Chapter 1

Introduction

1.1 Overview of audio watermarking

Digital watermarking [14] is the process of imperceptibly embedding a watermark message into a host signal. The resulting signal is called a watermarked signal. The message should introduce only tolerable distortion to the host signal (cover signal) and it should be recoverable by the intended receiver after signal processing operations on the watermarked data. Watermarks can be embedded into audio, image, video, and other formats of digital data. There are different concerns to different formats of data in watermarking system design. In this thesis, we narrow our scope to audio watermarking systems.

Watermarking is closely related to steganography in that they are both concerned with covert communication and belong to a broader subject known as information hiding [65]. Steganography, derived from Greek, literally means “covered writing”. It is the art of hiding information inside other data in ways that prevent the detection of hidden message. A steganographic system is typically not required to be robust against intentional removal of the hidden message. On the other hand, the watermarking requires that the hidden message should be robust to attempts aimed at removing it. In the case of copyright protection the copyright information should resist any modifications by pirates intending to remove it. This is a significant step forward compared to common steganography.

Watermarking is either “visible” or “invisible” [47]. Perceptible mark (“visible watermark”) of ownership or authenticity has been around for centuries in the form of stamps, seals, signatures or classical watermarks. Nevertheless, for known data manipulation technologies the imperceptible digital watermarks are mandatory in most of the applications.

There are several ways to categorize audio watermarking systems [47]. They can be classified as time-domain and frequency-domain watermarking. Time-domain watermarking algorithms embed watermarks into host signals in their time domain. Frequency-domain watermarking algorithms embed watermarks in certain transform domain, such as Fourier transform domain, cosine transform domain, wavelet domain, or cepstrum domain.

Watermarking algorithms can also be categorized as fragile and robust [47]. A fragile watermark will be changed if the host audio is modified. On the contrary, watermarks in robust algorithms cannot be removed by common signal processing operations. In this thesis, we will follow this classification.

In recent years, digital watermarking algorithms [47] boomed rapidly, especially in the image watermarking field. Compared to image, the research on audio watermarking is not as mature. The most important reason is the difference between the human visual system (HVS) and human auditory system (HAS). In general, HAS is more sensitive to distortions than the visual system. Therefore, it is challenging to implement imperceptible audio watermarks.

1.1.1 Audio watermarking system

A simple yet complete model of the audio watermarking system [19] is shown in Fig. 1.1. The embedding process takes the cover signal S, a secret key K and the copyright message (watermark) W to be embedded and produces the stego-signal (watermarked signal) $S^{'}$. The embedding function is defined as

$$S^{'} = E(S, W, K) \tag{1.1}$$

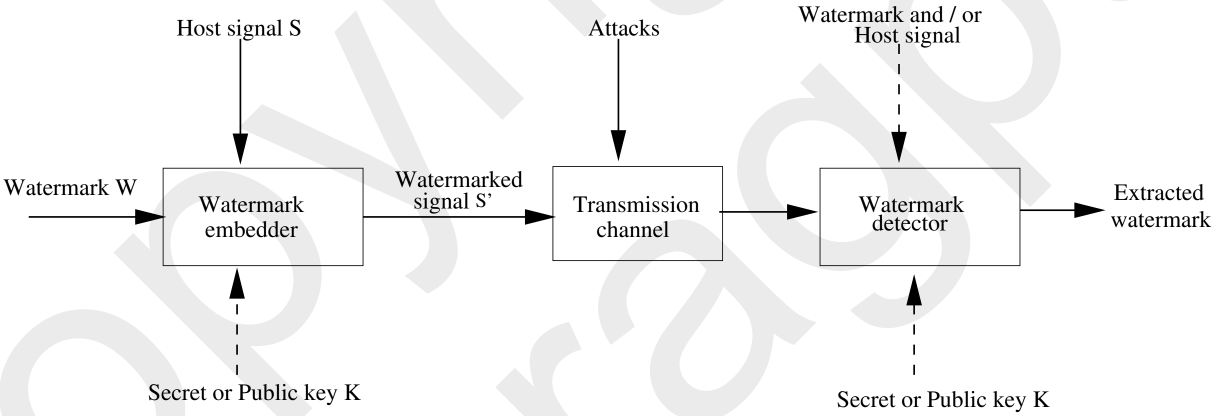

Figure 1.1: Audio watermarking system

The decoding process is completely analogous to embedding process, wherein we take the stego-signal $S^{'}$, a secret key K, and attempt to detect if the signal contains a watermark, and if it does, we attempt to decode the message. With respect to the watermark decoding process, there are two types of detectors: informed or non-blind detector, where the host signal is available during decoding phase, and blind or oblivious detector, which has no access to the host signal. In practice oblivious recovery is an important requirement as otherwise a large search in a database of signals may be required to find the original signal.

1.1.2 Requirements and algorithm design issues

According to IFPI (International Federation of the Phonographic Industry) [35] audio watermarking systems are required to satisfy a number of properties. In a later section, several applications of audio watermarking will be introduced. We will find that different applications may have different properties. Therefore, there is no unique set of properties that all watermarking techniques must satisfy. The importance of each property varies in different applications. In this part, we highlight seven properties.

1. **Imperceptibility**

   The embedded watermark should not be noticed by the user, and it cannot destroy the quality of the original signal. In other words, the data embedding process should not introduce any perceptible artifacts into the host data. Unlike robustness, this property should be achieved in all watermarking algorithms. The subjective and objective tests are usually used to evaluate the signal quality.

2. **Robustness**

   A second important requirement of watermarking algorithm is robustness. It refers to the ability to detect the watermark after common signal processing operations as well as malicious attacks. Different host signals may face certain types of transformations. For example, image watermarks may need to be robust against rotation, which is never a concern in audio watermarking algorithms. Even for one certain type of host signal, a watermarking algorithm need not be robust to all possible signal processing operations. It is very much application dependent.

3. **Data Payload**

   Data payload is the number of bits of watermark that is embedded within a host signal. For audio, data payload refers to the number of watermark data bits that may be reliably embedded within a host signal per unit of time, usually measured using bits per second (bps). There should be more than 20 bps data payload for watermark. Different applications may require different data payload. On-line applications like broadcast monitoring require embedding a segment of watermark in every short-time interval. The total amount of watermark bits required for such applications is then huge. On the contrary, an application for copyright protection may need only a small amount of information to be incorporated in the signal.

4. **Security**

   The security of a watermarking system means that an unauthorized user can neither embed a watermark, nor detect if a given signal contains a watermark. If security is required in a watermarking system, at least one secret key has to be used for the embedding and extraction process. For example, in many algorithms, the embedded watermarks are pseudo-random signals. In this case, the seed of the pseudo-random number generator may be used as a secret key. According to the Kerchhoff's principle, the security of watermarking algorithm should not rely on the secrecy of its algorithm but on the knowledge of the key only.

5. **Oblivious vs. non-oblivious watermarking**

   In some applications, like copyright protection, watermark extraction algorithms can use the original un-watermarked data to find the watermark. This is called non-oblivious watermarking. In most other applications, e.g. copy protection, the watermark extraction algorithms do not have access to the original un-watermarked data. This renders the watermark extraction more difficult. Watermarking algorithms of this kind are referred to as public or blind or oblivious watermarking algorithms. It turns out that blind methods are more secure than non-blind methods.

6. **Low complexity**

   Different applications require the watermark detection to be done at different speeds and complexity. In broadcast monitoring, the detector is required to work in real time. The computational cost should be low enough to make the decoder keep up with the real time broadcasts. However, speed is not an issue when using watermarks to track illegal copies. For real-time applications, watermarking algorithms should not be excessively time consuming.

7. **Reliability**

   Data contained in the watermark should be extracted with acceptable error rates.

The relative importance of a particular property is application dependent, and in many cases, even the interpretation of a watermark property varies with the application. In fact, it is not possible to achieve all of these properties in one watermarking system at the same time, since there are some trade-offs among them. In order to make a watermark difficult to remove, the watermark must be placed in the perceptually "important" parts of the host signal. For example, the audio watermark should be placed in the portions of the audio that affect human hearing the most. However, placing the watermark in the "important" parts is against the goal of reducing the perceptual effect of the watermark. Thus, the robustness and the imperceptibility of a watermark cannot be maximized at the same time. The similar conflict occurs between the data payload and the imperceptibility of the watermark. The more bits we embed into the signal, the more likely people will notice the presence of the watermark. Therefore, we should optimize these properties according to the specific application. The mutual dependencies between each of these basic requirements are shown in Fig. 1.2. In fact the first three requirements (imperceptibility, robustness, payload) can form sort of a magic triangle as shown in Fig. 1.3, which means that if one is improved, the other two might be affected.

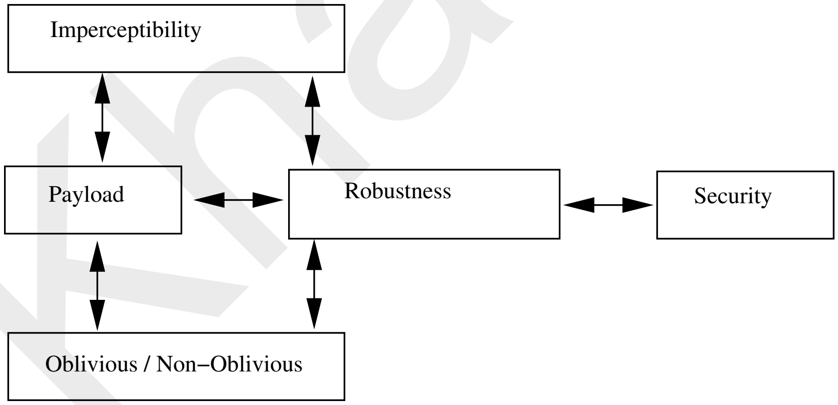

Figure 1.2: Mutual dependencies between the basic requirements

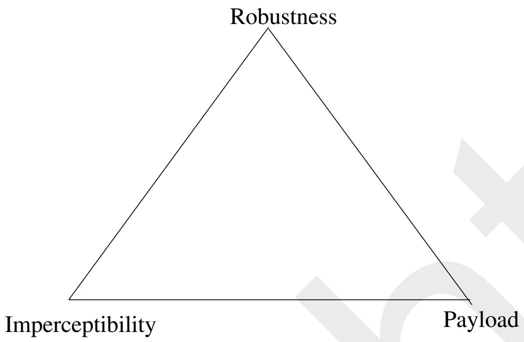

Figure 1.3: Three contradictory requirements of watermarking

The watermarking system design process involves balancing several conflicting requirements of watermark, like robustness, stego-signal fidelity, watermark data payload, and security. In the case of fragile watermarking algorithms, robustness to selective manipulations or attacks is desired. The fidelity and robustness criteria are generally competing, as greater robustness requires more watermark energy and more manipulation of the cover-signal, which in turn lead to distortion of the original content. In fact, when stego-signal fidelity parameter is fixed, robustness and data payload cannot be increased simultaneously [19]. A trade-off also exists between watermark security and stego-signal fidelity as evidenced by the fact that visible watermarking algorithms tend to be less secure. The computational complexity of watermarking algorithms is also a factor for consideration, especially for real-time applications.

1.1.3 Attacks against watermarking algorithm

In this section we briefly review some common attacks on robust audio watermarking system [74]. Any procedure that can decrease the performance of the watermarking algorithm may be termed as an "attack". Testing the robustness and security of a audio watermarking system against attacks is as important as the design process. To achieve the high reliability of watermark detection, the watermark detection process has to be robust to the alterations in the host signal caused from both common signal processing operations and malicious attacks. Depending on the application and watermarking requirements, the list of attacks to be considered includes, but not limited to:

- additive and multiplicative noise;
- linear and nonlinear filtering, for example, low-pass filtering;
- data compression, for example, MPEG audio layer 3;
- local exchange of samples, for example, permutations;
- quantization of sample values;
- temporal scaling, for example, stretch by 10%;
- equalization, for example, +6 dB at 1 kHz and -6 dB at 4 kHz;
- removal of insertion of samples;
- averaging multiple watermarked copies of a signal;
- D/A and A/D conversions;
- frequency response distortion;
- group-delay distortions.

The aim of these attacks is not always to completely remove or destroy the watermark but usually to disable its detection. Attacks are limited to those not producing excessive degradations. Otherwise, the transformed watermarked object would be unusable. These distortions could also introduce degradation in the performance of the system.

1.1.4 Applications

There are quite a number of watermarking systems developed based on different applications [14]. In recent years, some new applications of audio watermarking have been discovered. For example, audio watermarking algorithm has been designed for audio and speech quality evaluation. In this section, we list out six main applications of audio watermarking: copyright protection, copy protection, content authentication, broadcast monitoring, fingerprinting and steganography.

1. **Copyright protection**

   One of the motivations of introducing audio watermarking is copyright protection. The idea is to embed information of the copyright or the owner(s) into the data to prevent other parties from claiming to be the rightful owner(s) of the data. A watermark used for this purpose is known only by the author of the digital source and is supposed to be very robust against various attacks intended to remove it. It also has to be unambiguous and still resolve rightful ownership after other parties embed additional watermarks. The idea of using digital watermarks for copyright protection was introduced in 1994 by Brassil et al. [8]. Since then digital watermarking has gained a lot of attention and has evolved quickly. A lot of practical working methods and systems have been developed.

2. **Copy protection**

   This application is used to prevent illegal copies of record-able CDs, record-able DVDs, or any other digital recording technologies. A watermark is embedded in the content. It is used to tell the recording equipment whether this content could be recorded. If the recording device were fitted with a watermarking detector, the device could be made to prohibit recording whenever a never-copy watermark is detected in the content.

3. **Content authentication**

   Fragile watermarks can be used to check the authenticity of the data. A fragile watermark indicates whether the data has been altered and supplies localization information as to where the data was altered. In content authentication, the signature information is embedded in the source, and later is used to verify whether the content has been tampered with or not. In this application, the robustness of the watermark is not a concern. If the source is modified, the watermark along with it is also modified. We call this kind of watermark a fragile watermark. The difference between the original and extracted watermarks is used to authenticate the source.

4. **Broadcast monitoring**

   In audio broadcast, advertisers want to make sure that they receive all of the air-time for which they are paid. Owners of copyright programs need to know about any illegal broadcast. Audio watermarking is one convenient technique for the purpose of broadcast monitoring. In this application, the program or the advertisement to be monitored is embedded with a watermark before broadcast. During broadcast period, a computer monitor broadcasts and detects the watermark in certain time intervals. This application requires high data payload and low computational cost at the decoder. Imperceptibility of the watermark is also a major concern.

5. **Fingerprinting**

   To trace the source of illegal copies, the owner can use a fingerprinting technique. In this case, the owner can embed different watermarks in the copies of the data that are supplied to different customers. Fingerprinting can be compared to embedding a serial number that is related to the customers' identity in the data. It enables the intellectual property owner to identify customers who have broken their license agreement by supplying the data to third parties.

6. **Steganography (secret communication)**

   This application requires that the embedded message should be undetectable by an attacker. In this case the embedder does not just want to hide the conveyed information, but he/she also tries to hide the process of communication itself, in such a way that an observer of the watermarked signal should not be able to state whether a watermark was hidden or not. Steganography can be used in military communications to present the enemy from knowing that the communication is taking place. In that sense, steganography resembles the use of low-powered communication techniques, such as spread spectrum, since an observer unaware of the spreading sequence (equivalently, a secret parameter for steganography) will not be able to decide whether a communication is taking place.

1.2 Motivation

The thesis is motivated by several emerging applications of audio watermarking. With the rapid growth of the Internet, the digital format of media is becoming more and more ubiquitously used. A variety of software makes it convenient for consumers to create, manipulate, and store digital multimedia data. Internet and wireless network provide channels to transmit and exchange multimedia information. However, they also pose the danger of illegal copying, redistribution, and various malicious attacks. Therefore, the protection of ownership and the prevention of unauthorized tampering of digital multimedia data become important concerns.

The most common method to achieve security is to use cryptographic techniques [75]. In cryptography, data is encrypted before transmission, and can be retrieved by decryption. After decryption, the digital signature is removed, and there is no further proof of the ownership of the data. In other words, cryptography can protect digital data only during transmission, but once decrypted, the data forfeits protection. Moreover, cryptographic tools like encryption are meant for point-to-point communication and are not applicable to copy protection straightaway.

Watermarking is a new technique that has the potential to protect digital data even after decryption. A watermark is a data stream imperceptibly embedded into the host signal. Once a watermark is embedded, it is never removed during normal use. When the host signal is modified, the watermark undergoes transformations.

The watermark can be used to identify the copyright holder, to prevent illegal copying, and to verify whether the content is modified. The main advantages of watermarks over other techniques include the following.

- They are imperceptible.
- They are not removed when the data are converted to other file formats.
- They undergo the same transformations as the data in which they are embedded.

The music industry claims annual multi-billion-dollar loss in its revenues due to piracy. This poses a challenge to the watermarkers to improve upon existing copyright protection schemes and come up with robust and secure technologies.

A watermark can be embedded in several domains: time domain [21], discrete Fourier transform domain (DFT) [81], discrete cosine transform domain (DCT) [96], discrete wavelet transform domain (DWT) [88], singular value decomposition domain (SVD) [62], cepstrum transform domain [50], Quantization index modulation (QIM) [10] and so on. Watermarking in the time domain exhibits very good performance in terms of imperceptibility and capacity. At the same time, it shows lack of robustness against compression and common signal processing attacks. On the other hand, watermarking in the transform domain results in improved robustness.

The advantages of using SVD in audio watermarking arises from the fact that small changes in the singular values (SVs) of an audio signal do not affect the audio quality significantly, and the SVs are invariant under common signal processing operations.

Only a few audio watermarking algorithms using QIM techniques have been proposed in the literature [19]. There remains a need to study watermark embedding techniques using QIM. Some advantages of QIM method are good rate-distortion-robustness tradeoffs, blind detection, simplicity, and low complexity of encoding and decoding. QIM algorithms have been found to perform better than traditional spread spectrum watermarking. In contrast with the low capacity problem inherent in spread spectrum based watermarking techniques, quantization based watermarking techniques normally have high capacity.

1.3 Objective of the thesis

Audio watermarking techniques must be designed to meet a number of objectives. These can be broadly summarized under three categories: imperceptibility, robustness and functionality.

- **Imperceptibility:** The most important aspect of a successful watermarking algorithm is that the embedded watermark must be imperceptible to the end user. This requirement means that not only should the addition of the watermark be undetectable by the user, but also that the existing carrier signal should, for all practical purposes, be unaffected by the watermark. In audio, the watermark must be added in such a way that the HAS is unable to distinguish between the marked and the unmarked versions of the carrier signal.

- **Robustness:** Any successful watermarking algorithm must be robust. The embedded watermark signal must be resistant to general noise addition (during copying and transmission). It must also withstand a variety of signal processing and malicious attacks.

- **Functionality:** The functionality of the watermarking algorithm is also of high concern. If the detection of the watermark is unreliable or inconsistent, it is of little use. The watermarking algorithm must also be able to embed a variety of message types at as high a data rate as possible.

Watermarking algorithms using QIM appear to be the best known candidates to meet the above objectives. In view of this, we propose several QIM-based algorithms in the thesis.

1.4 Contributions of the thesis

In this thesis, audio watermarking algorithms using quantization techniques are studied. The basic aim is to develop novel audio watermarking algorithms providing performance enhancements over existing algorithms. This is accompanied by extensive experimentation to validate the claims of performance enhancements. Five novel algorithms proposed in the thesis are now elaborated.

1. **Improved Watermarking Based on Quantization in the Wavelet Domain**
   The first contribution is to develop an improved audio watermarking algorithm based on the quantization of wavelet coefficients. In the improved watermarking algorithm we have applied SVD instead of DWT. The watermark data is embedded by quantizing largest singular value coefficients. Experimental results show that the proposed algorithm has good imperceptibility, and robustness against common signal processing and stirmark attacks. The data payload of the algorithm is 196 bps. The false negative probabilities under the proposed algorithm are close to zero.

2. **Improved Watermarking Based on Mean Quantization in the Cepstrum Domain**
   The second contribution is devoted to design of an improved audio watermarking algorithm based on cepstrum domain transform. In the improved watermarking algorithm we have used SVD instead of cepstrum transform. This algorithm embeds the watermark data into the original audio signal using quantization of the norm of singular value coefficients. Experimental results show that our proposed algorithm is not only imperceptible, but also robust against common signal processing and stirmark attacks. The watermark payload of the proposed algorithm is 196 bps. The false negative error probabilities of our algorithm are close to zero. The algorithm can extract the watermark without the help of the original audio signal.

3. **Adaptive Watermarking using SVD and Quantization**

   The third contribution introduce a secure, robust, and blind adaptive audio watermarking algorithm based on SVD and quantization. In our algorithm watermark is embedded by applying a quantization process on the SVs in the SVD of the audio signal blocks. The watermarked signal is perceptually similar to the original audio signal and gives high-quality outputs. Experimental results show that the hidden watermark data is robust to signal processing and stirmark attacks. The data embedding rate of the proposed algorithm is 196 bps. The false negative error probabilities of the proposed algorithm are close to zero.

4. **Watermarking Based on SVD and Quantization**

   The fourth contribution is based on a blind audio watermarking algorithm using SVD and quantization. The watermark insertion and extraction process are based on the quantization process on the singular values of the blocks of the host audio signal. The quality of the watermarked signal is very high. Simulation results demonstrate that this algorithm is robust to signal processing and stirmark attacks. The data payload of the algorithm is 196 bps. The false negative probabilities are very low.

5. **Watermarking Based on SVD and Dither-Modulation Quantization**

   The last contribution deals with a new audio watermarking algorithm based on SVD and dither-modulation (DM) quantization. The watermark is embedded using DM quantization on the singular values of the blocks of the host audio signal. The watermark can be blindly extracted without the knowledge of the original audio signal. Subjective and objective tests confirm high imperceptibility of the watermark. Our algorithm highly robust against signal processing and stirmark attacks. The watermark data payload of the algorithm is 196 bps. The proposed algorithm achieves low false negative error probability rates.

6. **Performance Comparison of Proposed Audio Watermarking Algorithms**
   In this chapter performance comparison of proposed audio watermarking algorithms is presented. Finally performance is given with other audio watermarking algorithms available in the literature.

1.5 Organization of the thesis

The thesis is organized into nine chapters and the rest of the thesis is organized as follows.

Chapter 2 provides a survey of published audio watermarking algorithms.

Chapter 3 presents improved audio watermarking algorithm based on quantization in the wavelet transform domain.

Chapter 4 describes improved audio watermarking algorithm based on mean quantization in the cepstrum domain.

Chapter 5 introduces an adaptive blind audio watermarking algorithm based on singular value decomposition and quantization.

Chapter 6 deals with a new blind audio watermarking algorithm based on singular value decomposition and quantization.

Chapter 7 proposes an oblivious audio watermarking algorithm based on singular value decomposition and DM quantization.

Chapter 8 gives performance comparison of proposed watermarking algorithms.

Chapter 9 concludes the thesis by highlighting some important issues that are not addressed in this thesis and that may be taken up as avenues for further research in audio watermarking.