

# Speech Watermarking by using DWT and SVD

P.Hindhya Priya,17HU1D3805,DECS,Department of ECE, TECH,Tadipatri.

S.Mounika, Assistant Professor, Department of ECE, TECH,Tadipatri.

## ABSTRACT

Audio watermarking has been proved as a powerful tool against illegal manipulation of audio products. It is generally used as a multimedia copyright protection tool. In this paper, we propose an audio watermarking algorithm based on two mathematical functions: Discrete Wavelet Transform (DWT) and Singular Value Decomposition (SVD). These algorithms performance is validated in the presence of the standard watermarking attacks. A new audio signal framing, DWT matrix formation and embedding methods are proposed and successfully implemented to improve the quality of watermarked audio. Natural variability of speech features allows watermarking alterations to be imperceptible to the human observer. On the other hand, significance of these features makes the system robust to common signal processing operations.

## Keywords

Audio watermarking, Copyright Protection, DWT, SVD, Imperceptibility, Robustness.

## 1. INTRODUCTION

Information generation and transmission is a common concept from very early days. With due course of time, various techniques has been developed to improve this. Watermarks are used for information hiding purpose during transmission. Watermarking is the process of embedding information into a signal in a way that is difficult to remove. Digital watermarking came to be in great demand when sharing information on the Internet became a usual practice. Internet flexibility enables user to modify & access digital contents very easily. Today, with the revolution of information technology and Wide Area Networking, data has become less private[1].For this, a new method of protection is needed against unauthorized copying of digital multimedia files that includes image, audio or video components[2]. A digital watermark is a pattern of bits inserted into a digital file - image, audio or video. Such messages usually carry copyright information of the file. When the digital file is in the form of an audio signal, the embedding technique is called audio watermarking. The digital media that carries the watermark is called a cover signal or host signal. The watermark is embedded into the host signal by a watermark embedder and is detected by a watermark detector. Audio watermarking techniques slightly modify the usable signal to embed additional information into the original signal. This modification is so slight, that human hearing cannot perceive any difference between the watermarked audio signal and the original signal.

A watermarking scheme is said to employ informed detection if it requires the original host signal to be present at the watermark detector. Conversely, a watermarking scheme is said to employ Blind Detection [1] [3] if it does not require the original host signal to be present at the watermark detector.

Depending on the intention and the kind of watermark, watermarking techniques should possess certain properties. According to the International Federation of Phonographic Industry (IFPI), audio watermarking should meet the various requirements.

- a) Imperceptibility: The digital watermark should not affect the quality of original audio signal after it is watermarked.
- b) Robustness: The embedded watermark data should not be removed or eliminated by using common audio signal processing operations and attacks.
- c) Capacity: It refers to the number of bits that can be embedded into the audio signal within a unit of time.
- d) Security: It implies that the watermark can only be detected by the authorized person.

These requirements are often contradictory with each other and there is a need to make a trade-off among them.

## 1.1 Application

In the last decade, digital watermarking has attracted considerable research interest. Numerous applications in various domains (e.g. audio, video) have been proposed. It may be used for a wide range of applications, such as:

1. Copyright protection.
2. Source tracking (different recipients get differently watermarked content)
3. Broadcast monitoring (television news often contains watermarked video from international agencies).
4. Covert communication.
5. Automatic monitoring and tracking of copy-write material on WEB.
6. Automatic audit of radio transmissions.
7. Data augmentation - to add information for the benefit of the public.
8. Fingerprinting applications (in order to distinguish distributed data).

9. In health care, especially medical imaging systems.

The rest of the paper is organized as follows: In Section 2 theoretical background of digital audio watermarking is briefly described. Section 3 gives the idea about the proposed approach. Evaluation of proposed technique is depicted in the Section 4 followed by the conclusion in the last section.

## 2. THEORETICAL BACKGROUND

### 2.1 Classification

Many techniques of audio watermarking algorithms can be grouped into three categories[4]: patchwork in the frequency domain, echo hiding in the time domain and spread spectrum. Based on their embedding domain, watermarking algorithms can also be classified [1]. In time domain watermarking system, watermark is directly embedded into audio signal. Wide range of time domain embedding techniques for digital audio signal such as Least Significant Bit (LSB) alteration, Echo addition and phase coding methods have been developed [2].

In time domain algorithms, watermark is embedded without any transformation & watermark can be easily destroyed. Implementation of these methods is very easy & requires less computation. The watermark signal is shaped before embedding. This enables the system to maintain the audibility of the original audio signal. The robustness of the time domain algorithm is poor. Hence most of the research work is focused on transform based watermarking techniques. Because the audio signal sampling frequency is low, and the human auditory system (HAS) is more sensitive than the human visual system (HVS), so the amount of information to be embedded in the audio signal is much less than in the visual media [4][5][6].

Transform domain audio watermarking technique take the advantage of frequency masking property of Human auditory system. Transforming audio signal into frequency domain enables watermarking system to hide the watermark into perceptually significant component of an audio signal. It gives high robustness against signal processing operations. This technique includes the use of discrete cosine transform (DCT), Discrete Fourier Transform (DFT), and discrete wavelet transforms DWT. Transforming audio signal from time domain to frequency domain enables watermarking system to embed the watermark into perceptually significant part of a signal. This will develop the technique with a high robustness.

In spatial domain, the watermark can be embedded directly into a digital media and modified its values [1] [5].

The choice of the algorithm depends on several factors: the type of cover audio, the computational complexity of the algorithm and the application.

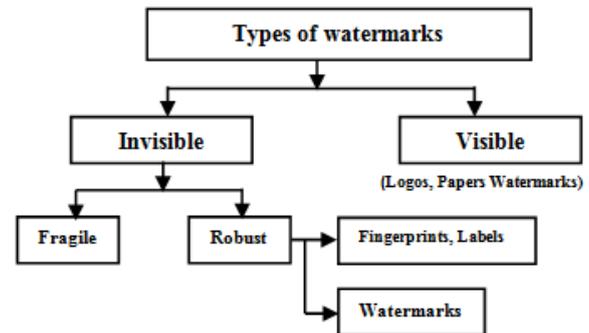


Fig 1: A classification of watermarking techniques

A basic classification of watermarking technique [7] is given by Fig.1. Speaking of digital image watermarking, watermarks can be divided into two main groups – *visible* and *invisible* watermarks. The visible one, like different logos and usually images that are superimposed upon a still picture or a moving picture. *Visible* watermarks are more robust against image transformation. Other, an *invisible* watermark is an embedded image which cannot be perceived with human's eyes. Only electronic devices (or specialized software) can extract the hidden information to identify the copyright owner. Invisible watermarks are used to mark a specialized digital content (text, images or even audio content) to prove its authenticity. These digital watermarks are created by embedding extra information, commonly in the form of digital patterns, into the computer files containing the images or sounds to be protected. The use of a fragile watermark is important when one wants to verify if the protected media was tampered with or not. Robust watermarking is designed to provide proof of ownership.

Depending upon the requirement of original signal at the receiver, watermarking schemes are classified as blind, semi-blind and non-blind [1]. Blind scheme require only the secret key(s), semi- blind scheme require secret key(s) and the watermarking bits and non-blind scheme require both the original signal and the secret key(s) for watermark embedding.

Static watermarks are those which do not change regardless of who opens and processes the watermarked document. They are used in the same as those on the banknote. Static watermarks also be used for copy preventing. Dynamic watermarks in the physical world were created by affixing a seal to a document, or stamping it. These kinds of watermarks are used to identify the institution or individual associating them with the authenticity of the document. Dynamic watermarks may be used as a form of copy resisting.

### 2.2 Transform Basics

In this section, the DWT and SVD transforms [2] has been described.

#### Discrete Wavelet Transform

Wavelets are functions defined over a finite interval and having an average value of zero [1]. Wavelet transform is used to represent any function as a superposition of a set of wavelet or basic functions. DWT is capable of giving a time-

frequency representation of any given signal [2]. Here, the wavelets are discretely sampled. As with other wavelet transforms, a key advantage it has over Fourier transforms is

temporal resolution. Functionally, it is very much like the Discrete Fourier Transform, in that the transforming function is orthogonal, a signal passed twice through the transformation is unchanged, and the input signal is assumed to be a set of discrete-time samples. The basic idea of DWT is to separate frequency details, which is multi-resolution decomposition [8]. The audio signal can be transformed into frequency domain ranging from low frequency to high frequency [9]. But, high frequency spectrum is less sensitive to human ear. Thus the high frequency component is usually discarded in the compression process. Therefore, information to be hidden can be embedded into the low frequency component to against the compression attack. The DWT is defined by the following equation:

$$W(j, k) = \sum_i \sum_k x(k) 2^{-j/2} \psi(2^{-j} n - k) \tag{1}$$

Where  $\Psi(t)$  is a time function with finite energy and fast decay called the mother wavelet.

From the original Audio signal S, DWT produces two sets of coefficients A(Approximation) and D(Detail) as in Fig.2 by passing through low pass filters g and y [2] .

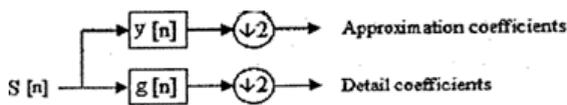


Fig. 2

Depending upon the application and length of the signal, the low frequency part might be further decomposed into two parts of high and low frequencies. Fig .3 shows a three –level DWT decomposition of signal S. The original signal S can be reconstructed using the inverse DWT process.

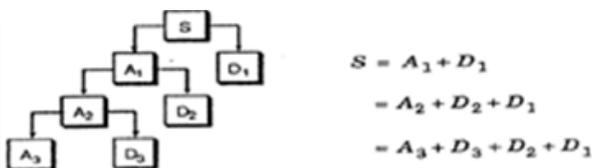


Fig.3

**Singular Value Decomposition**

The SVD is a numerical technique used to transform matrices into series of linear approximations that expose the underlying structure of the matrix [3]. Mathematically, for example, the SVD of an n x n matrix A is defined by the operation:

$$A = U * S * V^T \tag{2}$$

In matrix form, Eqn. (2) can be presented as [3] [5]

$$\begin{bmatrix} A_{1,1} & \dots & A_{1,n} \\ A_{2,1} & \dots & A_{2,n} \\ \dots & \dots & \dots \\ A_{n,1} & \dots & A_{n,n} \end{bmatrix} \begin{bmatrix} U_{1,1} & \dots & U_{1,n} \\ U_{2,1} & \dots & U_{2,n} \\ \dots & \dots & \dots \\ U_{n,1} & \dots & U_{n,n} \end{bmatrix} \begin{bmatrix} S_{11} & 0 & 0 & 0 \\ 0 & S_{2,2} & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & S_{n,n} \end{bmatrix} \begin{bmatrix} V_{1,1} & \dots & V_{1,n} \\ V_{2,1} & \dots & V_{2,n} \\ \dots & \dots & \dots \\ V_{n,1} & \dots & V_{n,n} \end{bmatrix}^T$$

The columns of U and V are mutually orthogonal unit vectors. The M x N matrix S is a pseudo-diagonal matrix, and its diagonal elements, which are arranged by descending gradation, are all non-negative values.

An image is regarded as a non-negative matrix generally and the SVs of the audio represent its intrinsic characteristic [4].

**3. PROPOSED WATERMARKING SCHEME**

This section provides an overview of our proposed watermarking scheme, which consists of the watermark embedding process and watermark extraction process. Watermark is embedded in lower frequency components of the audio signal, which are below the perceptual threshold of the human auditory system.

**3.1 Watermark Embedding**

Fig.4 illustrates the block diagram of the proposed watermark embedding process.

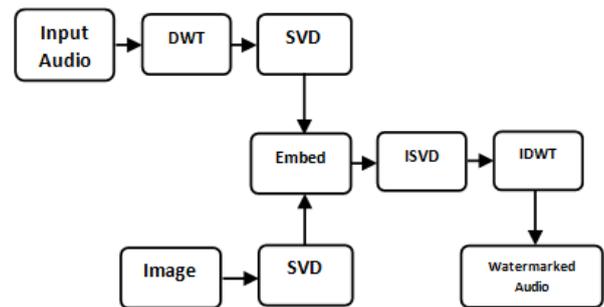


Fig 4: Watermark embedding process

The embedding process consists of certain steps:

- Step 1: Select one image to embed in the audio. Take it in grey form.
- Step 2: Convert the binary-image into a one dimensional vector.
- Step 3: Sample the original audio signal at a sampling rate(fs) and partition the sampled file into frames each having certain samples.[9]
- Step 4: Apply DWT transformation on original audio signal. This operation produces two sub-bands: A and D. The D represents Details sub-band and A represents the Approximation sub-band [9]. Select the sub- band A for embedding.

Step 5: Apply SVD to the DWT performed approximation Sub-band A. SVD Decomposes the DWT coefficients into three matrices namely, U, S,  $V^T$ . Where U is Unary matrix, S is Singular matrix.

Step 6: Perform the steps 4 and 5 to the image also.

Step 7: Embed the image bits into the DWT-SVD transformed original audio signal according to the equation.

$$S_{em} = S + k * S_w \quad (3)$$

Where  $S$  = singular matrix of original audio signal.

$S_w$  = singular matrix of image signal.

$S_{em}$  = singular matrix of watermarked audio signal.

Step 8: 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 Step7.

Apply the inverse DWT operation to obtain each watermarked audio frame [9]. The final watermarked audio signal is obtained by summing all Watermarked frames.

### 3.2 Watermark Extraction

The algorithm for extraction is similar as shown in Fig.5. Steps consist of:

Step 1: Perform steps 4 and 5 of the embedding procedure until the S matrix is obtained for all frames of the watermarked audio signal [9].

Step 2: Compose the singular matrix of watermark image in the DWT-SVD transformed watermarked audio signal according to the equation,

$$S_{ext} = (S_{em} - S) / 0.0001 \quad (4)$$

Where  $S_{ext}$  = singular matrix of extracted watermark signal.

Step 3: Produce the final image 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 (3).

Apply the inverse DWT operation to obtain each watermarked audio frame.

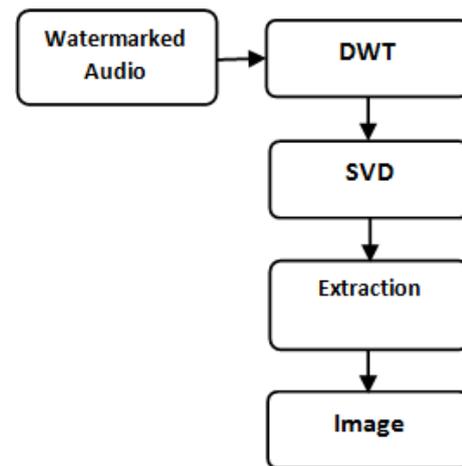


Fig 5: watermark extraction process

## 4. EVALUATION PARAMETERS

In order to evaluate the performance of watermarking algorithm, signal-to-noise ratio (SNR) is used as an objective measure and a listening test as a subjective measure. Various experiments were conducted to demonstrate the efficiency of the proposed audio watermarking scheme.

Robustness against attacks is a major watermarking requirement. Absolute robustness against all possible attacks may be impossible to achieve. Thus, the practical requirement is that, a successful attack must impair the host data to the point of significantly reducing its commercial value before the watermark is impaired so much that it cannot be recovered. It should be noted that robustness always has to be traded against watermark data rate and imperceptibility, and the optimum trade-off depends on the application.

### 4.1 Robustness Test

The watermarking scheme was subjected to the following robustness tests [10]: Additive Random Noise, Volume Scaling, Low pass Filtering, Re-sampling & Re-quantization.

**Additive Random Noise:** It was simulated by generating a scaled random sequence using MATLAB and saving this sequence as a noise pattern. This noise pattern was added to the watermarked sample, which was then subjected to watermark decoding.

**Volume Scaling:** The watermarked audio scaled over a range of 0.5 to 1.5 times the original amplitude.

**Low- Pass Filtering:** It was carried out by subjecting the watermarked audio to filtering.

**Re-sampling:** For this, the watermarked audio was subjected to a cycle of down sampling and up sampling. The audio was down-sampled to half the original sampling rate and then up-sampled back to the original sampling rate.

### 4.2 Imperceptibility

To evaluate the imperceptibility or the perceptual transparency of a watermarking scheme, watermarked

samples are subjected to listening tests. The procedure followed is given below:

- 1) The original sample was first identified and played to a group of listeners.
- 2) The watermarked samples were then played in random order.
- 3) The listeners were asked to evaluate each sample on a scale of 1 to 5 with 5 representing the quality of the original sample as shown in Table 1.

**Table 1: Subjective Grades**

Grade	Description	Quality
5	Imperceptible	Excellent
4	Perceptible	Good
3	Slightly annoying	Fair
2	Annoying	Poor
1	Very annoying	Bad

4) An Average Score is computed about 4.9. for different audio samples.

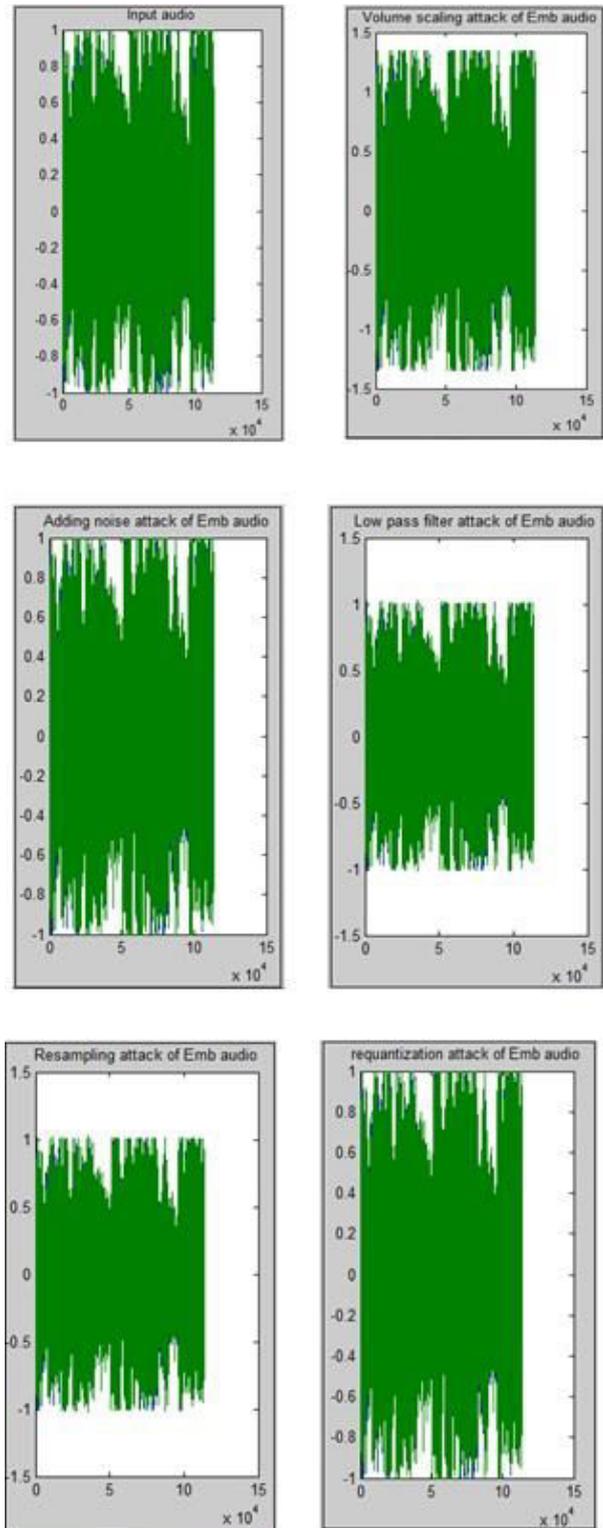
### 5. EXPERIMENTAL RESULTS

In this section the performance of proposed method being evaluated for various audio samples as shown in following figures.

The quality of the watermarked signal is evaluated in terms of Signal to noise ratio (SNR) using the equation.

$$SNR = 10 \log_{10} \frac{\sum_{n=1}^N w^2(n)}{\sum_{n=1}^N [w(n) - w^*(n)]^2}$$

Where  $w(n)$  &  $w^*(n)$  are original audio signal & watermarked audio signal respectively.



**Table 2: SNR Values in dB For Various attacks**

Audio File	Without Attacks	Volume Scaling	Additive Noise	LPF	Re-sampling	Re-quantization
Pop	39.5982	40.3020	39.8159	40.4907	41.2994	39.8722
Rock	41.5028	41.3142	40.4218	39.277	38.9659	40.9130
Classical	44.6522	43.1193	44.9837	45.3051	45.5294	44.1264
Movie	46.4307	43.2634	41.8588	47.1765	47.3899	45.8129

Table 2 shows the performance of our proposed method in terms of Signal to noise ratio. It gives improved SNR as compared to traditional watermarking methods.

## 6. CONCLUSION

In this paper, an efficient audio watermarking technique for copyright protection has been presented. The watermarking algorithm is based on DWT and SVD techniques. The scheme was subjected to a series of imperceptibility (audio fidelity) and robustness tests. The obtained test result showed improved performances. In order to attain higher hidden data density in the watermarked signal, more advanced techniques must be used such as spread spectrum, phase encoding, or echo hiding.

## 7. REFERENCES

- [1] S.M.Ramesh, A.Shanmugam, 2011, An Efficient Robust Watermarking Algorithm in Filter Techniques for Embedding Digital Signature into Medical Images Using Discrete Wavelet Transform , European Journal of Scientific Research, pp-33-44.
- [2] Ali Al-Haj, Ahmad Mohammad, 2010, Digital Audio Watermarking Based on the Discrete Wavelets Transform and Singular Value Decomposition , European Journal of Scientific Research, pp-6-21.
- [3] Ali AlHaj, Tuqa Manasrah,"Non-Invertible Copyright Protection of Digital Images Using DWT and SVD, Department of Computer Engineering,University of Jordan.
- [4] Sartid Vongpraphip and Mahaak Ketcham, An Intelligent Audio Watermarking Based on DWT-SVD using ATS , Department of Computer Engineering, Chulalongkorn University,Bangkok.
- [5] Musa S. Al-Yaman, Majid A. Al-Taee, Hemam A.Alshammas,"Audio-Watermarking Based Ownership Verification system Using Enhanced DWT-SVD Technique",2012-9th International Multi-Conference on Systems,signals and Devices.
- [6] Musa S. Al-Yaman, Majid A. Al-Taee, Anas T. Shahrour, Ismail A.Al-Husseini," BIOMETRIC BASED AUDIO OWNERSHIP VERIFICATION USING DISCRETE WAVELET TRANSFORM AND SVD TECHNIQUES", 2011-8th International Multi-Conference on Systems,signals and Devices.
- [7] Chunlin Song, Sud Sudirman, Madjid Merabti, Recent Advances and classification of Watermarking Techniques in Digital Images ,School of Computing and Mathematical Sciences, Liverpool John Moores University,UK.
- [8] Rakhi Dubolia , Roop Singh , Sarita Singh Bhadoria , Rekha Gupta, 2011, DIGITAL IMAGE WATERMARKING BY USING DISCRETE WAVELET TRANSFORM AND DISCRETE COSINE TRANSFORM AND COMPARISION ON PSNR, International Conference on Communication Systems and Network Technologies
- [9] Gulivindala Suresh, N.V.Lalitha,Ch.Srinivasa Rao, V.Sailaja, "An Efficient and Simple Audio Watermarking Using DCT-SVD".
- [10] F.Hartungand, M.Kutter, "Multimedia watermarking techniques," Proc.IEEE, vol. 87, no. 7, July 1999, pp. 1079–1107.
- [11] Rabia Khan, Abdul Ghafoor and Naveed Iqbal Rao, A Blind Image Adaptive Watermarking Scheme for Audio using Wavelet Transform, International Conference on Digital Image Processing, pp-67-71.
- [12] Pranab Kumar Dhar, Isao Echizen, Robust FFT Based Watermarking Scheme for Copyright Protection of Digital Audio Data, Seventh International Conference on Intelligent Information Hiding and Multimedia Signal Processing.