A Survey of Transform Coding for High-Speed Audio Compression

Zinah J. Jabbar[a] ,Dr. Loay E. George[b] *

[a] Informatics Institute for Postgraduate Studies,Iraqi Commission for Computers and Informatics,Baghdad, Iraq.Zenajamaljabar@gmail.com

[b] University of Information Technology And Communication ,Baghdad ,Iraq. Loayedwar57@uoitc.edu.iq

A R T I C L E   I N F O

A B S T R A C T

The necessity for audio compression continues to be a critical issue., as it is associated with lowering the data capacity of one of the most widely distributed digital files transmitted between remote parties, audio compression remains a critical topic. The effectiveness of the two sound pressure units in this article has been tested. Because of this   industry's importance,, which is reflected in storage capacity and transportation needs, Over the last two decades, sound pressure has been the topic of intense investigation. The rapid expansion of the computer industry is leading to a growing demand for high- quality audio data; As a result, the development of sound compression methods, of which there are two types: lossy and lossless, is critical. The purpose of this study is to explore sound compression techniques and to highlight the suitability and applications of each approach.

MSC.41A25; 41A35; 41A3

1. Introduction

 Voice is the most essential form of communication between people, and as a result, speech data compression is a critical issue that is viewed as a problem, particularly given the rapid advancement of communication technologies [1]. Voice compression is the process of compressing human speech signals into a compact representation, with the result of the decoding process being the most precise replication of the original signal. [2]. Any compression technique is motivated by the need to minimize the amount of data necessary to represent the original data, resulting in quicker data transfer and lower storage costs [3]. For every type of data, there are two types of compression algorithms: lossless and lossy compression. The signal that was sent into the encoding unit was created using a lossless approach. [4]. The loss type is present when the returned signal does not match the signal supplied in this scenario. The lossless technique must be used for medical and satellite text and photos, but the lossy method can be used for audio, video, and other sorts of images [5]. In general, the lossy kind has a higher gain in compression ratio, [6] but the quality of the returned signal is unquestionably lower. The core of any

compression technique is to remove repetition in the input signal [7]. Compression is a crucial signal processing method that is important because enormous volumes of data are frequently exchanged via a network's communication channel. To employ data compression, several forms of information, such as audio, video, pictures, and text, are required [8]. Voice compression is a technique for modifying human voice in a coded frame such that it can be returned to its original state, hence decreasing duplication between close samples and frames. The goal of audio compression is to use less storage space in order to reduce transmission speed by encoding audio data. To accomplish this, many compression methods have been devised. The two methods of audio compression are lossless and lossy compression, which may be classed similarly to any other digital data compression.. The three functional classes of audio compression techniques are as follows [9]:

1. Forms those are direct,

2. Forms for extracting parameters,

3. Types of transformation

When the signal's samples are directly treated to provide compression, the direct form is used; when a signal is preprocessed to extract certain features that are then utilized to reconstruct the signal, the parameter extraction form is used. In the realm of audio signal processing, transformation types Discrete Fourier Transform (DFT), discrete Cosine Transform (DCT), and discrete Wavelet Transform (DWT) are examples of discrete wavelet transforms [10]. DCT is frequently employed for signal compression, particularly where there is a high correlation and the signal can be quickly rebuilt with minimal fidelity error [10]. Because of its localization property in time-frequency space, DWT is suitable for signal compression. Audio is an electronic form that may be used to represent sound. Sound that is audible to humans is referred to as audio. The human ear can distinguish between frequency ranges of 20 Hz and 20 kHz. There are two factors that contribute to the popularity of audio compression [11].

People's unwillingness to throwing anything away and their need to collect data; regardless of how large a storage device is, it will eventually overflow; data compression appears advantageous since it delays this eventuality.

People despise having to wait a lengthy time for data to be sent. We consider anything more than a few seconds to be a long waiting time when waiting for a file to download or a web page to load [12].

The primary objective for speech/audio compression systems development aims to lower the amount of bits required to represent an audio signal in order to save money on memory storage and bandwidth. The basic principle of audio compression is to remove redundant signals while keeping signal quality and clarity. [13].

2. Types of Audio Files

Digital audio data is stored in a variety of file forms, just like other digital media. When choosing a file type, keep in mind the file format's universality and hence its reading by a range of software programs. File formats that are proprietary and unlikely to be supported in the future should be avoided. Here are a few examples of well-known audio file types [14]:

2.1. WAV: Microsoft created this file format. It is widely used and can be read by most audio software products. The WAV file format has established itself as a standard and is highly recommended. Furthermore, the WAV file type is also available in a professional format (i.e., broadcast WAV, BWF) that allows information to be included in the file header. Despite the fact that not all audio software packages can currently read or write to the information header, the BWF format is quickly becoming the WAV file type of choice for archiving audio projects. [15]

2.2. MP3: This is the file extension's name, as well as the type of MPEG-audio layer3 file it is. Layer 3 is one of three coding algorithms for compressing audio streams. (Layer 1, layer 2, and layer 3). To remove any undesirable information, perceptual audio coding and psychoacoustic compression are used (especially redundant and irrelevant aspects of a sound source). Things that the human ear can't hear in the first place. It also has a filter bank with an MDCT (Modified Discrete Cosine Transform) that boosts frequency resolution by

18 times throughout layer 2. Layer 3 can shrink the original audio-data on a CD (With stereo music at a bit rate of 1411.2 kilobits per second) by a factor of 12 (i.e., down to 112/128 kbps) without sacrificing sound quality [16].

2.3.WMA : (Windows-MEdia-Audio) is a Microsoft-file format for encoding digital audio files that likes MP3 but has the potential to compress data more quickly. WMA files, which end in ".wma," may be any size and compressed to accommodate a wide range of connection speeds and bandwidths [14].

2.4.Real Audio (rm-ram-raThe streaming audio format is a one-of-a-kind audio format that allows you to listen to digital audio files in real time. RealPlayer is required to open this sort of file (for Windows or Mac) [14-15].

3. Digital Sound Representation.

One of the first things to consider when establishing a data technique for sound waves, such as speech or music,this is how audio is represented digitally. The term "audio" signifies sound that is audible to humans (20Hz to 20 KHz) [17]. In nature, an audio signal is analog. Analog sound is made up of waves that are sensed by the human ear. Both in terms of time and amplitude, these waves continue. The height or (volumes) of the sound is represented by amplitude. The analog signal should be transformed to digital format so that it may be stored, processed, and were passed via computer networks. An A/D converter is required to convert analogue to digital signals. The sampling and quantization steps make up the A/D conversion process [18].

3.1. Sampling: Sampling is the process of taking periodic measurements of an analog signal and using the results (samples) rather than the original signal. Figure (2) shows a sampled wave example; Often, binary numbers are used to represent two samples, although They may also be saved in a variety of formats. Pulse Code Modulation is a well-known method of representing each sample with a series of pulses that encode the sample's binary code (PCM). There are other forms of modulation, but PCM is the most used in digital audio. Different modulation schemes are unimportant to a programmer. The consecutive binary values are simply recorded as integers in a computer's memory. For most programmers [19].
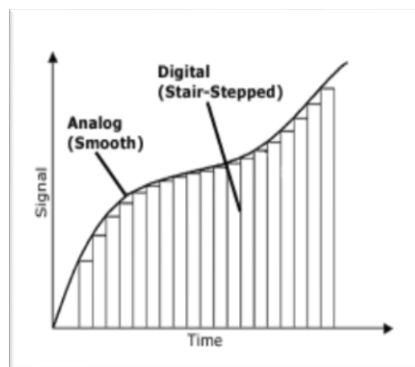

Fig (1): A typical example of Digital Sampled Wave [20]

3.2. Quantization: The phrase "signal quantization" refers to the process of accurately calculating the signal value. The digital representation is finite due to the finiteness of computer capability. If an 8-bit or 16- bit integer is utilized, a discrete integer sample value of 256 (2^8) or 65,536 (2^{16}) can be produced, Despite the fact that the original samples are not integers. Quantization is the conversion of a precise sample value to a less precise number. The quality and resolution of digital audio are influenced by two factors[21]:

A. The sampling rate is the number of times per second that the amplitude of the wave is sampled.

B. The "bit depth," or the range of integers utilized to record each measurement.

The first value, the "sampling rate," is expressed in kilohertz, which translates to thousands of samples per second. The sampling rate is set at 44.1 kHz to record consumer audio CDs. That implies each second of audio is recorded by 44,100 individual amplitude measurements, as seen in figure (3), which depicts the wave flows [22].
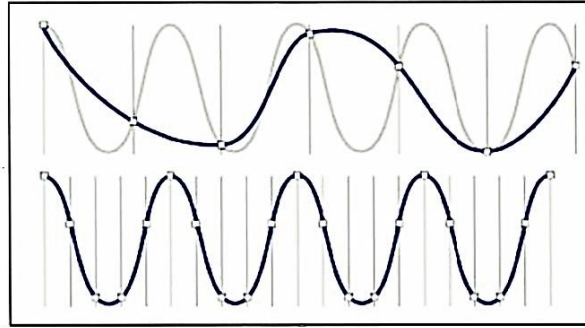

Fig (2): The visual depiction of two different sampling rates [22].

4. Compression of data.

The procedure of encoding information with fewer bits than the original representation is known as data compression, source coding, or bit-rate reduction in signal processing. The two types of compression are lossy and lossless compression[22].

By discovering and removing statistical redundancy, lossless compression decreases bits. When utilizing lossless compression, no data is lost. By removing unnecessary or less relevant material, lossy compression reduces the number of bits in a file. An encoder is a device that compresses data, whereas a decoder is a device that reverses the process (decompression) [23]. The phrase "data compression" refers to the process of shrinking a data set's size. Source coding is a document that is encoded at the source of the data before it is saved or delivered in the context of data transmission. Source coding is not to be confused with channel coding, which is used to detect and repair faults, and line coding, which is used to map information onto a signal [24]. Compression is advantageous since it reduces the amount of data storage and transmission necessary. The operations of compression and decompression use computer resources [25]. When compressing data, there is a trade-off between space and temporal complexity. A video compression approach, for example, may necessitate costly technology to decompress the ability to decompress the film in It may be inconvenient or demand extra storage if it's full before watching, thus it may be inconvenient or require additional storage to decompress the film in its entirety before watching it [26]. When building data compression systems, all aspects to consider include the compression level, while using lossy data compression, the amount of distortion produced, and the computational resources required to compress and de-compress the data [27].

5. DCT Transform

A collection of points of data is represented as the sum of cosine-functions pulsating at different frequencies. Nasir Ahmed proposed the NDCT conversion method in 1972, and it is now widely utilized in signal processing and data reduction [28]. Digital photos (such as JPEG and HEIF, which ignore minor It's used in digital video (including H.26x & MPEG), audio in digital format (including Dolby-Digital, MP3 & AAC), and television in digital format (such as Dolby Digital, MP3 & AAC) [29]. Examples include Standard Definition TeleVision (SDTV), High Definition TeleVision (HDTV), Video on Demand (VOD) , digital radio (e.g., AAC Plus, DAB Plus), and coding of speech (e.g, AAC- LD, Siren, Opus). Many additional applications in research and engineering rely on discrete cosine transformations, including processing of digital signals, communication devices, Network bandwidth reduction, & spectroscopic approaches for numerical resolution of PDE devices [30].

Compression requires the use of cosine rather than sine since it turns out (as shown below) that fewer cosine functions are necessary to attain the target. A typical sign, whereas the cosine conveys the sign in differential equations. A certain set of boundary conditions the DCT is a Fourier transform-related transformation that functions in the same way as the

Discrete Fourier Transform (DFT) but it only works with real numbers. The coefficients of periodic and symmetric extended Fourier series are often connected to DCT transformations, whereas the coefficients of periodic and symmetric extended Fourier series are generally related to DFT transformations [31]. DCTs are nearly twice as lengthy as DFTs and work with real data in a symmetrical manner (since the fourier transform of an even and real function is an even and real function), Certain transformation versions, on the other hand, offset the input or one or both by half a sample. There are 8 DCT versions in all, four of which are often used [32]. The second type DCT, sometimes known as "dct," is the most frequent discrete cosine transform. Nasir Ahmed proposed the DCT for the first time in 1972. Its polar opposite, type III DCT, is commonly referred to as type III, and it, like type II, is referred to as "inverted DCT" or "IDCT." The Modified Discrete Cosine Transform (MDCT), it is based on overlapping data DCT, and the Discrete Sine Transform (DST) , which is analogous to a DFT of single and real functions, are 2 of the DCTs linked with it. MDDCTS (Multidimensional DCT) was created to expand the DCT idea to MD signaling. MDDCT may be calculated using a variety of methods [33].

To lower the computational complexity of DCT implementation, a number of quick methods have been devised. Integer DCT (int DCT) is an acceptable approximation of the DCT standard that is used in several international standards, including ISO/IEC and ITU-T. Data is compressed into groups of independent DCT blocks using DCTcompression, frequently referred to as Block Compression. DCT blocks come in a variety of sizes, ranging from 8x8 pixels for a basic DCT to 4x4 to 3232x pixels for a real DCT. DCT has a lot of "power compression," which means it can achieve good quality while compressing a lot of data. When significant DCT pressure is applied, however, compression artifact distortions might develop. To recover audio data from its transform representation, an Inverse Discrete Cosine Transform (IDCT) can be used. The DCT and IDCT may be described using the formulae below [34]:

A. **For Forward Transform**:

$$DCT(U) = a(U) \sum_{t=0}^{N-1} Wav(t) \, cos(\frac{\pi(2t+1)U}{2N}), \; for \; U = 0,1,\dots N-1 \qquad (1)$$

B. **For Backward Transform**

$$Wav(t) = \sum_{U=0}^{N-1} a(U) \, DCT(U) \, cos\left(\frac{\pi(2t+1)U}{2N}\right), \; for \; U = 0,1,\dots N-1 \qquad (2)$$

$$where, a(0) = \sqrt{\frac{1}{N}}; \quad a(U) = \sqrt{\frac{2}{N}} \quad for \; U = 1..N-1$$

6. Wavelet Transform

Wavelets are based on the principle of analyzing data according to scale. Indeed, some wavelet experts believe that exploiting wavelets requires a completely new attitude or viewpoint when it comes to data processing. Wavelets are mathematical functions that are used to represent data or other functions by satisfying specific mathematical conditions. This concept isn't new. Since the early 1800s, when Joseph Fourier found that he could superpose sines and cosines to represent other functions, approximation utilizing superposition of functions has been used [35]. However, in wavelet analysis, the size at which we examine data is critical. Wavelet algorithms are used to handle data at various sizes and resolutions. We would detect gross characteristics if we looked at a signal with a huge "window." Similarly, we would see little details if we looked at a signal with a narrow "window." The upshot of wavelet analysis is the ability to perceive both the forest and the trees. A little wave is referred to as a wavelet. The term "smallness" denotes the fact that this function has a limited length. The condition that this function is oscillatory is referred to as the wave. The term "mother" denotes that the function with several support regions employed in the transformation process is descended from a single basic function known as the mother wavelet. The discrete cosine transform is Fourier-based, but the wavelet transform is not, and as a result, it is better at managing data discontinuities [36]. There is a push to utilize wavelet in signal processing and analysis instead of (or in addition to) the Discrete Cosine Transform (DCT), and numerous techniques to employ wavelet for picture or audio processing have recently been suggested. Wavelet transformations can benefit from the same strategies that are now utilized with audio. Wavelets have a wide range of applications, and their applications in signal processing appear to be limitless. A consistent time and frequency resolution is determined by the window length. In order to capture the fleeting behavior of a signal, a shorter temporal windowing is utilized. In this instance, the frequency resolution is sacrificed. Because actual signals (such as sound, picture, and video signals) are non-periodic and transitory in nature, they are difficult to analyze using traditional transforms . To extract the appropriate time-amplitude information from a signal, an alternate mathematical method (such as the wavelet transform) must be chosen. The wavelet transform can concurrently

provide time and frequency information, resulting in a time-frequency representation of the signal. Stationary signals are those whose frequency content does not fluctuate over time [36]. To put it another way, the frequency content of stationary signals remains constant across time. It is not necessary to know when frequency components appear in this scenario because all frequency components present at all times. However, to evaluate non-stationary signals, the Wavelet Transform (WT) is required (i.e., whose frequency response varies in time). Because the Fourier Transform (FT) is ineffective when dealing with non-stationary data. Wavelet transforms have been shown to be extremely efficient and successful in the analysis of a wide range of signals and phenomena. The following are the qualities that contribute to the effectiveness:

1.   A more precise local description and separation of signal properties is possible with the wavelet expansion. Temporary events must be defined by phase characteristics that allow cancellation and reinforcement across long time periods, as a Fourier coefficient reflects components that endure indefinitely. The coefficients of wavelet expansion indicate a component that is both local and understandable. The wavelet expansion may allow signal components that overlap in time and frequency to be separated .

2.   Wavelets may be adjusted and adapted to different situations. Because there isn't just one wavelet, they may be customized to meet certain systems that can adapt to the signal .

3.   The wavelet coefficient generation is particularly suited to digital computers. There are no derivatives or integrals in the digital computer; just multiplications and additions are used .

The four "lifting" phases and two "scaling" steps used to achieve bi-orthogonal (9/7) wavelet decomposition are described in the following equations:

- $Y(2n + 1) = x(2n + 1) + a[x(2n) + x(2n + 2)]$ (3)

- $Y(2n) = x(2n) + b[Y(2n - 1) + Y(2n + 1)]$ (4)

- $Y(2n + 1) = Y(2n + 1) + c[Y(2n) + Y(2n + 2)]$ (5)

- $Y(2n) = Y(2n + 1) + d[Y(2n - 1) + Y(2n + 1)]$ (6)

- $Y(2n + 1) = -kY(2n + 1)$ (7)

- $Y(2n) = \frac{1}{k}Y(2n)$ (8)

And this table shows the coefficient of determination a, b, c, d, and k:

Table 1 - 9/7 biorthogonal analysis filter lifting coefficients

Coefficients	Values
a	-1.58613434
b	-0.0529801185
c	-0.8829110762
d	-0.4435068522
k	-0.4435068522

7. Literature review of audio compression.

Previous techniques and algorithms are discussed in this section. Audio compression technologies are investigated and compared. This is an overview of the area. So far, it has been accomplished in this field. The table below show some comparisons between different approaches of audio compression :

Table 2 - REVIEW AND DISCUSSIONS

No.	Methods	Author's Name	Reference	Description
1	DCT and DWT	M.V.Patil , 2013 .	[37]	With the DWT-based approach, the compression factor and signal-to-noise ratio have been improved. It's also been discovered that Varied wavelets have different impacts on the represented spoken signal.
2	DCT and DWT	Fatma zohra Chelali , 2018.	[38]	PSNR=50 dB
3	DCT with TAM	Ayodeji Salau , 2019.	[39]	The recommended method achieves a 4:1 compression ratio (CR) of the original data.
4	DWT and RLE Techniques	Abebe Tsegaye ,2019.	[40]	With DWT and RLE based approach, the compression factor and signal to noise ratio of all daubechiles family (DB1, DB6, DB8, DB10) were enhanced.
5	wav, mp3 and Speex speech compression techniques , and MFCC and DWT parameters	A. Pramod Reddy,2020.	[41]	It is noticed that a larger bit-rate has higher recognition Accuracy with less Distortion by using these approaches and parameters.
6	bi-orthogonal (tab9/7),DCT transform and LZW coding technique	Zainab.J.Ahmed,2021.	[42]	The compressive system minimized the size of audio files and eliminated the need for a lot of storage.

8. CONCLUSION

In our study, we discussed generic audio compression algorithms. The (DWT and DCT) transform approaches is the most widely used in studies and perform well when combined with other technologies to increase the compression ratio, taking into account achieving a balance between pressure gain and reconstructed signal quality. This means that we must use hybrid methods to achieve good results because it is difficult to use a single method for

all types of audio files, Because of the difference in the compression ratio (in terms of sampling rate and sample accuracy).

References

[1] Weber, Philip, and Thomas Ludwig. "(Non-) Interacting with conversational agents: perceptions and motivations of using chatbots and voice assistants." In Proceedings of the Conference on Mensch und Computer, pp. 321-331. 2020.

[2] Hussain, Abir Jaafar, Ali Al-Fayadh, and Naeem Radi. "Image compression techniques: A survey in lossless and lossy algorithms." Neurocomputing 300 (2018): 44-69.

[3] K. Sayood, "Introduction to Data Compression", Elsevier Inc (2006).

[4] H. Baker, "speeding -up fractal audio compression using double moments descriptor", thesis, 2013.

[5] Wang, Jinzhen, Tong Liu, Qing Liu, Xubin He, Huizhang Luo, and Weiming He. "Compression ratio modeling and estimation across error bounds for lossy compression." IEEE Transactions on Parallel and Distributed Systems 31, no. 7 (2019): 1621-1635.

[6] Salau, A. O., Oluwafemi, I., Faleye, K. F., and Jain, S., "Audio Compression Using a Modified Discrete Cosine Transform with Temporal Auditory Masking," 2019 International Conference on Signal Processing and Communication (ICSC), NOIDA, India, 2019, pp. 135-142.

[7] L. A. Fitriya and T. W. Purboyo, "," International Journal of Applied Engineering Research, vol. 12, no. 19, p. 89, 2017

[8] Zohra Chelali, F., Cherabit, N., Djeradi, A., and Falek, L., "Wavelet transform for speech compression and Denoising," 2018 6th International Conference on Multimedia Computing and Systems (ICMCS), Rabat, Morocco, 2018, pp. 1-7.

[9] A. Tsegaye and G. Tariku, "Audio Compression Using DWT and RLE Techniques," American Journal of Electrical and Electronic Engineering, vol.7, no. 1, pp. 14-17, 2019.

[10] Abd-Elhafiez, W. M., Gharibi, W., and Heshmat, M., "An efficient color image compression technique," TELKOMNIKA Telecommunication, Computing, Electronics and Control, vol. 18, no. 4, p. 237, 2020.

[11] R. Mohammad and M. V. Kumar, "Audio compression using multiple transformation techniques," International Journal of Computer Applications, vol. 86, no. 13, pp. 9-14, 2014.

[12] Z. T. Drweesh and L. E. George, "Audio compression using biorthogonal wavelet, modified run length, high shift encoding," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 4, no. 8, pp. 63-73, 2014.

[13] S. Bousselmi, N. Aloui, and A. Cherif, "DSP real- time implementation of an audio compression algorithm by using the fast hartley transform," (IJACSA) International Journal of Advanced Computer Science and Applications, vol. 8, no. 4, pp. 472-477, 2017.

[14] Maithani S, Din M. Speech systems classification based on frequency of binary word features. In 2004 International Conference on signal processing and communications, 2004. SPCOM'04. 2004 Dec 11 (pp. 193–197). IEEE.

[15] Neyaz, Ashar, and Cihan Varol. "Audio Steganography via Cloud Services: Integrity Analysis of Hidden File." International Journal of Cyber-Security and Digital Forensics 7, no. 1 (2018): 80-87..

[16] Ware, Ryan. "File Extension Renaming and Signaturing." Digital Forensics (September 19, 2006) (2006).

[17] Kientzle, T., "Programmer's guide to sound", Addison Wesley Developers Press, 1998. p 98.

[18] Decamro, L., "New Technologies for audio copy Protection", Electronic Media Article, 1999. [Dav00] Davern, P., and Scott, M. p 42.

[19] Janssens, Jelle; Stijn Vandaele; Tom Vander Beken (2009). "The Music Industry on (the) Line? Surviving Music Piracy in a Digital Era". European, Journal of Crime, Criminal Law and Criminal Justice. 77 (96): 77–96.

[20] Abou Haidar, Gaby, Roger Achkar, and Hasan Dourgham. "Quality Verification of Audio and Image Modulation by the Simulation of PCM, DM and DPCM Systems." International Journal of Information and Communication Sciences 3, no. 4 (2019): 110.

[21] P. Cummiskey, Nikil S. Jayant, and J. L. Flanagan, "Adaptive quantization in differential PCM coding of speech", Bell Syst. Tech. J., vol. 52, pp. 110—111, Sept. 1973

[22] Noura Qusay Ebraheem, Audio-Hiding System Using Wavelet and DCT Transforms, (2008), p 44.

[23] Wade, Graham (1994). Signal coding and processing (2 ed.). Cambridge University Press. p. 34. Retrieved 2011-12-22. The broad objective of source coding is to exploit or remove 'inefficient' redundancy in the PCM source and thereby achieve a reduction in the overall source rate R.

[24] Mahdi, O.A.; Mohammed, M.A.; Mohamed, A.J. (November 2012). "Implementing a Novel Approach an Convert Audio Compression to Text Coding via Hybrid Technique" (PDF). International Journal of Computer Science Issues. 9 (6, No. 3): 53–59. Retrieved 6 March 2013.

[25] Pujar, J.H.; Kadlaskar, L.M. (May 2010). "A New Lossless Method of Image Compression and Decompression Using Huffman Coding Techniques" (PDF). Journal of Theoretical and Applied Information Technology. 15 (1): 18–23.

[26] Tank, M.K. (2011). "Implementation of Lempel- ZIV algorithm for lossless compression using VHDL". Implementation of Limpel-Ziv algorithm for lossless compression using VHDL. Thinkquest 2010: Proceedings of the First International Conference on Contours of Computing Technology. Berlin: Springer. pp. 275–283.

[27] Navqi, Saud; Naqvi, R.; Riaz, R.A.; Siddiqui, F. (April 2011). "Optimized RTL design and implementation of LZW algorithm for high bandwidth applications" (PDF). Electrical Review. 2011 (4): 279–285.

[28] Stephen, Wolfram (2002). New Kind of Science. Champaign, IL. p. 106.

[29] Stanković, Radomir S.; Astola, Jaakko T. (2012). "Reminiscences of the Early Work in DCT: Interview with K.R. Rao" (PDF). Reprints from the Early Days of Information Sciences. Tampere International Center for Signal Processing. P.60. Archived (PDF) from the original on 30 December 2021. Retrieved 30 December 2021 – via ETHW.

[30] Britanak, Vladimir; Yip, Patrick C.; Rao, K. R. (6 November 2006). Discrete Cosine and Sine Transforms: General Properties, Fast Algorithms and Integer Approximations. Academic Press. ISBN 978- 0123736246. LCCN 2006931102. OCLC 220853454. OL 18495589M. S2CID 118873224.

[31] Thomson, Gavin; Shah, Athar (2017). "Introducing HEIF and HEVC" (PDF). Apple Inc. Retrieved 5 August 2019.

[32] Ahmed, Nasir; Natarajan, T. Raj; Rao, K.R. (1 January 1974). "Discrete Cosine Transform". IEEE Transactions on Computers. IEEE Computer Society. C-23 (1): 90–93.

[33] Rao, K. Ramamohan; Yip, Patrick C. (11 September 1990). Discrete Cosine Transform: Algorithms, Advantages, Applications. Signal, Image and Speech Processing. Academic Press. doi:10.1016/c2009-0-22279-3. ISBN 978- 0125802031. LCCN 89029800. OCLC 1008648293. OL 2207570M. S2CID 12270940.

[34] Barbero, M.; Hofmann, H.; Wells, N. D. (14 November 1991). "DCT source coding and current implementations for HDTV". EBU Technical Review. European Broadcasting Union (251): 22–33. Retrieved 4 November 2019.

[35] James, Jithin, and Vinod J. Thomas. "Audio compression using DCT and DWT techniques." Journal of Information Engineering and Applications 4, no. 4 (2014): 119-124.R. Crandall, Projects in Scientific Computation, Springer-Verlag, New York, 1994, pp. 197-198, 211- 212.

[36] Kaur, Harmanpreet, and Ramanpreet Kaur. "Speech compression and decompression using DWT and DCT." International Journal of Computer Technology and Applications 3, no. 4 (2012): 1501-1503.G. Kaiser, A Friendly Guide to Wavelets, Birkhauser, Boston, 1994, pp. 44-45.

[37] Patil, M. V., Apoorva Gupta, Ankita Varma, and Shikhar Salil. "Audio and speech compression using dct and dwt techniques." International Journal of Innovative Research in Science, Engineering and Technology 2, no. 5 (2013): 1712-1719.

[38] zohra Chelali, Fatma, Noureddine Cherabit, Amar Djeradi, and Leila Falek. "Wavelet transform for speech compression and denoising." In 2018 6th International Conference on Multimedia Computing and Systems (ICMCS), pp. 1-7. IEEE, 2018.

[39] Salau, Ayodeji Olalekan, Ifetayo Oluwafemi, Kehinde Felix Faleye, and Shruti Jain. "Audio compression using a modified discrete cosine transform with temporal auditory masking." In 2019 International Conference on Signal Processing and Communication (ICSC), pp. 135-142. IEEE, 2019.

[40] Tsegaye, Abebe, and Girma Tariku. "Audio Compression Using DWT and RLE Techniques." American Journal of Electrical and Electronic Engineering 7, no. 1 (2019): 14-17.

[41] Reddy, A. Pramod, and V. Vijayarajan. "Audio compression with multi-algorithm fusion and its impact in speech emotion recognition." International Journal of Speech Technology 23, no. 2 (2020): 277-285.

[42] Ahmed, Zainab J., Loay E. George, and Raad Ahmed Hadi. "Audio compression using transforms and high order entropy encoding." International Journal of Electrical & Computer Engineering (2088-8708) 11, no. 4 (2021).