Multiplicative Watermarking of Audio in Spectral Domain

Jyotsna Singh, Parul Garg, and Alok Nath De

Division of Electronics and Communication Engineering,
Netaji Subhas Institute of Technology,
Sector 3, Dwarka, New Delhi 110075, India

Abstract. In this paper, watermark is multiplicatively embedded in discrete fourier transform magnitude of audio signal using spread spectrum based technique. A new perceptual model for magnitude of discrete fourier transform coefficients is developed which finds the regions of highest watermark embedding capacity with least perceptual distortion. Theoretical evaluation of detector performance using correlation detector and likelihood ratio detector is undertaken under the assumption that host feature follows Weibull distribution. Also, experimental results are presented in order to show the performance of the proposed scheme under various attacks such as presence of multiple watermarks, additive white gaussian noise and audio compression.

Keywords: audio, correlation detector, discrete fourier transform, log-likelihood detection, watermarking.

1 Introduction

Various watermarking embedding techniques have been proposed which embed watermark additively or multiplicatively in audio signal using the imperfections of human auditory system (HAS). These techniques explore the fact that the HAS is insensitive to small amplitude changes, either in the time [1]-[3] or frequency [4], [5] domains. Boney [1] generated the watermark by filtering a PN-sequence with a filter approximating the frequency masking characteristics of the human auditory system (HAS). This filtered watermark was then weighted in the time domain to account for temporal masking. Swanson [2] proposed audio dependent watermarking procedure which directly exploited temporal and frequency masking properties to guarantee that the embedded watermark is inaudible and robust. The shaping of watermark is performed using a masking curve computed on the original signal. This masking curve is obtained by psychoacoustic modeling of host audio signal. Bassia [3] presented an audio watermarking algorithm by adding a perceptually shaped spread-spectrum (SS) sequence in time domain.

In the other category watermark is embedded in frequency domain. Cox [4] suggested that a watermark should be constructed as an independent and identically distributed (i.i.d.) gaussian random vector that can be imperceptibly

K. Singh, A.K. Awasthi, and R. Mishra (Eds.): QSHINE 2013, LNICST 115, pp. 591–605, 2013.

inserted in the perceptually most significant spectral components of the data. An audio watermarking scheme based on frequency-selective spread spectrum (FSSS) technique in combination with the subband decomposition of the audio signal was presented by Malik et. al. [5]. Megias [6], Fujimoto [7] and Fallahpour [8] developed a high bit-rate audio watermarking technique with robustness against common attacks and good transparency. These algorithms are based on spline interpolation technique.

The embedding techniques in [4], [5], exploit psychoacoustic characteristics of HAS while embedding the watermark additively or multiplicatively in spectral domain. These techniques explored the fact that HAS is insensitive to small amplitude changes in spectral domain. Whereas, phase discontinuity of an audio signal causes perceptible distortion when the phase relation between each frequency component of the signal is changed. Hence Discrete Fourier Transform (DFT) magnitude would be a better option for inserting watermark. However, in literature no perceptual model is defined for DFT magnitude which can decide the location and strength of watermark to be embedded in audio spectrum. Also, these techniques have two major drawbacks. First, the psychoacoustic modeling used by existing techniques require rigorous complex computations. Second, the watermark embedding capacity of these schemes is low i.e. there is not much space to accommodate watermark in the host feature within the defined perceptual limits.

To overcome these two problems, a new method of evaluating masking threshold for DFT magnitude is proposed which requires lesser computations as compared to traditional psychoacoustic model based thresholds. The technique finds best possible locations in spectra for watermark embedding and finds scaling factor of watermark to achieve high watermark embedding capacity.

In this paper, we present a blind robust watermarking system based on pseudorandom signals embedded in the magnitude of the DFT coefficients of an audio signal. Blind detection systems requires only the secret key for data extraction or detection. The cover data or watermark is not required during the extraction process. The scheme obviates the use of complex HAS calculations. Also, it allows us to build a model which can decide the location and strength of watermark in DFT spectra. The paper is organized as follows.

The watermarking system model is presented in section II. In the next section, the signal model is presented and the distribution of DFT magnitude coefficients is shown. Then, in section IV, the construction of the optimal detector is depicted. In sections V and VI, the experimental results and the conclusions are presented.

2 Description of Watermarking Model

A watermarking system encompasses three major functionalities, namely, watermark generation, watermark embedding, and watermark detection. The aim of watermark generation is to construct a sequence $\mathbf{W}$ using an appropriate function f. Hence the watermark vector $\mathbf{W} = [W(0), W(1), \cdots, W(N-1)]$, such that $W(i) \in \mathcal{R}$, where $\mathcal{R}$ is real number, is given as

$$\mathbf{W} = f(\mathbf{K}, N) \tag{1}$$

here $\mathbf{K}$ is the watermark key, N is the length of watermark. Watermarked feature $\mathbf{F}'$ is obtained by multiplicatively embedding watermark $\mathbf{W}$ in host feature $\mathbf{F}$ given as

$$\mathbf{F}' = \mathbf{F}(1 + a\mathbf{W}) \tag{2}$$

here $\mathbf{F}' = [F'(0), F'(1), \cdots, F'(N-1)]$ and a is the scaling factor lying between 0 and 1. The scaling factor is introduced to maintain imperceptibility of the distortions caused to the host signal due to watermarking.

Watermark detector is used to examine whether the signal under test $\mathbf{F}_t$ contains a watermark $\mathbf{W}$ or not under a binary-decision hypothesis test framework. Each module is now discussed in detail, in the following subsections.

2.1 Watermark Generation

The steps required for generation of watermark are as follows:

- To construct watermark $\mathbf{W}$, a white PN or pseudo-random noise sequence $\mathbf{W}_0$ is generated such that $\mathbf{W}_0 = [W_0(0), W_0(1), \cdots, W_0(N_w - 1)]$, where $W_0(i) \in (-1, 1)$. The sequence is generated using secret key $\mathbf{K}$ such that they are mutually independent with respect to the host signal.
- The magnitude nature of host feature needs to be preserved implying that $\mathbf{F}'$ given in (2), should always be greater then zero. Such condition is obtained when $aW(i)'s\ \forall\ 0 \leq i \leq N-1$ take the value in the finite interval [-1,1] keeping scaling factor $a \leq 1$.
- The N point DFT region hosting the watermark is usually split in number of subregions, which in our case are the critical bands. The start location (m) and end location (n) of watermark embedded in these critical bands is decided by a pre-defined masking threshold. Hence the length of watermark N_w is evaluated as

$$N_w = \lceil (n - m)N \rceil \tag{3}$$

- To maintain the symmetry of DFT magnitude a reflected version of $\mathbf{W}_0$ is required to be generated as

$$W_0'(i) = W_0(N_w - i - 1), \qquad 0 \leq i \leq N_w - 1 \tag{4}$$

  The reflected chip $\mathbf{W}_0'$ is embedded in the frequency components around coefficient $N-1$. This is essential to obtain real valued audio in time domain.

2.2 Masking Threshold for DFT Magnitude

In this paper, the magnitude of DFT coefficients of host audio signal are modified by adding watermark, such that the modified spectra is always below the predefined masking threshold, termed as maximum amplitude spread (MAS).

The MAS is defined as the maximum of all amplitude spreads (AS) of DFT components at a particular frequency location within a frame. Following steps are involved to find MAS.

STEP-I: **Finding Amplitude Spread (AS)**
The AS of DFT components is evaluated from the energy spreading function given by Schroeder [9] and its effect is seen at all the N frequency locations of a frame. Schroeder presented a real nonnegative energy spreading function which approximated the basilar spreading as a triangular spreading function and is given as

$$SF_{dB}(i,j) = 15.81 + 7.5(\Delta_z + 0.474) \\ -17.5\sqrt{1 + (\Delta_z + 0.474)^2} \tag{5}$$

here $SF_{dB}(i,j)$ is the energy spread in decibels (dB) from i^{th} to j^{th} frequency location. The bark separation between these two points is given as $\Delta_z = z_j - z_i$, where z_i and z_j denote the bark frequencies of i^{th} and j^{th} frequency locations respectively.
Let the audio signal $\mathbf{s}$, given by (6), is sampled at frequency, f_s Hz, is given as

$$\mathbf{s} = [s(0), s(1) \cdots, s(N-1)] \tag{6}$$

The k^{th} component of DFT, $S(k)$, of signal $s(n)$ is given as

$$S(k) = \sum_{n=0}^{N-1} s(n)e^{-j2\pi kn/N} \tag{7}$$

The samples of discrete time signal $s(n)$ is recovered using the $IDFT$ of $S(k)$ as,

$$s(n) = \frac{1}{N}\sum_{k=0}^{N-1} S(k)e^{j2\pi nk/N} \tag{8}$$

Since audio is real valued signal, its DFT will satisfy the symmetry property i.e. $S(k) = S(N-k)^*$, where $k = 1, ..., N/2 - 1$. The DFT coefficients $S(k)$ corresponds to frequencies f_k given as

$$f_k = f_s \times k/N\,, \tag{9}$$

here $0 \le k \le N-1$, N being a power of 2. Considering the duplication in the spectra for $k \ge N/2$, we evaluate the masking spread $A_1(i,j)$ for amplitude of $N/2$ components only, given as

$$A_1(i,j) = \sqrt{SF(i,j)}\,, \qquad 0 \le i \le N/2 - 1 \tag{10}$$

where $SF(i,j)$ is the inverse decibel of $SF_{dB}(i,j)$. The square root is to convert the masking spread from energy scale to amplitude scale. Now respecting the symmetry property of DFT components, we define $A(i,j)$ as,

$$A(i,j) = \begin{cases} A_1(i,j), & 0 \le j \le N/2 \\ A_1(i, N-j), & N/2+1 \le j \le N-1 \end{cases} \tag{11}$$

The amplitude spread of i^{th} DFT component is then defined as,

$$A'(i,j) = |A(i,j)S(i)|,\ 0 \le i \le N/2-1\,, 0 \le j \le N-1\,, \quad (12)$$

where $S(i)$ is given by (7). This gives $N/2 \times N$ matrix showing amplitude spread of each of the $N/2$ DFT components at N frequency locations. Figure 1 shows a plot of Amplitude spread $A'(i,j)$ of $i = 17^{th}$ and 20^{th} frequency components at all the frequency location f_j for $0 \le j \le N-1$ given in (9) where $N = 512$ and $fs = 44.1kHz$.

STEP II: Evaluation of Maximum Amplitude Spread

The amplitude spreads of neighboring DFT components overlap each other. Maximum amplitude spread (MAS) is the maximum of all the overlapping amplitude spreads at f_i frequency due to DFT coefficients $S(j)$, $\forall\ 0 \le j \le N/2-1$ and $j \ne i$. MAS, $Y(i)$, at location i can therefore be evaluated as

$$Y(i) = \max(A'(i,j)) \quad for \quad 0 \le j \le N/2-1 \quad (13)$$

Now maximum amplitude spread MAS for a critical bands z will be the minimum of all $Y(i)$ in that critical band. From (13), we evaluate the Maximum Amplitude Spread $Y(z)$ for critical bands $z = 1, 2, \cdots, z_t$ as

$$Y(z) = \min(|Y(i)|) \quad for \quad LB_z \le i \le HB_z\,, \quad (14)$$

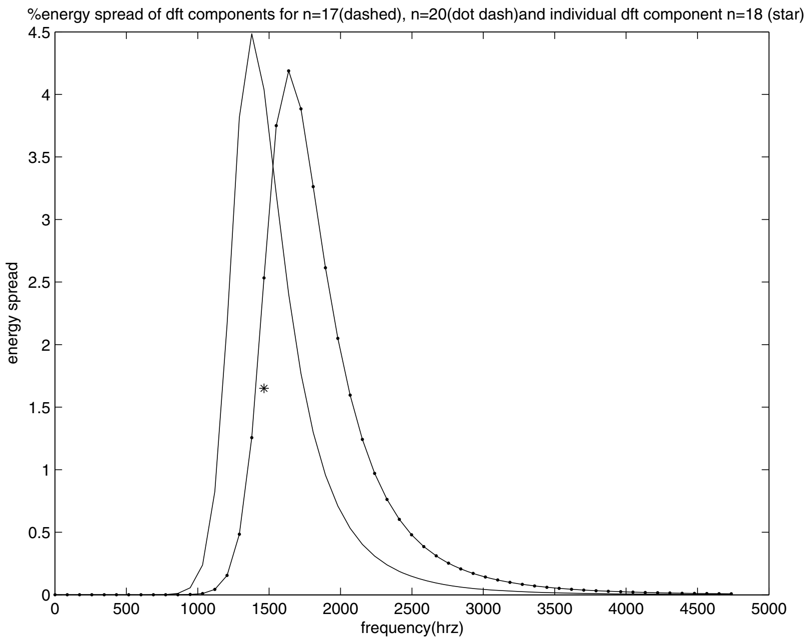

Fig. 1. Overlapping of Amplitude Spread of 17^{th} and 20^{th} DFT components and magnitude of 19^{th} DFT component

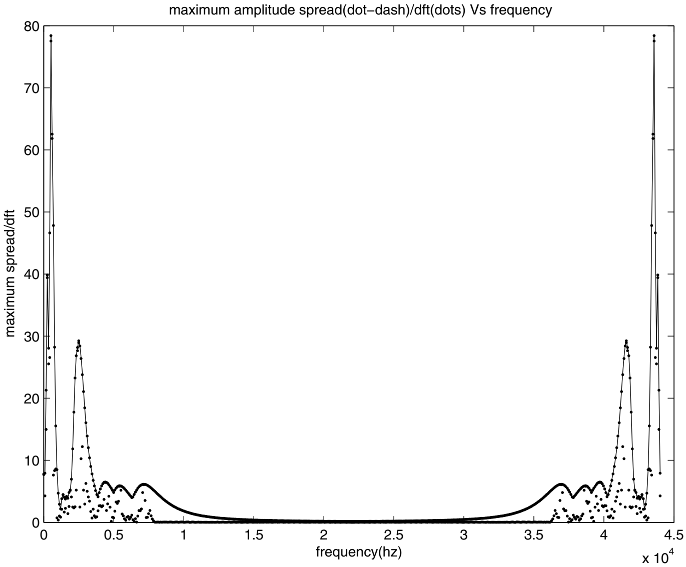

Fig. 2. Maximum amplitude spread of DFT magnitude for a given audio of frame length $N = 512$

where LB_z and HB_z are lower and upper frequency components of z^{th} critical band. Figure 2 shows the plot between maximum amplitude spread $Y(i)$ and the magnitude of DFT coefficients $F(i)$ at all the frequency locations f_i for $i = 0, 1, ..., N - 1$.

2.3 Watermark Embedding

In watermark embedding the watermark **W** is added to host signal **F** in a way that the symmetry of **F** is not disturbed. Also, the DC component and Nyquist component of DFT spectrum should remain unchanged. This is essential in order to retrieve real valued audio signal after watermarking process. The magnitude of DFT coefficients of host audio signal are modified by multiplicative watermarking, such that the modified spectra is always below the maximum amplitude spread of original signal. Hence, the DFT magnitudes are modified only in certain critical bands to maintain the transparency of audio signal. The embedding steps are described as follows

- The magnitude $F(k) = |S(k)|$ and phase $\phi(k) = \angle S(k)$ of the spectral coefficients are evaluated for $k = 0, 1, \cdots, N - 1$, where $S(k)$ is given by (7).
- The distribution of magnitude of DFT coefficients per critical band $F_z(k)$, for $LB_z \leq k \leq HB_z$ is found by translating frequency into bark scale. Here $z = 1, 2, \cdots, z_t$ are the critical bands, z_t is total number of critical bands and LB_z and HB_z are the respective lower and higher frequencies in the critical band z.

- The watermark is embedded in critical bands in which magnitude of DFT coefficients is less than the defined masking threshold, $Y(z)$.
- The final watermark is now generated as

$$W(k) = \begin{cases} W_0(i), & if \quad mN \leq k \leq nN \\ W_0'(i), & if \quad (1-n)N \leq k \leq (1-m)N \\ 0, & otherwise \end{cases} \tag{15}$$

  here $0 \leq i \leq N_w$ and $0 \leq k \leq N-1$ and $(0 < m < n < 0.5)$ to maintain symmetry of final watermark.
- Once location of embedding is decided, the watermark scaling factor a has to be calculated for each critical band to ensure inaudibility of the embedded watermark. The scale factor a_z of z^{th} critical band is obtained by dividing masking threshold $Y(z)$ by the maximum magnitude component of the DFT coefficient in each critical band as

$$a_z = A\frac{Y(z)}{max(|F(k)|)}, \qquad for \quad z = 1, 2, \cdots, z_t \tag{16}$$

  Here A is the gain factor that controls the overall magnitude of the watermarked signal $F'(k)$ given in (2). The value of A varies from 0 to 1. The scaling factor a_z decides how much the amplitude of watermark is to be suppressed in the selected critical band before adding it to the spectrum of host signal.
- The scaled watermark is now added according to rule

$$\begin{aligned} F'(k) &= F(k), \qquad\qquad\qquad if\, F(k) \geq Y(z) \\ &= F(k)(1 + a_z W(k)),\ if\, F(k) < Y(z) \end{aligned} \tag{17}$$

  here $0 \leq k \leq N-1$.
- The modified amplitude of DFT coefficient $F'(k)$ is now combined with their corresponding phases $\phi(k)$, to get watermarked DFT coefficients $S'(k)$.
- The corresponding time domain watermarked signal $s'(n)$ is obtained by calculating $IDFT$ of $S'(k)$ given by (8).

2.4 Optimal Watermark Detection

The aim of watermark detection is to verify, whether or not the given watermark $\mathbf{W}_d$ at receiver end resides in the test signal $\mathbf{F}_t$. The detection is blind i.e. secret key is the only information that detector has at the receiver end. The detector uses salient points for synchronizing the embedded information, so that audio can be analyzed for salient point extraction. Watermark detection can be considered as a binary hypothesis test, solved by means of a correlation detector [10] and log-likelihood ratio detector [11], [12]. However, few assumptions are done before performing the detector tests.

The Host signal $\mathbf{F}$ and the watermark $\mathbf{W}$ are independent and identically distributed i.i.d random variables, hence the detector is optimum. DFT magnitude is wide sense stationary process and for large number of samples likelihood ratio and correlation coefficient attain Gaussian distribution due to central limit theorem.

Likelihood Ratio Detector. The watermarked signal $\mathbf{F}'$ given in (2), may undergo various signal processing or noisy channel attacks before reaching the receiver end. The received signal $\mathbf{F}_t$ is now used for watermark detection, by using log-likelihood ratio test. The best suited distribution for magnitude of DFT coefficients $F = [f(1), \cdots, f(N)]$ is two parameter Weibull distribution [13] which is defined for positive real axis only. The parameter estimation problem consists of finding the underlying distribution parameters by observing samples of random variable described in [14]. Given N sample values $[f(1), \cdots, f(N)]$, from the random variable F, which can be modeled by a two parameter Weibull distribution, the maximum likelihood estimators are utilized to find the values of shape and scale parameters respectively [15].

Since decoding is done without resorting to the original audio, the decoder has no access to the original coefficients. Hence the distribution of the non-watermarked coefficients f_i needs to be approximated by the distribution of the watermarked coefficients f_i'. As long as the embedding strength and thus the watermark power is kept small, the difference between the two distributions will be negligible.

Having identified a suitable model for host feature, we now find the likelihood ratio, as given in [16]. Also, the performance of a log-likelihood based technique is shown by Receiver Operating Characteristic (ROC) curve drawn between probability of false alarm P_f and probability of misdetection P_m.

Correlation Detector. The correlation detector, which is the Maximum Likelihood (ML) optimal detector, is applied to additive or multiplicative watermarking system. These detectors give optimal results while considering Gaussian distribution for the host signals. The correlation detection can be performed by computing the correlation c between pseudo-random sequence $\mathbf{W}$ and watermarked signal $\mathbf{F}_t$ in time or frequency domain given as

$$c = \mathbf{F}_t\mathbf{W} = [\mathbf{F}(1 + \alpha\mathbf{W})]\mathbf{W} = \mathbf{F}\mathbf{W} + \alpha\mathbf{F}\mathbf{W}\mathbf{W} \tag{18}$$

The correlation of watermark (pseudo-random sequence) is compared to a predefined threshold to determine whether watermark is present in the signal or not. The received signal $\mathbf{F}_t$ is used for watermark detection, by using correlation test.

3 Experimental Results

To generate experimental results, a total of 9 standard audio test sequences are taken which are listed in table I. These test sequences are adopted to analyze the performance of the proposed watermarking algorithm. Each signal was sampled at 44.1 kHz, represented by 16 bits per sample, and eight seconds in length. The DFT magnitude of audio signal was assumed to follow Weibull distribution. The shape parameter, $\beta = 0.6833$ and scale parameter, $\alpha = 2.9369$ of Weibull distribution was evaluated using maximum likelihood estimation method.

Table 1. Audio test sequences (44.1 kHz, 16 bit)

Drums	Clarinet	Flute
speech(mono)	speech(stereo)	waltz
Synth	jazz	violin

3.1 Experimental Performance Evaluation

The value of scaling factor a is changed and its effect is seen on performance of likelihood ratio detector and correlation detector respectively. For this the values of a for various critical bands are obtained using (16).

- **Effect on detection threshold**
  In case of LLR detector, the effect of scale factor a is observed on detection threshold Λ. First we have shown the curves between Λ and P_f keeping the value of a fixed. The upper and lower portion of figure 3 shows the variations of Λ with respect to P_f for two values of a, 0.0024 and 0.8 respectively. The first value, $a = 0.0024$, is obtained from MAS threshold and second value, 0.8 is selected close to the maximum limit of a to show the effects clearly visible. As can be seen from figure, for the same range of P_f the variations in Λ is only $0 < |\Lambda| \leq 0.08$ when $a = 0.0024$. Whereas the variations are quite high $(0 \leq |\Lambda| \leq 20)$ for $a = 0.8$. Next we analyse the variations of Λ with respect to

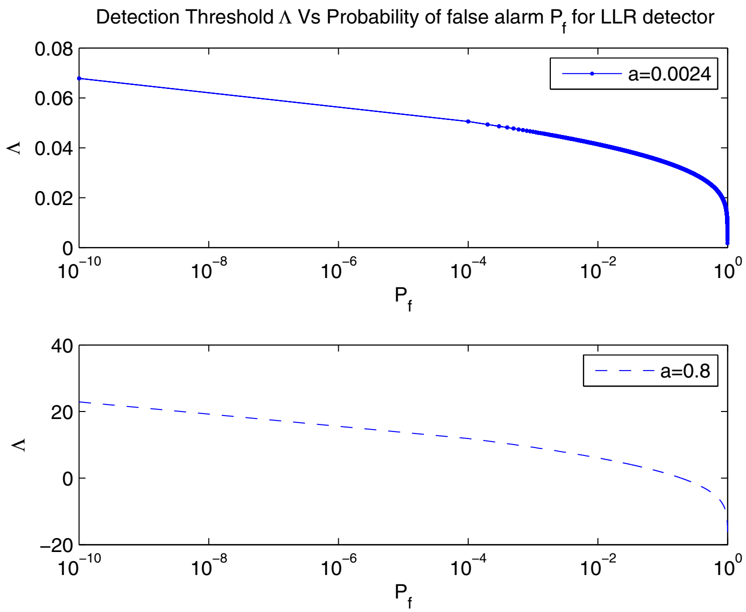

Fig. 3. Threshold versus probability of false detection for LLR detector for two values of scaling factor a

a for all the values in the range of $0 < a \leq 1$. Figure 4 shows the variation of Λ with respect to a for a fixed value of P_f ($\simeq 10^{-6}$). From the plot we observe that the value of Λ remains constant for $a \leq 0.04$. However, as the value of a is increased beyond 0.04 a steep rise in Λ is obtained. Another observation from figure is that with decreasing a, the value of detection threshold Λ also decreases which in turn degrades the detector response.

In case of correlation detector the effect of a is observed on detection threshold T_c. A plot between T_c and P_f for two different values of a (i.e. 0.0024 and 0.8) is shown in upper and lower portion of figure 5. From the figure we observe that $T_c = 0.27$ when $P_f = 10^{-3}$ for both the values of a. Also the value of threshold lies within the range of $0 \leq T_c \leq 0.5$ for a wide variation of a (i.e. $0 \leq a \leq 1$). Hence it can be inferred from the curves that the output of correlation detector is not much effected by scaling factor a.

– **Effect on ROC**

The Receiver Operating Characteristic (ROC) curve is obtained from likelihood ratio and correlation watermark detectors, as shown in figure 6 respectively. The results are compared with actual experimental curve for both detectors with two different values of a, i.e. 0.0024 and 0.8. It is observed that for a=0.0024 the three curves nearly coincide with each other, whereas the same is not true for the case a=0.8. For the proposed value of a, the statistical detectors give optimum results which are close to actual experimental value. Further we observe that LLR detector gives better approximation to experimental results as compared to correlation detector, for all values of scaling factor.

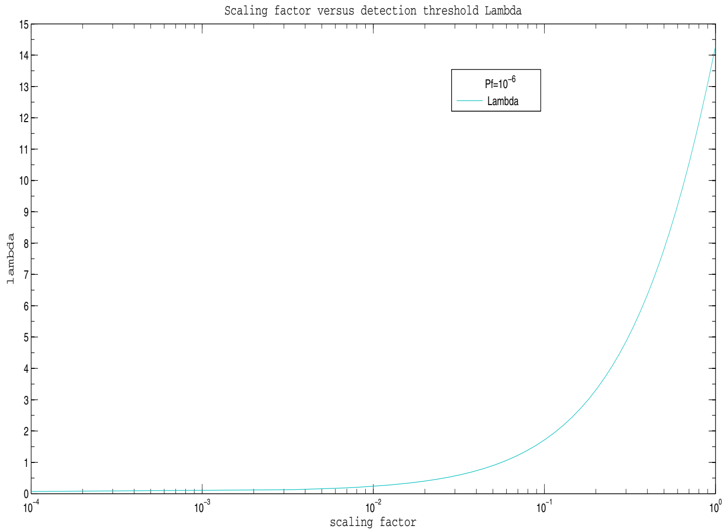

Fig. 4. Threshold versus scaling factor for Log-LLR detector for $P_f = 10^{-6}$

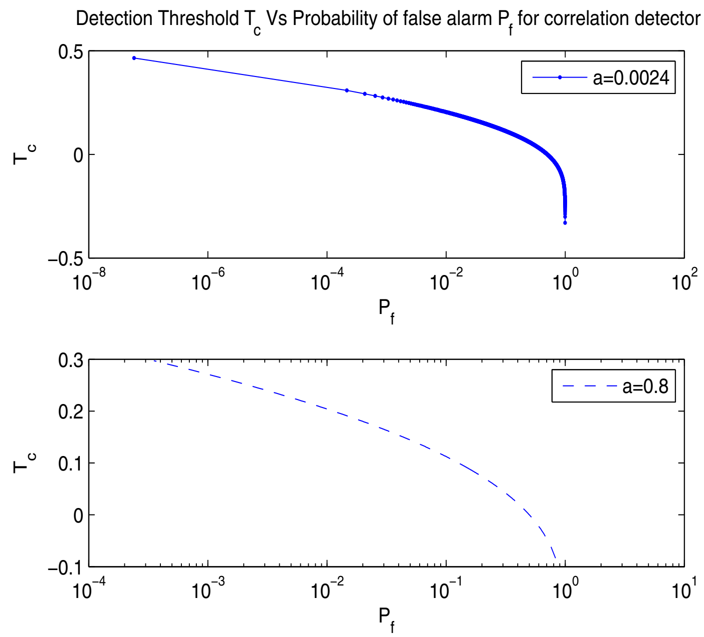

Fig. 5. Threshold versus probability of false detection for correlation detector for two values of scaling factor a

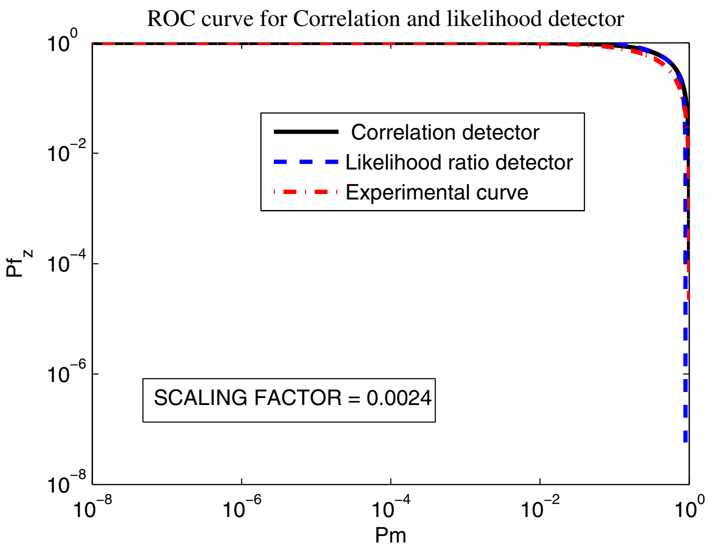

Fig. 6. Receiver operating characteristic curve for scaling factor α=0.0024

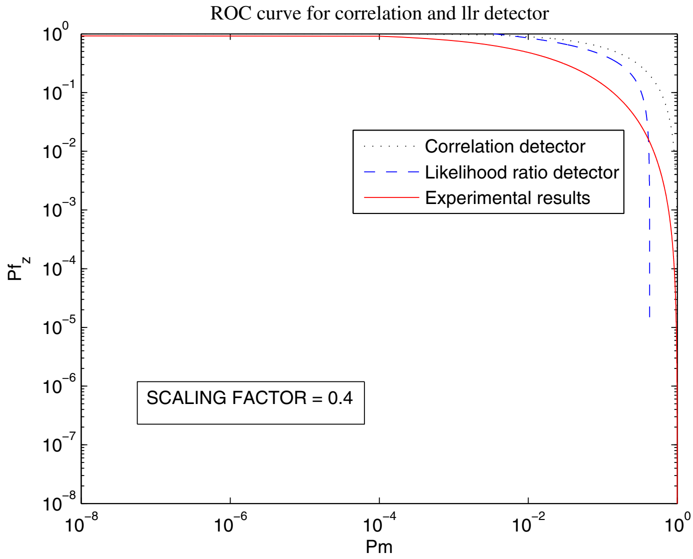

Fig. 7. Receiver operating characteristic curve for scaling factor α=0.0024 (Upper curve) and α=0.4 (lower curve) respectively

3.2 Objective and Subjective Quality Evaluation

Subjective and objective quality tests are performed to evaluate the quality of watermarked audio signal [17]. It is observed from the above results that the quality degradation of the proposed watermarking scheme is very small for the vast majority of the test items, given in table II. For all test items the Subjective difference grade (SDG) is within -0.7 to -0.065 which indicates that there is no significant distortion introduced by this scheme. For objective quality measure, software $'PQevalAudio'$ for perceptual evaluation of audio quality ($PEAQ$) is utilized to evaluate an objective difference grade (ODG), which is an objective measurement of SDG. Table 2 lists the average value of PEAQ/ODG with the give test items for varying value of a. It shows that as value of scaling factor a decreases, perceptual quality of watermarked audio becomes better. However, if the value of scaling factor is lowered below the value obtained from MAS (0.0024), ODG obtained is positive which is not acceptable, as per ITU recommendations. The value of ODG obtained from watermarked audio is -0.065 for the optimum value of $a = 0.0024$. From ROCs plotted in figure 6 and the objective quality given in table II we observe that for small values of a the detector response is poor, but perceptual quality is within acceptable limits. On the contrary, for larger values of scale factor ($a \simeq 0.04$) the detector response improves, but then the perceptual transparency is deteriorated. It can be inferred from these results that proposed technique gives a good tradeoff between perceptual transparency and detector performance.

Table 2. ODG for varying scaling factor and threshold

S.No.	*ODG*	a	Λ
1	-1.624	0.8	11.8741
2	-1.436	0.4	6.2374
3	-0.999	0.1	1.9353
4	-0.710	0.05	1.0094
5	-0.641	0.005	0.1050
6	-0.065	0.0024	0.0505
7	+0.045	0.001	0.0211

3.3 Watermark Embedding Capacity

The proposed scheme provides high watermark embedding capacity with least perceptual distortions. The embedding capacity of proposed scheme was found to be 1.4kbps, with ($ODG = -0.065$), when embedding was done in only one critical band. Table 3 compares the embedding capacity and perceptual quality of proposed scheme with other schemes present in literature. The average watermark capacity increased to 4kbps, with $ODG = -0.7$), when embedding was performed in more then one critical bands (*i.e.* 3).

Table 3. Comparison of ODG and watermark embedding capacity between available literature schemes

Technique	*ODG*	*EmbeddingCapacity*
Megias [6]	-0.5 to -2	61bps
Fujimoto [7]	–	1 Kbps
Fallahpour [8]	-0.5	3kbps
Proposed	-0.065 to -0.7	1.42 to 4kbps

3.4 Robustness to Attacks

The other major issue in watermarking is robustness to various attacks. We will now present the robustness of watermark against additive white gaussian noise (AWGN) noise, presence of multiple watermarks and MP3 compression.

Addition of AWGN Noise. More then 99 percent of watermark recovery is achieved for SNR value of $6dB$ and above. This implies high robustness of watermark against AWGN noise.

Presence of Multiple Watermark. To see the effect of presence of multiple watermark both types of detectors i.e. likelihood ratio and correlation detectors, are used. In case of LLR detector the value of Λ obtained is 9.8 for $P_f = 10^{-6}$ with scaling factor $a = 0.8$. The LLR detector output is shown for high value of

a, as the response of this detector is poor for small values of a, as can be seen from figure 6. Log-likelihood ratio Λ of correct watermark obtained is above 9.8 whereas the LLR ratio of other watermarks is well below the threshold. Similarly in case of correlation detector the value of threshold T_c obtained statistically was 0.271.

4 Conclusion

The proposed multiplicative spread spectrum based blind audio watermarking technique embeds watermark in DFT magnitude of audio signal. In order to improve two parameters, the embedding capacity and the computational complexity, a new perceptual model for magnitude of DFT coefficients is developed. This model finds the regions of highest watermark embedding capacity with least perceptual distortion. Also the proposed method reduces computations by bypassing the complex psychoacoustic modeling, required for fulfilling the condition of transparency. Theoretical evaluation of detector performance using correlation detector and likelihood ratio detector is undertaken under the assumption that host feature (DFT magnitude) follows Weibull distribution. The experimental and statistical results shown that proposed scheme gives higher embedding capacity as compared to existing watermarking techniques keeping the perceptual quality well within limits. Also, it was observed from experimental results that proposed scheme is robust to various signal processing attacks like presence of multiple watermarks, AWGN and MP3 compression.

References

1. Boney, L., Tewfik, A.H., Hamdy, K.N.: Digital watermarks for audio signal. In: Proc. IEEE Int. Conf. Multimedia Comput. Syst (ICMCS), Hiroshima, Japan, pp. 473–490 (June 1996)
2. Swanson, M.D., Zhu, B., Tewfik, A.H., Boney, L.: Robust audio watermarking using perceptual masking. Signal Process 66(3), 337–355 (1998)
3. Paraskevi Bassia, P., Ioannis Pitas, I., Nikos Nikolaidis, N.: Robust audio watermarking in the time domain. IEEE Trans. Multimedia 3, 232–241 (2001)
4. Cox, I.J., Kilian, J., Leighton, T., Shamoon, T.: Secure spread spectrum watermarking for multimedia. IEEE Transactions on Image Proc. 6(12), 1673–1687 (1997)
5. Malik, H., Ansari, R., Khokhar, A.: Robust audio watermarking using frequency-selective spread spectrum. IET Information Security 2(4), 129–150 (2008)
6. Megías, D., Herrera-Joancomartí, J., Minguillón, J.: Total Disclosure of the Embedding and Detection Algorithms for a Secure Digital Watermarking Scheme for Audio. In: Qing, S., Mao, W., López, J., Wang, G. (eds.) ICICS 2005. LNCS, vol. 3783, pp. 427–440. Springer, Heidelberg (2005)
7. Fujimoto, R., Iwaki, M., Kiryu, T.: A Method of High Bit Rate Data Hiding in Music Using Spline Interpolation. In: Proceedings of the 2006 International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP 2006), pp. 11–14 (2006)

8. Fallahpour, M., Megias, D.: High capacity audio watermarking using FFT amplitude interpolation. IEICE Electronics Express 6(14), 1057–1063 (2009)
9. Schroeder, M.R., Atal, B.S., Hall, J.L.: Optimizing digital speech coders by exploiting properties of the human ear. Journal Acoust. Soc. America 66(6), 1647–1652 (1979)
10. Hyun, K.W., Dooseop, C., Hyuk, C., Taejeong, K.: Selective correlation detector for additive spread spectrum watermarking in transform domain. Signal Processing 90(8), 2605–2610 (2010)
11. Barni, M., Bartolini, F., De Rosa, A., Piva, A.: A new decoder for the optimum recovery of nonadditive watermarks. IEEE Trans. Image Processing 10(5), 755–766 (2001)
12. Cheng, Q., Huang, T.S.: Robust optimum detection of transform domain multiplicative watermarks. IEEE Trans. Signal Processing 51(4), 906–924 (2003)
13. Weibull, W.: A statistical distribution function of wide applicability. Journal of Applied Mechanics 18(3), 293–297 (1951)
14. van Trees, H.L.: Detection, Estimation and Modulation Theory, Part I. Wiley, New York (1968)
15. Stone, G.C., Van, H.G.: Parameter estimation for the Weibull distribution. IEEE Transactions On Electrical Insulation EI-12(4) (August 1977)
16. Barni, M., Bartolini, F.: Watermarking systems Engineering: Enabling Digital Assets Security and Other Applications. Marcel Dekker, New York (2004)
17. Neubauer, C., Herre, J.: Digital watermarking and its influence on audio quality. In: Proceedings of 105th Audio Engineering Society Convention, San Francisco, CA (September 1998)