Annals of Emerging Technologies in Computing (AETiC)
Vol. 6, No. 2, 2022

Research Article

High Synthetic Audio Compression Model Based on Fractal Audio Coding and Error-Compensation

Ahmed Hussain Ali[1,*] and Loay Edwar George[2]

[1]Ministry of Higher Education and Scientific Research, Baghdad, Iraq
ahmedhussainali78@gmail.com
[2]University of Information Technology and Communications, Baghdad, Iraq
loayedwar57@cs.uobaghdad.edu.iq
*Correspondence: ahmedhussainali78@gmail.com

Received: 27th November 2021; Accepted: 3rd March 2022; Published: 1st April 2022

Abstract: This study presented a model for improving audio files quality using fractal coding specifically when a high compression ratio is required. The proposed high synthetic audio compression model which can be called (HSACM) is based on conventional fractal coding and lifting wavelet transform. Various lifting wavelet transform families and levels are used and their effects on the reconstructed audio files are discussed as well. Audio files from GTZAN dataset and standard measurements for data compression are used in the evaluation of the proposed model. The results reveal that using block length 50 samples which is the worst case, PSNR is increased, on average, from 34.1 to 44.8 dB and from 34.1 to 40.5 dB using lifting wavelet transform with 3 and 2 levels, respectively. Thus, the PSNR is improved by 10 and 5 dB with slightly reducing the compression ratio by 6.2 and 12.5%, respectively. Moreover, it can be noticed that adopting lifting wavelet transform with basis Haar, db1, db4, db5, cdf1.1 and cdf2.2 provide higher audio quality while db6, db8, sym7 and sym8 give the worst audio quality. Furthermore, the performance of HSACM is compared with that of existing work to highlight its performance.

Keywords: ***Audio Compression; Audio Quality; Fractal Audio Coding; Lifting Wavelet Transform***

---

1. Introduction

Audio compression is still the main issue in digital audio processing. Speech compression represents transforming the signal speech to a compressed form, which thereafter can be transmitted with a smaller size because of the limited bandwidth [1]. Three main requirements should be satisfied in each proposed data compression technique which are compression ratio, uncompressed audio quality or fidelity and encoding/decoding time [2]. Various compression techniques were used to compress the multimedia files such as Discrete Wavelet Transform (DWT), Fractal Coding (FC), Vector Quantization (VQ) and Discrete Cosine Transform (DCT) [3].

Fractal coding is a promising lossy technique that is used for compression files with a high compression ratio, asymmetric coding/decoding and accepted reconstruction quality [4]. However, it still has an issue related to the quality of the reconstructed audio files in addition to the encoding process time using fractal coding [5]. A high compression ratio can be achieved by enlarging block length, however, this leads to degrading quality of the audio file. An inverse relationship is shown with the audio quality and block length. Increasing block length leads to decreasing in quality of the audio file and conversely. The quality of the files after the decoding depends mainly on the block length which is an inverse relationship. Since fractal coding has a high compression ratio, it can be compromised to improve the audio quality [5].

Ahmed Hussain Ali and Loay Edwar George "High Synthetic Audio Compression Model Based on Fractal Audio Coding and Error-Compensation", Annals of Emerging Technologies in Computing (AETiC), Print ISSN: 2516-0281, Online ISSN: 2516-029X, pp. 1-12, Vol. 6, No. 2, 1st April 2022, Published by International Association for Educators and Researchers (IAER), DOI: 10.33166/AETiC.2022.02.001, Available: http://aetic.theiaer.org/archive/v6/v6n2/p1.html.

In this study, a high synthetic compression model for audio files is proposed based on conventional fractal coding and error-compensation using wavelet transform. Lifting Wavelet Transform (LWT) is utilized in the proposed model to improve the quality of audio files. The error between the original audio and the reconstructed using conventional Fractal Audio Coding (FAC) is compressed and compensated with the fractal parameters in the compressed file. After applying LWT on the error, the approximate coefficients of the error are adopted and the detail coefficients are ignored since the first has the most important information about the audio signal on the opposite of the second. Different audio files, block length and wavelet families such as (Haar, Daubechies, Symlet, Cohen–Daubechies–Feauveau and 9.7) are used and discussed their effect on the audio quality. Moreover, the results are compared with the related work. The contribution of this study is improving uncompressed audio files quality using conventional FAC when a high compression ratio is required. Moreover, the effect of the 16 LWT basis on the audio quality and WT levels are discussed.

The structure of the paper is as follows: basic concepts are explained in section 2. Section 3 highlights the related works that were used fractal audio as a compression technique for audio files. Section 4 presents the details of the HSACM model and its processes. Section 5 shows the results and discussion of two experiments and comparison with related works. Finally, the conclusion and future work are included in Section 6.

2. Preliminary

This section is dedicated to explaining briefly the concepts of the techniques that are adopted in the proposed model which are fractal audio coding and lifting wavelet transform.

2.1. Fractal Coding

Fractal coding is an approach adopted to compress data since it produces high compression ratio and the accepted fidelity of the reconstructed data. The idea of the fractal is based on fact that most real-world objects have similarities and redundancies. Barnsley and Sloan [6] were the first to employ fractal coding for image compression in 1988. Later, utilizing the mathematics of IFSs, Jacquin adopted Sloan and Barnsley's work, and he ultimately established fractal coding using Partition Iteration Function Systems (PIFS) [7]. When compared to other compression techniques like DWT and DCT, FC has a lower computational complexity because it does not require any transformation. However, the encoding process is time-consuming [8].

Fractal coding is based on the generation of blocks which are domain and range. The encoding process begins by selecting the most similar domain block, each range is compared with domain blocks. The matching process uses Root Mean Square (RMS) to select most similar domain block for each range block with minimum RMS. Most similarity means lower value of RMS. It computes using Eqs. (1)-(4) [10]:

$$RMS_j{}^2 = \sigma_r^2 + Scl_j\left[Scl_j\sigma_d^2 + 2\bar{d}\bar{r} - \frac{2}{n}\sum_{n=0}^{n-1} d_i r_i\right] \tag{1}$$

$$Scl_j = \frac{\frac{1}{n}\sum_{i=0}^{n-1} d_i r_i - \bar{d}\bar{r}}{\sigma_d^2} \tag{2}$$

$$\bar{r} = \frac{1}{n}\sum_{i=0}^{n-1} r_i, \quad \bar{d} = \frac{1}{n}\sum_{i=0}^{n-1} d_i \tag{3}$$

$$\sigma_r^2 = \frac{1}{n}\sum_{i=0}^{n-1} r_i^2 - \bar{r}^2, \qquad \sigma_d^2 = \frac{1}{n}\sum_{i=0}^{n-1} d_i^2 - \bar{d}^2 \tag{4}$$

Where Scl_j is the range block scale value of index the j^{th}; (0<j<N). N represents the number of range blocks. *d* is the domain block and *r* is the range. $\bar{d}$ is the domain block mean and $\bar{r}$ is the range block mean. σ_d^2 is the domain block variances and σ_r^2 is the range block variances. Each range block at the end is represented by four parameters termed fractal parameters (FPs). Each FPs consists of scale (Scl_j), mean ($\bar{d}$), index of the domain (i), and affine mapping (direction of the domain bock, original or rotate by 180^{O}) [5] [9]. Each block is represented by a 1-D array in fractal audio coding.

The decoding process is dedicated to reconstructing the range blocks by simply performing the appropriate affine transform on d_i then applying Eq. (5):

$$r_k' = Scl_j(d_i - \bar{d}) + \bar{r} \tag{5}$$

Where r'_k is range retrieved block of index the k^{th}. Scl_j is the range block scale parameter of index j^{th}. $d_i$is the arbitrary block sample value of index i^{th}. The arbitrary block is a block with samples of value 0. $\bar{d}$ is the arbitrary domain mean and $\bar{r}$ is the range blocks mean [11].

2.2. Lifting Wavelet Transform

Any given signal can be represented in time-frequency frequency transform such as WT. The audio signal is divided into two sets of coefficients using WT which are the approximation coefficients by passing the signal through a low-pass filter and the detail coefficients by using a high-pass filter. Depending on the application and the length of the signal, the approximate coefficients can be decomposed into more levels [12].

The LWT is proposed to minimize the wavelet transform's computation time and memory requirements. The lifting wavelet simplifies the problem by analyzing it in the integer domain directly. Furthermore, the lifting wavelet saves time and provides a frequency localization characteristic that overcomes the standard wavelet's shortcoming. Split/merge, prediction, and update are the three processes of the lifting wavelet method. LWT is described in detail in [13, 14]. The inverse process of decomposition is used to recreate the lifting wavelet transform.

3. Related Works

This section includes a discussion of the related works that were proposed using fractal audio coding.

Wannamaker and Vrscay proposed fractal audio coding in 1997 using a WT and the FC. Since audio signals are smoother than visuals, they devised a hybrid technique combining FC and WT. With this method, they were able to obtain compression ratio of 6:1 and maintaining the quality of reconstructed audio signal [15].

Xiao [16] implemented fractal coding on an audio signal. He conducted studies using audio signals rather than images to see how this technique worked. Bit allocation, range partitioning, domain creation, and others parameter are proven to affect performance in this study. In this study, the compression factor was around 3 to 6.

George and Salih [17] suggested a PCM wave file-based audio compression technique based on fractal coding and affine transformation. The relationship of block length, encoding time, jumping step, PSNR, compression ratio and MSE were highlighted in the study. Furthermore, the results show an inverse relationship between PSNR and block length and jump step for audio quality. Using 10 and 40 samples block length, the PSNR are 45.6 and 31.53, respectively.

FAC approach is proposed by Bedan and George [18] that uses a moment descriptor-based filtering mechanism to reduce the domain blocks and accelerate the matching process. For classification, the method employs first and third order moments. The findings of this study are expected to provide a significant decrease in the time of the encoding process for the PIFS without a noticeable reduction in audio quality. The adopted block length in the experiments was not indicated in this study.

An investigation is performed by authors in [5] to highlighting the effect of various block lengths on audio quality using conventional FC. The achieved compression ratio was roughly 93% using a block length of 40 samples and audio quality on average PSNR 35 dB. While the audio quality was PSNR 32 dB on average using block length 50 and the achieved compression ratio of 96%.

In sum, the related works in the area of FAC are dedicated to improving the encoding process time. Up to our knowledge, improving the quality of the uncompressed audio file using conventional FAC did not discuss and need further research.

4. The HSACM Model

The HSACM model consists of four processes: encoding, error-compensation, decoding and audio improving. These processes are shown in Figure 1. The next subsections describe the details of each process.

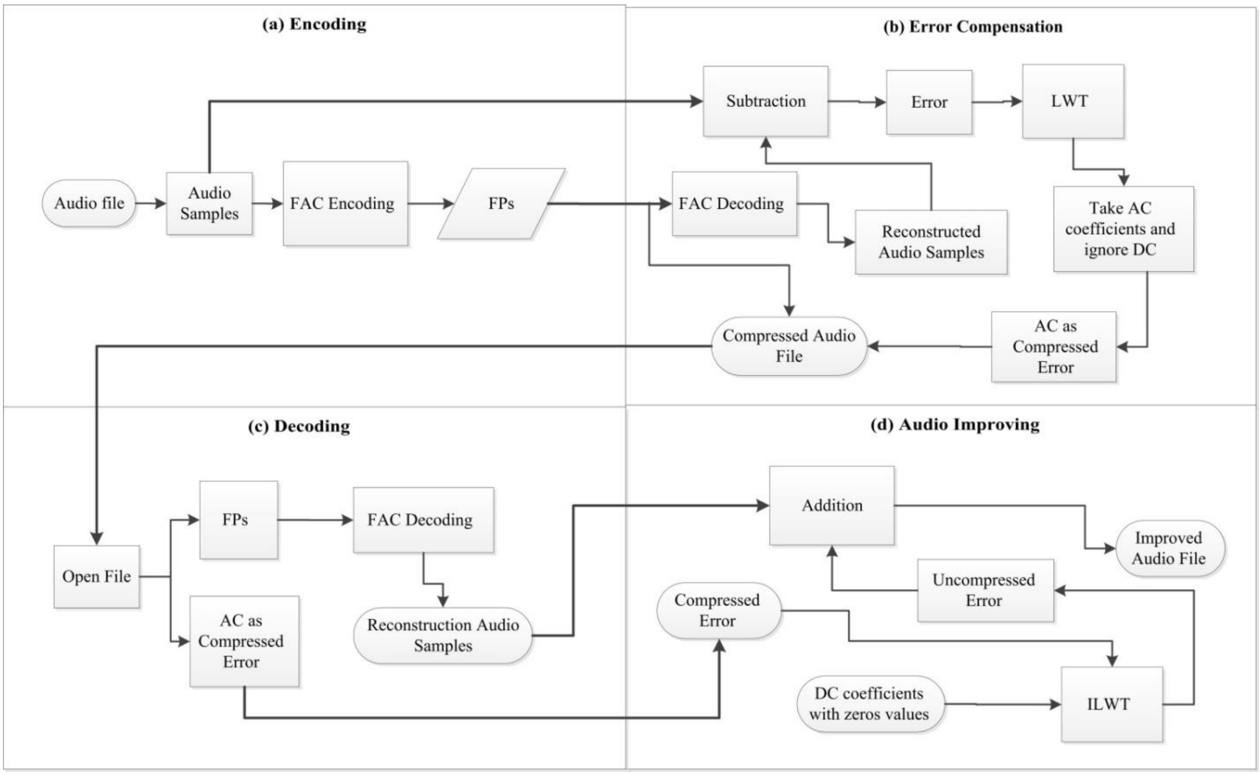

Figure 1. HSACM model, (a) Encoding, (b) Error-Compensation, (c) Decoding and (d) Audio Improving

4.1. Encoding

Encoding is the first process in the proposed model and responsible for encoding the uncompressed audio file and producing a set of FPs in order to include and generate compressed audio files later in the second process. The steps are as follow:

1- Select the audio file as input, open and read the data audio samples.
2- Perform FAC encoding by generating the range and domain blocks by partitioning the samples into blocks of a specific length, this length can be ranged from 2 to N samples which is based on the required comparison ratio and uncompressed audio quality.
3- Find the domain block which is the most similar to the particular range block using Eqs. (1)-(4).
4- Set the FPs of each range block.
5- Reconstruct the compressed audio file by including the FPs with the error that got from the error-compensation process.

The flowchart of this process is shown in Figure 1 (a).

4.2. Error-Compensation

This process is dedicated to obtaining the error which is the difference between the original audio samples and the uncompressed audio samples using conventional FAC and the compressed it and then include it with the PFs in the compressed audio file. The steps are as follow:

1- The input to this process is the original audio samples and the reconstructed uncompressed audio samples after applying the conventional FAC decoding.
2- Compute the error by subtraction every two samples, the original and the reconstructed samples.
3- Apply LWT on the error using different levels and families. The selection of the level and wavelet family is ranged from 2 to L levels which is based on the required audio quality and compression ratio.
4- Keep the approximate coefficients, which are the first set of the coefficients from LWT that represents the significant information about the error and consider it a compressed error and ignore the detailed coefficients since they are not important.
5- Include compressed error along with the FPs and reconstruct the compressed audio file.

Figure 1 (b) shows the flowchart of this process.

4.3. Decoding

The third process in HSACM model is decoding. This process starts with the compressed audio file and finishes with constructing the uncompressed audio and getting the compressed error as follow:

1- Select the compressed audio file and open it.
2- Read the samples from the file byte by byte to get the FPs.
3- The samples after the PFs represent the compressed error which are getting by reading them byte by byte.
4- Select a random set of samples the same size as the original audio samples.
5- Perform the FAC decoding using the random set of samples and FPs to get the initial uncompressed audio samples using Eq. (5).
6- Repeat step 4 several iterations by replacing the random set of samples with the initial uncompressed audio samples until getting the audio samples which are approximately similar to the original audio samples.

The flowchart of this process is shown in Figure 1 (c).

4.4. Audio Improving

This process is used to improve the quality of the uncompressed audio samples that got from the conventional FAC which is the consequence of the decoding process. The input to this process is the uncompressed audio samples and the errors got from the decoding process and the output is the improved uncompressed audio samples and the steps of this process are as follow:

1- Generate a set of detail coefficients same size as the compressed error with zero values.
2- Perform ILWT by applying Eq. (5) using the compressed error which represents the approximate coefficients and the detail coefficients from step 1 using the same level and family that were used in process 4.2 to get the uncompressed error.
3- Addition is performed to the uncompressed error and the uncompressed audio samples to get the improved uncompressed audio samples.
4- Reconstruct the uncompressed audio file using the improved uncompressed audio samples and header information.

The flowchart of this process is shown in Figure. 1 (d).

5. Results and Discussion

The effect of the HSACM model on the improvement of audio files is demonstrated in this section using a variety of audio files, block lengths, LWT levels and families. The evaluation metrics used in the experiments, the results of the experiments and dataset are included in this section as well. The proposed model's prototype was created in MATLAB R2013a on an Intel® Core TM i7-7500U CPU @ 2.50 GHz with 8GB RAM. The audio files duration for each experiment is specified by a particular software, Adobe Audition 1.5. Compressed audio quality is measured using Signal to Noise Ratio (SNR), Mean Square Error (MSE) [19] and Peak Signal to Noise Ratio (PSNR) [20], while compression performance is measured using compression ratio, which represents the ratio between the size of original and compressed audio file and compression factor, which means the inverse of the compression ratio [21]. The audio dataset is based on the GTZAN [22]. This dataset is used to evaluate the HSACM model since it contains speech and music audio file types. 16 bits per sample, 44100 sample rate, mono, speech and music with 1-10 second duration are the characteristics of the audio files utilized in the studies. Two experiments are conducted in this section which highlights the improvement of the quality using various block lengths and LWT families. Moreover, a comparison with related work is included as well.

5.1. Block Length and Quality

This experiment presents the impact of block length on the quality of the audio files using different LWT levels. Three speech and four music audio file types are used. Table 1 depicts the effect 10, 20, 30, 40 and 50 samples block length on the PSNR using LWT with L=2. The results tabulated show that the improvement in the PSNR is ranged from 2.1to 12.2 dB using the above block lengths. This improving produce decreases in the CR by 12.5%. it can be noticed that the improvement in the PSNR is increased proportionally with the increase of the block length. The results are obtained using different LWT families and represent the average of these results.

Table 1. Effect of using LWT with 2-Levels on audio quality and compression ratio

LWT	*2L*				
Block Length	*10*	*20*	*30*	*40*	*50*
CR (%)	81.2-68.7	90.6-78.1	93.7-81.2	95.3-82.8	96.2-83.7
	Decreased CR (%)				
	12.5				
Audio	*Improved PSNR dB*				
Kalimba	3	5.4	6.6	9.1	10.1
Bagpipe	*4.4*	*6.6*	*8.3*	*9.8*	*9.7*
Female	2.1	4.6	6.6	8.4	10.3
Jazz	*4.1*	*8*	*9*	*11.1*	*12.2*
news1	2.7	3.8	4.9	7.2	6.6
Vlobos	*2.5*	*4.9*	*6.7*	*8.3*	*9.3*
Voice	3.2	3.9	7.2	8.2	8.3

The decrease in the CR using LWT L=3 is less than that got using LWT L=2 as shown in Table 1. At the same time, the improvement in the PSNR is less than the result using LWT L=2 using 10, 20, 30, 40 and 50 samples. Moreover, the results in Table 2 show that the improvement in the PSNR is between 0.4 to 7.9 dB.

Table 2. Effect of using LWT with 3-Levels on audio quality and compression ratio

LWT	*3L*				
Block Length	*10*	*20*	*30*	*40*	*50*
CR (%)	*81.2-75*	*90.6-84.4*	*93.7-87.5*	*95.3-89.1*	*96.5-90*
	Decreased CR(%)				
	6.2				
Audio	*Improved PSNR dB*				
Kalimba	*1*	*2*	*3.2*	*5.5*	*5.5*
bagpipe	*1.4*	*2.8*	*3.3*	*5.4*	*4.2*
female	*0.5*	*1.9*	*3.2*	*4.9*	*5.9*
Jazz	*1.3*	*4.6*	*5.4*	*7.9*	*7.8*
news1	*0.7*	*1.3*	*2*	*4.1*	*3.4*
vlobos	*0.4*	*1.4*	*4.3*	*5.6*	*5.8*
Voice	*1*	*1.5*	*3.4*	*4.6*	*4.2*

It can be concluded from this experiment that the improvement of the PSNR has proportional with the increase of the block length and inverse relationship with the LWT level. Furthermore, when the level of the LWT is increased, the CR is decreased.

5.2. LWT Families and Quality

This experiment presents how the quality of the audio files is affected by the LWT families. Kalimba and Jazz are the audio files that are used. Haar, Daubechies (db), Symlet (sym), Cohen–Daubechies–Feauveau and 9.7. Female as a speech file and Jazz as music audio file are used to highlight the effect of LWT families. Block length of 10 to 50 samples, SNR, PSNR and LWT level=2 are used in this experiment. SNR, SNR_NEW, PSNR and PSNR_NEW represent the quality of the audio using conventional FAC and HSACM model, respectively. Figure 2 depicts the effect of LWT families on the Female audio. It seems that using 10, 20 and 40 block length, haar, db1, db4, db5, cdf1.1, cdf2.2 provide better results in terms of audio quality. While using 30 and 50 block lengths, in addition to the above, sym3 and sym4 give better quality as well. On the other hand, the worst audio qualities are got using db6, db8, sym6 and sym7 with all block lengths as shown in Figure 2.

Figure 3 highlights the effect of LWT families on the music audio file type which is Jazz. It seems that in addition to haar, db1, db4, db5, cdf1.1, cdf2.2, sym3 and sym4 provide better result using all block lengths and the worst audio qualities is got using db6, db8, sym6 and sym7 as well in Figure 2.

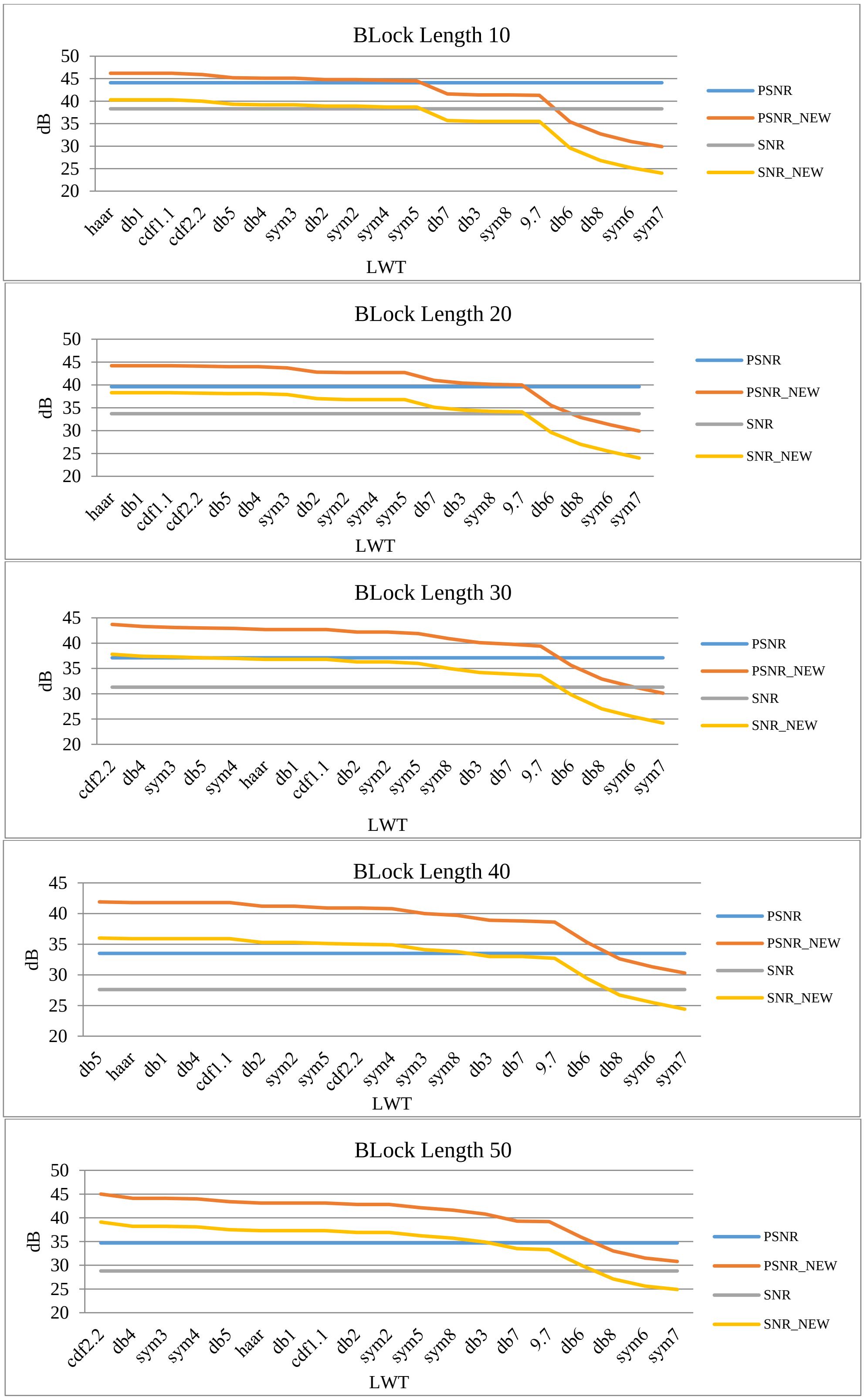

Figure 2. Female LWT Level 2 with Different Block Length

To show the visual effect of the LWT families on the quality of the audio files, Figure 4 include the waveform of the audio file using Kalimba music audio type with block length 10 samples and LWT L=2. It can be noted that all basis of the families produces acceptable improvement except db6, db8, sym6 and sym7.

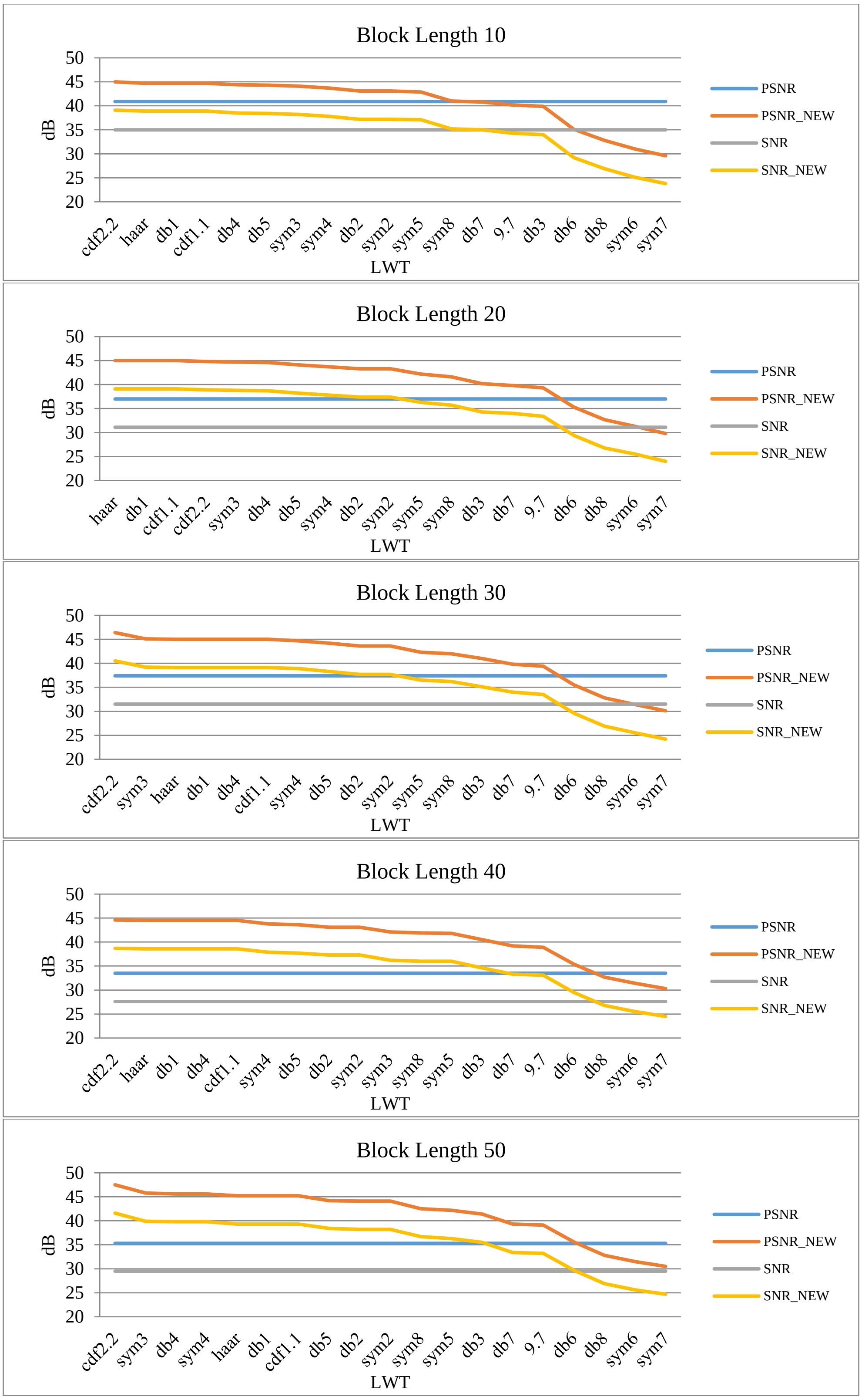

Figure 3. Jazz LWT Level 2 with Different Block Length

Figure 5 depicts the effect of LWT families on the Kalimba audio file with a block length of 50 samples. It can be concluded that when the level of the LWT is 3 the improvement on the quality is limited and more basis such as db3, db6, db8, sym6, sym7, 9.7 give the worst quality.

Figure 4. Waveform of Kalimba audio file with block Length 10 samples and LWT Level 2 with different families

5.3. Comparison with Related Works

Comparison with related work is important to highlight the contribution of the proposed HSACM model. As mentioned in section 3, authors in [5] conducted the conventional FAC and discussed the effect of the block length on the compression ratio and audio quality using a standard dataset. Therefore, in this section comparison results are tabulated in Table 3 in terms of MSE, PSNE, and CR. In this comparison, the results of the HSACM using L=2 and L=3 are included.

As seen from the Table 3, comparison between [5] and HSACM using L=2 achieves improving the quality from 0.8 to11.6 dB, 1.9 to 10.6 dB and 0.3 to 10.1using using block length from 10 to 50 samples and audio files Female, Jazz and Voice, respectively and slightly reducing the CR by 12.5%.

Moreover, HSACM using Female and block length 10 to 50 samples and L=3, the quality improvement is between 0.2 and 7.2 dB. The improvement of Jazz with the same LWT level is 0.1 to 6.2 and using Voice audio file is between 0.1 to 6. The CR is slightly reduced by 6.2%.

In summary for this comparison, using HSACM with LWT L=2 provides better improvement of quality PSNR than using L=3, however, the reduction of CR is decreased when the LWT level is increased.

Table 3 Comparison HSACM with related work Ref. [5]

		Ref [5]			*HSACM LWT L=2*			*HSACM LWT L=3*		
Audio	*Block Length*	*MSE*	*PSNR*	*CR*	*MSE*	*PSNR*	*CR*	*MSE*	*PSNR*	*CR*
Female	10	1.8	45.4	81.2	1.6	46.2	68.7	1.7	45.6	75
	20	3.5	42.6	90.6	2.5	44.2	78.1	3.5	42.6	84.4

	30	14.7	36.4	93.7	2.8	43.7	81.2	6.1	40.3	87.5
	40	31.5	33.1	95.3	4.2	41.9	82.8	9.3	38.4	89.1
	50	29.5	33.4	96.2	2.1	45	83.7	5.6	40.6	90
Jazz	10	3.1	43.1	81.2	2.1	45	68.7	3	43.2	75
	20	6.3	40	90.6	2.1	45	78.1	4.5	41.6	84.4
	30	6.8	39.7	93.7	1.5	46.4	81.2	3.4	42.8	87.5
	40	17.9	35.5	95.3	2.3	44.6	82.8	4.7	41.4	89.1
	50	13	36.9	96.2	1.2	47.5	83.7	3.2	43.1	90
Voice	10	2.8	43.5	81.2	2.6	43.8	68.7	2.7	43.6	75
	20	11.2	37.6	90.6	5.2	41	78.1	8.9	38.6	84.4
	30	22	34.6	93.7	4.6	41.5	81.2	11.1	37.7	87.5
	40	39	32.2	95.3	4.5	41.6	82.8	10.4	38	89.1
	50	40	32	96.2	4	42.1	83.7	10.4	38	90

Figure 5. Waveform of Kalimba audio file with block Length 50 samples and LWT Level 3 with different families

6. Conclusion and Future Work

Fractal coding is an emerging technique used to encoding the audio files with a high compression ratio, asymmetric encoding/decoding process and acceptable audio quality. However, the quality is degraded with a high compression ratio as the block length should be large. Therefore, this study proposed a compression model based on using conventional FAC to improve the quality of the audio files. Conventional fractal coding and lifting wavelet transform with various LWT families and levels are adopted in the HSACM model. The results show improvement in the audio quality when block length 50 samples are used. Moreover, some LWT families produce better PSNR while other are not. Comparison of HSACM with related work highlights the outperforming of the PSNR of the proposed model.

Further research is required to include the details coefficients of the error rather than ignoring them. Moreover, adopting variable length for the approximate coefficients that are compensated as a error with the compressed audio file instead of fixed length to improve the compression ratio.

Acknowledgement

This research was supported by the Ministry of Higher Education and Scientific Research and the University of Information Technology and Communications, Iraq.

References

[1] Abdulmawla Najih, Abdul Rahman bin Ramli, Veeraraghavan Prakash and Abdul Rahman Syed, "Speech compression using discreet wavelet transform", in *Proceedings of the 4th National Conference of Telecommunication Technology (NCTT)*, 14-15 January 2003, Shah Alam, Malaysia, DOI: 10.1109/NCTT.2003.1188289, pp. 1-4, Published by IEEE, Available: https://ieeexplore.ieee.org/abstract/document/1188289.

[2] Khalid Sayood, *Introduction to data compression*, 4th ed. Waltham, USA: Morgan Kaufmann Elsevier, 2012.

[3] Mingwei Tang, Zeng Shenke, Xiaoliang Chen, Jie Hu and Yajun Du, "An adaptive image steganography using AMBTC compression and interpolation technique", *Optik-International Journal for Light and Electron Optics*, Online ISSN: 0030-4026, pp. 471-477, Vol. 127, No. 1, 2016, DOI: 10.1016/j.ijleo.2015.09.216, Published by Elsevier, Available: https://www.sciencedirect.com/science/article/abs/pii/S0030402615012978.

[4] Keyvan Jaferzadeh, Inkyu Moon and Samaneh Gholami, "Enhancing fractal image compression speed using local features for reducing search space", *Pattern Analysis and Applications*, Online ISSN: 1433-7541, Vol. 20, No. 4, pp. 1-10, 2016, DOI: 10.1007/s10044-016-0551-1, Published by Springer, Available: https://link.springer.com/article/10.1007/s10044-016-0551-1.

[5] Ahmed Hussain Ali, Loay Edwar George, Omar Saad Saleh, Mohd Rosmado Mokhtar and Qusay Al-Maatouk, "Investigating the Effect of Block Length on the Performance of Fractal Coding Using Audio Files", *Periodica Polytechnica Electrical Engineering and Computer Science*, Online ISSN: 2064-5279, pp. 303-312, Vol. 64, No. 3, 2020, DOI: 10.3311/PPee.14896, Available: https://pp.bme.hu/eecs/article/view/14896.

[6] Michael Fielding Barnsley, *Fractals everywhere*, 2nd ed. London, UK: Academic Press Professional, 1993.

[7] Arnaud Jacquin, "Image coding based on a fractal theory of iterated contractive image transformations", *IEEE Transactions on Image Processing*, Online ISSN: 1057-7149, pp. 18-30, Vol. 1, No. 1, 1992, DOI: 10.1109/83.128028, Published by IEEE, Available: https://ieeexplore.ieee.org/document/128028.

[8] Tarek Sheltami, Muhammad Musaddiq and Elhadi Shakshuki, "Data compression techniques in Wireless Sensor Networks", *Future Generation Computer Systems*, Online ISSN: 0167-739X, pp. 151-162, Vol. 64, 2016, DOI: 10.1016/j.future.2016.01.015, Published by Elsevier, Available: https://www.sciencedirect.com/science/article/abs/pii/S0167739X16000285.

[9] Ahmed Hussain Ali, Ali Nihad Abbas, Loay Edwar George and Mohd Rosmadi Mokhtar, "Image and Audio Fractal Compression: Comprehensive Review, Enhancements and Research Directions", *Indonesian Journal of Electrical Engineering and Computer Science*, Online ISSN: 2502-4752, pp. 1564-1570, Vol. 15, No. 3, 2019, DOI: 10.11591/ijeecs.v15.i3.pp1564-1570, Published by Institute of Advanced Engineering and Science, Available: http://ijeecs.iaescore.com/index.php/IJEECS/article/view/16323.

[10] Ayman Ibaida, Dhiah Al-Shammary and Ibrahim Khalil, "Cloud enabled fractal based ECG compression in wireless body sensor networks", *Future Generation Computer Systems*, Online ISSN: 0167-739X, pp. 91-101, Vol. 35, 2014, DOI: 10.1016/j.future.2013.12.025, Published by Elsevier, Available: https://www.sciencedirect.com/science/article/abs/pii/S0167739X13002896.

[11] Chong Sze Tong and Minghong Pi, "Fast Fractal Image Encoding Based on Adaptive Search", *IEEE Transactions on Image Processing*, Online ISSN: 1057-7149, pp. 1269-1277, Vol. 10, No. 9, 2001, DOI: 10.1109/83.941851, Published by IEEE, Available: https://ieeexplore.ieee.org/document/941851.

[12] Ali Al-Haj, "An imperceptible and robust audio watermarking algorithm", *EURASIP Journal on Audio, Speech, and Music Processing*, Online ISSN: 1687-4722, pp. 1-12, Vol. 2014, No. 37, 2014, DOI: 10.1186/s13636-014-0037-2, Published by Springer, Available: https://asmp-eurasipjournals.springeropen.com/articles/10.1186/s13636-014-0037-2.

[13] Wim Sweldens, "The lifting scheme: A custom-design construction of biorthogonal wavelets", *Applied and computational harmonic analysis*, Online ISSN: 1063-5203, pp. 186-200, Vol. 3, No. 2, 1996, DOI:

10.1006/ACHA.1996.0015, Published by Elsevier, Available: https://www.sciencedirect.com/science/article/pii/S1063520396900159.

[14] Baiying Lei, Yann Soon, Feng Zhou, Zhen Li and Haijun Lei, "A robust audio watermarking scheme based on lifting wavelet transform and singular value decomposition", *Signal Processing*, Online ISSN: 0165-1684, pp. 1985-2001, Vol. 92, No. 9, 2012, DOI: 10.1007/978-3-642-32205-1_9, Published by Elsevier, Available: https://www.sciencedirect.com/science/article/abs/pii/S0165168411004658.

[15] Robert A. Wannamaker and Edward R. Vrscay, "Fractal wavelet compression of audio signals", *Journal of the Audio Engineering Society*, Online pp. 540-553, Vol. 45, No. 7/8, 1997, Published by Audio Engineering Society, Available: https://www.aes.org/e-lib/online/browse.cfm?elib=7848.

[16] Henry Xiao, "Fractal Audio Coding", Master's Thesis, *Queen's University*, Canada, 2005, Available: https://research.cs.queensu.ca/home/xiao/doc/Thesis.pdf.

[17] Loay Edwar George and Huda Mahdi Salih, "Audio Compression Using Fractal Coding", *Diyala Journal For Pure Science*, Online ISSN: 2518-9255, pp. 1-10, Vol. 6, No. 1, 2010, Available: https://www.iasj.net/iasj/download/5440e530cd1221b0.

[18] Azhaar Kareem Bedan and Loay Edwar George, "Speeding-up Fractal Audio Compression Using Moment Descriptors", *International Journal of Scientific & Engineering Research*, Online ISSN: 2229-5518, pp. 1623-1627, Vol. 4, No. 7, 2013, Available: https://www.ijser.org/researchpaper/Speeding-up-Fractal-Audio-Compression-Using-Moment-Descriptors.pdf.

[19] Diego Renza and Camilo Lemus, "Authenticity verification of audio signals based on fragile watermarking for audio forensics", *Expert Systems with Applications*, Online ISSN: 0957-4174, pp. 211-222, Vol. 91, 2018, DOI: 10.1016/j.eswa.2017.09.003, Published by Elsevier, Available: https://www.sciencedirect.com/science/article/abs/pii/S0957417417305997.

[20] Xiaoxiao Ma, Zhibin Pan, Sen Hu and Lingfri Wang, "Reversible data hiding scheme for VQ indices based on modified locally adaptive coding and double-layer embedding strategy", *Journal of Visual Communication and Image Representation*, Online ISSN: 1047-3203, pp. 60-70, Vol. 28, 2015, DOI: 10.1016/j.jvcir.2015.01.009, Published by Elsevier, Available: https://www.sciencedirect.com/science/article/abs/pii/S1047320315000103.

[21] David Salomon and Giovanni Motta, *Handbook of data compression*, 5th ed. New York, USA: Springer-Verlag, 2010.

[22] George Tzanetakis and Perry Cook, "Musical genre classification of audio signals", *IEEE Transactions on speech and audio processing*, Online ISSN: 1558-2353, pp. 293–302, Vol. 10, No. 5, 2002, DOI: 10.1109/TSA.2002.800560, Published by IEEE, Available: https://ieeexplore.ieee.org/document/1021072.