Model of Audio Watermarking for improving Robustness/Imperceptibility

Dr.Mrs.M.R.Patil[1], Mr.M.B.Bhilavade[2]
[1(]*Professor, Electronics and Communication Department AGMRCET Varur Hubli-581207, India)*
[2]*(Assistant Professor Electrical Department, JJMCOE Jaysingpur-416101, India)*

ABSTRACT : *The techniques implemented in this paper are proposed to embed the data in transform domain. The watermark is embedded adaptively meaning that discrimination factor used to embed each bit is varied according to the segment characteristics. Robustness is significantly improved by cyclic encoding and decoding of watermark. Diversity in time is used to increases the robustness of the system. The synchronization pattern is added to trace the start of watermark in the watermarked file. The models we proposed are able to embed the watermark bit adaptively and perform the blind extraction of watermark successfully. The model proposed is applicable in the areas like copy protection, piracy control, fingerprinting applications where robustness is preferred.*

Keywords - *Encoder, Decoder, audio watermark, cyclic codes, Synchronization pattern.*

I.INTRODUCTION

In order to describe the link between watermarking and standard data communications, the traditional model of a data communications system is often used to model watermarking systems. The basic components of a data communications system, related to the watermarking process, are highlighted. One of the most important parts of the communications models of the watermarking systems is the communications channel, because a number of classes of the communications channels have been used as a model for distortions imposed by watermarking attacks. The other important issue is the security of the embedded watermark bits, because the design of a watermark system has to take into account access that an adversary can have to that channel.

In this paper we highlight on the communications models of watermarking. The first section of this paper introduces the basic model of watermarking. The second section proposes the encoder decoder model of audio watermarking based on the Spread spectrum based implementations of watermarking[1] and third section of this paper proposes the encoder decoder model of audio watermarking based on the GOS(Group of segment) based implementations of watermarking[2].

II. BASIC MODEL OF WATERMARKING

The process of watermarking is viewed as a transmission channel through which the watermark message is being sent, with the host signal being a part of that channel [6]. In Fig.1, a general mapping of a watermarking system into a communications model is given. After the watermark is embedded, the watermarked signal is usually distorted after watermark attacks. The distortions of the watermarked signal are, similar to the data communications model, modeled as additive noise.

A watermark message **m** is embedded into the host signal **x** to produce the watermarked signal **s**. The embedding process is dependent on the key k and must satisfy the perceptual transparency requirement, i.e. the subjective quality difference between **x** and **s** must be below the just noticeable difference threshold. Before the watermark detection and decoding process takes place, **s** is usually intentionally or unintentionally modified. The intentional modifications (n) are usually referred to as attacks; an attack produce attack distortion at a perceptually acceptable level. After attacks, a watermark extractor receives attacked signal **r**.

Depending on a watermarking application, the detector performs informed or blind watermark detection. The term **attack** requires some further clarification. Watermarked signal **s** can be modified without the intention to impact the embedded watermark (e.g. dynamic amplitude compression of audio prior to radio broadcasting). Why this kind of signal processing is called an attack? The first reason is to simplify the notation of the general model of digital watermarking. The other, an even more significant reason, is that any common signal processing impairing an embedded watermark drastically will be a potential method applied by adversaries that intentionally try to remove the embedded watermark. The watermarking algorithms must be designed to endure the worst possible attacks for a given attack distortion which might be even some common signal processing operation (e.g. dynamic compression, low pass filtering etc.). Furthermore, it is generally assumed that the adversary has only one watermarked version **s** of the host signal **x**. In fingerprinting applications, differently watermarked data copies could be exploited by collusion attacks. It has been proven that robustness against collusion attacks can be achieved by a sophisticated coding of different watermark messages embedded into each data copy. However, it seems that the necessary codeword length increases dramatically with the number of watermarked copies available to the adversary.

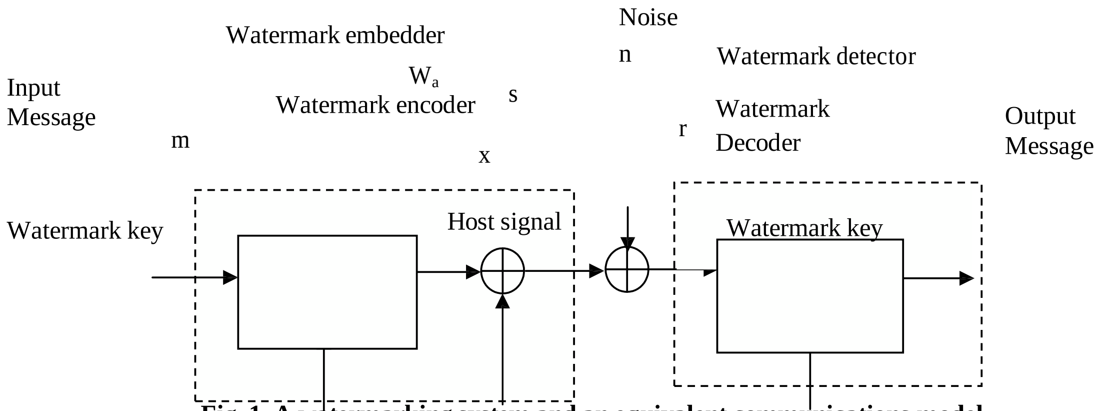

Fig. 1. A watermarking system and an equivalent communications model.

The importance of the key k has to be emphasized. The embedded watermarks should be secure against detection, decoding, removal or modification or modification by adversaries. Kerckhoff's principle [5], stating that the security of a crypto system has to reside only in the key of a system, has to be applied when the security of a watermarking system is analyzed. Therefore, it must be assumed that the watermark embedding and extraction algorithms are publicly known, but only those parties knowing the proper key are able to receive and modify the embedded information. The key k is considered a large integer number, with a word length of 64 bits to 1024 bits. Usually, a key sequence k is derived from k by a cryptographically secure random number generator to enable a secure watermark embedding for each element of the host signal.

In order to properly analyze digital watermarking systems, a stochastic description of the multimedia data is required. The watermarking of data whose content is perfectly known to the adversary is useless. Any alteration of the host signal could be inverted perfectly, resulting in a trivial watermarking removal. Thus, essential requirements on data being robustly watermarkable are that there is enough randomness in the structure of the original data and that quality assessments can be made only in a statistical sense.

Let the original host signal **x** be a vector of length Lx. Statistical modeling of data means to consider **x** a realization of a discrete random process ***x.*** In the most general form, ***x*** is described by probability density function (PDF). A further simplification is to assume independent, identically distributed (IID) data elements of x. Most multimedia data cannot be modeled adequately by an IID random process. However, in many cases, it is possible to decompose the data into components such that each component can be considered almost statistically independent. In most cases, the multimedia data have to be transformed, or parts have to be extracted, to obtain a component-wise representation with mutually independent and IID components. The watermarking of independent data components can be considered as communication over parallel channels.

Watermarking embedding and attacks against digital watermarks must be such that the introduced perceptual distortion - the subjective difference between the watermarked and attacked signal to the original host signal is acceptable. In the previous section, we introduced the terms embedding distortion and attack distortion, but no specific definition was given. The definition of an appropriate objective distortion measure is crucial for the design and analysis of a digital watermarking system.

Watermark extraction reliability is usually analyzed for different levels of attack distortion and fixed data features and embedding distortion. Different reliability measures are used for watermark decoding and watermark detection. In the performance evaluation of the watermark decoding, digital watermarking is considered as a communication problem. A watermark message **m** is embedded into the host signal **x** and must be reliably decodable from the received signal **r**. Low decoding error rates can be achieved only using error correction codes. For practical error correcting coding scenarios, the watermark message is usually encoded intoa vector **b** of length L_b with binary elements $b_n = 0; 1$. Usually, **b** is also called the binary watermark message, and the decoded binary watermark message is $\hat{b}$. The decoding reliability of **b** can be described by the bit error rate (BER)

The BER can be computed for specific stochastic models of the entire watermarking process including attacks. The number of measured error events divided by the number of the observed events defines the measured error rates, word error rate, WER. The capacity analysis provides a good method for comparing the performance limits of different communication scenarios, and thus is frequently employed in the existing literature[6]. Since there is still no solution available for the general watermarking problem, digital watermarking is usually analyzed within certain constraints on the embedding and attacks. Additionally, for different scenarios, the watermark capacity might depend on different parameters (domain of embedding, attack parameters, etc.).

If an informed watermark detector is used, the watermark detection is performed in two steps. In the first step, the unwatermarked host signal may be subtracted from the received signal **r** in order to obtain a received noisy added watermark pattern $\mathbf{w}_n$. It is subsequently decoded by a watermark decoder, using the same watermark key used during the embedding process. Because the addition of the host signal in the embedder is exactly canceled by its subtraction in the detector, the only difference between $\mathbf{w}_a$ and $\mathbf{w}_n$ is caused by the added channel noise. Therefore, the addition of the host signal can be neglected, making watermark embedding, channel noise addition and watermark extraction equivalent to the data communications system.

III. PROPOSED MODEL FOR ROBUST AND SECURE AUDIO WATERMARKING BASED ON SPREAD SPECTRUM

The main aim of the present work is to make the system robust to all kinds of attacks. In this section we propose the model of encoding and decoding based on spread spectrum method [1]. To make the system intelligent we propose to add the following features.

1) Add synchronization patterns to indicate the start of the file from where the watermark is embedded in the host signal.
2) To improve the robustness of the system through diversity, add multiple watermarks at different locations in a file using time diversity.
3) Encode the watermark to reduce the bit error rate and to improve the robustness further.
4) Make the system secure by using SS technique.

In Fig. 2 we propose the spread spectrum based model of watermarking system incorporating all above mentioned features. The start of the signal is first identified and marked. The host audio signal x is decomposed in to smaller segments of user defined size from the marked point. Then the synchronization pattern is added to the host audio signal at various points of audio signal from where the watermark is embedded in host signal. A 6 bit zero mean synchronous pattern with +1 and -1 values alternatively is added in the music signal before embedding the watermark. The pattern is very small and it does not affect the signal quality. Due to the continuous higher value of amplitude the pattern is easily recognizable. Watermark embedding in time domain directly modifies the sample in time domain and hence the added distortion creates a smaller hum in the watermarked signal [4]. Transforming the signal in to any suitable transform domain modifies the transformed samples of the signal and does not affect much on the imperceptibility. Each segment is therefore transformed into DWT-DCT transform domain. Design of the technique to recover the watermark without using the host audio signal is required so the PRN (Pseudo random number) sequence is generated using any secret key k using cryptographic methods.

The watermark message to be transmitted is mapped into an added pattern, $\mathbf{w}_a$, of the same type and dimension of the host signal (one dimensional patterns). The watermark bit stream is then encoded using (7,4) cyclic encoder. Each resulting watermark bit is scaled and multiplied by PN sequence to embed in each segment of host audio. The encoded message pattern is then perceptually weighted in order to obtain the added pattern $\mathbf{w}_{a.}$ After that, $\mathbf{w}_a$ is added to the host signal, to construct the watermarked signal. If the watermark embedding process does not use information about the host signal, it is called the blind watermark embedding; otherwise the process is referred to as an informed watermark embedding.

To find the imperceptibility between the two watermarked signal and original signal SNR is computed. While watermark embedding we modify each segment adaptively to embed the binary data. The computation of α (the scaling parameter) is made based on the energy of the signal. This helps for keeping the audibility of added watermark below the masking threshold. After the added pattern is embedded, the watermarked work y is usually distorted during watermark attacks. We model the distortions of the watermarked signal as added noise, as in the data communications model. The types of attacks may include compression and decompression, low pass filtering, resampling, requantization, cropping, time scaling, etc.

When the signal ‘r’ reaches to the destination it is required to recover the embedded watermark signal from ‘r’. The watermark decoder model is shown in Fig 3. First start of the signal is identified. The synchronization pattern is tracked. Then ‘r’ is decomposed in to smaller segments of the same size used while embedding the watermark. Each segment is then transformed to the DWT-DCT domain. Every segment is then multiplied with the watermark key used. Watermark bit is then recovered from each segment as explained in the

[1]. Decode the watermark using (7,4) cyclic decoder. Concatenate each watermark bit to get one dimensional watermark and then dimension transformation is used to convert back the one dimensional watermark into its original two dimensions.

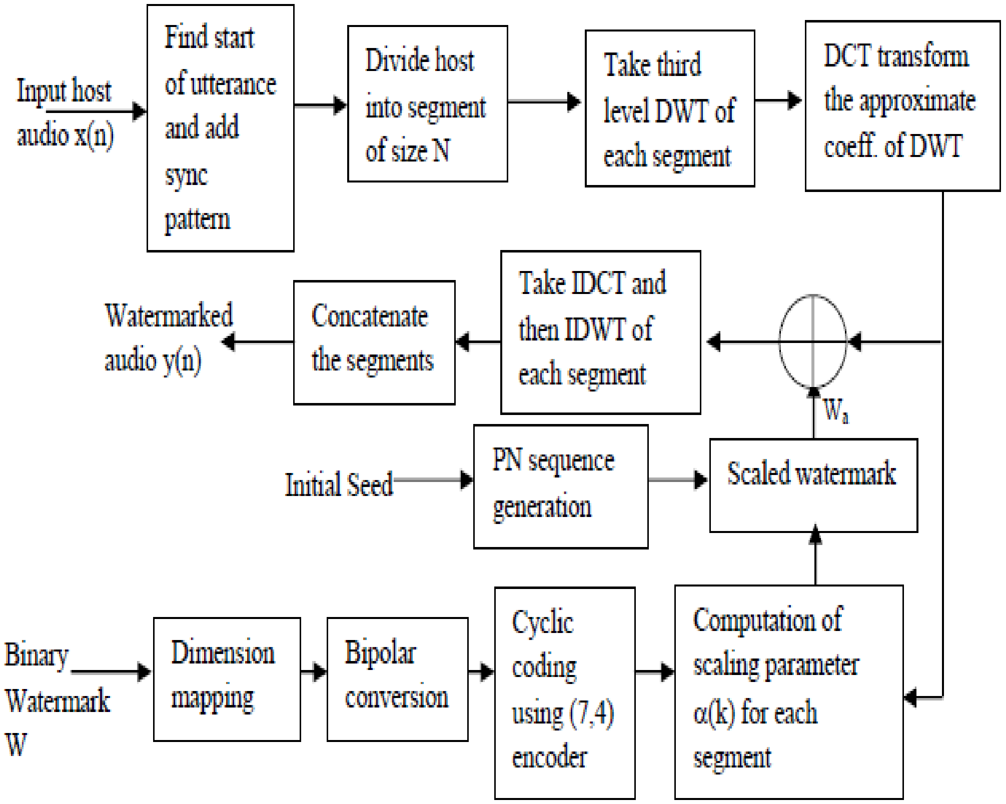
Input host audio x(n)
Find start of utterance and add sync pattern
Divide host into segment of size N
Take third level DWT of each segment
DCT transform the approximate coeff. of DWT
Watermarked audio y(n)
Concatenate the segments
Take IDCT and then IDWT of each segment
W_a
Initial Seed
PN sequence generation
Scaled watermark
Binary Watermark W
Dimension mapping
Bipolar conversion
Cyclic coding using (7,4) encoder
Computation of scaling parameter α(k) for each segment

Fig. 2 Encoder model for proposed adaptive SNR based blind technique in DWT-DCT domain

As in most applications, the watermark system cannot perform its function if the embedded watermark cannot be detected; robustness is a necessary property if a watermark is to be secure. Generally, there are several methods for increasing watermark robustness in the presence of signal modifications. Some of these methods aim to make watermarks robust to all possible distortions that preserve the perceptual quality of the watermarked signal.

Received
audio r(n)

Find start
of utterance

and search Sync
pattern

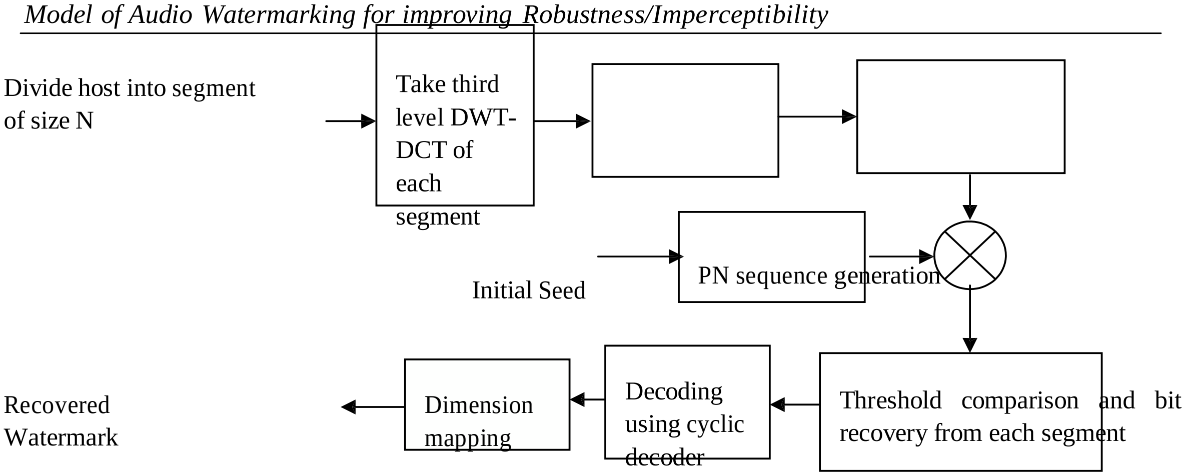
Divide host into segment of size N
Take third level DWT-DCT of each segment
Initial Seed
PN sequence generation
Recovered Watermark
Dimension mapping
Decoding using cyclic decoder
Threshold comparison and bit recovery from each segment

Fig. 3 Decoder model for adaptive SNR based blind technique in DWT-DCT domain.

One of the earliest methods of attack characterization consisted of diversity. Diversity is employed through watermark repetition. Although it is well known that the repetition can improve the reliability of robust data hiding schemes, it is traditionally used to decrease the effect of fading. If properly designed, a repetition can often significantly improve performance and may be worth the apparent sacrifice in the watermark bit rate.

If the repetition is viewed as the application of communication diversity principles, it can be shown that a proper selection of an appropriate watermark embedding domain with an attack characterization can notably improve reliability.

A communication channel can be broken into independent sub channels, where each sub channel has a certain capacity. Since, in a fading environment, some of these channels may have a capacity of zero in a particular time instant, diversity principles are employed. Specifically, the same information is transmitted through each sub channel with the hope that at least one repetition will successfully be transmitted. For watermarking, it is referred to as *coefficient diversity* because different coefficients within the host signal are modulated with the same information.

The host audio signal x with longer length is selected to achieve the purpose. The signal x is then split into smaller signals of duration 7 sec. The watermark embedding is then applied to each of the 7 sec duration signal. Once embedding the watermark in all signals is over they are all concatenated to form a one signal and sent on communication channel. At the reception the received signal is again split in to smaller and 7sec duration signals and the watermark recovery is done from each signal of 7sec duration.

TABLE 1 Robustness results after attack characterization using time diversity

	SNR	Lowest BER after diversity	BER after removing isolated error pixels
Without attack	27.5764	0	0
downsampled	26.6478	0	0
upsampled	26.8021	0	0
Mp3compressed	26.8021	0	0
requantization	26.8021	0	0
cropping	23.3522	0.0029	0.0009
Lpfiltered fc=22050	24.8021	0	0
Time scaling -10%	20.4113	0.0105	0.0038
Time scaling +10%	20.3347	0.0052	0.0029
Echo addition	20.6543	0.0019	0.0009
Equalization	24.6745	0	0

The robustness results of this scheme implemented through time diversity are provided in Table 1. In a 50 sec signal 5 watermarks are embedded after every 7 sec and the results are observed. The 5 watermarks of five 7 sec duration signals are recovered and their BER with the original watermark is computed. The recovered watermark which results in lowest BER is then identified and reported as a valid mark. The recovered watermark is then enhanced to remove the isolated error bits. The window of 3* 3 is used to identify the isolated pixel in the neighborhood of the center pixel. The values of neighborhood pixels are compared with the center pixel and if the value of center pixel is different than the maximum values of its neighborhood pixels then it is replaced with the opposite one. As the watermark image is binary image it contain either 0 or 1 value.

IV. PROPOSED MODEL FOR ROBUST AND SECURE AUDIO WATERMARKING BASED ON GOS MODIFICATION

We also propose the model based on patchwork algorithm [2,4] which performs blind detection of watermark. The proposed encoder model is shown in Fig. 4 which modifies the group of segments (GOS) to embed the watermark. The start of the signal is first identified and marked. The host audio signal x is decomposed in to smaller segments of user defined size from the marked point. Then the synchronization pattern is added to the host audio signal at various points of audio signal from where the watermark is embedded. Now decompose the original signal x into smaller segments of user defined size. Transform the signal in to DWT-DCT domain. Each segment is again decomposed in to two part of group of samples (GOS) of equal/unequal size. Then compute the mean of each GOS and define them as A and B.

The watermark message to be transmitted is mapped into an added pattern, **w**$_a$, of the same type and dimension of the host signal (one dimensional patterns). To provide the security the watermark is encoded by error control coding (Cyclic encoder (7,4)) and then embedded into each segment. As explained in [2] one bit of binary watermark is embedded in one segment of host signal by modifying each GOS to satisfy the required condition of data embedding.

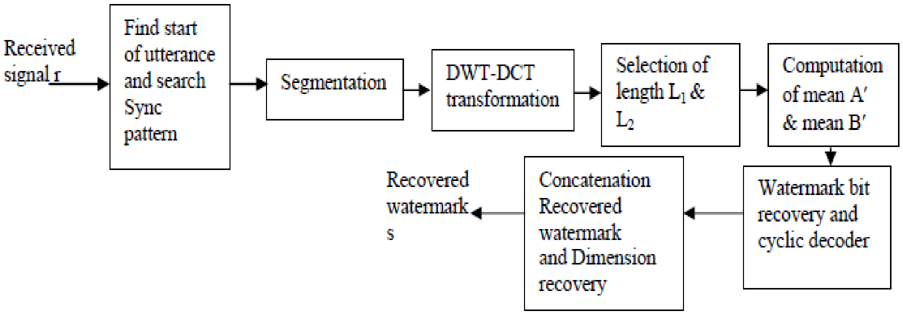
Received signal r
Find start of utterance and search Sync pattern
Segmentation
DWT-DCT transformation
Selection of length L1 & L2
Computation of mean A′ & mean B′
Watermark bit recovery and cyclic decoder
Concatenation Recovered watermark and Dimension recovery
Recovered watermark s

Fig. 4 Block schematic of GOS based encoder of watermark in DWT-DCT domain

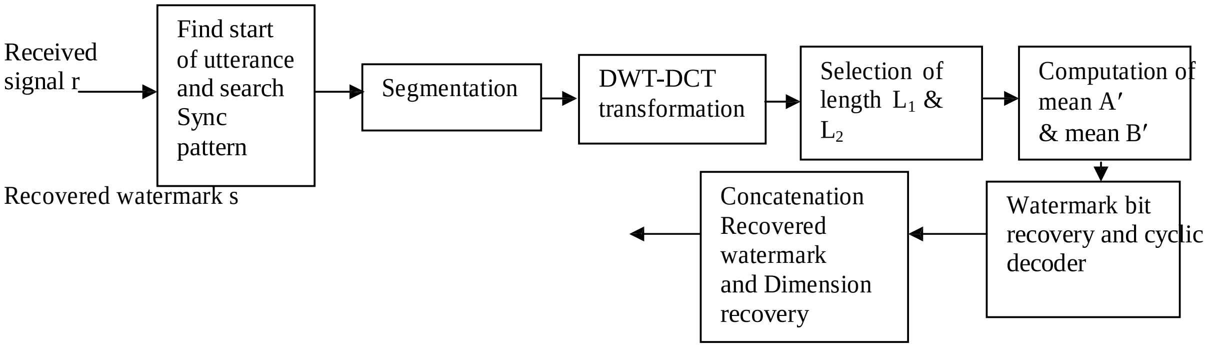
Received signal r
Find start of utterance and search Sync pattern
Segmentation
DWT-DCT transformation
Selection of length L1 & L2
Computation of mean A′ & mean B′
Watermark bit recovery and cyclic decoder
Concatenation Recovered watermark and Dimension recovery
Recovered watermark s

Fig. 5 Block schematic of GOS decoder of watermark in DWT-DCT domain.

The watermarked signal y is then transferred from source to destination through transmission channel. While traveling the signal on transmission channel it undergoes various intentional and unintentional signal processing attacks. The proposed decoder model based on GOS modification is shown in Fig. 5. The received signal r is decomposed in to smaller segments. The r is transformed into DWT-DCT domain and the watermark s recovered by observing the mean of each GOS according to the rules.

The robustness results of this scheme implemented through time diversity are provided in Table 2. The results presented in table 1, and 2 it is clear that the diversity in time increases the robustness of the system. The system implemented using GOS modification increases the imperceptibility where as the spread spectrum based technique is more robust. So we propose to use the model-1 based SS when robustness is more concerned than the imperceptibility and use model-2 when imperceptibility is the requirement.

TABLE 2 Robustness results after attack characterization using time diversity.

	SNR	Lowest BER	BER after removing isolated pixels
Without attack	36.4356	0	0
downsampled	35.2543	0.0009	0
upsampled	35.2543	0.0009	0
Mp3compressed	35.2543	0.0009	0
requantization	35.2543	0.0009	0
cropping	33.2541	0.0029	0.0009
Lpfiltered	30.8021	0.0009	0
Time scaling -10%	23.6543	0.0185	0.0068
Time scaling +10%	22.4532	0.0193	0.0078
Echo addition	20.7645	0.0124	0.0048
Equalization	24.8437	0.0098	0.0019

V. CONCLUSION

In this paper we modeled the audio watermarking techniques using data communication principles. The robustness of adaptive techniques implemented here is significantly improved by cyclic encoding and decoding of watermark. Diversity in time is used to increase the robustness of the system further. We have also added the synchronization pattern to trace the start of watermark in the watermarked file. The models we proposed are able to embed the watermark bit adaptively and perform the blind extraction of watermark successfully. Finally it is observed that the model proposed in method-1 is more robust compare to the method-2. The max value of BER for method-1 is 0.0038 for time scale modification and for method-2 is 0.0078**.**Hence is applicable in the areas like copy protection, piracy control, fingerprinting applications where robustness is preferred. The mehod-2 model is more imperceptible than method-1 and therefore applicable in applications where high degree of imperceptibility is the requirement. The maximum value of SNR for method-1 is 27.5764 and for method-2 is 36.4356.

REFERENCES

Journal Papers:

[1] M. R. Patil, S. D. Apte, Adaptive Spread spectrum audio watermarking, *ICFAI Journal on Information technology,Vol.1.* September 2009.

Proceedings Papers:

[2] M. R. Patil, S. D. Apte, Adaptive Audio Watermarking for Indian Musical Signals by GOS Modification, *Proc. IEEE International conference TECNCON 2009* held on 23-26 November 2009 at Singapore

[3] N. Sriyingyong and K. Attakimongeol, "Wavelet based Audio watermarking using adaptive Tabu search", *Proc. of IEEE Int. Symp. Wireless Pervacive computing 2006 pp 1-5 available at http://sutir.sut.ac.th:8080/sutir/bitstream/123456789/1821/1/BIB990_F.pdf.*

[4] I.K. Yeo and H.J. Kim, "Modified patchwork algorithm: a novel audio watermarking scheme*", Proc. ICITCC 2001,* p.p. 237-242.

[5] A. N. Lemma, J. Aprea, W. Oomen and L. V. D. Kerkhof, "A Temporal Domain Audio Watermarking Technique", *IEEE Transaction on Signal Processing, Vol. 51,* No. , April 2003, p.p. 1088-1097.

Theses:

[6] N Cvejic, "*Algorithms for audio watermarking and steganography*", academic dissertation report, available on line http://hekules.olu.fi/isbn9514273842.

ngpur