# Enhanced Audio Watermarking Algorithm using Discrete Wavelets Transform and Singular Value Decomposition

Ali Al-Haj[1], Christina Twal[2], Ahmad Mohammad[3]
[1,3]Princess Sumaya University for Technology
[2]The University of Jordan
Amman, JORDAN

***ABSTRACT:*** *Audio watermarking is gaiing momentum in the recent years to embed copyright information in audio signals as a proof of their ownership. The significance of audio watermarking has enabled to protect the illegal manipulation of genuine audio sources. An efficient and scalable audio watermarking algorithm is designed and proposed in this study. The discrete wavelets transform (DWT) and the singular value decomposition (SVD) are applied Suggestive evidences are generated through the experimental results.*

**Keywords:** Digital audio watermarking, copyright protection, ownership verification, discrete wavelets transform, singular value decomposition, robust and inaudible watermarking.

**Received:** 11 June 2010, Revised 29 July 2010, Accepted 3 September 2010



## 1. Introduction

Unauthorized copying and distribution of digital media has become a dilemma for the multimedia industry. Traditional methods, such as encryption, offered limited solution to the problem [1], leaving the door open for new technologies to offer better solutions. A promising technology that has been receiving a considerable attention for broad range of multimedia applications is digital watermarking [2,3]. In principle, digital watermarking techniques embed a small copyright information, called a watermark, in the digital media such that the watermark is imperceptible and robust against attempts to degrade it or remove it from the digital object. In fact, digital watermarking has been successfully applied to enforce copyright protection in digital images and video clips. Unfortunately, audio watermarking received less attention from the research community due to its inherit difficulty.

Existing audio watermarking systems exploit the irrelevant properties of the human auditory system (HAS) [4]. In particular, HAS is insensitive to small amplitude changes in the time domain and frequency domains, allowing the addition of weak noise signals (watermarks) to the host audio signal such that the changes are inaudible [5-10]. Frequency-domain techniques, in particular, have been more effective than time-domain techniques since watermarks are added to selected regions in the transformed domain of the host audio signal, such that inaudibility an robustness are maintained. In this paper, will report on hybrid-transform audio watermarking algorithm. Two powerful transforms are applied on the original audio signal; discrete wavelets transform (DWT) [11], and singular value decomposition (SVD) [12,13].

In section 2 we describe in details the watermarking embedding and extraction procedures of the proposed algorithm. In section 3 we evaluate the performance of the algorithm and present our results with regard to the fidelity, inaudibility and robustness requirements of effective watermarking. We conclude in section 4 with some remarks.

## 2. Proposed Hybrid DWT-SVD Algorithm

The proposed algorithm employs a cascade of two transforms; the discrete wavelet transform and the singular value decomposition transform. The algorithm is described in this section by outlining the major steps in its two procedures; the watermark embedding procedure and the watermark extraction procedure.

### 2.1 Watermark Embedding Procedure

***Step # 1:*** Convert the binary-image watermark into a one-dimensional vector ***W*** of length ***m x n***.

$$W_t = \{[0,1],\ 1 \leq t \leq (m \times n)\} \qquad (1)$$

***Step # 2:*** Sample the original (WAV) audio signal at a sampling rate of 44100 samples per second. Then, partition the sampled file into frames each having 50,000 samples. The summation of ***N*** frames makes up the overall sampled audio signal as illustrated in the following equation:

$$A = \Sigma_{t=1}^{N} A_t \qquad (2)$$

***Step # 3:*** Perform a four-level DWT transformation on each frame ***Ai***. This operation produces five multiresolution sub-bands: ***D1, D2, D3, D4*** and ***A4.*** The ***Ds*** represent the details sub-bands and ***A4*** represents the approximation sub-band.

***Step # 4:*** Arrange the four details sub-bands ***D1, D2, D3,*** and ***D4*** in a matrix form as shown in figure 1 below. The matrix, named ***DC*** thereafter, has the size ***4 x ( L/2 )***, where ***L*** is the length of each frame.

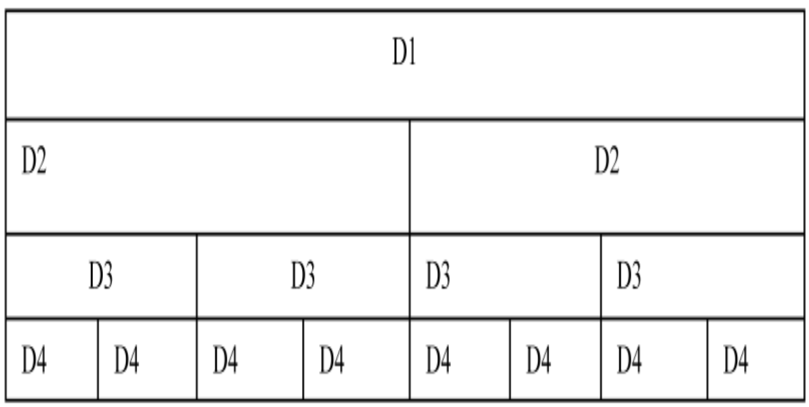


Figure 1. Matrix formulation of the of details *D* sub-bands

***Step #5***: Decompose the ***DC*** matrix using the SVD operator. This operation produces the three orthonomal matrices ***S***, ***U*** and $\boldsymbol{V^T}$ as follows:

$$DC = U \times S \times V^T \qquad (3)$$

where the ***S*** is the following ***4 x 4*** diagonal matrix:

$$S = \begin{bmatrix} S_{11} & 0 & 0 & 0 \\ 0 & S_{22} & 0 & 0 \\ 0 & 0 & S_{33} & 0 \\ 0 & 0 & 0 & S_{44} \end{bmatrix} \qquad (4)$$

The diagonal $s_{ii}$ entries are the non-zero singular values of the ***DC*** matrix. The $s_{11}$ value is used for embedding as will be shown later, and therefore it needs to be stored for later use in the watermark extraction procedure.

***Step # 6:*** Embed the binary-image watermark bits into the DWT-SVD-transformed audio signal according to the following formula:

$$S_{11w} = S_{11} \times (1 + \alpha \times w(n)) \qquad (5)$$

where $w(n)$ is the watermark bit: 0 or 1, $\alpha$is the watermark intensity, $s_{11}$ is the top left value in the Smatrix, and $s_{11w}$ is the watermarked $s_{11}$. If $\alpha$ was set to 0.2, then $s_{11w}$ will equal (1.2 $s_{11}$ ) for $w(n)$ is 1, and to ($s_{11}$) for $w(n)$ is 0.

***Step # 7:*** Produce the final watermarked audio signal as follows:

• Apply the inverse SVD operation using the $U$ and $V^T$ matrices, which were unchanged, and the $S$ matrix, which has been modified according to Equation 5. The $CD_W$ matrix is the watermarked $DC$ matrix of Equation 3.

$$CD_w = U \, x \, S_w \, x \, V^T \qquad (6)$$

• Apply the inverse DWT operation on the $CD_W$ matrix to obtain each watermarked audio frame $A_{iw}$. The overall watermarked audio signal $A_W$ is obtained by summing all watermarked frames.

$$A = \Sigma_t A_{tw} \qquad (7)$$

### 2.3 Watermark Extraction Procedure

The watermark extraction procedure requires the watermarked audio signal and the singular values of each frame of the original audio signal. The major steps of the procedure are described as follows:

***Step # 1:*** Perform steps 1 through 5 of the embedding procedure until the $S$ matrix is obtained for all frames of the watermarked audio signal.

***Step # 2:*** If the original audio signal has been used instead of the original singular values of each frame, then perform steps1 through 5 of the embedding procedure until the $S$ matrix is obtained for all frames of the original signal.

***Step # 3:*** Determine the watermark bit $w(n)$ by recalling that for embedding a bit into a frame, $s_{11}$ was modified according to Equation 5. Therefore, the watermark bit $w(n)$ will have a value of 1 if $s_{11w} / s_{11}$ is 1.2, and 0 if $s_{11w} / s_{11}$ is 1.0

***Step # 4:*** Assemble the extracted bits from the individual frames and construct the original binary image watermark.

## 3. Experimental Results

A pop music file of length 600,000 samples (13.6 seconds sampled at 44100 sample per second ) has been used to evaluate the performance of the proposed algorithm. The pop file (pop.wav) is a stereo type signal with left and right channels The quantized audio signal was watermarked with a 6 x 4 binary image. The waveform of the original pop audio signal is shown in Figure 2(a), before watermarking, and in Figure 2(b), after watermarking.

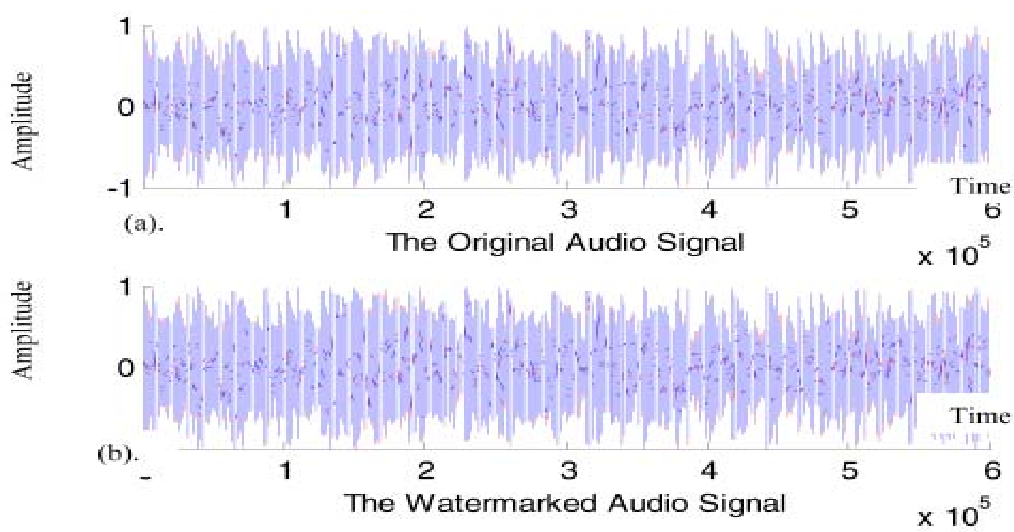


Figure 2. (a). The original pop signal, and (b). the watermarked pop signal

Performance of audio watermarking algorithms is usually evaluated with respect to fidelity, imperceptibility (inaudibility) and robustness [14]. In what follows we give a brief description of each metric and the results we obtained.

***Fidelity.*** Fidelity refers to the similitude between the undistorted original audio signal and the distorted watermarked audio signal. In this paper, we used the SNR metric which has been widely used in evaluating fidelity of watermarking algorithms. An SNR value of 28.55 was obtained, which indicates a high similarity between the un-watermarked and watermarked signals. This is evident by the obvious similarity between the two waveforms shown in Figure 2.

***Imperceptibility (Inaudibility)***. To get an accurate estimation of the audio quality or the audibility of the watermarked signal, a well-known objective measurement method was performed. The method is called "Perceptual Evaluation of Audio Quality" (PEAQ) [15]. It has been developed in order to get a perceptual measurement scheme that estimates the results of real world listening tests as faithfully as possible. PEAQ uses five subjective grades to indicate inaudibility of watermarked signal. The grades ranges from *Grade 1* which indicates *Very Annoying Audio* up to *Grade 5* which indicates *Imperceptible Audio*. We conducted listening tests repeatedly and obtained an approximate grade of 5.0,which corresponds to '*Imperceptible (Inaudible)*' watermarked audio signal.

***Robustness.*** To evaluate robustness of the proposed audio watermarking algorithm we implemented a set of attacks that commonly affect audio signals. Most of implemented attacks have been defined by Stirmark® watermarking benchmark [16]. The watermarks which were extracted after applying various Stirmark® attacks are given in Figure 3. As shown in the figure, the quality of the extracted watermarks varies from one attack to another, however, we were able to extract recognizable watermarks after most attacks.

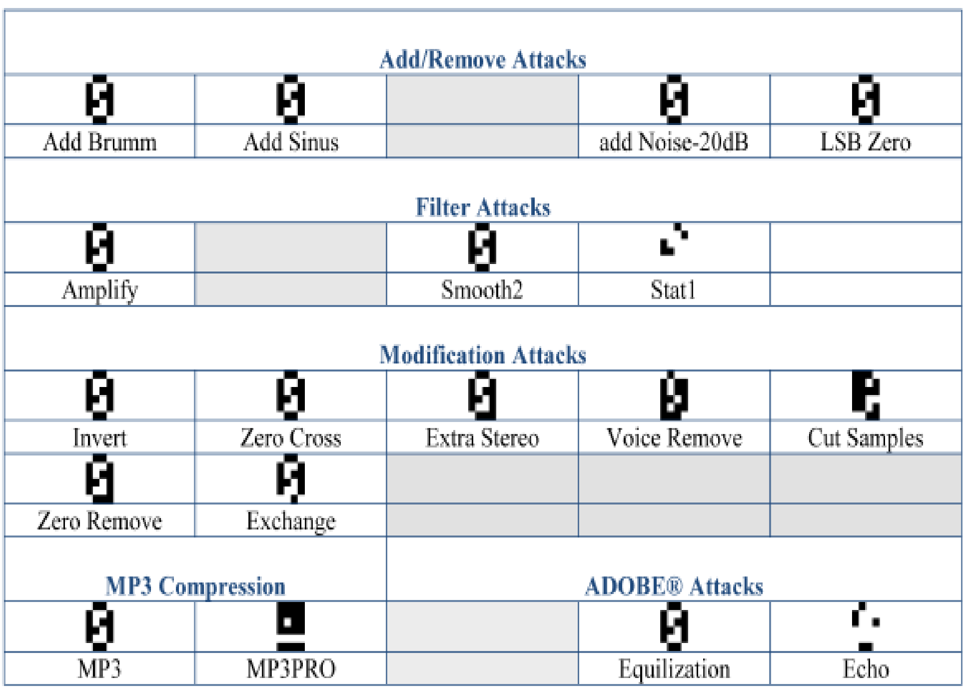


Figure 3. Extracted watermarks after attacks to the watermarked pop signal

To measure the robustness of a watermarking algorithm to the Stirmark® attacks, we computed the bit error rate (BER) after each attack. BER reflects the certainty of detection of the embedded watermark and is expressed by Equation 8 given below. The equation defines the ratio of incorrect extracted bits to the total number of embedded bits.

$$BER = \frac{100}{l} \sum_{n=0}^{i-1} \begin{cases} 1, & W'n = Wn \\ 0, & W'n \neq Wn \end{cases} \tag{8}$$

where $l$ is the watermark length, $W_n$ corresponds to the $n^{th}$ bit of the embedded watermark and $W'_n$ corresponds to the $n^{th}$ bit of the extracted watermark. The BER for most attacks is 0, as shown in the Figure 3. However, the overall BER, which is a summation of the BERs due to all attacks, has been found to be 13.82 %. An overall BER value of 8.33 % was obtained when experiments were repeated for speech audio signals, indicating the effect of the audio signal type on watermarking performance.

## 4. Conclusions

Transform-domain digital audio watermarking has a performance advantage over time-domain watermarking by virtue of the fact that frequency transforms offer better exploitation of the human auditory system (HAS). In this paper we utilized the advantages of two powerful transforms; the discrete wavelets transform (DWT) and singular value decomposition (SVD) to achieve an inaudible and robust audio watermarking. The effectiveness of the DWT-SVD based algorithm was demonstrated by watermarking a pop audio clip. Ongoing research is concentrated on enlarging the payload capacity of the algorithm.

## References


[1] Furht ,B., Kirovski, D.(2006). Encryption and Authentications: Techniques and Applications. Auerbach, USA.

[2] Katzenbeisser, S., Petitcolas, F (2000). Information Hiding: , Artech House, USA.

[3] Potdar,V. Han, S., Chang, E. (2005). A Survey of Digital Image Watermarking Techniques, *In*: Proc. of the 2005 IEEE International Conference on Industrial Informatics, p. 709-716.

[4] Arnold, M. (2000). Audio watermarking: features, applications and algorithms, *In*: Proc. of IEEE International Conference on Multimedia and Expo, V. 2, p. 1013-1016.

[5] Bassia, P., Pitas, L., Nikolaidis, N. Robust Audio Watermarking in the Time-Domain, *IEEE Transactions on Multimedia*., 3 (2) 232-242.

[6] Bassia, P., Pitas, L. , Nikolaidis, N.(2003).A temporal-domain audio watermarking technique, *IEEE Trans. on Signal Proc.*, 51 (4) 1088-1097.

[7] Kirovski, D., Malvar, H (2003). Spread-spectrum watermarking of audio signals, *IEEE Transactions on Signal Processing*, 51 (4) 1020-1033.

[8] Swanson, M., Zhu, B. Tewfik, A.,Boney, L. (1998).Robust audio watermarking using perceptual coding, *Signal Processing*, 66 ( 3) 337-356.

[9] Seok, Jing Won., Hong, Jin Woo (2001). Robust Audio Watermarking in the Time-Domain, *IEEE Electronic Letters*, 37 (1). 60-61.

[10]Yeo, In-Kwon., Kim, Hyoung , Jong (2003). Modified Patchwork Algortihm: A Novel Audio Watermarking Scheme, *IEEE Transactions on Speech and Audio Processing*, 11 ( 4) 381-382.

[11] Mallat, S. (1989). A theory for multi-resolution signal decomposition: The wavelet representation, *IEEE Trans. Pattern Anal. And Machine Intell*., 11 (7) 674-693.

[12] Deprettere ,E (1988). *Singular Value Decomposition and Signal Processing*. Elsevier, Holland.

[13] Liu, R., Tan, T.(2005). An SVD-Based Watermarking Scheme for Protecting Rightful Ownership, *IEEE Trans. Multimedia*, 7 ( 2) 121-128.

[14] Sihirli, M., Gurgen,F., Ikizoglu, S (2005) Performance evaluation of digital audio watermarking techniques designed in time, frequency and cepstrum domains, *Lecture Notes in Computer Science*, V. 3261, p. 430-440.

[15] Thielde, T. et al., (2000). PEAQ-The ITU standard for objective measurement of perceived audio quality, *Journal of Audio Engineering*, 48 (1/2/3) 3-29.

[16] Lang, A., Dittmann, J., Spring, R., Vielhauer, C. (2005). Audio watermark attacks: from single to profile attacks, *In*: Proc. of 7th ACM workshop on multimedia and security, p. 39-50, USA.