Frequency Hopping Method for Audio Watermarking

Ana Anastasijević and Duška Čoja

Abstract — **This paper evaluates the degradation of audio content for a perceptible removable watermark. Two different approaches to embedding the watermark in the spectral domain were investigated. The frequencies for watermark embedding are chosen according to a pseudo-random sequence making the methods robust. Consequentially, the lower quality audio can be used for promotional purposes. For a fee, the watermark can be removed with a secret watermarking key. Objective and subjective testing was conducted in order to measure degradation level for the watermarked music samples and to examine residual distortion for different parameters of the watermarking algorithm and different music genres.**

Keywords — **Audio watermarking, copyright protection, frequency hopping method, pseudo-random noise.**

I. Introduction

THE addition of information which is not inherent to content itself into audio, video or multimedia for various different purposes is called watermarking. This additional information is called a watermark and can be extracted as a code or its presence can be detected enabling either secret data transmission or various kinds of content verification. For these reasons, digital watermarking is applied in a number of fields ranging from copyright protection to confidential data transmission and has gained significant attention in research as well as development. With Internet expansion, data integrity and copyright ownership becoming increasingly important, the area of digital watermarking is growing more appealing for engineers and researchers than ever before [1].

Copyright protection and/or control as well as content fingerprinting and ownership assertion are just some of the applications of watermarking multimedia. Digital watermarking can also be used for tracking the use of audio and/or multimedia or monitoring broadcasting or even, with developing additional services, automatic billing for viewing. Depending on specific needs, different types of watermarking methods should be applied.

Requirements impose whether the technique (or techniques) used should result in perceptible or imperceptible watermarking, the level of content quality degradation, etc. [1]. Adding pseudo-noise (PN) into media has become a fairly common procedure in audio watermarking in the near past [2] with the possibilities for adding PN in time domain [3], DCT (Discrete Cosine Transform) domain [4], wavelet domain [5], [6], FFT (Fast Fourier Transform) coefficients [7]-[9], etc.

Perceptible audio watermarking in the spectral domain for adding copyright owner's authentication to media is the main focus of this paper. The approaches tested expand on [7] by modifying various parameters of implementation and testing as well as suggesting an alternative technique in applying this watermarking method.

Perceptible watermarking of audio content produces media of lower quality, yet good enough for promotional services. Therefore, in the advent of Internet, authors can insert pseudo noise into original content to deliberately decrease sound quality and circulate the noisy media as a demo or teaser. For a fee, the authors can remove the added noise and the users can enjoy excellent music quality.

To make all this possible, certain conditions have to be met: added noise has to have 'a method to it', meaning that it has to be pseudorandom noise so that the process of watermarking can be reversible (extraction of the original of high quality from the media of lower quality), but such that the audio with added noise has to be of noticeably lower quality (so that the users of the demo are not satisfied with it in its current state) yet the features of the original audio should not be missed by a less discerning listener. In addition, added noise has to be robust enough to disable any unauthorized watermark removal. Based on the algorithm presented in [7], a procedure for varying parameters of the watermark process was developed in order to guarantee that the degradation of watermarked content is small but noticeable. Degradation levels were measured for the watermarked audio content for different parameters and different genres. Afterwards, the quality of the recovered music samples was tested.

The paper is organized as follows. Section II describes the watermarking method and implementation used to protect audio. Results obtained experimentally are presented in Section III. Lastly, Section IV concludes the paper.

II. Explaining the Method

The main idea of this project is protection from unauthorized use of audio signals and thus perceptible watermark is embedded into the host audio.

Protection scheme of this method is proposed in Fig. 1 [7]. Fig.1. explains the main purpose of embedding the perceptible watermark. Only the listeners who buy the watermark key are able to remove the watermark from the watermarked audio and listen to the high quality audio.

Ana Anastasijević is a student at the School of Electrical Engineering, University of Belgrade, Serbia (e-mail: aanastasijevic@gmail.com).

Duška Čoja is a student, School of Electrical Engineering, University of Belgrade, Serbia (e-mail: duskacoja@gmail.com).

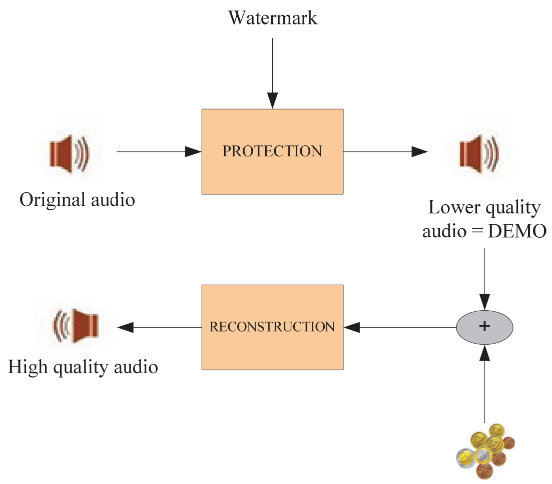

Fig. 1. Protection scheme.

Methods proposed here are based on frequency hopping technique and use pseudo-random noise in order to prevent users from trying to crack the watermarked signal by themselves.

Watermark is embedded into the spectrum of the audio signal and positions for embedding are chosen based on the pseudo random sequence a.k.a. the watermark key and this “randomness” provides robustness to the method [10].

A. Watermark Embedding

The embedded watermark is a pseudorandom noise sequence. The basic embedding scheme for both techniques is shown in Fig. 2 [7].

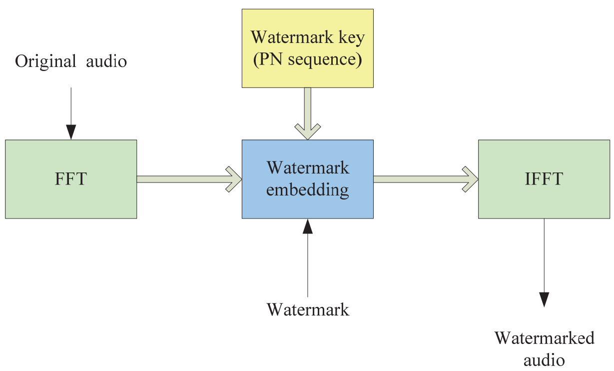

Fig. 2. Basic embedding scheme.

The watermark is first embedded similarly to [7] by dividing the signal into non-overlapping time segments consisting of 1024 samples. FFT is performed on each of the non-overlapping time segments and based on the pseudo-random sequence generated using the secret watermark key, two FFT coefficients are chosen for watermark embedding. The procedure of embedding the watermark is the same for these two methods and will be explained later on.

Afterwards, inverse FFT is performed to acquire non-overlapping time segments with embedded watermark. This is done for the entire audio signal and the acquired time segments are concatenated into watermarked audio.

The other technique entails decomposing host data (audio signal into which the watermark is embedded) into coefficients of an N-point FFT and dividing into sub-bands of equal number of coefficients. As in the previous technique, a sub-band consists of 1024 samples and a pseudo-random sequence is generated using the secret watermark key. Again, based on this sequence in each of the sub-bands two FFT coefficients are chosen for watermark embedding. The difference is that with this method the division of host data is done in the spectral domain rather than in the time domain.

The following step is applied for both techniques: the mean value of each sub-band is calculated and watermark is embedded so that the mean value doesn’t change.

Depending on which bit is to be embedded the process varies – if bit 1 is to be embedded the value of the coefficient with lower frequency is set to be k dB above the mean value, while the value of the other coefficient is set to be k dB below the mean value. If bit 0 is to be embedded the operation is done vice versa [7]. The fact that mean value of each sub-band remains unchanged is used when removing the watermark. Embedding process is shown in Fig. 3 [7].

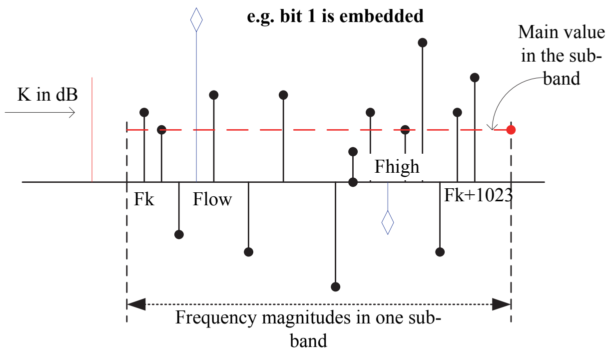

Fig. 3. Embedding process.

The value of distortion – k is not chosen randomly, it must be such that the quality of audio is not extremely destroyed, but lowered for the purposes of demo. K is taken as a percentage of the distance between the mean value and the hearing threshold approximation.

This is the value that defines whether the watermark embedded into the host audio is perceptible or not. If a different value of distortion had been chosen, the listeners wouldn’t have been able to distinguish the difference between the original signal and the watermarked one which is crucial to this method.

B. Watermark Extraction

Extraction and removal procedures are illustrated in Fig. 4 [7].

By calculating the difference – D of the magnitudes of FFT coefficients that were changed when embedding the watermark in every subband ($D = A_{lower} - A_{higher}$,) it is possible to extract the embedded information before removing it. If D is positive the embedded bit was 1 and if D is negative the embedded bit was 0. This information can be used in order to check whether the appropriate removal process was applied by comparing if the extracted watermark and the original one are identical. Given the different approaches of the two techniques, this was only possible for the second one and presented additional verification.

C. Watermark Removal

For watermark removal the same frequency hopping pseudo random sequence a.k.a. watermark key is used. Values of the FFT coefficients in each sub-band that were changed when watermark was embedded are now set to mean value.

Although these methods are not distortion free it is theoretically possible to alter the method with the purpose of completely recovering the original signal by memorizing previous values of FFT coefficients when embedding the watermark and setting the changed values to them respectively while removing the watermark. However, this requires additional information for watermark removal which lessens security.

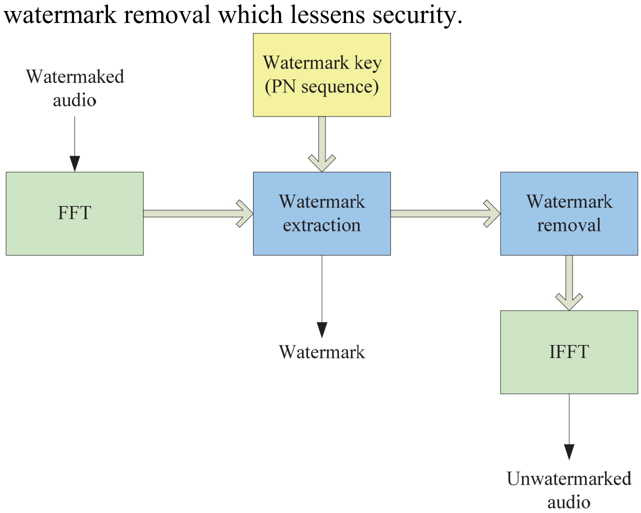

Fig. 4. Extraction and removal scheme.

III. Experiments

A. Spectral Domain Features

Original signal and its watermarked version produced by both techniques in frequency domain are shown in Fig. 5(a), (b) and (c). The signal shown in Fig. 5(a) is randomly chosen from test signals used as the process of watermarking doesn't differ in method per sample.

It is noticeable that it is impossible to extract watermark without some additional information (such as the correct positions where watermark was embedded). Although on the face of it, it might seem that using the second technique distorts the host data more vigorously for higher frequencies, this is actually not the case- the distortion appears higher because of the characteristics of HAS (the threshold for hearing increases with frequency) and simple high- pass filtering wouldn't extract most of the noise (the method is such that it evenly distributes the distortion throughout the hearing range). Recovered signals (the same music sample used previously) for both techniques in frequency domain are shown in Fig. 6(a) and (b).

While the reconstructed signals seemingly look identical to the original one it isn't the case because the methods are not distortion free. The new content has some residual distortions and both objective and subjective experiments are used to determine the level of distortion and the way it affects the listeners.

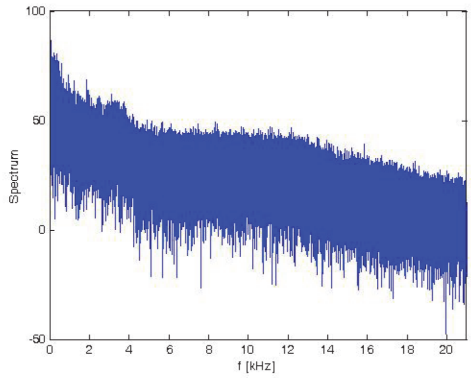

Fig. 5(a). Original signal in frequency domain.

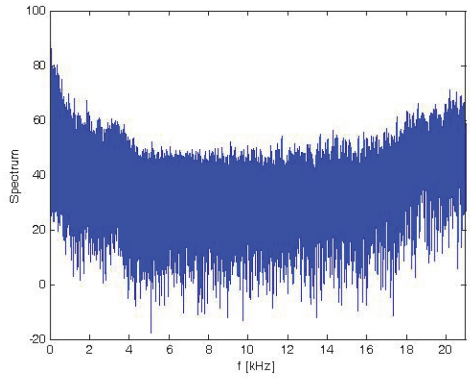

Fig. 5(b). Watermarked signal using technique 1 in frequency domain.

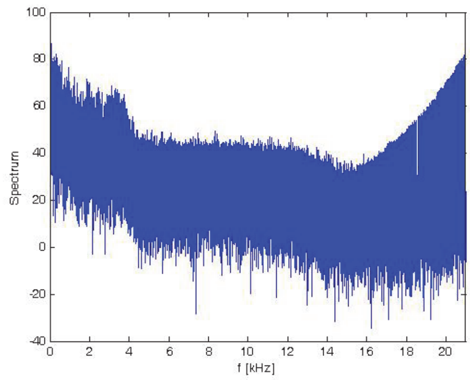

Fig. 5(c). Watermarked signal using technique 2 in frequency domain.

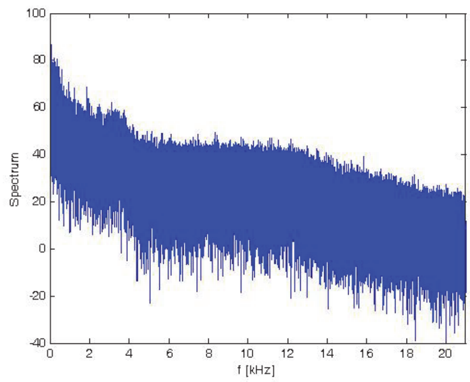

Fig. 6(a). Recovered signal using technique 1 in frequency domain.

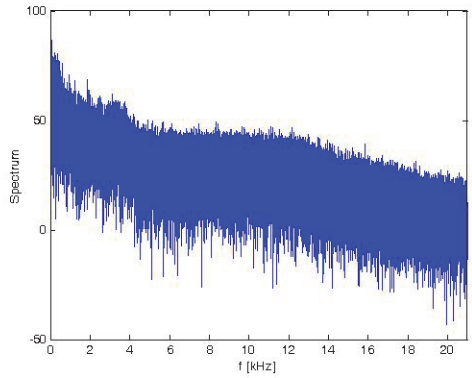

Fig. 6(b). Recovered signal using technique 2 in frequency domain.

B. Testing

Experiments consisted of measuring various signal to noise ratio (SNR) (so-called objective testing) and subjective testing for the test samples. Test samples are audio file stereo clips of different lengths sampled with the sampling frequency of 44.1 kHz. These included classical music, as well as popular national and international music. Rock and roll and folkloric music as well as slow and fast paced music was incorporated in the samples in order to make the results as objective as possible.

Tables I and II list the results of SNR measurements for the test samples for technique 1. Results of SNR measurements using technique 2 are presented in Tables III and IV. The third column in Table III (IV) shows SNR for the test samples watermarked with the value of distortion of 40% (30%), which was deemed optimal for watermarking for this set of samples. Additionally, the samples were watermarked with varying the value of distortion and the results for some of these are shown in Tables III and IV. The fifth column in Tables III and IV shows SNR after the watermark is removed. Since values in the fifth column are well over 20 dB (generally accepted minimum SNR values for watermarked audio content [7]), watermarked audio after watermark removal is considered to be of high quality. All tables show that for both techniques and for each set of test sets (and even samples) a different k should be used for optimal results. Furthermore, the results of objective tests demonstrate that the second technique shows better SNR for the test samples used.

TABLE I: SIGNAL TO NOISE RATIO FOR WATERMARKED SAMPLES - TECHNIQUE 1(DB).

Song	*SNR wm, k=30%*	*SNR wm removed*
Track 1	5.45	22.73
Track 2	5.00	24.22
Track 3	6.86	23.98
Track 4	1.95	22.64
Track 5	6.09	23.11
Track 6	8.83	24.29
Track 7	5.88	24.54

Tables II and IV demonstrate results for classical music with frequent changes of tempo and intensity.

As seen in before mentioned tables, signal to noise ratio for the reconstructed signal is independent of the value of distortion k, as expected.

TABLE II: SIGNAL TO NOISE RATIO FOR WATERMARKED SAMPLES (DB), TECHNIQUE 1- CLASSICAL MUSIC (MOZART) ONLY.

Song	*SNR wm k=20%*	*SNR wm removed*
The Marriage of Figaro	6.84	23.49
The Clemency of Titus	5.16	22.15
Symphony No. 40 in G minor	4.53	25.24
Piano Concerto No. 21 in C major	3.74	24.05
Piano Sonata No.11 in A major “Alla turca”	3.19	21.35
Eine kleine Nachtmusik	4.56	23.89

TABLE III: SIGNAL TO NOISE RATIO FOR WATERMARKED SAMPLES – TECHNIQUE 2(DB).

Song	*SNR wm, k=30%*	*SNR wm, k=40%*	*SNR wm, k=50%*	*SNR wm removed*
Track 1	18.41	12.89	-5.73	25.05
Track 2	18.97	6.89	-15.96	29.45
Track 3	17.68	2.59	-19.54	30.20
Track 4	20.28	11.25	-12.15	29.74
Track 5	16.22	6.94	-14.13	28.04
Track 6	15.99	1.23	-20.41	25.79
Track 7	17.29	-0.87	-22.64	26.38

TABLE IV: SIGNAL TO NOISE RATIO FOR WATERMARKED SAMPLES (DB), TECHNIQUE 2- CLASSICAL MUSIC (MOZART) ONLY.

Song	*SNR wm k=20%*	*SNR wm k=30%*	*SNR wm k=40%*	*SNR wm removed*
The Marriage of Figaro	23.82	15.04	-18.06	28.67
The Clemency of Titus	23.44	13.84	-20.03	27.58
Symphony No. 40 in G minor	24.07	13.17	-21.07	29.68
Piano Concerto No. 21 in C major	22.29	12.85	-21.31	27.68
Piano Sonata No.11 in A major “Alla turca”	24.78	7.5	-27.51	27.89
Eine kleine Nachtmusik	24.16	10.89	-24.04	28.23

The other part of the experiments was conducting subjective tests- test subjects evaluated watermarked audio files after watermark removal. There were 6 test subjects and most of them were average users with an exception of one who has some musical background (sings in a choir). A special graphical users’ interface (GUI) was developed in Matlab to help gather results from the test subjects.

The scale for rating audio quality was as follows: 0: imperceptible, -1: perceptible but not annoying, -2: slightly annoying, -3: annoying, -4: very annoying [11].

The same test samples were used for both subjective and objective testing. According to [7], Subjective Difference Grade (SDG) was calculated using the test results acquired as:

$$SDG = Score_{signal\_under\_test} - Score_{reference\_signal}. \quad (1)$$

The resulting SDG scores, for the samples with watermark removed are displayed in Tables V, VI, VII and VIII. They are coherent with objective testing in confirming that the quality of audio after watermark removal is excellent. Moreover, subjective tests show that classical music with frequent changes in tempo and intensity is more vulnerable to watermarking.

TABLE V: SUBJECTIVE DIFFERENCE GRADE OF THE WATERMARKED SAMPLES- TECHNIQUE 1.

Song	*SDG wm removed*
Track 1	-0.167
Track 2	-0.67
Track 3	-0.167
Track 4	-0.167
Track 5	-0.167
Track 6	-0.33
Track 7	-0.167

TABLE VI: SUBJECTIVE DIFFERENCE GRADE OF THE WATERMARKED SAMPLES- TECHNIQUE 2.

Song	*SDG wm removed*
Track 1	-0.50
Track 2	-0.33
Track 3	-0.33
Track 4	0
Track 5	0
Track 6	-0.66
Track 7	0

The results for SDG scores for the selected works of Mozart are shown in Tables VII and VIII.

As part of subjective testing, specific tests for classical music were conducted with frequent changes of tempo and intensity (slow – fast and loud - quiet) using a specially designed GUI in Matlab, with the only difference being that after listening the test subjects had to guess which version of audio was the watermarked one and which one was original. This additional testing was carried out to determine whether this kind of music is more vulnerable to the embedding and removal process. Most of the test subjects were able to guess the original. When listening to signals watermarked with the first technique test subjects noticed slight residue of added noise. Setting certain values in the spectral domain to mean value results in a very faint occasional crackling sound noticed by the test subjects. They found this to be more noticeable and irritating with the classical music test set. On the other hand, for the second technique, their grading and choosing of the original wasn't based on the low quality of watermarked audio or on revealing characteristics in the specific moments where tempo and intensity of the composition changes but in the slight difference between the beginning of the original signal and the reconstructed one. Most of the listeners noticed that the silence at the beginning of the original signal was replaced by a specific sound in its reconstructed version. This is a consequence of the removal process in which magnitudes of the FFT coefficients were replaced by mean values of the sub-bands instead of their values before embedding.

TABLE VII: SUBJECTIVE DIFFERENCE GRADE OF THE WATERMARKED SAMPLES, TECHNIQUE 1- CLASSICAL MUSIC ONLY.

Song	*SDG wm removed*
The Marriage of Figaro	-1
The Clemency of Titus	-1
Symphony No. 40 in G minor	-0.83
Piano Concerto No. 21 in C major	-0.83
Piano Sonata No.11 in A major “Alla turca”	-1,167
Eine kleine Nachtmusik	-0.67

TABLE VIII: SUBJECTIVE DIFFERENCE GRADE OF THE WATERMARKED SAMPLES, TECHNIQUE 2- CLASSICAL MUSIC ONLY.

Song	*SDG wm removed*
The Marriage of Figaro	-0.6
The Clemency of Titus	-1
Symphony No. 40 in G minor	-0.5
Piano Concerto No. 21 in C major	-0.8
Piano Sonata No.11 in A major “Alla turca”	-0.5
Eine kleine Nachtmusik	-1

IV. CONCLUSION

Various genres of music (classical music, pop and rock) were used for the host audio and varied the coefficient of distortion in order to determine whether it is possible to achieve an acceptable quality level of a reconstructed audio signal (taking k to be everywhere from 25% to 50%). Both objective and subjective tests revealed that residual distortions that exist are not annoyingly noticeable and reconstructed audio has satisfactory quality for the perceptible removable watermark algorithm based on frequency hopping. Although both techniques provide satisfactory results, the results of objective and subjective tests indicate that the second technique provides better quality. For further comparison of these two techniques, robustness of watermarked audio should be tested.

ACKNOWLEDGMENT

The mentorship of dr Jelena Ćertić and help from dr Nataša Nešković and the Faculty of Electrical Engineering, especially its Audio and Video technologies department, is gratefully acknowledged.

REFERENCES

[1] R. Chandramouli, Nasir Memon, and Majid Rabbani, "Digital Watermarking," Available online from: http://onlinelibrary.wiley.com/doi/10.1002/0471443395.img010/abstract (Opened last on March 22nd 2012).

[2] N. Cvejic and T. Seppänen, "Spread spectrum audio watermarking using frequency hopping and attack characterization," *Signal Processing*, vol. 84, no. 1, pp. 207-213, 2004, also available at http://www.mediateam.oulu.fi/publications/pdf/465.pdf (Opened last on March 22nd 2012).

[3] P. Bassia, I. Pitas, and N. Nikolaidis, "Robust audio watermarking in the time domain," *IEEE Transactions Multimedia 3*, pp. 232–241, 2001.

[4] Y. Yang, G. Chen, and Y. Q. Wang, "A Digital Audio Multi-Watermarking Algorithm Based on Improved Quantization," *Third Pacific-Asia Conference on Circuits, Communications and System (PACCS 2011)*, July 2011.

[5] C. Qian and C. Yi, "A WAV Format Audio Digital Watermarking Algorithm Based on HAS," *International Conference on Control, Automation and Systems Engineering (ICCASE 2011)*, July 2011.

[6] X. Li and H.H. Yu, "Transparent and robust audio data hiding in subband domain," *Proceedings of the Fourth IEEE International Conference on Multimedia and Expo*, pp. 397–400, New York, 2000.

[7] M. Löytynoja, N. Cvejic, and T. Seppänen, "Audio Protection with Removable Watermarking," *6th International Conference on Information, Communications and Signal Processing, ICICS*, art. no. 4449635, 2007.

[8] J. Haitsma, M. Van Der Veen, T. Kalker, and F. Bruekers, "Audio watermarking for monitoring and copy protection," *Proceedings of the Eighth ACM Multimedia Workshop*, pp. 119–122, Los Angeles, 2000.

[9] D.Kirovski and H. Malvar, "Spreadspectrum watermarking of audio signals," *IEEE Trans. Signal Process*. vol. 51, pp. 1045–1053, 2003.

[10] J. Proakis and M. Salehi, *Digital Communications*, McGraw-Hill, pp. 729-743, 2008.

[11] International Telecommunication Union, "Subjective Assessment of Sound Quality," ITU-R BS. 1284, 2003.