QIM - Based Audio Watermarking with Combination Technique of DCT-QR-CPT

RENDY DWI RENDRAGRAHA, GELAR BUDIMAN, IRMA SAFITRI

Fakultas Teknik Elektro Universitas Telkom
Email: rendydr23@gmail.com

ABSTRAK

Audio watermarking adalah teknik memasukkan informasi ke dalam file audio dan untuk melindungi hak cipta data digital dari distribusi ilegal. Makalah ini memperkenalkan audio stereo watermarking berdasarkan Quantization Index Modulation (QIM) dengan teknik gabungan Discrete Cosine Transform (DCT) - QR- Cartesian Polar Transform (CPT). Host audio dibagi menjadi beberapa frame, selanjutnya setiap frame ditransformasi oleh DCT, kemudian output DCT diuraikan menjadi matriks orthogonal dan matriks segitiga menggunakan metode QR. Selanjutnya, CPT mengubah dua koefisien kartesian dari matriks segitiga (R) pada posisi (1,1) dan (2,2) menjadi koefisien polar. Setelah itu, penyisipan dilakukan pada koefisien polar oleh QIM. Hasil simulasi menunjukkan bahwa imperseptibilitas audio terwatermark berkualitas baik dengan Signal to Noise Ratio (SNR)> 20, Mean Opinion Score (MOS)> 4 dan tahan terhadap serangan seperti Low Pass Filter (LPF) dan Band Pass Filter (BPF) dengan cut off 25-6k, resampling, Linear Speed Change (LSC) dan MP3 Compression dengan rate diatas 64 kbps.

Kata kunci: Audio Watermarking, CPT, DCT, QIM, QR

ABSTRACT

Audio watermarking is a technique for inserting information into an audio file and to protect the copyright of digital data from illegal distribution. This paper introduces a stereo audio watermarking based on Quantization Index Modulation ... — Cartesian ... DCT, then DCT output is decomposed using QR method. Next, CPT transform two cartesian coefficients from triangular matrix (R) in position (1,1) and (2,2) to polar coefficients. After that, embedding is executed on polar coefficients by QIM. The simulation result shows that the imperceptibility is good with Signal to Noise Ratio (SNR)>20, Mean Opinion Score (MOS)>4 and it is robust against attacks such as Low Pass Filter (LPF) and Band Pass Filter (BPF) with cut off 25-6k, Resampling, Linear Speed Change and MP3 Compression with rate 64 kbps and above.

Keywords: Audio Watermarking, CPT, DCT, QIM, QR

1. INTRODUCTION

Nowadays, sending and receiving some information is easier especially through internet. Anyone can communicate with each other and exchange information via the internet. Behind the easiness, there are some people who commit crimes. As example, stealing ownership of the copyright. Watermarking is one of the techniques to prevent theft of a copyright ownership. One of the effective methods is a digital watermarking, to prevent theft of copyright by hiding digital data in the form of multimedia data such as sound, images, and video without damaging the quality of the data to be inserted **(Aparna J R & Ayyappan, 2014) (Bahi & Adib, 2014)**. Audio watermarking is one of hiding or embedding methods of certain data or information in the form of public or secret information into digital audio, but its presence is not recognized by the human hearing, and able to encounter impairment or attacks. Audio watermarking requirement is imperceptibility, robustness and capacity. The requirement is needed to control audio watermarking performance.

There are several papers published about audio watermarking with Quantization Index Modulation (QIM). Wonel and Chen were the first to propose QIM in **(Chen & Wornell, 2001)**. In **(Khademi, Akhaeet, Ahadi, Moradi, & Kashi, 2007)** it is explained that QIM can survive against attacks such as White Gausian Noise (WGN) but there is no evaluation for several attacks, such as : LPF above 2k, BPF and Linear Speed Change. Merging the QIM method with DWPT on **(Bahi & Adib, 2014)** produce a high capacity and subband that can increase embedding capacity without altering the quality of the audio host and maintaining high extraction requirements. In **(Agradriya, Perdana, Safitri, & Novamizanti, 2017)**, authors used QIM embedding after processing host audio by DWT-SVD with Arnold Transform. In **(Aminullah, Budiman, & Safitri, 2017)**, authors used QIM with SWT-DCT-SVD combination method, but watermarking robustness against MP3 compression was good only for high rate MP3 compression or rate>128 kbps.

Discrete Cosine Transform (DCT) used in audio watermarking as published by **(Yan, Rong, & Mintao, 2009)** and **(Zhou & Zhou, 2007)** were robust against several attacks but the attack were only LPF, noise, smooth, and MP3 compression. DCT was also used by **(Adiwijaya, Novraditya, Baihaqi, & Wisesty, 2013)** and compared with other methods in audio watermarking, but the performance indicator described were only SNR and time computation. In **(Budiman, Suksmono, & Danudirdjo, 2016)**, authors described DCT and FFT comparison in audio watermarking based on Fibonacci sequence, but they described only the robustness without attack.

In **(Li & Wu, 2015)**, authors designed the audio watermarking scheme based on LWT and QR decomposition, but in several attacks, such as LPF and resampling, it was not too robust with value of BER>10%. Watermarking implementation in **(Dhar & Shimamura, 2012)** was using FFT for transforming signal to frequency domain, then it was decomposed by SVD. Next, Cartesian to Polar Transformation (CPT) transformed to polar coefficients before it was embedded by QIM method, but the BER on MP3 compression is still high. Different methods will obtain different result in audio watermarking, depending on the advantages and disadvantages of each method **(Bajpai & Kaur, 2016)**. Similar topic with this paper is found in **(Budiman, Suksmono, Danudirdjo, & Pawellang, 2018)**, where the authors proposed audio watermarking with SWT-DST-QR-CPT-QIM combination method. QR-CPT-QIM combination method in this paper is the same as the combination in **(Budiman et al., 2018)**, but in this paper DST is replaced by DCT and we do not use wavelet decomposition. In **(Dhar & Shimamura, 2015)**, authors proposed FFT-SVD-CPT-QIM-based audio watermarking with high payload and good robustness, but he only described MP3 attack with rate 128 kbps.

In this paper, we propose a watermarking application on the host audio by using methods that have been combined with the intention of taking advantage and reducing the weaknesses of each method. The proposed method is a combined technique of QIM, DCT, QR, and CPT. We use DCT for transforming time domain signal to frequency domain signal in order to get more robust domain for watermarking. The output of DCT is applied by a QR Decomposition that the frame will be decomposed into orthogonal matrix (Q) and triangular matrix (R). Next, we transform the components using CPT transform obtaining magnitude and phase component. We use QIM method for embedding process on phase component of CPT coefficient. We use phase component of CPT coefficients for embedding in order to be robust against signal processing attacks. After that, we transform back the component using ICPT from polar to cartesian. QR reconstruction the component and IDCT transform frequency domain signal into time domain signal. Finally, we can calculate performances consisting of ODG, SNR and payload from watermark.

This paper is described as follows : Section 1 describes introduction, research background and basic theory, section 2 explains the research methodology in this method of audio watermarking, section 3 describes simulation results and discussion, and section 4 describes the conclusion of this paper.

1.1 Audio Watermarking

Watermarking is a method or technique for inserting information transparently into a signal carrier. This technique is used to resolve problems such as a copyright retrieval, because the internet nowadays is very popular for transferring data and building communications. Watermarking techniques are used either in images, video or audio to provide copyright and authentication protection by hiding secret information or special features into a file. This makes watermarking appropriately applicable where the knowledge of hidden messages leads to potential dangers of manipulation **(Baranwal & Datta, 2011)**.

Watermarking audio is a technique of concealment of data or secret information into audio data to be "hosted" (audio host) but humans are not aware of the existence of additional data on its host data, thus there is no difference between the original host audio and the watermarked audio. Hiding secret information into an audio file is more difficult than into an image file because the human auditory system is more sensitive than the human visual system **(Baranwal & Datta, 2011)**.

Based on the insertion domain, the audio watermarking procedure can be divided into two methods, temporal watermarking and spectral watermarking. Temporal watermarking performs the insertion of data in the audio host in the time domain, while spectral watermarking transforms the time domain into the frequency domain before embedding process, so the insertion is performed on the frequency domain **(Warkar, More, & Waghole, 2015)**.

Embedding process is a process to insert a data watermark into the audio. Watermark will be extracted from watermarked audio which is called as extraction process. However, a watermarked audio could be attacked by several signal processing attacks. There are 2 kinds of watermarking domain **(Bajpai & Kaur, 2016)**: (1) Spatial Domain, which is described as a time domain where the embedding process is applied without using a transformation. This technique is very simple, but it has less robustness than the robustness in embedding process based on frequency domain. The various techniques in time domain are Least Significant Bit (LSB), echo hiding, phase coding, spread spectrum, and patchwork. (2) Frequency Domain, which is a domain with the more difficult technique and better robustness than the spatial

domain technique. The various techniques of how to transform the host signal into frequency domain are such as : Fast Fourier Transform (FFT), DCT, and Discrete Sine Transform (DST).

1.2 Quantization Index Modulation
QIM is one method that has a good level of robustness. Chen and Wornell firstly proposed the QIM method for watermarking. They applied quantization on the host audio samples in the embedding process **(Khademi et al., 2007)**. QIM in this paper is different than QIM in **(Safitri & Ginanjar, 2017)** which used round and floor operation in both embedding and extraction.The equation is described as follows **(Budiman et al., 2018)**:

$$F'(0) = \begin{cases} A_k, \text{if } w = 1 \text{ dan } arg\,min\,|F(0) - A_k| \\ B_k, \text{if } w = 0 \text{ dan } arg\,min\,|F(0) - B_k| \end{cases} \tag{1}$$

Where:

$$A_k = \left(2k + \frac{1}{2}\right)\Delta; B_k = \left(2k - \frac{1}{2}\right); \text{dengan } k = 0, \pm1, \pm2\ldots \tag{2}$$

For the process of extraction, the equation is as follows:

$$\tilde{V}(k) = mod\left(ceil\left(\frac{F'(0)}{\Delta}\right), 2\right) \tag{3}$$

where:

$$\Delta = \frac{1}{2^{(nbit-1)}} \tag{4}$$

Δ = quantization step size
nbit = quantization bits on QIM
$F(0)$ = sample host audio before quantized
$F'(0)$ = sample host audio after quantized

1.3 Discrete Cosine Transform
The DCT is a transformation with real coefficients that is obtained from the sum of multiplication between time domain signal and cosine formula. For one-dimensional discrete signal with *N* length, the DCT equation is describe as follows **(Nikmehr & Hashemy, 2010)**:

$$C(m) = a(m)\sum_{n=0}^{N-1} f(n)cos\left[\frac{\pi(2n+1)m}{2N}\right] \tag{5}$$

where *C(m)* is the coeficient DCT from the transformed signal with $m = 0,1,2\ldots, N-1$ and *n* is the real number. The inverse DCT (IDCT) can be expressed as follows.

$$f(n) = \sum_{m=0}^{N-1} a(m)C(m)cos\left[\frac{\pi(2m+1)n}{2N}\right] \tag{6}$$

In Equation (5) and (6), for $m = 0,1,2,\ldots,N-1$, And for $a(m)$ as follows.

$$a(m) = \begin{cases} \sqrt{\dfrac{1}{N}}, if \ m = 0 \\ \sqrt{\dfrac{2}{N}}, if \ m \neq 0 \end{cases} \tag{7}$$

The ability to compress the energy of the signal into coefficients, is one of the criteria for comparing the performance of the transformation. DCT has a compression capability. This transformation can ignore a low amplitude coefficient without reducing the accuracy in the signal reconstruction process from the coefficient **(Nikmehr & Hashemy, 2010)**.

1.4 QR Decomposition
QR Decomposition or QR factorization is also called a Gram Schmidt procedure which is used to decompose a matrix into orthogonal matrix and triangle matrix. *A* matrix is decomposed using QR decomposition with the following equation **(Hemis, Boudraa, & Merazi-meksen, 2015) (Kaur, Dutta, Soni, & Taneja, 2014)** :

$$A = Q * R \tag{8}$$

where:
Q = Orthogonal Matrix ($Q^2.Q = I$)
R = Upper Triangular Matrix

1.5 Cartesian Polar Transform
Cartesian-Polar Transform (CPT) is a conversion from cartesian coordinate to polar coordinate. System of polar coordinates (r, θ), r is indicated the distance between the point of origin and θ is denoted the angle between a reference line. The conversion from cartesian to polar coordinates is given by the following equation **(Dhar & Shimamura, 2012)**:

$$r = \sqrt{x^2 + y^2}, \theta = tan^{-1}(\frac{y}{x}) \tag{9}$$

where (x, y) is a point in a cartesian coordinate system. Transformation of polar coordinates to cartesian coordinates can be written as follows **(Dhar & Shimamura, 2012)**:

$$x = r \ cos\theta, y = r \ sin\theta \tag{10}$$

2. METHODS

This paper aims to design and analyze the performance of audio watermarking with the combination method of DCT-QR-CPT method with QIM embedding method. Designing and analyzing this audio watermarking method have passed series methodology stages in the following explanation.

2.1 Research Problem Identification
As described in Section 1, the problems of audio watermarking in the previous research published in several literatures are classified into several points, such as :
    a. The system robustness is still not too high.
    b. Watermark payload is still below 100 bps.
    c. Several papers only described a few attacks for testing the system.

Because of that problem, we identify that DCT-QR-CPT-QIM method for audio watermarking design is a good solution. DCT has an energy compaction which compresses the signal energy to a certain region in frequency domain. It also makes the signal more robust before it is extracted by QR for getting stronger coefficients. CPT transforms QR coefficients into magnitude and phase component. In this proposed method, we embed the watermark using only the phase component of CPT output due to its robustness against amplitude scaling attack.

In order to design this audio watermarking method, we have studied several literatures containing :
  a. Audio watermarking in similar techniques
  b. Discrete Cosine Transform (DCT) and its properties
  c. QR Decomposition
  d. Cartesian to Polar Transformation (CPT)
  e. Quantization Index Modulation (QIM)
  f. Audio watermarking attacks
  g. Audio watermarking performance

2.2 Data Preparation
In this paper we prepare watermarking data, such as watermark data and host audio. There are 3 host audio files for audio watermarking testing, such as: voice.wav, jazz.wav, and rock.wav with audio duration 2-3 minutes, 44.1 kHz sampling rate and 16 bits audio quantization. This audio files have different genre, thus we can analyze how much the robustness and the imperceptibility are influenced by the genre of each audio file. Watermark image file is binary image with resolution 16x16 pixels displaying in Figure 1.

Figure 1. Watermark Image

2.3 Embedding Process
Embedding is a process or stage where an audio file will be inserted with a watermark data into host audio file. Figure 2 displays a block diagram of embedding process.
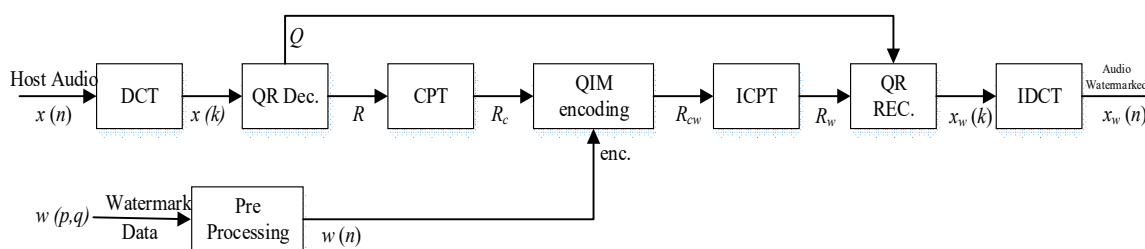
Figure 2. Embedding Process

Step 1:   Use image for data watermark with 2 dimensional matrix. The 2 dimensional matrix will be converted into a matrix one dimension as $x(n)$ in pre-processing and the output is $x_w(n)$.
Step 2:   Read host audio into one dimension matrix $x(n)$.
Step 3:   Use DCT for dividing host audio into several frame. Transform spatial domain into frequency domain. The transformed audio file $x(n)$ will be transformed again into

frequency domain by DCT using Equation (5). The output is $x(k)$ where the host file is already in frequency domain.

Step 4: Apply QR decomposition to each frame of DCT coefficients after reshaping DCT coefficients into square matrix. The square matrix is decomposed using QR into orthogonal matrix (Q) and triangular matrix (R). We embed the watermark only in R matrix at position (1,1) and (2,2). The matrix elements in both positions are processed by CPT using Equation (9) to produce magnitude and phase components. The phase component or θ is used for embedding process by QIM.

Step 5: The θ value is used for QIM embedding method using Equation (1) to produce $\hat{\theta}$.

Step 6: After the embedding process is done, $\hat{\theta}$ value and unchanged magnitude value are transformed into Cartesian coordinate using ICPT or Equation (10). Thus, we get $\hat{x}$ and $\hat{y}$ value that are used for (1,1) and (2,2) elements in $\hat{R}$ matrix.

Step 7: Using Equation (8), Q matrix and $\hat{R}$ matrix is reconstructed to obtain $x_w(k)$.

Step 8: Using IDCT in Equation (6), $x_w(k)$ is transformed back to time domain signal to get watermarked audio $x_w(n)$.

Step 9: Evaluate performances consisting of ODG as described in Subsection 2.5, payload using Equation (11), and SNR using Equation (12).

2.4 Extraction Process

Extraction process is a process to take back the watermark from a watermarked audio. The following figure is a block diagram of the stages in the extraction process.
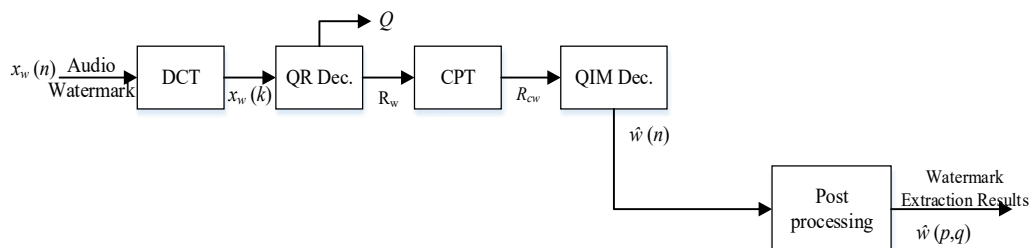

Figure 3. Extraction Process

Step 1: Change the watermarked audio $x_w(n)$ to one dimensional matrix.

Step 2: Apply DCT transformation for change time domain to frequency domain from $x_w(n)$ into $x_w(k)$.

Step 3: Each frame in the form of a matrix is composed to obtain a singular matrix. Use QR decomposition for decomposed matrix into orthogonal matrix and triangular matrix (R).

Step 4: Apply the CPT transformation just like the embedding process.

Step 5: Extract watermark using QIM method with Equation (5).

Step 6: Change the 1 dimensional matrix into 2 dimensions in post-processing and then take the watermark $\hat{w}(p,q)$.

Step 7: Calculate value of BER using Equation (13).

2.5 Experiment

Since embedding and extraction design are well prepared, the experiment of audio watermarking is ready to be executed. The experiment is executed by Matlab simulation in PC with i7 specification and RAM 16 GB. With the existing parameter, the first experiment is to test the simulation for obtaining robustness, imperceptibility and capacity of watermark without

attack to understand the effect of each parameter to the performance. The payload or capacity of watermark formula in bps (C) is described as following equation.

$$C = \frac{L_w F_s}{L} \tag{11}$$

where L_w is the length of a watermark in bits, L is the watermarked audio length in sample and F_s is the watermarked audio sampling frequency in sample/s. Objective measurement for evaluating the watermarked audio imperceptibility consists of ODG and SNR. ODG is a standard perceptual evaluation of audio quality (PEAQ) which is usually used for measuring the compressed or watermarked audio quality **(ITU-R, 1998)**. SNR (in dB unit) is a simple formula to measure audio quality as following equation.

$$SNR = 10 log_{10} \frac{\sum_{i=0}^{L-1} x^2(n)}{\sum_{i=0}^{L-1}(x(n) - x_w(n))^2} \tag{12}$$

BER is a formula to measure the robustness between original watermark and extracted watermark in binary digit based. The BER formula is displayed in following equation.

$$BER = \frac{correct\ extracted\ watermark\ (bits)}{all\ watermark\ (bits)} \tag{13}$$

The second experiment is to know the system performance with attack. Then, adjusting parameter in attack environment is needed for optimizing the performance result. Adjusting parameter is executed on several attacks. Adjusting parameter aims to have optimal performance which consisting of high imperceptibility with SNR>20 dB or MOS>4 or ODG>-2, robust with BER<10% and high capacity with payload>100 bps. After adjusting parameter, system with the optimal parameters is re-tested by the same attacks. Finally, we get the final result from the optimal parameter. Detail results of this experiment are described in next section. In this paper, we use several attacks for experiment, such as : LPF with cut off frequency 3k, 6k and 9k, BPF with cut off frequency 100 Hz-6k, 50 Hz-6k, 25 Hz-6k, Resampling with frequency 22k, 11k, and 16k, Linear Speed Change with percentage change 1%, 5%, 10%, and MP3 compression with rate 32k, 64k, 128k, and 256k. Subjective testing is conducted on 30 respondents to recognize watermarked audio quality. The subjective testing obtains an average of mean opinion survey (MOS) result from each audio with optimal parameter.

3. EXPERIMENTAL RESULT AND DISCUSSION

In this section, we describe the result of simulation experiment which consists of the analysis of audio watermarking performance without attack, the analysis of optimized audio watermarking parameter with attack, and the analysis of robustness result using optimized parameter.

3.1 Watermark Robustness
In Table 1 shows that the watermark inserted in 3 host audio has robust against several attacks like LPF, BPF, Resampling, Linear Speed Change and Mp3 Compression. Based on result, based on result from table 1, 3 parameter samples were taken for adjustment:
  1. Parameter 1: Attack MP3 Compression (32k) on host that shows BER value is 0.429
  2. Parameter 2: Attack BPF (25-6k) on Jazz that shows BER value is 0.32
  3. Parameter 3: Attack Mp3 Compression (64k) on Rock that shows BER value is 0.043

All of 3 parameters are re-adjusted until obtaining the best parameter and optimal performance. The parameter adjustment or optimization is conducted because usually the robustness from initial parameter is not good enough. The parameters need to be changed or adjusted empirically for a better robustness than a robustness from initial parameters.

Table 1. Robustness Performance before Parameter Adjusted

Attack	Type	BER		
		Host	**Jazz**	**Rock**
LPF	3k	0.03	0.20	0.06
	6k	0.00	0.05	0.00
	9k	0.01	0.02	0.00
BPF	100-6k	0.44	0.51	0.51
	50-6k	0.50	0.48	0.51
	25-6k	0.21	0.32	0.51
Resampling	22.05k	0.00	0.05	0.00
	11.02k	0.02	0.00	0.00
	16k	0.01	0.27	0.00
Linear Speed Change	1%	0.01	0.04	0.00
	5%	0.02	0.07	0.00
	10%	0.03	0.04	0.00
Mp3 Compression	32k	0.43	0.51	0.53
	64k	0.46	0.08	0.04
	128k	0.02	0.03	0.29
	192k	0.02	0.02	0.01
Average		**0.14**	**0.17**	**0.15**

3.2 Adjusting Audio Watermarking Parameter Using Attack

After we test audio watermarking using unadjusted parameter with attacks, the adjusted parameter will be tested by several attacks. This procedure is useful for decrease BER values while the capacity still tolerable. In this test we used 3 samples of audio to be tested. The attack that we used are LPF, BPF, Resampling, Linear Speed Change and MP3 Compression. Table 1 displays the robustness result for 3 audio files. The highlight cells in Table 1 are selected for parameter adjustment. The adjustment result is displayed in Table 2. Table 2 shows several optimal parameters such as: decomposition level (N), sample per frame (N frame) and quantization bit (*nbit*) with the robustness.

1. Parameter 1
   The audio used is Host.wav and attacked by MP3 Compression 32k. The result of this parameter is SNR value of 25.58 dB, BER value of 0.056 and payload 269.16 bps.
2. Parameter 2
   The audio used is Jazz.wav and attacked by BPF 25-6k. The result of this parameter is SNR value of 24.29, BER value of 0.046 and payload 269.16 bps.
3. Parameter 3
   The audio used is Rock.wav and attacked by MP3 with rate 64k result of this parameter is SNR value of 24.39, BER value of 0.062 and payload 269.16 bps.

As a result, we choose parameter 3 for the robustness test, because the robustness result of optimized parameter on parameter 3 has the lowest average of BER.

Table 2. Optimal Parameter Using Attack and Its Robustness

Host	Parameter	*N*	*NFrame*	*nbit*	Average BER
Host	1	3	1024	4	0.218
Jazz	2	1	1024	4	0.221
Rock	3	2	1024	4	0.217

3.3 Robustness with Adjusted Parameter

We report the result of robustness testing after parameter adjustment using parameter 3. The following table displays the result of watermark using optimized parameter with several attacks. Based on result from Table 3, from attacks from LPF, Resampling, Linear Speed Change and Mp3 Compression with rate 64 kbps and above, watermarking obtain good robustness with BER<10%.

Table 3. Robustness Performance after Parameter Adjusted

Attack	Type	BER		
		Host	Jazz	Rock
LPF	3k	0.00	0.02	0.00
	6k	0.02	0.02	0.00
	9k	0.02	0.02	0.00
BPF	100-6k	0.39	0.46	0.47
	50-6k	0.15	0.17	0.38
	25-6k	0.09	0.07	0.09
Resampling	22.05k	0.00	0.04	0.00
	11.02k	0.02	0.04	0.00
	16k	0.02	0.01	0.00
Linear Speed Change	1%	0.01	0.00	0.00
	5%	0.01	0.05	0.00
	10%	0.00	0.05	0.00
Mp3 Compression	32k	0.16	0.11	0.23
	64k	0.08	0.05	0.06
	128k	0.01	0.02	0.02
	192k	0.01	0.04	0.00
Average		**0.06**	**0.07**	**0.08**

3.4 Watermark Quality Degradation

The BER value of watermark robustness result shows the quality of watermark image extraction from watermarked audio was attacked. The higher of BER value then the quality watermark image extraction is bad. But there is a maximum limit BER value on watermark image that is still acceptable to human if the image still can be recognized although the image was damaged. The BER maximum limit depends on the image resolution, in this test we inserted watermark image with a 16x16 resolution as displayed in Figure 1 where the original image can be seen on Table 4 when BER=0. When the watermark image is damaged and

obtained BER=0.01 to BER=0.09 it can still be recognized by human vision, but when the BER is above 0.09, the extracted watermark image cannot be recognized by human vision because it is damaged very badly. This shows that if watermark robustness has BER below 0.1 at 16x16 image resolution, the image is still acceptable or it is still robust to attack if the quality of the watermark has BER value below 10%.

Table 4. Watermark Degradation

BER	Watermark	BER	Watermark	BER	Watermark
0		0.05		0.15	
0.007		0.06		0.16	
0.01		0.08		0.17	
0.02		0.09		0.23	
0.04		0.11		0.38	

3.5 Watermarked Audio Quality

There are two assessments for audio quality, first is by objective which is calculate value of SNR and ODG and second is subjective which is giving test the watermarked audio to respondent and the respondent give rate for MOS value of the audio watermark with range 1 to 5. In Table 5, objective test obtains good imperceptibility of the watermarked audio files with SNR>20 dB and ODG>-1. For subjective test we requested 30 respondents to compare 3 original host audios with 3 watermarked host audios. In the same audio files, subjective test also obtains good imperceptibility with MOS>4. The watermarked audio is obtained from the embedding process from parameter 3.

Table 5. Watermarked Audio Quality

Host	MOS	SNR	ODG
Host.wav	4.03	24.01	-0.19
Jazz.wav	4.23	24.85	-0.3
Rock.wav	4.26	24.39	-0.12

4. CONCLUSION

In this paper, we conclude that there are several parameters affecting SNR and BER, such as level decomposition (N), sample per frame (*Nframe*) and quantization bit (*nbit*) have any effects to SNR and BER value. From subjective measurement, MOS obtained from adjusted parameter is still acceptable by the human sense of hearing. With adjusted parameter, audio watermarking in this paper is robust against several attack such as LPF, BPF 25-6k,

Resampling, Linear Speed Change and Mp3 Compression above 32k with BER <10% and its imperceptibility is good with ODG>-1, SNR>20 dB, MOS>4 and payload 269.16 bps.

ACKNOWLEDGEMENT

REFERENCES

Adiwijaya, Novraditya, E., Baihaqi, I. F. N., & Wisesty, U. N. (2013). Copyright Protection in Audio File Using Watermarking Approach Based on Wavelet-SVD. *Applied Mechanics and Materials*, *321–324*, 1191–1195. https://doi.org/10.4028/www.scientific.net/AMM.321-324.1191

Agradriya, B. A. F., Perdana, F. K., Safitri, I., & Novamizanti, L. (2017). Watermarking Technique Based On Arnold Transform. In *2017 2nd International Conference on Automation, Cognitive Science, Optics, Micro Electro-Mechanical System, and Information Technology* (pp. 17–21).

Aminullah, R., Budiman, G., & Safitri, I. (2017). Implementation of QIM Based Audio Watermarking Using Hybrid Transform Of SWT-DCT-SVD Methods Optimized With Genetic Algorithm. *Jurnal Penelitian Dan Pengembangan Telekomunikasi, Kendali, Komputer, Elektrik, Dan Elektronika*, *2*(1).

Aparna J R, & Ayyappan, S. (2014). Comparison of digital watermarking techniques. *2014 International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC)*, 87–92. https://doi.org/10.1109/ICCPEIC.2014.6915345

Bahi, A. El, & Adib, A. (2014). A High Capacity Quantization-Based Audio Watermarking Technique Using the DWPT. In *International Conference on Multimedia Computing and Systems (ICMCS)* (pp. 1–5). IEEE. https://doi.org/10.1109/ICMCS.2014.6911215

Bajpai, J., & Kaur, A. (2016). A literature survey - Various audio watermarking techniques and their challenges. In *Proceedings of the 2016 6th International Conference - Cloud System and Big Data Engineering, Confluence 2016* (pp. 451–457). https://doi.org/10.1109/CONFLUENCE.2016.7508162

Baranwal, N., & Datta, K. (2011). Comparative Study of Spread Spectrum Based Audio Watermarking Techniques. In *IEEE-International Conference on Recent Trends in Information Technology, ICRTIT 2011* (pp. 896–900).

Budiman, G., Suksmono, A. B., & Danudirdjo, D. (2016). Fibonacci Sequence – based FFT and DCT performance comparison in Audio Watermarking. In *The International Conference*

on Science, Engineering, Built Environment, and Social Science (ICSEBS) (pp. 2–8).

Budiman, G., Suksmono, A. B., Danudirdjo, D., & Pawellang, S. (2018). QIM-based Audio Watermarking with Combined Techniques of SWT-DST-QR-CPT Using SS-based Synchronization. In 2018 6th International Conference on Information and Communication Technology (ICoICT) (pp. 286–292).

Chen, B., & Wornell, G. W. (2001). Quantization index modulation: A class of provably good methods for digital watermarking and information embedding. IEEE Transactions on Information Theory, 47(4), 1423–1443. https://doi.org/10.1109/18.923725

Dhar, P. K., & Shimamura, T. (2012). Audio Watermarking in Transform Domain Based on Singular Value Decomposition and Quantization. In Asia-Pacific Conference on Communications (APCC) (pp. 516–521). IEEE. https://doi.org/10.1109/APCC.2012.6388247

Dhar, P. K., & Shimamura, T. (2015). Advances In Audio Watermarking Based on SVD. New York, London: Springer. https://doi.org/10.1007/9783319148007

Hemis, M., Boudraa, B., & Merazi-meksen, T. (2015). Intelligent Audio Watermarking Algorithm using Multi-objective Particle Swarm Optimization. In International Conference on Electrical Engineering (ICEE) (pp. 0–4). IEEE. https://doi.org/10.1109/INTEE.2015.7416777

ITU-R, R. RECOMMENDATION ITU-R BS.1387-1 Method for objective measurements of perceived audio quality, BS.1387-1 International Telecommunications Union-Recommendation , Geneva (1998). Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Method+for+object ive+measurements+of+perceived+audio+quality+1#0%5Cnhttp://scholar.google.co m/scholar?hl=en&btnG=Search&q=intitle:1387,+Method+for+objective+measureme nts+of+perceived+audio+qualit

Kaur, A., Dutta, M. K., Soni, K. M., & Taneja, N. (2014). A High Payload Audio Watermarking Algorithm Robust against Mp3 Compression. In 2014 Seventh International Conference on Contemporary Computing (IC3) (pp. 531–535). https://doi.org/10.1109/IC3.2014.6897229

Khademi, N., Akhaeet, M. A., Ahadi, S. M., Moradi, M., & Kashi, A. (2007). Audio Watermarking based on Quantization Index Modulation in the Frequency Domain. In 2007 IEEE International Conference on Signal Processing and Communications (ICSPC 2007) (pp. 1127–1130).

Li, J., & Wu, T. (2015). Robust Audio Watermarking Scheme via QIM of Correlation Coefficients

using LWT and QR decomposition. In *International Conference on Informative and Cybernetics for Computational Social System (ICCSS)* (pp. 1–6).

Nikmehr, H., & Hashemy, S. T. (2010). A new approach to audio watermarking using Discrete Wavelet and Cosine Transforms. In *International Conference on Communications Engineering* (pp. 1–10). Retrieved from http://seminars.usb.ac.ir/Files/iccome/en-us/Document/1394_IEEE_manuscript0.pdf

Safitri, I., & Ginanjar, R. R. (2017). Multilevel Adaptive Wavelet BCH Code Method for Copyright Protection in Audio Watermarking System. In *2017 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)* (pp. 30–35). https://doi.org/10.1109/ICCEREC.2017.8226712

Warkar, R., More, P., & Waghole, D. (2015). Digital Audio Watermarking and Image Watermarking for Information Security. In *2015 International Conference on Pervasive Computing (ICPC)* (Vol. 3).

Yan, Y., Rong, H., & Mintao, X. (2009). A Novel Audio Watermarking Algorithm for Copyright Protection Based on DCT Domain. *2009 Second International Symposium on Electronic Commerce and Security*, 184–188. https://doi.org/10.1109/ISECS.2009.141

Zhou, Z., & Zhou, L. (2007). A Novel Algorithm for Robust Audio Watermarking Based on Quantification DCT Domain. In *Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing*. https://doi.org/10.1109/IIH-MSP.2007.46