17

Digital Watermarking Techniques for AVS Audio

Bai-Ying Lei[1], Kwok-Tung Lo[1] and Jian Feng[2]
[1] *The Hong Kong Polytechnic University, Hong Kong, China*
[2] *Hong Kong Baptist University, Hong Kong, China*

1. Introduction

One of the biggest technological events of the last two decades was the invasion by digital media in every aspect of our life. The proliferation of the Internet, which has enabled audio material (copyrighted or not) to be widely disseminated, has also made wiser the use of many compressed audio formats such as MP3 (MPEG Layer 3), AAC (Advanced Audio Coding), WMA (Windows Media Audio) and AVS (Audio Video Standard) at a global scale. AVS is an emerging coding standard fully developed by China (AVS Audio Expert Group, 2005). As digital data can easily be copied and distributed, protection of the copyright of media data becomes one of the most important topics in the Internet world. Digital watermarking has been identified as one of the feasible solutions for content protection and received considerable attention in recent years. Watermarking is the process of embedding hidden content protection information into digital data by making small modifications on the data. The biggest challenge of watermarking techniques is to embed the copyright information without affecting the human perception.

Digital audio watermarking technology helps to prevent or reduce unintentional and intentional copyright infringements by either notifying a recipient of any copyright or licensing restrictions or inhibiting or deterring unauthorized copying (He, 2008; Nedeljko & Seppanen, 2008). However, once the user downloads the music with the valid key, the embedded watermark can be easily removed by some software and hence it infringes on the purpose of audio watermarking. How to protect audio copyright efficiently and effectively has become a great challenge. The conflicts with MP3 copyrights make AVS audio copyright protection a problem which must resolve as AVS audio doesn't take copyright protection scheme into consideration. The development of effective watermarking schemes in the AVS compressed domain will remain an important research area in the future.

In order to achieve copyright protection of audio, the watermarking scheme needs to meet the following requirements:

- The watermark should be inaudible to human ears;
- Watermark detection should be done without referencing the original audio signals;

- The watermark should be undetectable without prior knowledge of the embedded watermark sequence;
- The watermark is directly embedded in the audio signals, not in a header of the audio;
- The watermark is robust to resist common signal processing manipulations such as filtering, compression, filtering with compression, and so on.

The organization of this chapter is as follows. Following the Introduction section, in Section 2, the emerging Chinese Audio Video Standard (AVS) for digital audio will be described. An overview of AVS audio coding will be given in this section. The state of the art audio watermarking techniques and related works are surveyed in Section 3. The psychoacoustic modelling for AVS and audio watermarking is described in Section 4. Section 5 introduces chaotic sequence and presents a comparative study with PN sequence in the field of watermarking. The new perception-based watermarking scheme for AVS audio is introduced in Section 6. The proposed algorithm embeds the chaotic signal as watermark in parts of the audio data that are masked or not perceptible because of psycho-acoustic laws. Finally, the chapter is concluded in Section 7.

2. Overview of AVS Audio

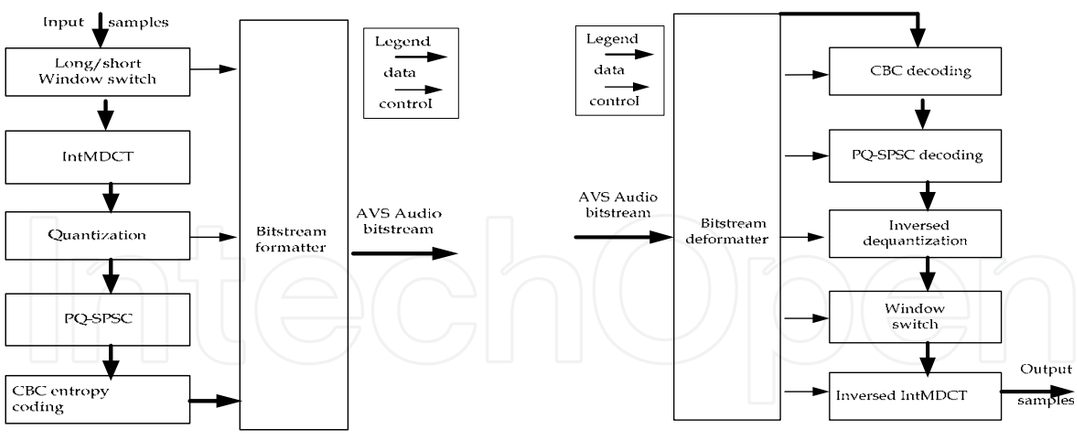
AVS is an emerging coding standard fully developed by China (AVS Audio Expert Group, 2005). AVS includes four main technical standards: system, video, audio and digital right management and support standard-consistency test. AVS has great local advantages with its own independent intellectual property rights in China. Currently, AVS is used by China Unicom as its IPTV standard. Meanwhile, it is also adopted in the CMMB (China Multimedia Mobile Broadcasting) and TD-SCDMA (Time Division - Synchronous Code Division Multiple Access) which is the third generation mobile telecommunication system developed by China and is currently deployed by China Mobile. It is evident that AVS audio will be given the first priority by IPTV and mobile TV operators in China. This indicates that the prospective future and broad development of the AVS standard.

The primary goal that the AVS Audio workgroup wanted to achieve, on the premise of developing their own intellect property, is to establish an advanced Chinese audio coding and decoding standard with general performance equivalent or superior to MPEG AAC. AVS Audio codec supports mono, dual and multichannel PCM audio signal and sampling rate of the input signal ranges from 8 KHz to 96 KHz, and the output bitrate is from 16kbps to 96kbps per channel. AVS audio can be applied in different fields of information industry such as high-resolution digital broad-cast, high-density laser and digital storage media, wireless broadband multimedia communication, internet broadband streaming media and portable media player. Compared with the MPEG AAC, AVS audio has its own characteristics and merits. For example, it introduces the Context-dependent Bit-plane Coding (CBC) for entropy coding, which has higher coding efficiency than the existing ones.

The AVS audio codec is based on a perceptual audio coding model, which has similar architecture and characteristics with most of perceptual audio coders. The lossy compression systems exploit both perceptual irrelevancy and statistical redundancy to

achieve the coding gain. But many new algorithms are designed for the core modules, which are different from other solutions. The block diagram of the AVS audio codec is shown in Fig. 1. As shown in Fig.1(a), the encoder manipulates digital audio signal and outputs the compressed bitstream. When the input audio samples enter the encoder, the window switch determines the length of analysis block depending on the transients. IntMDCT is introduced as the time-frequency analysis. Post Quantization Square Polar Stereo Coding (PQ-SPSC) is adopted to improve stereo audio signals' quality. CBC (Context-dependent Bit-plane Coding) performs entropy coding of quantized spectrum data and scale factors. Finally, the bitstream formatter outputs the bitstream in a suitable packet format. The decoding process is the inverse of encoding process as illustrated in the Fig. 1(b). The decoder recovers and decodes the quantized audio spectra of the bitstream with the process of the reconstructed spectra. At first, AVS audio bitstream is decoded in the CBC decoding part, and then the bitstream is fed into PQ-SPSC module before the dequantization module according to the header information. The frequency domain audio signal in the window switch module is transformed into the spatial domain in the Inversed IntMDCT module. At last, it outputs the PCM signal.

In order to meet different demands, AVS Audio coding technology adopted two profiles: main profile and scalable profile. Main profile is high quality and high complexity, while scalable profile has scalable bitrate and quality. In scalable profile, the coded bitstream is composed of base layers and some enhanced layers, and it can dynamically adapt variable bandwidth and the terminal decoding capability. The coding quality of AVS Audio main profile is equivalent or superior to that of MPEG 2 AAC LC profile. The coding and decoding complexity of AVS Audio main profile is higher than that of MPEG 2 AAC LC profile. Moreover, AVS Audio supports scalable coding.



a) Encoder b) Decoder

Fig. 1. AVS audio codec diagram (AVS Audio Expert Group, 2005)

3. Review on Audio Watermarking Techniques

Digital data protection and copyright issues have become more and more important in the face of today's technology. As a solution to copyright protection issues, digital

watermarking technology (Swanson et al., 1998; Wang 1998; Bassia et al., 2001; Podilchuk & Delp, 2001; He, 2008; Nedeljko & Seppanen, 2008) is gaining attention as a new method of protecting copyrights for digital data. The basic idea to avoid the unauthorized use and distribution of proprietary data is to sign every signal by means of an imprint characteristic of the source owner or distributor. This signature should be hidden in the signal and should not change its quality. The hiding procedure can take into account the human auditory imperfect detection. This is the basic idea of watermarking. Over the years, watermarking had been applied in various applications. In general, the applications areas of watermarking include the following:

- Copyright protection;
- To prove the digital media ownership;
- The tamper-proofing to check the data integrity;
- The covert communications;
- To exchange messages secretly embedded within multimedia data and the captioning;
- To embed descriptive and useful information within audio or video for applications like audio indexing.

For copyright-related applications, the embedded watermark is expected to be immune to various kinds of manipulations to some extent and acceptable in terms of perceptual quality. Therefore, watermarking schemes for copyright-related applications are typically robust (Seok et al., 2002), i.e. they are designed to ignore or remain insensitive to manipulations. Although a watermark is designed to be imperceptible to humans, the embedding is certainly intrusive and incurs distortion to the content. In some authentication applications (Zmudzinski & Steinebach, 2009) where any tiny changes to the content are not acceptable, the embedding distortion has to be compensated perfectly. In an attempt to remove the watermark so as to completely recover the original media after passing the authentication process, reversible watermarking schemes have been proposed in the last few years. Reversibility with media-independent embedding capacity will also be in the research agenda in the future for authentication applications. Although some perceptual models have been proposed to ensure low embedding distortion, how distortion and robustness could be optimized is still an open question and we expect new models will be proposed in the future. Requirements of digital watermarking vary across applications. The main requirements are low distortion, high capacity, and high security. However, meeting all the three requirements simultaneously is usually infeasible. Thus, trade-offs are frequently made to optimize the balance for each specific application. In many applications where original media is not available at the watermark decoder, blind detection of the watermark without any prior knowledge about the original is desirable. Although many solutions have been proposed by researchers in the past decade, psychoacoustic modelling needs to be further explored as it is the fundamental problems for most watermarking systems.

Depending on the embedding domain of the watermark, the existing audio watermarking techniques can be generally classified into two main categories: time domain techniques and compressed domain techniques. For time domain techniques, the watermark is inserted directly in original audio signal without a loss of sound quality. Common techniques

include the phase or amplitude modulation, quantization scheme and the echo hiding scheme. In (Bender et al. 1996), the information is embedded by modulating the phase of the audio signal. In (Bassia et al. 2001), it usually replaces the least significant bit (LSB) of the original audio signal with pseudorandom sequences, which can be viewed as a user identification number. They found a set of schemes that embed watermarks by adding predefined very weak noise signals to audio such that the changes are inaudible. The shortcoming of these sorts of watermarks is that they are fragile to most signal processing attacks including audio compression. Another technique exploits the insensitivity of the human ear to very short delay echoes. Although there are a few variations within this group, we collectively refer to them as echo hiding techniques (Gruhl et al. 1996; Seob et al. 2005; Chen et al. 2008). Echo hiding is the technique that embeds the watermark by introducing an echo which can effectively embed imperceptible information into an audio signal. One outstanding shortcoming of these approaches is that the watermark embedding success is signal and/or delay dependent. Moreover, resolving this problem tends to make the system unacceptably complex. But the watermark embedding process is signal dependent and detection rules are very lenient due to the fact that the information can be detected by anyone—even those without a special key. The echo hiding technique proposed by (Gruhl et al. 1996) is based on human insensitivity to audio distorted with a fused "echo" signal. Basically, the echo is defined with two delay times (offsets) to specify the embedding of binary "1" and "0". Their experimental tests showed that the relative volumes of the original and the echo signals control the rate of recovery of the watermark, and that the offset largely determines the perceptibility of the modifications. Echo hiding, in general, can be a guaranteed to an extent for a limited range of audio types. Further research is still under development to improve its performance on other types of audio data. The spread spectrum (SS) technique spreads the watermark all over the host signal and makes the attackers hard to locate the watermarks. The SS scheme (Kirovski & Malvar, 2003) requires psycho-acoustic adaptation for inaudible noise embedding. This adaptation is rather time-consuming. Another disadvantage of SS scheme is its difficulty of synchronization.

Since most multimedia products are distributed in compressed format, numerous methods in compressed domain have been reported. In (Lan et al. 2007), they proposed a perceptual based scalable AVS-DRM encryption model for AVS audio in order to protect audio content. The perception classification and multiple security levels are adopted and implemented in AVS audio codec. In (Li et al. 2006), a scalable lossless watermarking built on Advanced Audio Zip was developed with watermarking scalability and adaptiveness which address the problem of nonadptiveness of lossless watermarking and irreversible distortion problems. Tachibana, et al. (2002) proposed to use a two-dimensional pseudorandom array (PRA) as watermarking identification key to embed and extract watermarks in MPEG-2 AAC bitstream. The multiple watermarking scheme with PRA can not only achieve the synchronization, but also detect watermark correlating the amplitude, therefore, it is robust to cropping, pitch shifting and other attacks. A watermarking algorithm to embed watermark into low frequency MDCT coefficients during compression was proposed by Wang et al. (2004). In their scheme, the watermark is embedded into the MDCT-transformed permuted block with embedding locations chosen quite flexibly. A pseudorandom sequence is used as the watermark embedded by modifying the LSBs of AC coefficients. In the watermark extraction process, the stego audio is first transformed into the MDCT domain

and then the watermark can be obtained by the AC coefficients modulo 2. Embedding watermarks in the MDCT domain is very practical because MDCT is widely used in audio coding standards. The watermark embedding process can be easily integrated into the existing coding schemes without additional transforms or computation. Therefore, this method introduces the possibility of industrial realization in our everyday life. Petitcolas et al. (1999) exploited watermarking software named MP3Stego for MP3 audio. The software achieves watermark based on the parity of error length after audio quantizing and coding. But it cannot meet the real-time requirement and its capacity is too low. Digital watermark is embedded to the compressed domain of AAC audio in (Neubauer &Herre, 2000). In this scheme, watermark could be extracted with ancillary data calculated by psychoacoustic model. So it would increase computational complexity greatly. Qiao & Nahrstedt (1999) modified the scale factor of MPEG audio bitstream to embed watermark. However, the results for testing the robustness were not introduced in this scheme. Quan & Zhang (2006) employed wet paper codes and hide data directly in the MPEG audio bitstream by modifying the MPEG audio quantization process. The drawback is that the scheme could change the audio file size. However, as a recent developed China's standard, there is rarely report on watermarking schemes for AVS audio based on chaos in the literature.

Besides, there are some watermarking scheme based on audio content and HAS (Garcia, 1999). Boney et al. (1996) proposed an algorithm to make use of the MPEG psychoacoustic model in order to obtain frequency-masking values necessary to achieve prime imperceptibility, which generates watermarks by filtering a PN sequence with a filter that approximates the HAS frequency masking characteristics. An enhanced technique was developed by Swanson et al. (1998), in which they examine perceptual coding techniques in order to embed the watermark. Using only temporal masking or frequency masking will not be good enough for watermark inaudibility, thus both masks are used in order to achieve minimum perception distortion in the stego audio. Experimental results showed that this technique is both transparent and robust. The similarity values proved that the watermark can still survive under cropping, resampling and MP3 encoding. Therefore, this technique is an effective one under all the considerations for audio watermarking.

Cheng et al. (2002) proposed enhanced spread spectrum watermarking for compressed audio in MPEG-2 AAC format which embedded watermarks in the discrete cosine domain with the psychoacoustic model. A spectral envelope filter was introduced in the detection phase to reduce the noise variance in the correlation, thus improving the detection bit rate. They transformed the original host signal into a new host signal with smaller variance before adding the watermark and they used a heuristic estimation on perceptual weighting for embedding the watermark. All of these features result in both low structural and computational complexities. Seok et al. (2002) proposed an audio watermarking based on traditional direct sequence spread spectrum approach and achieved a watermark insertion rate of 8 bits per second. The inaudibility of the watermark was maintained by incorporating the psychoacoustic model derived from the MPEG I layer 1 audio coding standard. Their experiments showed the audio watermarking system to be robust to several signals transformations and malicious attacks.

In summary, the watermarking systems are designed to embed a hidden robust watermark into digital media. These systems have to satisfy two conflicting robustness and good fidelity requirements. To accomplish this task, variety of techniques has been exploited, and different domains are involved to enhance a certain application of watermarking and/or improve fidelity and robustness of watermarked signal. However, watermarking systems have a number of differences. These differences can be considered in evaluating the performance of watermarking systems and suitability of these systems for a specific application. Digital watermarking work mainly concentrated on the design of watermark algorithm, which is generally divided into watermark generation, embedding and detection in the spatial, frequency, cepstrum or mixed domain using symmetric and public key. Under this background, digital watermarking has received much attention recently and has been a focus in network information security. There is unprecedented development in the audio watermarking field. On the other hand, attacks against watermarking systems have become more sophisticated. In general, these attacks can be categorized into common signal processing techniques, such as AVS, MP3 and MPEG compression, low-pass filtering, noise addition, requantization, resampling and so on. Apart from security-oriented applications, which will continue to attract research interests, digital watermarking has been proved to be useful for broadcast monitoring and we believe that it can be useful for other non security-oriented applications such as error concealment and metadata hiding within multimedia content for legacy systems so that the metadata can survive format conversions. The latter is particularly useful for document identification as it allows us to re-associate medical images with patients' records and linking multimedia to the World Wide Web.

4. Psychoacoustic Modelling

Psychoacoustic modelling is important in audio coding and watermarking, which ensures that the changes of the original signal remain imperceptible. Compared with HVS, HAS is much more sensitive, which makes audio watermarking more challenging than image watermarking. HAS can detect signals with a range of frequency greater than 10^3 and with a power greater than 10^6. Understanding how HAS perceives sound is important for the development of a successful audio watermarking system. The main property of audio perception lies in the masking phenomena, which includes pre-masking and post-masking. Psychoacoustic modelling has made important contributions in the development of recent high quality audio compression methods and has enabled the introduction of effective audio watermarking techniques (Swanson et al., 1998; Robert& Picard, 2005). In audio analysis and coding, it strives to reduce the signal information rate in lossy signal compression, while maintaining transparent quality. This is achieved by accounting for auditory masking effects, which make possible to keep quantization and processing noises inaudible. In speech and audio watermarking, the inclusion of auditory masking has made possible the addition of information that is unrelated to the signal in a manner that keeps it imperceptible, transparent and can be effectively recovered during the identification process. Furthermore, no psychoacoustic model is available in the AVS compressed domain to enable the adjustment of the watermark to ensure inaudibility.

HAS can be modelled as a frequency analyzer consisting of a set of 25 bandpass filters (called critical bands) with logarithmically widening bandwidth for higher frequencies. The

critical band rate (CBR) specifies the correspondence between frequencies pooled for perception and locations on the basilar membrane. The mapping between CBR and frequency has been approximated as:

$$z = 13\arctan(0.76) + 3.5\arctan(f/7.5)^2 \quad (1)$$

where z is CBR in Bark and f is frequency in kHz. In psychoacoustics, the intensity of a sound is measured in terms of Sound Pressure Level (SPL). The HAS cannot perceive the sound if SPL is below a threshold. Such threshold is called absolute threshold of hearing(ATH), which determines the energy for a pure tone that can be detected by HAS in noiseless environment. ATH is a nonlinear function approximated by (2) and is illustrated in Fig. 2.

$$T(f) = 3.64(f/1000)^{-0.8} - 6.5e^{-0.6(f/1000-3.3)^2} + 10^{-3}(f/1000)^4(dB) \quad (2)$$

where T is the SPL in dB and f is the frequency in Hz.

The basic idea underlying perceptual watermarking schemes is to incorporate the watermark into the perceptually insignificant region of an audio signal in order to ensure transparency. An important characteristics of HAS is auditory masking that has been applied in the area of perception-based compression audio coding and in the watermarking field too. If a sound has any frequency components with power levels below the ATH, then these components can be discarded without degrading the perceptual quality of the sound. The AVS audio algorithm compresses audio data in large part by removing the acoustically irrelevant parts of the audio signal. In fact, it takes advantage of the HAS inability to hear quantization noise under conditions of auditory masking. This masking occurs whenever the presence of a strong audio signal makes a temporal or frequency neighbourhood of weaker audio signals imperceptible. The psychoacoustic model analyzes the audio signal and computes the amount of noise masking available as a function of frequency. The encoder uses this information to decide how best to represent the input audio signal with its limited number of code bits. The psychoacoustic model defined in AVS audio algorithm exploits the frequency masking: let two simultaneously occurring signals be close together in frequency, the lower-power frequency components may be inaudible in the presence of the higher-power frequency components. Note that the frequency-masking model defined in AVS audio is to obtain the spectral characteristics of a watermark based on the inaudible information of the HAS. The perceptual model is based on the psychoacoustic phenomena of masking.

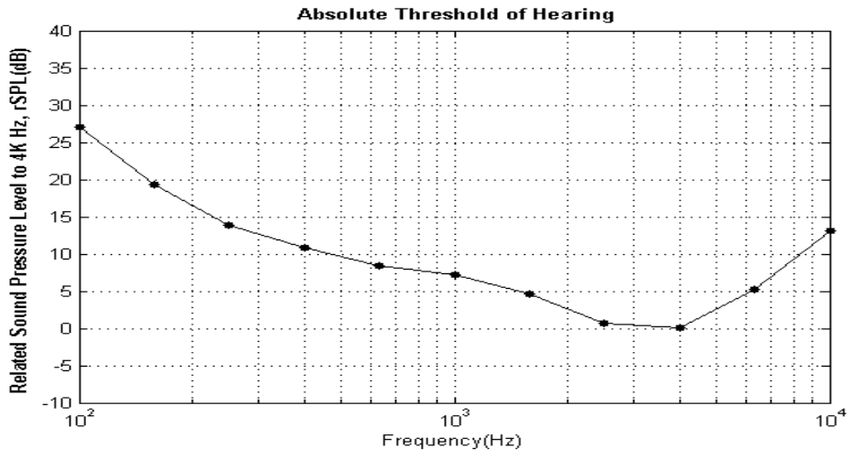


Fig. 2. Absolute threshold of hearing

The masking model estimates the amount of (noise-like) energy that can be added to the original audio signal without becoming perceptible. This energy varies over frequency. Therefore, the implementation of the masking model introduces frequency band partitions. The maximum energy allowed in a particular band is referred to as the masking threshold. ATH at any frequency is the minimum intensity of the sound at that frequency a normal human ear can perceive which is similar to overall threshold. A portion of an audio with 44.1 kHz sampling rate, which is expressed by the power spectrum, ATH, overall threshold and the audio power spectrum density (PSD) is illustrated in Fig.3. In the psychoacoustic model, the frequency masking threshold for each frequency component is computed from specific audio signal. Besides, modifications to the audio frequency components whose magnitudes are less than the masking threshold create no audible distortions to the audio piece by normalizing the added chaotic signal based on these threshold values.

Perceptual weighting in audio watermarking algorithm considers the human audio perception by means of adapting the introduced watermark energy to the current masking threshold. On the one hand, this can lead to audible distortion for the case that the amount of embedded watermark energy is too large. On the other hand, from a psycho-acoustical view, in some cases, more noise energy would be allowed. In other words, less watermark energy than possible is introduced and only suboptimal extraction performance is achieved. Psychoacoustic modelling strives to reduce the signal information rate in lossy signal compression while maintaining transparent quality. This is achieved by accounting for auditory masking effects, which makes it possible to keep quantization and processing noise inaudible. In speech and audio watermarking, the inclusion of auditory masking has made possible the addition of information that is unrelated to the signal in a manner that keeps it imperceptible. Perceptual weighting of the added chaotic sequence is performed with the frequency characteristic similar to the threshold in quiet curve of the HAS.

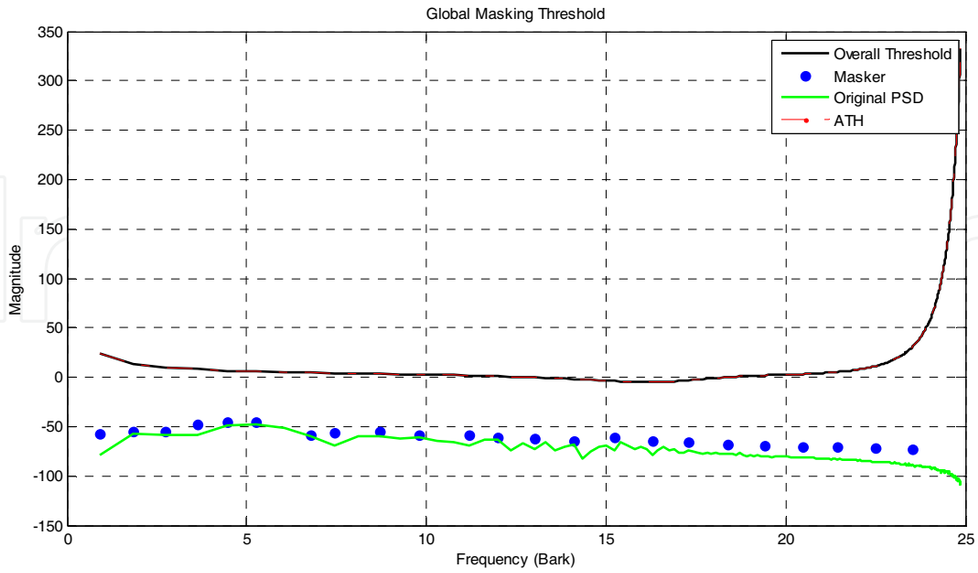


Fig. 3. ATH, overall threshold and audio PSD

5. Chaotic Watermarking Scheme for Audio Signal

5.1 Properties of Chaotic Function

Mathematically, chaos means deterministic behavior which is very sensitive to its initial conditions, in other words, infinitesimal perturbations of initial conditions for a chaotic dynamic system lead to large variations in behavior. A chaotic map exhibiting some sort of chaotic behavior is fully described by (3).

$$\{x_n : x_n = F(x_{n-1,} r)\} \tag{3}$$

where r is a function seed and $F()$ is a nonlinear transformation that maps scalars to scalars condition. A discrete-time chaotic signal x_n can be generated from a chaotic system with a single state variable by applying the recursion:

$$x_n = F(x_{n-1}) = F^n(x_0) \tag{4}$$

where x_0 is the system initial condition. A chaotic sequence may be easily reproduced given the same initial conditions and initial value x_0. A slight change in the initial conditions of a chaotic function will lead to significant changes in the resultant mapping but their correlation may change slightly which can be seen in Fig. 4. Chaotic maps often occur in the study of dynamical systems and often generate fractals. A fractal may be constructed by an iterative procedure studied as sets rather than in terms of the map that generates them. This is often because there are several different iterative procedures to generate the same fractal. The simplest chaotic map is logistic map which is defined as:

$$x_{n+1} = rx_n(1 - x_n) \tag{5}$$

where r is the function seed and x_n is the current value of the mapping at time n with an initial value x_0. For the logistic map, there are two distinct regimes, namely, the periodic or bifurcation regime and the chaotic regime. When r=3, x_n oscillates between two values and never converges. As r increases, x_n goes through bifurcations and eventually becomes chaotic. When $r \approx 3.5699$, x_n becomes random. When $r > 3.5699$, the behavior is in chaotic state. A bifurcation diagram summarizes this in Fig. 5. The horizontal axis shows the values of the parameter r, while the vertical axis shows the possible long-term values of x.

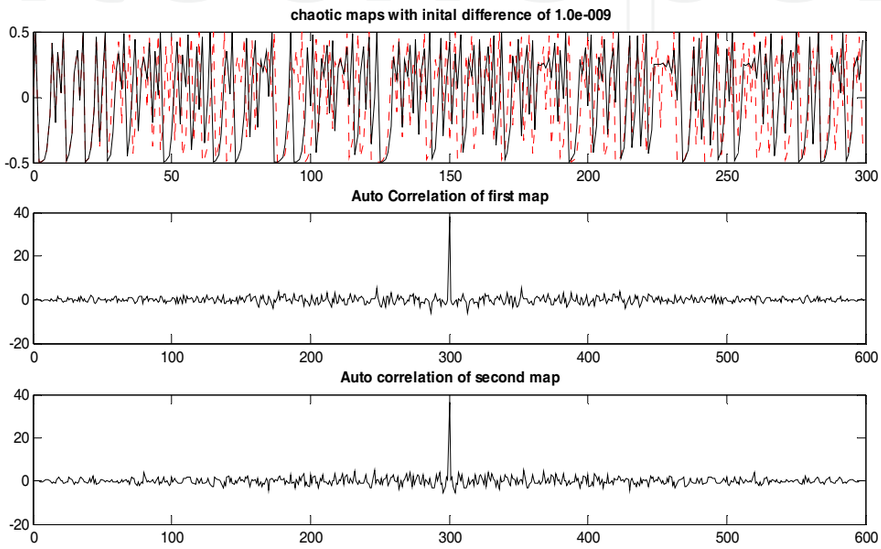


Fig. 4. The chaotic sequence with slight change and auto correlation

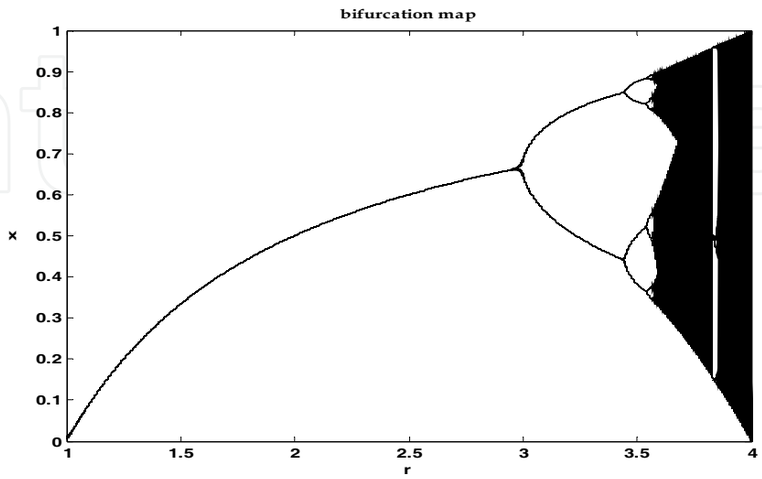


Fig. 5. Bifurcation diagram of logistic map

To date, a number of chaotic functions have been proposed in the literature for the purpose of watermark generation (Giovanardi et al., 2003; Tsekeridou et al.,2001), the most prominent and popular ones are the tent map, Bernoulli map, logistic map and quadratic map. Pseudo random sequences generated by chaotic dynamic system can be used to randomize coefficients in the watermarking field. A chaotic function can produce almost uncountable random sequences that have good auto and cross correlation characteristics and are extremely sensitive to the initial secret keys. From Fig. 6, we can see different time series of different chaotic sequences and their auto and cross correlation properties, where ACF denotes the auto correlation and CCF means cross correlation.

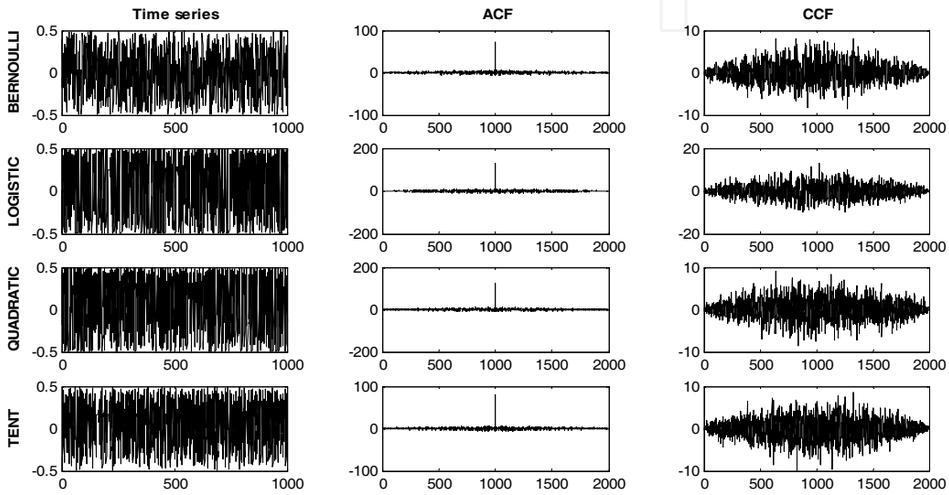


Fig. 6. Time series, ACF and CCF of different chaotic maps

The degree of the randomness of any map, *F(x),* can be seen from the Lyapunov Exponent variation of the map with respect to the value of r. We can calculate the Lyapunov Exponent λ for the chaotic function as follows:

$$\lambda = \lim_{N\to\infty} \frac{1}{N} \log \left| \frac{d}{dx} F^n(x) \right|_{x_0} \tag{6}$$

By applying the chain rule, the term reduces to

$$\lambda(x_0) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N} \log(|F'(x_i)|) \tag{7}$$

After N iterations, if Lyapunov Exponent is negative, then the orbits converge in time(periodic), and if Lyapunov Exponent is positive, then the distance between nearby orbits grows exponentially in time, and the system exhibits sensitive dependence on initial conditions(chaotic). The dynamics of Lyapunov exponents in terms of time values are displayed in Fig. 7.

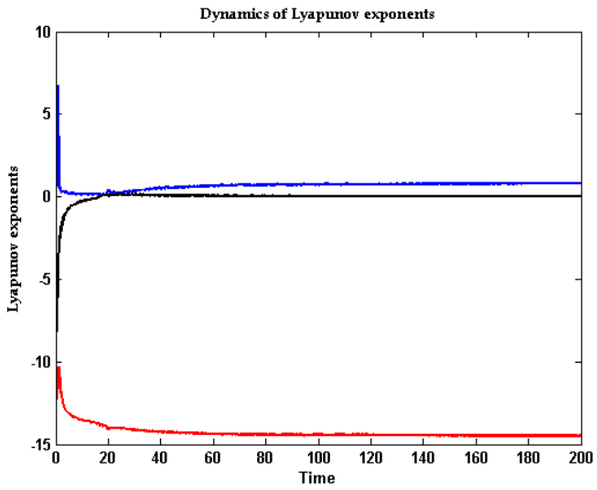


Fig. 7. Dynamics of Lyapunov exponents

5.2 Advantages of Chaotic Watermarks over Pseudorandom Watermarks

The chaotic map has been shown to produce lowpass watermarks and also white watermarks. These white watermarks are similar to those generated by a PN generator. Since there are irrational numbers close to every rational number, the map exhibits sensitive dependence on initial conditions. As iteration times increase, chaotic sequences are characterized by white spectrums. The control of spectral properties of the generated sequences of these chaotic functions offers a distinct advantage over the sequences generated by the pseudorandom generators. For example, if we know that the watermarked audio will be subjected to manipulations which are lowpass in general, we can generate a lowpass watermark which will be more robust to these attacks. By simply altering the function seed in these chaotic functions, one can generate watermarks with different spectral properties. In applications where no severe distortions are expected, e.g. in captioning or indexing applications, highpass spectrum watermarks can be used since they guarantee superior performance. Watermark signals generated by iterating of a chaotic function have an advantage over signals generated by coloring white noise in that these signals are much easier to create and re-create. Rather than have to seed a pseudorandom number generator and then apply a filter to the resultant signal to generate colored noise, a single seed can determine the properties of the generated sequence from the chaotic function. Highpass sequences are typically less robust to lowpass filtering and small geometric deformations of the image than lowpass sequences. Moreover, we can see the differences and advantages from their auto and cross correlation shown in Fig. 8 too. Chaotic have better auto and cross correlation characteristics than PN sequence. These characteristics make chaotic maps excellent candidates for watermarking and have been shown to have superior robustness than the widely used PN sequences in watermarking applications (Giovanardi et al., 2003; Tefas et al., 2003; Tsekeridou et al., 2001). Therefore, chaotic maps have recently been used for digital watermarking to increase the security.

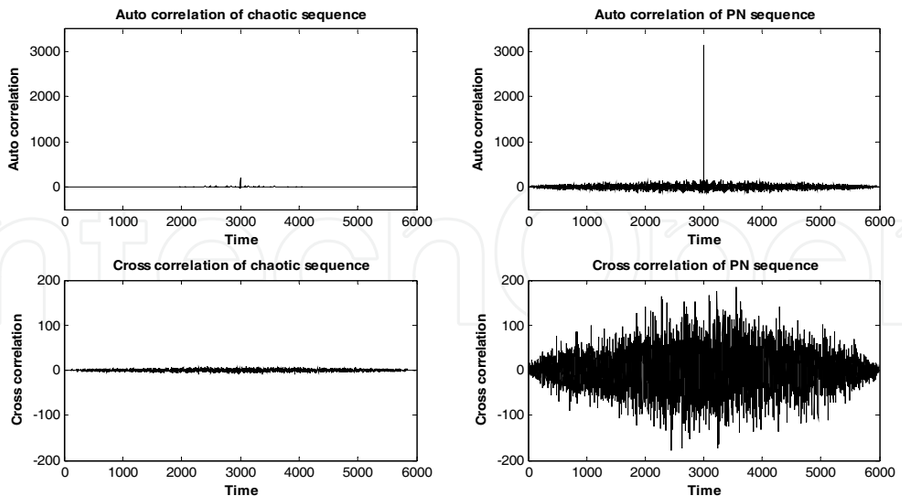


Fig. 8. Comparison of auto and cross correlation of chaotic sequence and PN sequence

Therefore, on the basis of the characteristics of the human auditory system (HAS) and the techniques of chaotic map, in next section, we present a perception-based audio watermarking algorithm for AVS audio files that works in the compressed domain, makes its manipulations in the frequency domain. Our algorithm overcomes the problem of the algorithms that operate in the uncompressed domain, which are vulnerable to compression/recompression attacks, as it embeds the chaotic signal as watermark in parts of the audio data that are masked or are not perceptible because of psycho-acoustic laws.

6. New Perceptual Watermarking Scheme for AVS Audio

6.1 Introduction

In this section, a new perceptual audio watermarking approach is developed for AVS audio based on the HAS and chaos. The copyright information is embedded into the IntMDCT (integer modified discrete cosine transform) coefficients in the AVS compressed audio data based on perceptual model of frequency masking. The proposed AVS-WM scheme is shown in Fig. 9.The IntMDCT is an integer approximation of MDCT with perfect reconstruction and has good characteristics of energy compaction. Chaotic sequences are adopted to improve the security of the proposed watermarking scheme. The low frequency components after IntMDCT transform are chaotically spread and encrypted to protect the audio copyright. The good property after various attacks demonstrates that the proposed audio watermarking scheme based on chaos is able to provide a feasible and effective copyright protection scheme for AVS audio signal efficiently.

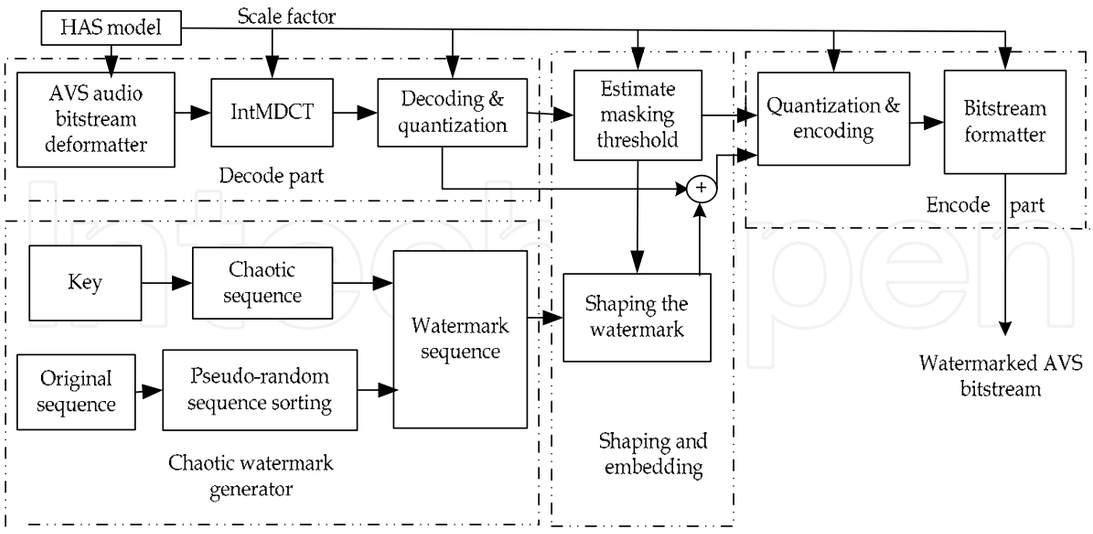


Fig. 9. The diagram of AVS-WM scheme

6.2 Chaotic Watermark Generation

Chaotic watermarks have been proposed as an alternative to the more commonly used pseudo random watermarks. The process of generating a watermark derived from a chaotic map involves several steps. A value for the function seed and an initial starting value must be first selected. The chaotic function is iterated n times. To increase the watermark security, we introduced a chaotic sequence generated by chaotic dynamic system to randomize the coefficients. A chaotic function can produce almost uncountable random sequences that have good autocorrelation characteristics and are extremely sensitive to the initial secret keys. The copyright holder can use the arbitrary real numbers between 0 and 1 as the initial secret keys in order to get satisfactory auditory quality in embedding. Without correct key, knowing the algorithm itself is far from enough to extract watermarks. The sequence x_n generated by the chaotic map is composed of real numbers, so the output sequence $c(n)$ is quantized into binary stream by the perceptual auditory masking threshold T:

$$c(n) = \begin{cases} 1, & x_n \geq T \\ 0, & x_n < T \end{cases} \tag{8}$$

where T is chosen as the perceptual masking threshold in assigning the binary values. Then we use Exclusive OR operation to generate the watermark $w(n)$:

$$w(n) = \sum_{n=1}^{N_w} b(n) \oplus c(n) \tag{9}$$

where N_w is the size of the watermark, $b(n)$ is the copyright information, $\oplus$ denotes the Exclusive-OR operation. Use the chaotic maps to generate chaotic digital watermark signal

with zero mean and one variance. The embedded watermark sequence $I'(n)$ can be obtained as follows.

$$I'(n) = I(n) + a(n)w_k(n) \tag{10}$$

where $I(n)$is the original audio sequence and a is strength parameter that can be controlled by characteristics of HAS or some variables of audio signal such as the average power of amplitude of signal for ensuring inaudibility and robustness of watermarked signal. By varying a, the embedded watermark intensity can be modulated and hence the hidden effects can be adjusted. An appropriate a is chosen to enhance the robustness of the watermark and renders the watermark imperceptible and yet with good watermark quality.

6.3 Watermark Embedding

This section proposes an enhanced watermarking scheme that not only embeds the watermark below masking threshold area, but also takes advantage the HAS properties and insert the watermark into the audible spectrum areas and still keep the introduced distortion imperceptible. The details of the watermark embedding process are as follows:

Step 1. The input original audio is segmented into overlapped frames with N samples long. Each frame is decomposed into 25 subbands.

Step 2. Compute the power spectrum of the audio segment. Each segment of the signal $s(n)$ is weighted with a Hanning window. The maximum is normalized to a reference sound pressure level of 96 dB.

Step 3. The special and noticeable audio blocks to the embedders with larger energy are selected as salient features blocks. Salient features where the audio signal energy is climbing fast to a peak value are robust against attacks. Psychoacoustic model is applied to determine the masking thresholds for each subband. High threshold value is suitable for audio with sharp energy variation. Such selection may generate audible high frequency noise. Careful shaping can reduce the noise to a hardly audible level.

Step 4. Calculation of the individual masking thresholds. An embedding frequency point in the selectable frequency range is chosen based on a chaotic secrete key. The interleaved data is spread by the chaotic sequence.

Step 5. After segmenting and choosing the frequency points, we quantize the frequency coefficients to embed the watermark. The digital watermark in the watermark generation section can be embedded into audio signal by quantizing the selected transformation coefficients.

Step 6. For each selected audio block, IntMDCT transform is performed. It is common to utilize the HAS masking effects for keeping good inaudibility and robustness. The watermark is only embedded into salient features of audio blocks which have a high energy value in our scheme. The data to be embedded (hidden data) is spread by chaotic sequence and interleaved to enhance watermarking robustness.

Step 7. The available coefficients are used to form the audio blocks by inverse IntMDCT. Finally, the watermarked audio blocks combine with the unselected audio blocks to form the watermarked audio signal.

6.4 Watermark Extraction

The extraction process should be implemented before reconstructing the IntMDCT coefficients because the information is embedded in the quantized coefficient. The incoming audio is first segmented into overlapped frames. The block diagram of extracting hidden watermark is the reverse processing of watermark embedding. The watermark detection procedure, the extraction rule and the detailed step are as follows:

Step 1. Perform segmentation of the watermarked signal. The frame is decomposed into 25 subbands.

Step 2. Find the first frame that contains watermark. According to the embedding algorithm, the first bit of watermark is embedded in the first non-zero frame. Therefore, the first non-zero frame should be the first detection candidate.

Step 3. Merge two neighboring frames to one frame. Apply Gaussian distribution analysis on all the macro frame for watermark detection. Watermark is extracted based on Gaussian distribution analysis function on watermarked frame.

Step 4. Carry out the inverse transformation on segmented watermarked signal by secret key to get the coefficients

Step 5. The same HAS model is applied on the data in frequency domain to determine the masking thresholds.

Step 6. The appropriate data are de-spread and de-interleaved in order to detect and recover any hidden data. Extracted watermark by quantization rule in the chosen frequency point. Calculate the cross-correlation coefficients ρ and τ and compare ρ and τ to get watermark $\hat{w}$

$$\rho = \sum_{i=1}^{N_w} w(i) * \hat{w}(i)) \Bigg/ \left[\sum_{i=1}^{N_w} | w(i) |^2 * \sum_{i=1}^{N_w} | \hat{w}(i) |^2 \right]^{1/2} \tag{11}$$

$$\tau = \sum_{i=1}^{N_w} \hat{w}(i) * \sqrt{N_w} \Bigg/ \left[\frac{1}{N_w} \sum_{i=1}^{N_w} (\hat{w}(i) - \hat{\mu}_{\hat{w}})^2 \right]^{1/2} \tag{12}$$

$$\begin{cases} \hat{w} = 1, & if\ \rho \geq \tau \\ \hat{w} = 0, & if\ \rho < \tau \end{cases} \tag{13}$$

6.5 Experimental Results and Performance Analysis

As to test files, in this experiment, the 19 well known SQAM files are selected. The files have a sampling frequency of 44.1 kHz and are 16 bit quantized. In our experiments, the files are reduced to a mono signal, For ease of expression hereafter, the host audio signals are marked with a number in ascending order, i.e. 1) Music: 1.Bach, 2.Pop, 3.Rock, 4. Jazz; 2) Percussive instruments: 5.Hihat, 6.Castanets, 7.Glockenspiel1, 8.Glockenspiel2; 3) Tonal instruments: 9. Harpsichord, 10.Violoncello, 11.Horn, 12.Pipes, 13.Trumpt, 14.Electronic tune; 4) Vocal: 15.Sopranor, 16.Bass, 17. Quartet; 5) Speech: 18.Female speech, 19.Male speech.

6.5.1Time Domain Waves and Spectrum

Tests were run to evaluate whether the watermarked data in an AVS audio can be detected through more qualitative methods, and the characteristics of the changes. This study only analyzed the sound waveforms in the time and spectrums in the frequency domains. The original AVS audio waveforms are compared with the watermarked audio containing the hidden data. The resulting waveforms in the time domain and spectrums in the frequency domain are shown in Fig.10 and 11 respectively. Both analyses do not show any distinguishing differences between the original audio and the watermarked audio. However, there seems to be a slight amount of signal loss and a slight increase in distortion due to the hidden data, but the overall distortion is acceptable. Besides, from the time waveform and spectrum of the watermarked and unwatermarked signal and their residue difference, it is obvious that the scheme is imperceptible and inaudible.

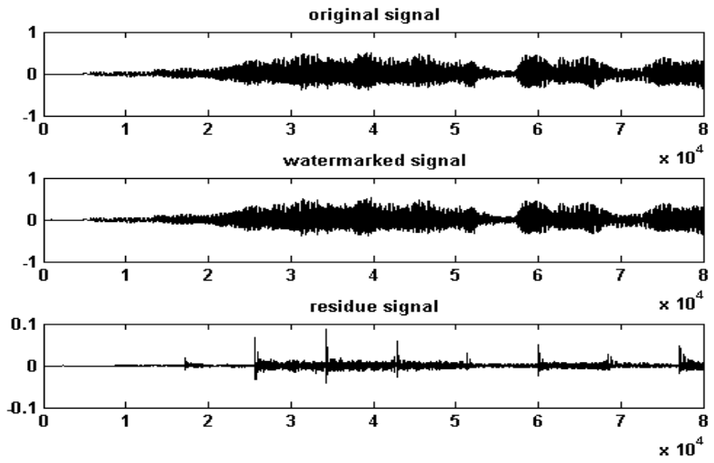


Fig. 10. Original, watermarked and residue audio waveform in time domain

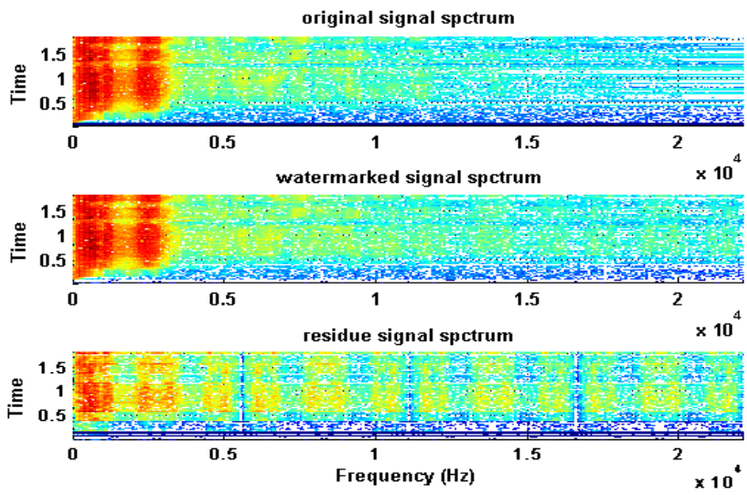


Fig.11. Spectrum of original, watermarked and residue signal

6.5.2 Robustness test

In order to illustrate the robustness nature of our watermarking scheme, in our experiment, common audio signal processing include re-quantization, re-sampling, additive noise, low-pass filtering, echo addition, equalization, mp3 compression, pitch shifting, time-scale modification and jittering are used to estimate the robustness of our scheme. Table 1 summarizes the watermark detection results against various common signal processing attacks. Watermark detection results after the attacks described above are shown in Table 1. It is evident that correlation values after attacks are very high which indicate the robustness of the audio watermarking scheme. The attacks are described in detail as follows:

- *Additive noise:* White noise with 10% of the power of the audio signal is added.
- *Amplitude variation*: The watermarked signal is attenuated up to 120% and down to120%.
- *AVS compression*: The coding/decoding is performed using a software implementation of the AVS Audio coder with several different bit rates (32kbps, 48 kbps, 64 kbps and 96 kbps).
- *Echo addition:* An echo signal with a delay of 50 ms and a decay of 50% is added to the original audio signal.
- *Expanding*: Expand the watermarked signal with increment of 3 dB and -3dB respectively.
- *Jittering*: Jittering is a random cropping performed evenly.
- *Low-pass filtering*: The low-pass filter with the 4 kHz cutoff frequency is applied to watermarked audio signal.
- *Pitch shifting:* Tempo-preserved pitch shifting is a difficult attack for audio watermarking algorithms as it causes frequency fluctuation. In this experiment, the pitch is shifted 1 degree higher and 1 degree lower.
- *Random cropping*: 10% samples are cropped at each of three randomly selected positions (front, middle and back).
- *Re-quantization*: A 16-bit watermarked audio signal is quantized 8-bit and back to 16-bit.
- *Re-sampling*: The original audio signal is sampled with a sampling rate of 44.1 kHz. Watermarked audio signal is down-sampled to 11.025 kHz, 22.05 kHz, and then up-sampled back to 44.1 kHz; up-sampled to 88.2 kHz, and then down-sampled back to 44.1 kHz.
- *Reverse amplitude*: Reverse the plus or minus of amplitude of samples.
- *Smoothness filtering*: Smoothly filter the watermarked audio signal.
- *Time-scale modification*: The watermarked audio signal is lengthened by 4% while preserving the pitch.

Attacks	Corr.value	Attacks	Corr.value
Without attack	1	Jittering	0.95
Additive noise	0.96	Low-pass filtering	0.93
Amplitude variation	0.86	Pitch shifting	0.92
AVS (96kbps)	0.96	Random cropping	0.82
AVS (48kbps)	0.93	Re-sampling 22.05kHz	0.83
AVS(64kbps)	0.89	Re-sampling 11.025kHz	0.79
AVS(32kbps)	0.87	Re-sampling 88.2kHz	0.73
Echo addition	0.83	Smoothness filtering	0.94
Expanding	0.88	Time-scale modification	0.86

Table 1. Robustness test results

6.5.3 StirMark Benchmark for Audio (SMBA) Test

SMBA is a standard and common robustness evaluation tool for examining audio watermarking technique(Lan et al.,2005). In this experiment, different attacks are applied to the watermarked and un-watermarked test set by using the Stirmark software. The parameters of the benchmark software were the default parameters included in the version of the tool available on the web. In addition, only the left channel has been marked in the experiments, thus stereo attacks do not apply here either. The attacks considered for this test are summarized in table 2. In this table, a total of thirty six different attacks are performed. From the results of the SMBA test, the high correlation values of after attacks denote that the watermarked can be extracted and this scheme is robust and secure.

Attacks	Corr. value	Attacks	Corr. value	Attacks	Corr. value
AddBrumm	0.98	Normalizer1	0.87	FFT Invert	0.88
AddNoise	0.94	Normalizer2	0.84	CopySample	0.49
AddSinus	0.76	Compressor	0.95	FlippSample	0.51
AddFFTNoise	0.58	BassBoost	0.92	CutSample	0.54
NoiseMax	0.77	RC-HighPass	0.93	ZeroCross	0.67
Denoise	0.92	RC-LowPass	0.90	ZeroLength1	0.71
LSBZero	0.85	FFT HLPass	0.71	ZeroLength2	0.69
Echo	0.89	Stat1	0.89	ZeroRemove	0.85
Exchange	0.54	Stat2	0.85	PitchScale	0.70
Resampling	0.52	FFTStat1	0.82	DynamicPitchScale	0.68
ExtraStereo	0.95	Smooth1	0.67	TimeStretch	0.62
VoiceRemove	0.57	Smooth2	0.76	DynamicTimeStretch	0.59

Table 2. Correlation values for SMBA test

6.5.4 Subjective Listening Tests

In order to evaluate the audibility of the watermarked audio, we have selected the “Double blind, triple stimulus, with hidden reference” test methodology according to ITU-BS.1116 listening test(ITU,1993). In this test, all clips and trials are randomized. This test method is

used for evaluating systems that cause only small degradations in audio quality. Ten listeners both experienced and familiar with the set of critical audio items participated in the test. The SQAM test items contain excerpts of single and multiple instruments, speech and complex sound sources. This test set was also extensively used for assessment of the subjective audio quality in the AVS audio development process. Fig. 12 presents the results from the ITU-BS.1116 listening test. It can be seen that the quality degradation of the bit stream watermarking system is very small for the vast majority of the test items. For all items the confidence intervals of both signals overlap which indicates that there is no significant distortion introduced by this scheme. The test results indicate that there is no statistical difference between the audio quality of the watermarked items and the original audio quality.

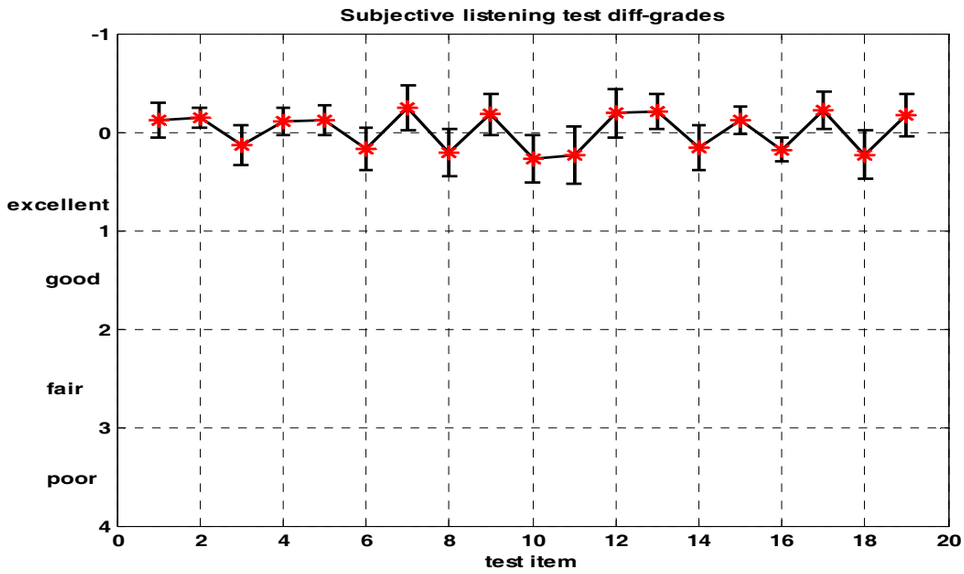


Fig. 12. Subjective listening test results of AVS audio watermarking scheme

From the above test results, we can come to some performance analyses:

1) Comparative robustness against common signal processing operations: the frequency domain selectable frequency range is chosen as the tradeoff of inaudible and robustness against normal operations. The embedded frequency points are all from this range in the AVS audio watermarking scheme. Consequently, this watermarking algorithm has the comparative robustness against common signal processing operations;
2) High robustness against malicious attack: the selectable frequency range brings the possibility to be hidden; and the chosen embedding point based on chaos secret key determines hidden and random embedding position. The two measures have guaranteed the privacy of the embedding position, which has higher robustness against malicious attack algorithm than fixed embedding positions. Therefore, the intruder without secret key must be difficult to implement malicious attack on the audio watermarking scheme.
3) High systematic security: chaotic systems have high security due to extreme sensitivity to initial value. In the proposed method, chaotic map and HAS model is used to choose

embedding position, the initial value of map is considered as systematic secret key. Therefore, the security of whole system only relies on secret key, which makes it have higher systematic security.

7. Conclusion

In this chapter, a literature review on audio watermarking techniques is first given. Different issues related to audio watermarking are described. A robust audio watermarking scheme with high robustness against malicious attack is then presented for AVS audio. Salient features are adopted to select embedding points for random and hidden embedding position, then the selected data are spread by chaotic sequence to increase robustness against watermark attacks during the embedding process. Besides, the masking characteristics of psychoacoustic model are used to avoid introducing audible noise in the AVS compression. Consequently, the proposed method has higher robustness against malicious attack. The performance analyses and simulation results show that the proposed scheme not only guarantees robustness against common operations, but also keeps the watermark imperceptible and inaudible.

References

AVS Audio Expert Group (2005). Information Technology —Advanced Audio Video Coding Standard Part 3: Audio, Audio Video Coding Standard Group of China (AVS).

Bassia, P., Pitas, I., & Nikolaidis, N. (2001). Robust audio watermarking in the time domain. *IEEE Transactions on Multimedia*, vol. 3, pp. 232-241.

Bender, W., Gruhl, D., Morimonto, N. & Lu, A. (1996). Techniques for data hiding, *IBM System Journal*, vol. 35, no. 3&4, pp. 313-336.

Chen, B., & Wornell, G. W. (2001). Quantization index modulation: a class of provably good methods for digital watermarking and information embedding. *IEEE Transactions on Information Theory*, vol. 47, pp. 1423-1443.

Chen, B., Zhao, J., & Wang, D. (2008). An Adaptive Watermarking Algorithm for MP3 Compressed Audio Signals, *Proceedings of IEEE Conference on Instrumentation and Measurement Technology*, pp. 1057-1060.

Chen, O. T. C., & Wen-Chih, W. (2008). Highly Robust, Secure, and Perceptual-Quality Echo Hiding Scheme. *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 16, pp. 629-638.

Chen, X.-S., Yang, Y.-T., Zhang, H., & Lu, X.-J. (2008). An audio blind watermark algorithm in wavelet domain based on chaotic encryption, *Proceedings of the International Conference on Wavelet Analysis and Pattern Recognition*, pp. 470-473.

Cheng, S., Yu, H., & Xiong, Z. (2002). Enhanced spread spectrum watermarking of MPEG-2 AAC audio, *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. IV/3728-IV/3731.

Chung, T.-Y., Hong, M.-S., Oh, Y.-N., Shin, D.-H., & Park, S.-H. (1998). Digital watermarking for copyright protection of MPEG2 compressed video, *IEEE Transactions on Consumer Electronics*, vol. 44, pp. 895-901.

Daniel Gruhl, Anthony Lu, & Bender, W. (1996). Echo hiding, *Lecture Notes in Computer Science*, vol. 1174, pp. 295-315.

Garcia, R. A. (1999). Digital watermarking of audio signals using a psychoacoustic auditory model and spread spectrum theory, *Audio Engineering Society*.

Giovanardi, A., Mazzini, G., & Tomassetti, M. (2003). Chaos based audio watermarking with MPEG psychoacoustic model I, *Proceedings of the 2003 Joint Conference of the Fourth Pacific Rim Conference on Multimedia & the Fourth International Conference on Information, Communications and Signal Processing*, vol. 1603, pp. 1609-1613.

Gruhl, D., Lu, A., & Bender, W. (1996). Echo hiding, *Information Hiding, ser. Spring Lecture Notes in Computer Science,* vol. 1174, pp. 295–315.

He, X. (2008). *Watermarking in Audio: Key Techniques and Technologies*, Cambria Press.

ITU-R. Recommendation BS.1116 (1993). Methods for the Subjective Assessment of Small Impairments in Audio Systems including Multichannel Sound Systems, *ITU Technical report*.

Kirovski, D., & Malvar, H. S. (2003). Spread-spectrum watermarking of audio signals, *IEEE Transactions on Signal Processin*g, vol.51, pp.1020-1033.

Lan, J., Huang, T., & Qu, J. (2007). A perception-based scalable encryption model for AVS audio, *Proceedings of the IEEE International Conference on Multimedia and Expo,* pp. 1778-1781.

Lang, A., Dittmann, J., Spring, R., & Vielhauer, C. (2005). Audio watermark attacks: from single to profile attacks. *Proceedings of the 7th workshop on Multimedia and Security*.

Li, Z., Sun, Q., & Lian, Y. (2006). Design and Analysis of a Scalable Watermarking Scheme for the Scalable Audio Coder, *IEEE Transactions on Signal Processing*, vol. 54, pp. 3064-3077.

Nedeljko, C., & Seppanen, T. (2008). *Digital Audio Watermarking Techniques and Technologies: Applications and Benchmarks*. Information Science Reference.

Neubauer, C., & Herre, J. (2000). Audio watermarking of MPEG-2 AAC bitstream, *Proceedings of 108th AES Convention*.

Petitcolas, F. A. P., Anderson, R. J., & Kuhn, M. G. (1999). Information hiding-a survey, *Proceedings of the IEEE,* vol. 87, pp. 1062-1078.

Podilchuk, C. I., & Delp, E. J. (2001). Digital watermarking: algorithm and application. *IEEE Signal Processing Magazine*, vol. 18, pp. 33-46.

Qiao, L., & Nahrstedt, K. (1999). Non-invertible watermarking methods for MPEG encoded audio, *Proceedings of the International Society for Optical Engineering,* pp. 194-202.

Quan, X., & Zhang, H. (2006). Data Hiding in MPEG Compressed Audio Using Wet Paper Codes, *Proceedings of 18th International Conference on Pattern Recognition,* pp. 727-730.

Robert, A., & Picard, J. (2005). On the use of masking models for image and audio watermarking, *IEEE Transactions on Multimedia*, vol. 7, pp. 727-739.

Seob, K.B., Nishimura, R., & Suzuki, Y. (2005). Time-spread echo method for digital audio watermarking, *IEEE Trans. Multimedia*, vol. 7, no. 2 pp. 212-221.

Seok, J., Hong, J., & Kim, J. (2002). A Novel Audio Watermarking Algorithm for Copyright Protection of DigitalAudio, *ETRI Journal*, vol.24.

Song, X., Su, Y., Hu, J., & Ji, Z. (2008). Information hiding in AVS compressed stream. *Proceedings of International Conference on Audio, Language and Image Processing*, pp. 988-992.

Swanson, M. D., Zhu, B., Tewfik, A. H., & Boney, L. (1998). Robust audio watermarking using perceptual masking, *Signal Processing*, vol. 66, pp. 337-355.

Tachibana, R., Shimizu, S., Kobayashi, S., & Nakamura, T. (2002). An audio watermarking method using a two-dimensional pseudo-random array, *Signal Processing*, vol. 82, pp. 1455-1469.

Tefas, A., Nikolaidis, A., Nikolaidis, N., Solachidis, V., Tsekeridou, S., & Pitas, I. (2003). Performance analysis of correlation-based watermarking schemes employing Markov chaotic sequences, *IEEE Transactions on Signal Processing*, vol. 51, 1979-1994.

Tsekeridou, S., Solachidis, V., Nikolaidis, N., Nikolaidis, A., Tefas, A., & Pitas, I. (2001). Bernoulli shift generated chaotic watermarks: Theoretic investigation, *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1989-1992.

Wang, C. T., Chen, T. S., & Chao, W. H. (2004). A new audio watermarking based on modified discrete cosine transform of MPEG/audio layer III, *Proceedings of IEEE International Conference on Networking, Sensing and Control*, pp. 984-989.

Wang, Y. (1998). A new watermarking method of digital audio content for copyright protection, *Proceedings of Fourth International Conference on Signal Processing*, pp. 1420-1423.

Wang, X.-Y., & Zhao, H. (2006). A novel synchronization invariant audio watermarking scheme based on DWT and DCT, *IEEE Transactions on Signal Processing*, vol. 54, pp. 4835-4840.

Xiang, S., Kim, H. J., & Huang, J. (2008). Audio watermarking robust against time-scale modification and MP3 compression, *Signal Processing*, vol. 88, pp. 2372-2387.

Zmudzinski, S., & Steinebach, M. (2009). Perception-based authentication watermarking for digital audio data, *Proceedings of the International Society for Optical Engineering*.

Multimedia

Edited by Kazuki Nishi

ISBN 978-953-7619-87-9

Hard cover, 452 pages

Publisher InTech

Published online 01, February, 2010

Published in print edition February, 2010

Multimedia technology will play a dominant role during the 21st century and beyond, continuously changing the world. It has been embedded in every electronic system: PC, TV, audio, mobile phone, internet application, medical electronics, traffic control, building management, financial trading, plant monitoring and other various man-machine interfaces. It improves the user satisfaction and the operational safety. It can be said that no electronic systems will be possible without multimedia technology. The aim of the book is to present the state-of-the-art research, development, and implementations of multimedia systems, technologies, and applications. All chapters represent contributions from the top researchers in this field and will serve as a valuable tool for professionals in this interdisciplinary field.

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen