Audio Watermarking using Group Amplitude Quantization

Sugumar D
Karunya University Coimbatore

Jismy N Zacharias
Karunya University Coimbatore

P.T. Vanathi, PhD.
P.S.G. Tech Coimbatore

ABSTRACT
This paper presents an optimization-based audio watermarking scheme using group amplitude quantization. In order to improve the robustness, the watermark is embedded in the lowest discrete wavelet transform frequency coefficients. The performance of this watermarking scheme is analyzed in terms of audio quality (signal-to-noise ratio) and robustness (bit error rate). As there is a trade off relationship between the audio quality and robustness, this study presents an optimization-based group amplitude quantization scheme for audio watermarking. First, SNR is rewritten as a watermarking cost function in matrix form. Then an equation connecting the watermarking cost function and the group amplitude quantization equation is proposed. Second, the Lagrange principle is used to derive the optimization solution. The optimal results are then applied to embed the watermark. Finally, the performance of the proposed scheme is tested. As per the experimental results, the SNR (audio quality) of the proposed scheme is 25.21 dB and the BER (robustness) is 0.4720. The hidden data are robust against most common attacks, such as re-sampling, low-pass filtering and amplitude scaling.

General Terms
Audio Watermarking, Group Amplitude Quantization

Keywords
Audio Watermarking, DWT, Group Amplitude Quantization, SNR, BER.

1. INTRODUCTION
Audio watermarking plays a significant role in multimedia security. Many audio watermarking methods have been proposed recently. Watermarking has the following requirements: (i) the watermark should be imperceptible in the embedded audio. (ii) The embedding technique should offer at least 20 dB signal-to-noise ratio (SNR) for watermarked audio against the original one. (iii) The embedded watermark should be able to prevent common attacks, such as re-sampling, low filtering and amplitude scaling.

A watermark, which usually consists of a binary data sequence, image, and audio, is inserted into the host signal in the watermark embedder. Thus, a watermark embedder has two inputs, one is the watermark message and the other is the host signal (e.g. .image, video clip, audio sequence etc.). The output of the watermark embedder is the watermarked signal. In the detection side the watermarked signal and the watermarking signal are given as inputs and watermark is detected. Without using host signal, detect the watermark in the detection side.

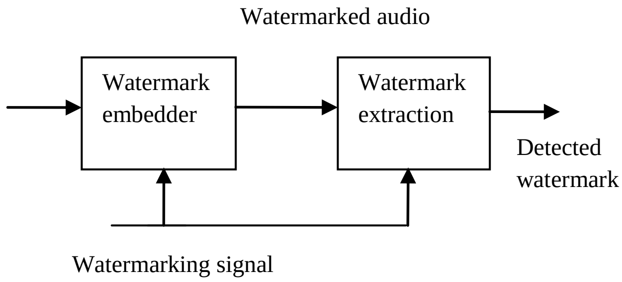

Fig 1 Generic audio watermarking block diagram

The methods for audio watermarking are grouped into two, time domain and frequency domain. The main weakness of time domain analysis is low robustness. To overcome this problem, in [2] adopted group amplitude modification to improve robustness in the time domain. But it has extremely low SNR. In the frequency domain, Huang [9] embed the watermark into discrete cosine transform (DCT) coefficients and hides the synchronization code in the time domain. Because the time domain has low embedding strength, synchronization codes are not robust enough. However embedding synchronization codes in DCT coefficients will increases the computation cost. The embedding in lowest frequency coefficients will improve the robustness of the watermarking scheme. Therefore watermark embedding in frequency domain using optimization based group amplitude quantization is employed in this paper.

In this paper, an audio watermarking algorithm that optimizes the trade-off relationship between audio quality and robustness is proposed. Audio quality and robustness can be measured in terms of SNR and BER respectively. The audio watermarking scheme used here is the optimization based group amplitude quantization. Lagrange principle is used to obtain the optimized solution. The performance of this scheme can be evaluated in terms of SNR and BER. The watermarked audio has good audio quality and strong robustness against attacks like re-sampling, low pass filtering and amplitude scaling.

The rest of this paper is organized as follows. Watermark embedding is described in section 2. Watermark extraction procedure is explained in section 3. Section 4 deals with the experimental results and analysis in terms of attacks are drawn in section 5. Section 6 deals with conclusion.

2. PROPOSED WATERMARK EMBEDDING PROCESS
In this section, optimization based watermarking is introduced. The audio watermarking scheme used here is group amplitude quantization. The watermark is embedded in to lowest frequency coefficients of audio signal.

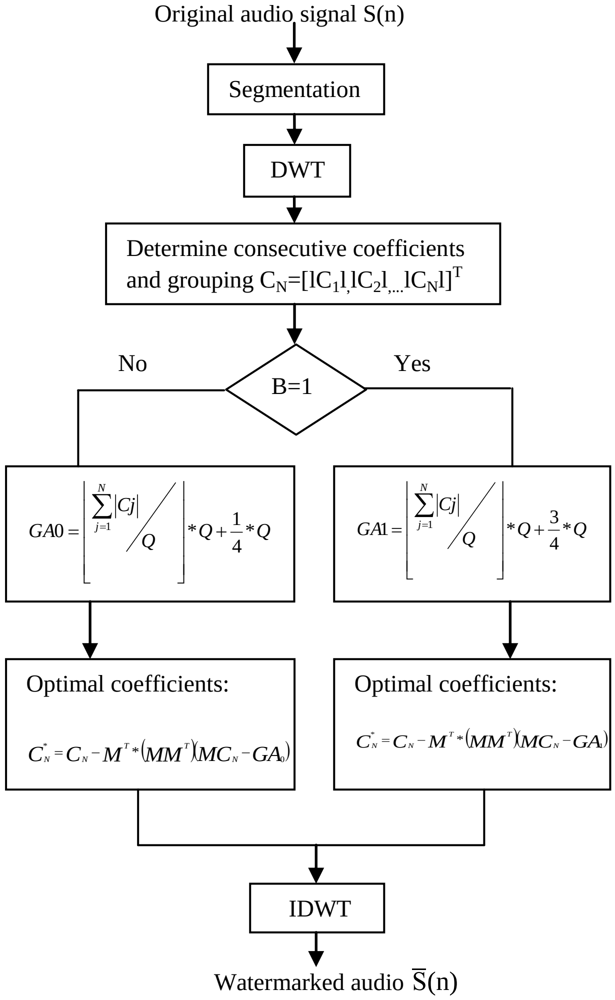

Fig 2 Flowchart for watermark embedding

2.1 Segmentation

For this paper, an audio file is taken and sampled at a sampling rate of 44100 samples per second. Then, partition the sampled file into segments without overlapping. Each segment contains the same amount of samples; the number of samples per segment depends on the used sample rate. Segmentation is shown in the following figure. Audio source is divided into different segments S1, S2, S3 etc. Each segment has *n* samples. In these paper 256 samples per frame is considered.

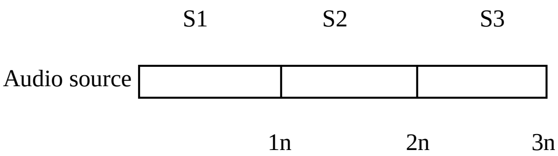

Fig 3 Segmentation

2.2 Discrete Wavelet Transform (DWT)

Audio watermarking in frequency domain is more robust than time domain. In this paper, watermark is embedded into the lowest frequency wavelet coefficients. DWT is applied on each audio segment up to seven levels and the wavelet filter used here is db4. The audio signal is decomposed into low frequency and high frequency components. The watermarking sequence is embedded into lowest frequency components in the seventh level.

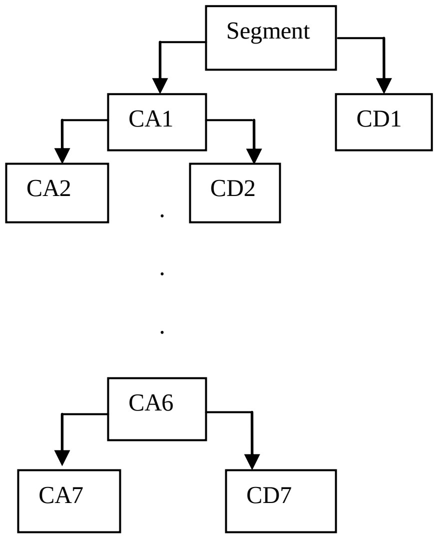

Fig 4 Seven level decomposition

To achieve robustness, the watermark is embedded only into the lowest frequency coefficients. Therefore only the approximation coefficients are decomposed seven times. The same seven level filter banks are used for the reconstruction

2.3 Watermarking Sequence

In this paper, a binary PN sequence is used as the watermark. The watermarked audio may undergo shifting or cropping. To achieve strong robustness, synchronization codes are embedded together with the watermark. The Gold sequence is used as the synchronization code in this paper. These synchronization codes are used to locate the positions where the watermark is embedded. The watermark and synchronization codes are combined to form a binary watermarking sequence B.

...	Sync	PN	Sync	PN	...

Fig 5 Structure of watermarking sequence B

Every N consecutive coefficient are grouped into a vector form,

$$C_N=[\,|C_1|,|C_2|,...|C_N|\,] \qquad (1)$$

For the consideration of high embedding capacity here the N value is set to be 2. The seventh level coefficients are grouped and the watermarking sequence has been embedded into these coefficients.

2.4 Amplitude quantization

Quantization is the process of mapping a large set of input values to a smaller set such as rounding values to some unit of precision. Here two sets of quantization process is taking place, depending upon the binary bit value B.

If the embedding bit B=1, then the amplitude of the coefficient is quantized to,

$$GA0 = \left\lfloor \frac{\sum_{j=1}^{N} |Cj|}{Q} \right\rfloor * Q + \frac{1}{4} * Q \qquad (2)$$

If B=0 then, the amplitude of the coefficient is quantized to

$$GA1 = \left\lfloor \frac{\sum_{j=1}^{N} |Cj|}{Q} \right\rfloor * Q + \frac{3}{4} * Q \qquad (3)$$

Where Q is the quantization parameter, $\lfloor\ \rfloor$ indicates the low floor function.

The audio quality depends on the value of quantization parameter. If the value of Q is very large then the noise in the watermarked audio will be high. The Q value should be small. In this paper amplitude quantization is used as the embedding process. The group amplitude value (GA) will decide the quality of the signal.

2.5 Optimization based embedding

The key attribute of this paper is the optimization of the trade off between the SNR and BER. For the optimization, first, SNR is rewritten using watermarking cost function in matrix form. Then an equation connecting the watermarking cost function and group amplitude quantization is formed. Finally, Lagrange principle is used to find the optimal solution. This optimal solution is then applied to embed the watermark.

Optimization of SNR is as follows:

$$\begin{aligned} SNR &= -10\log_{10}(\|\bar{S}(n)-S(n)\|_2^2/\|S(n)\|_2^2) \\ &= -10\log_{10}(\|\bar{C}_N-C_N\|_2^2/\|C_N\|_2^2) \end{aligned} \qquad (4)$$

Where S(n) is the original audio signal,
$\bar{S}(n)$ is the watermarked signal,
C_N is the lowest frequency coefficient,
$\bar{C}_N$ is the watermarked coefficient

For the optimization SNR is rewritten as watermarking cost function,

$$f(\bar{C}_N) = (\bar{C}_N - C_N)^T (\bar{C}_N - C_N) / C_N^T C_N \qquad (5)$$

In the case B=1, the optimization based quantization is in the following form,

$$\text{Minimize} \quad f(\bar{C}_N) = (\bar{C}_N - C_N)^T (\bar{C}_N - C_N) / C_N^T C_N, \qquad (6a)$$

$$\text{Subject to} \quad g(\bar{C}_N) = M\,\bar{C}_N - GA1 \qquad (6b)$$

Where M is scaling factor matrix.

To embed B, need to solve (6a) and (6b).To derive the optimal results, Lagrange principle is used. Lagrange parameter is used to compute optimized coefficients.

$$\lambda = 2*(MM^T)^{-1}*(MC_N - GA1)/(C_N C_N^T) \qquad (7)$$

This λ value can be used to compute the optimized coefficients and these coefficients are used for embedding.

$$C_N^* = C_N - (C_N C_N^T M^T \lambda)/2 \qquad (8)$$

3. WATERMARK EXTRACTION

In this section, extraction of embedded bits is described.

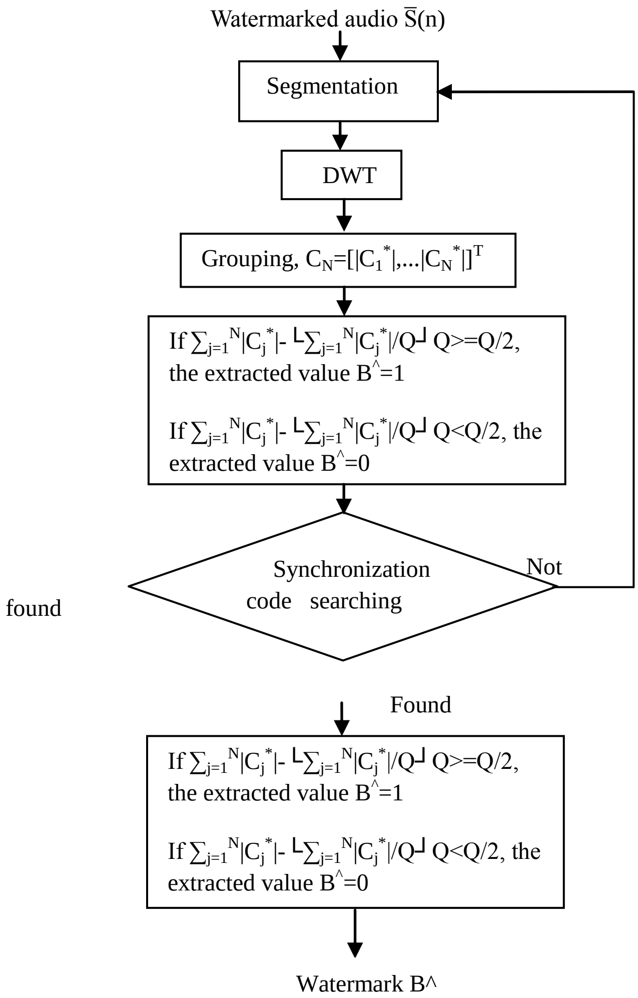

Fig 6 Flowchart for watermark extraction

First, watermarked audio is segmented into different sections. DWT is applied on each segment. Seven level DWT is used here. Seventh level coefficients are grouped to form a vector. Then the watermark extraction will take place.

The watermark will get extracted by checking the following conditions:

$If\ \sum_{j=1}^{N}|C_j^*| - \left\lfloor \sum_{j=1}^{N}|C_j^*| \big/ Q \right\rfloor * Q \geq \frac{Q}{2},$ then the extracted value will be $B^\wedge=1$.

$If\ \sum_{j=1}^{N}|C_j^*| - \left\lfloor \sum_{j=1}^{N}|C_j^*| \big/ Q \right\rfloor * Q < \frac{Q}{2},$ then the extracted value will be B^=0.

Where B^ is the watermarked audio.

Now extract the synchronization code and binary PN sequence from the extracted watermark B^. Synchronization codes are embedded together with the watermarking binary PN sequence to locate the positions where the watermark is embedded. If the synchronization code is not detected correctly then, repeat the steps again.

4. PERFORMANCE EVALUATION

In this paper, an audio file named music.wav with 16 bit resolution and mono channel is taken as the input audio signal for duration of 11.6 seconds. The sampling rate of the audio signal is 44.1 kHz. The input audio is shown in the above figure. The audio is then divided into 2000 segments and each segment consists of 256 samples. PN sequence of 1000 bits is used as the watermark and gold sequence of 1000 bits is used as the synchronization code. The PN sequence, Gold sequence and the combined binary sequence B are shown in the following figures 8,9,10 respectively.

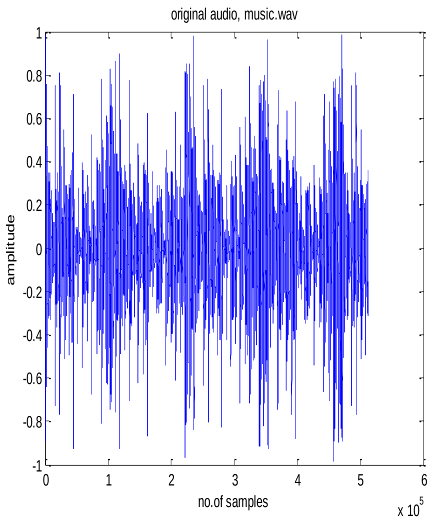

Fig 7 Original audio signal, music.wav

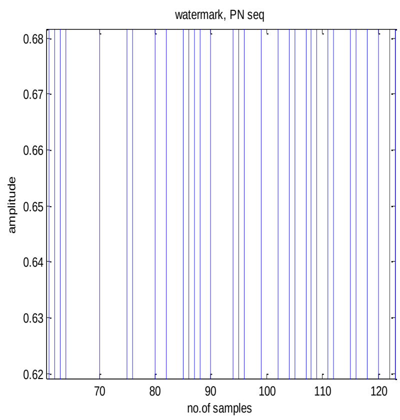

Fig 8 Watermark, PN sequence (1000 bits)

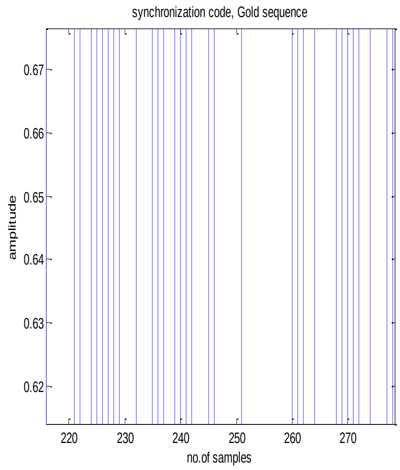

Fig 9 Synchronization code, Gold sequence (1000 bits)

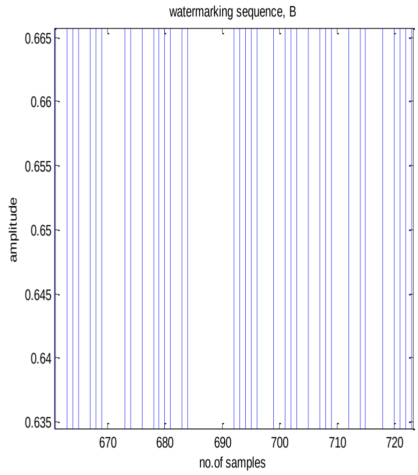

Fig 10 Watermarking sequence, B (2000 bits)

To achieve strong robustness, synchronization codes are embedded together with the watermark, into the lowest frequency coefficients. According to the binary bit value of the watermarking sequence, the low frequency coefficient will quantise and the quantization parameter used is 50.

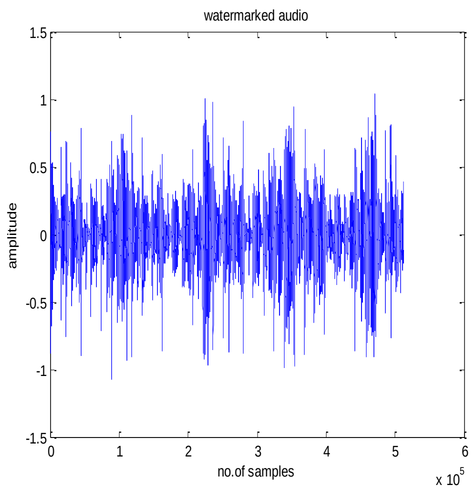

Fig 11 Watermarked audio

The performance of the audio watermarking scheme is analysed in terms of audio quality and robustness. The audio quality is measured in terms SNR and robustness is measured in terms of BER. The experimental results are given in following tables 1 and 2.

To check the performance of the watermarked audio, different wavelet filters are employed. The filters used for comparison are db1,db2,db4 and orthogonal filters. The SNR and BER of each filter has been calculated and made a comparative study. When considering both SNR and BER, db4 has better performance than the other filters.

Table 1 Comparison with different wavelet filters

	db1	db2	db4	orthogonal
SNR(dB)	0.0023	25.7203	25.2172	25.1995
BER	0.4195	0.4890	0.4720	0.5075

Table 2 SNR (dB) measurements

SNR(dB) of i/p audio signal	SNR(dB) of watermarked audio	SNR(dB) of extracted audio
41.1571	25.2172	37.2863

The robustness obtained in terms of BER is 0.4720.The experimental results show that the watermarked audio has good audio quality and better robustness.

5. ANALYSIS IN TERMS OF WATERMARKING ATTACKS

To test the robustness of the watermarked audio, some attacks have been performed on the watermarked audio. These attacks include three types (i) re-sampling, (ii) amplitude scaling and (iii) low pass filtering.

5.1 Re-sampling

The watermarked audio was down sampled from 44.1 kHz to 22.05 kHz, and then back to 44.1 kHz. Similarly the sampling rate was varied from 44.1 kHz to 11.025,8 and 6 kHz, and then back to 44.1 kHz. The results show that the watermarked audio has better robustness even in the presence of attacks. The results are tabulated and are shown below:

Table 3 Re-sampling

Re-sampling rate (Hz)	BER
6000	0.4775
8000	0.4870
11025	0.4970
22050	0.5180

5.2 Amplitude scaling

Under amplitude scaling, both amplification and attenuation attack are applied on the watermarked audio. For this, set the scaling factors for amplification as 2,3,4,5 and for attenuation as 0.2, 0.8, 1.1 and 1.2. The results are tabulated below:

Table 4 Amplitude scaling

Scaling factor	BER
0.2	0.5025
0.8	0.4850
1.1	0.4955
1.2	0.5125
2	0.5095
3	0.5065
4	0.4930
5	0.4815

5.3 Low pass filtering

In low pass filtering, it shows the effect of adopting low pass filtering with the cut off frequency of 12 kHz. The low pass filter is designed using filter design and analysis (FDA) tool. The BER obtained is 0.5075.

6. CONCLUSION

The illegal distribution of digital audio products and music files in particular, has been a major problem for industry for more than a decade. In this paper, an imperceptible (inaudible) and robust audio watermarking technique based on Group Amplitude Quantization is proposed. To balance the trade off between SNR and BER, optimization based group amplitude quantization technique is used. In order to improve the robustness, the watermark is embedded into the lowest frequency coefficients of DWT. Synchronization code and watermarking PN sequence are together embedded into these lowest frequency coefficients. The performance of this audio watermarking scheme is analysed in terms of audio quality and robustness. Audio quality and robustness is measured in terms of SNR and BER. The measured SNR and BER are 25.2172 dB and 0.4720. The db4 wavelet filter has better performance than the other wavelet filters. The embedding data are robust against attacks like re-sampling, amplitude scaling and low pass filtering. The simulation results obtained verify the effectiveness of the audio watermarking as a reliable solution to the copyright protection which is facing in the music industry.

7. ACKNOWLEDGEMENT

I would like to thank my guide Mr.Sugumar D for his sincere support and being approachable and available whenever I needed any assistance. Above all I render my gratitude to the Almighty God who bestowed self confidence, ability and strength in time to complete this task

8. REFERENCES

[1] S.T Chen, G.D Wu, H.N. Huang, IET Signal Processing 2010, Wavelet domain audio watermarking scheme using optimization based quantization

[2] LIE W.N., Chang L.C., IEEE Trans. Multimedi., 2006, Robust and high quality time domain audio watermarking based on low frequency amplitude modification

[3] W.U S, HUANG J, HUANG D, SHI Y.Q., IEEE Trans. Broadcast, 2005, Efficiently self synchronized audio watermarking for assure audio data transmission

[4] KIRVOSKI D., MALVAR H.S., IEEE Trans. Signal Process, 2003, Spread spectrum watermarking of audio signals

[5] CHEN B., PITAS I., WORNELL G.W., IEEE Trans. Inf. Theory, 2001, Quantization index modulation-a class of provably good methods for digital watermarking and information embedding

[6] BASSIA P., PITAS I., NIKOLAIDIS N., IEEE Trans. Multimed. 2001, Robust audio watermarking in the time domain

[7] ALARYANI H., YOUSSEF A., Proc. Seventh IEEE Int. Symp. on Multimedia 2005, A novel audio watermarking technique based on frequency components

[8] KIM H.J., Pacific Rim workshop on Digital Steganography Kyushu Institute of Technology, Kitakkyushu, Japan, 2003 Audio watermarking techniques

[9] HUANG J., WANG Y., SHI Y.Q., Proc. IEEE Int. Symp. Circuits and Systems, 2002, A blind audio watermarking algorithm with self synchronization

[10] KIM H., Proc. Third IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), 2001, Stochastic model based audio watermarking whitening filter for improve detection

[11] LI X., YU H.H., Proc. IEEE Int. Conf. on Information Techniques: Coding and Computing, 2000, Transparent and robust audio data hiding in sub-band domain

[12] KATZENBEISSER S., PETITCOLAS F.A.P., Artech House Inc., 2000, Information hiding techniques for steganography and digital watermarking