceptually error-free compression. There are two different types of quantization: uniform and non-uniform quantization[5].The encoding method is used to remove data that repeat themselves. The encoding method is used to removedata that repeat themselves. By removing superfluous datafrom the encoding process, we may also reduce the numberof coefficients. This aids in reducing the signal's bandwidth so that compression may be accomplished. The decoding unit entails doing the opposite procedures from those used in the encoding process, as well as performing these operations in the opposite sequence. For the transform-based audio compressionapproach, the performance is measured in terms of NMRSE, SNR, RSE, PSNR, and compression ratio (Factor)[5].

VII. FRACTAL COMPRESSION THEORY

The mathematician Benoit B. Mandelbrot is responsible for the history of fractal compression. In his book The Fractal Geometry of Nature from 1977, he used the term "fractal" for the first time in the context of geometry. The fundamental ideabehind FIC is to use a series of contractive transformationsto describe a picture by taking advantage of how similar the individual image pieces are to one another using PIFS.The decoding process is a simple process, but the encoding process has three subprocesses: first, dividing the image into non-overlapping range and overlapped domain blocks; second,looking for similar range-domain with the least amount of error; and third, storing the IFS coefficients in a file that represents a compressed file and using it decompression pro- cess.The asymmetric method, high compression ratio, superb reconstructed picture quality, and quick, simple decoding are only a few of the qualities of FIC. The thorough search methodneeded for sophisticated computations makes the encodingprocess time-consuming. Numerous methods to get around thelengthy encoding time have been suggested in literature.It is important to note that while numerous research in the literature revealed improvements in FIC, only a small number of them used the conventional fractal image model on the audio input, as demonstrated in Section 6. This restriction is brought on by a number of challenges that the FIC must overcome, rendering it impractical. The inability to detect self-similarity in audio signals using conventional fractal coding is another issue, asis the magnitude of the audio signal in relation to visuals. However, a number of methods were used in conjunctionwith fractal coding to reduce the amount of time needed for encoding audio[6].

VIII. CASCADE CONNECTION OF PREDICTION METHOD

This uses a cascade connection of distinct five phases (blocks) based on prediction models with backward adaptation. The first block utilizes the DPCM constant predictor, but the second block, for example, employs the RLS technique with an order. The outputs of the next blocks rely on thevalues estimated in the preceding blocks. The final three blocks employ the NLMS approach, using lower prediction orders in each succeeding block.The prediction order in the middle NLMS block has been raised in the solution that is being offered here, which is an extension of the codec from work. An upgraded version of OLS+, whose streamlined form was previously successfully employed in lossless image compression, has been added in place of the OLS block. The OLS+ block has been included in place of the DPCM and ALCM+ blocks in the version with increased complexity. The prediction errors e(n) are effectively coded into the resultant binary data stream using the final two blocks of the cascade audio data compression scheme that we propose[8].

XI. CONCLUSION

When applied to text and picture data, lossless compression and audio data compression provide different results. Thus, neither method is necessarily the best choice for a given set of data. It still has to be improved in terms of properly allocating bits. The goal of this research is to compare all lossless compression techniques with the same data in the

ISSN 2581-9429 IJARSCT

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, July 2023

future in order to provide a clearer picture of the outcomes of each approach. A redesigned audio compression method utilising DCT is presented in the audio compression system work. The outcomes of this work will significantly enhance audio compression systems, particularly for upcoming uses that need for highly compressed audio signals of excellent quality. In steganography Although LSB coding is a useful technique, it lacks sufficient security. Therefore, a strategy is presented to offer data security in which data is first encoded before being hidden. Signals are encoded through the process of encryption so that a third party cannot interfere with the data. After carefully examining several strategies, it has been determined that a strategy has to be implemented in order to more securely hide data and prevent other parties from being able to do so. Privacy issues might be resolved by combining the Qualatent and Encoder AE modules. We anticipate that the results of this study will be valuable in putting the DL- based AVSE model into practise. We'll look at time-domain compression in the future to improve the online computing efficiency of the suggested LAVSE system. We have discovered that the compression method, such as lossy and lossless, and their coding techniques are the best in their respective domains after completing audio compression using a variety of transform coding approaches. These strategies accomplish the goals of expanding storage capacity and decreasing noise and bandwidth. We draw the conclusion that the effectiveness of compression methods like wavelet depends on the audio quality and computational complexity. In the future, several transform coding algorithms can be merged to enhance the audio file's compression ratio and PSNR. This work has reviewed fractal image and audio compression as well as their boosting methods. For these two approaches to stand side by side with the counterpart, more investigation is neededin the fields of fractal audio compression in particular and fractal picture compression broadly. The cascading audio data encoder is enhanced in the multistage coder work. Arithmetic encoders and adaptive variants of Rice are both commonly employed in traditional lossless audio coding systems. The suggested technique makes use of an adaptive Golomb code,a generalized version of the Rice code that may be moreefficient in compressing data because it is better adapted to theprobability distribution of the data that is now being encoded

REFERENCES

[1]. Moreno-Alvarado, R., Rivera-Jaramillo, E., Nakano, M., Perez-Meana, H. (2020). Simultaneous audio encryption and compression using compressive sensing techniques. Electronics (Switzerland), 9(5). https://doi.org/10.3390/electronics9050863

[2]. Ali, A. H., Abbas, A. N., George, L. E., Mokhtar, M. R. (2019). Image and audio fractal compression: Comprehensive review, enhancements and research directions. Indonesian Journal of Electrical Engineering and Computer Science, 15(3), 1564–1570. https://doi.org/10.11591/ijeecs.v15.i3.pp1564-1570

[3]. Chuang, S.-Y., Tsao, Y., Lo, C.-C., Wang, H.- M. (2020). Lite Audio-Visual Speech Enhancement. http://arxiv.org/abs/2005.11769

[4]. Ahmed, Z. J., George, L. E., Hadi, R. A. (2021). Audio compression using transforms and high order entropy encoding. International Journal of Electrical and Computer Engineering, 11(4), 3459–3469. https://doi.org/10.11591/ijece.v11i4.pp3459-3469

[5]. Al-azawi, R. J., Drweesh, Z. T. (2019). Compression of audio using transform coding. Journal of Communications, 14(4), 301–306. https://doi.org/10.12720/jcm.14.4.301-306

[6]. Ulacha, G., Wernik, C. (2019). A high efficiency multistage coder for lossless audio compression using OLS+ and CDCCR method. Applied Sciences (Switzerland), 9(23). https://doi.org/10.3390/app9235218

[7]. Mohan, J., Gupta, A., Jaypee Institute of Information Technology University, Institute of Electrical and Electronics Engineers. Uttar Pradesh Section. SP/CS Joint Chapter, Insti- tute of Electrical and Electronics Engineers. (n.d.). 2019 Inter- national Conference on Signal Processing and Communication (ICSC): 07-09 March 2019, Jaypee Institute of Information Technology, NOIDA.

[8]. Verma, M., Saini, H. S. (2019). Analysis of Vari- ous Techniques for Audio Steganography in Data Security. www.ijsrnsc.org

[9]. Tsegaye, A., Tariku, G. (2019). Audio Compres- sion Using DWT and RLE Techniques. American Jour- nal of Electrical and Electronic Engineering, 7(1), 14–17. https://doi.org/10.12691/ajeee-7-1-3

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.301

Volume 3, Issue 1, July 2023

[10]. Hidayat, T., Zakaria, M. H., Pee, A. N. C. (2018). Com- parison of lossless compression schemes for WAV audio data 16-bit between huffman and coding arithmetic. International Journal of Simulation: Systems, Science and Technology, 19(6), 36.1-36.7. https://doi.org/10.5013/IJSSST.a.19.06.36

[11]. G. Ulacha and C. Wernik, "A High Efficient Cascade Coder with Predictor Blending Method for Lossless Audio Compression," 2020 Data Compression Conference (DCC), Snowbird, UT, USA, 2020, pp. 395-395, doi: 10.1109/DCC47342.2020.00058.

[12]. S. R. Quackenbush and J. Herre, "MPEG Standards for Compressed Representation of Immersive Audio," in Proceedings of the IEEE, vol. 109, no. 9, pp. 1578-1589, Sept. 2021, doi: 10.1109/JPROC.2021.3075390