PRINCIPLES OF AUDIO WATERMARKING

M. Hrnčár [1)], J. Krajčovič [2)]

[1)] Department of Control and Information Systems, Faculty of Electrical Engineering, University of Žilina, Univerzitná 1, 010 26 Žilina, Slovak Republic, e-mail: martin.hrncar@fel.uniza.sk,
[2)] Siemens Program and System Engineering s .r. o, Bytčická 2, 010 01 Žilina, tel.: +421 907 898 215, mail: jan.krajcovic@siemens.com

Summary: The article contains a brief overview of modern methods for embedding additional data in audio signals. It could have many reasons - for the purposes of access control or identification related to particular type of audio. This secret information is not “visible” for a user. This concept utilizes the imperfection of human auditory system. Simple data hiding into audio file has been proved in MATLAB.

1. INTRODUCTION AND A BRIEF HISTORY

The motivations for embedding hidden markers or messages in any kind of media content are essentially self-evident. Putting a ‘secret’ mark in an object can enhance security or enable more reliable identification. This process is commonly known as “watermarking”. One of the earliest documents describing the transmission of a hidden message comes from Herodotus, 440 BC. Later, it became widely used during WW II [1] At that time, steganographic technology consisted almost exclusively of invisible inks.
In the audio industry today there is much interest in copyright management and protection. Embedding some form of ‘hidden signal’ or watermark in the audio stream is seen as a potential method for managing the use of the material. However, there are many serious problems with this concept. Some of them ale listed here:
a) Audio quality - some people argue that any alteration of the original sound is unacceptable and that there are still a lot of improvements needed before it could be described as a perfect digital sound,
b) Robustness - to prevent unauthorized access, a watermarking system has to be resistant to significant degradation of the material, such as filtering or noise addition,
c) Reliability - any watermarking system used for access control would have to detect the watermark quickly and reliably, especially if it serves for commercial purposes.
The robustness of some watermarking algorithms is closely discussed in [3].
The rest of this paper is concerned with the technology of audio watermarking.

2. MODERN AUDIO WATERMARKING METHODS

The principles of commercial watermarking systems are secret and can only be deduced from the promotional material [1]. In some cases, the published results of tests on specific commercial systems have put into the public significant insights into the operating principles of those systems. Very brief summary is written here. The principles of watermarking can be divided into three categories, according to domains in which the transformation is performed.
Some of **time-domain methods**:

- Adding noise

The watermark may be added to the audio stream simply as a low-level signal. The most common way of implementing such a system is by using spread spectrum modulation; however, other modulation systems are possible,

- Adding echoes

It has been long established that the human auditory system is insensitive to ‘echoes’. This was firs established by Mr. Haas in early 1950´s [6]. On a short time-scale, around 20ms, a repetition of the signal is not perceptible with the amplitude somewhat lower than the original signal,

- Modifying phase

An audio is said to be phase insensitive, in that the Human Auditory System (HAS) is not aware of either absolute or relative phase of the signal. That is why phase encoding has been introduced for purpose of watermarking. This method is included under “time-domain methods”, but it requires the use of Fourier transformation for its implementation,

- Amplitude modulation – described later

Some of **frequency-domain methods**:

- Adding modulated carriers

At least one modern watermarking system uses the principle of adding modulated carrier frequencies to an audio signal. To be successful, such a system needs to have a system for calculating and making use of psycho-acoustic audibility thresholds. On the other hand, such a system will not be probably very secure,

- Adding noise in a transform domain
- Subtracting frequency bands (adding notches)
- The combination of last two methods.

The example of “**coded” domain methods** is:

- Adding noise-like signals to coding coefficients.

The watermarked data can simply be added to the coding coefficients as in any other additive system working in a transform domain. This method has also an alternative in modifying of coding coefficients by biased error distribution. Both are used in MPEG, AC3, etc.

In all of mentioned methods, account may be taken of the HAS sensitivities (temporal of frequency masking techniques have to considered) to embed the watermark signal in the balance between added watermark energy and perceptibility. Any practicable watermarking system represents a workable compromise between these requirements.

2.1 Spread-spectrum modulation method

This is one of well-known watermarking techniques. The method is based on spreading the watermarked signal over the entire audible frequency spectrum so that it approximates white noise, at power level as to be inaudible (below the ambient noise) [4]. A pseudorandom sequence called chip is used to modulate a carrier wave which creates the signal spread watermark code, as shown in Figure 1.

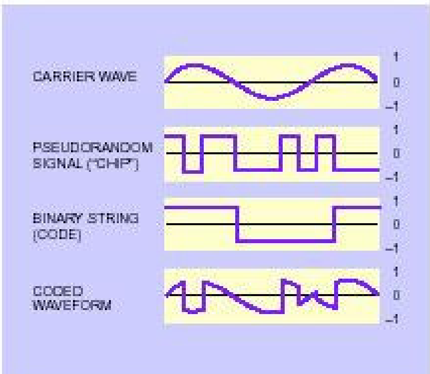

Fig. 1 Spread spectrum

The message bit rate is always less then chip rate.
Generally, the watermark extractor may have two purposes: deciding weather the tested signal contains a watermark and extracting a message that watermark may carry. The data can be recovered by appropriate filtering of the combined signal. In this case, cross-correlation function with the known pseudo-random pattern is usually used.
This modulation was created in Germany many years ago for secure military communications, but the principles of it are used today in many audio or video watermarking systems.

2.2 Amplitude modulation method

In principle, the time-domain envelope of the signal may be modified to carry watermark data. The envelope may be modulated on a longer time-scale. Although no audio system is definitely known to work in this way, a commercial video system intended for the cinema may work by introducing an artificial 'flicker' or frame-rate variation in the overall brightness. It is possible to prove, that an audio system (with necessarily a very low watermark data rate) could work in a similar fashion.

The principle of this method is simple. It was slightly modified for our purpose of simple data-hiding.

3. PROOF OF CONCEPT OF SIMPLE DATA-HIDING IN AN AUDIO SIGNAL

The process of embedding additional data in audio signals could be divided into several steps. This section details the implementation of simple digital watermarking scheme, which is shown in Figure 2. Although it serves for educational purposes, it provides an overview of some common techniques for audio watermarking.

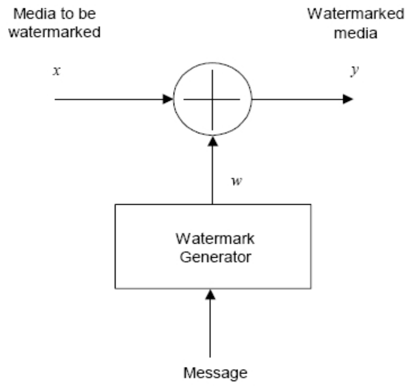

Fig. 2 Watermark embedding scheme

The concept of simple watermarking application was developed and proved in MATLAB. The watermarking hides the data in lower frequency components of the audio signal, which are below the perceptual threshold of human auditory system.
First take a look at the structure of one "word" in watermark message, which is shown in Figure 3. For easier detection, the length of one "word" is set to 20 signs. The header has one sign and the checksum at the end has two signs. The checksum give us some information whether message bits have changed during transmission. The useful information has up to seventeen signs.

Header (1 sign)	Information (17 signs)	Checksum (2 signs)

Fig. 3 Structure of watermark "word"

Every sign is encoded in 5-bits binary alphabet. Every bit is equal to one frame, which is 100 milliseconds in duration. This frame size has been chosen so that the embedded watermark does not introduce any audible distortion into file.
With a 100 milliseconds frame size, our bit rate for watermark message is 1 / 0,100 = 10 bits per second. Our "word" then lasts 20 · 5 · 0,1 = 10 seconds. The

amplitude of watermark signal is lower than the original signal amplitude.
In the next step, watermark message is attached to an audio file and sent to the output. The application in MATLAB shown in Figure 4 allows us to perform this without any audible difference to the original file. This application also allows successful decoding of watermark message.

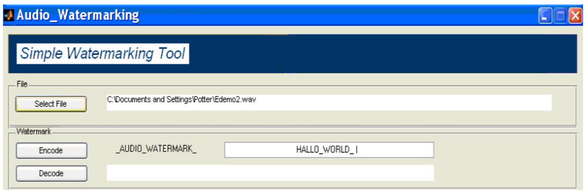

Fig. 4 Application in MATLAB

The difference between an original file and file containing one watermarked word, both in time domain, is shown in Figure 5. As you can see, the time-domain envelope of a signal has changed.

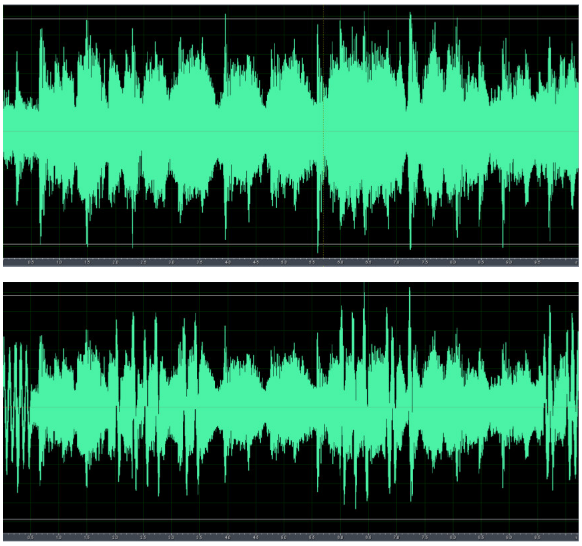

Fig. 5 Comparison between original (top) and watermarked (bottom) audio files in time domain

The process of extracting the digital watermark from an audio file is similar to one for inserting. The computer requirements were slightly higher because it uses correlation detection technique.

4. CONCLUSION

Although the principles of audio watermarking could be proved easily, it becomes particularly difficult, when the file is a lossy compression format such as MP3 and we would like to meet the robustness and reliability requirements. In case of broadcasting of some watermarked audio, we have to deal not only with reliability and robustness of such commercial system, but also with other problems, like modulation or synchronization.
Digital audio watermark technology is an active research area in industry and at the post-graduate level because there is a great need of managing an audio content.

Acknowledgement

This paper is supported by cultural and educational grant agency within project KEGA n. K-057-06-00: Innovation of methodology of laboratory education, based on modeling and simulation in MATLAB, with educational models through e-learning.

REFERENCES

[1] Walker, R.: *Audio watermarking, R&D White Paper WHP 057*, British Broadcasting Corporation, 2004.
[2] Psutka, J.: *Komunikace s počítačem mluvenou řečí*, Academia Praha 1995.
[3] Franeková, M., Janota, A., Kapusta, P.: *Odolnosť vodoznakov vložených na báze DCT transformácie*, Slaboproudý obzor, december 2002, ročník 59, číslo 4, ISSN 0037-668X.
[4] http://www.ece.uvic.ca/~aupward/w/watermarking.htm
[5] Seok, J., Hong, J., Kim, J.: *A Novel Audio Watermarking Algorithm for Copyright Protection of Digital Audio*, ETRI Journal, Vol. 24, June 2002.
[6] Haas, H.: *Über den Einfluss eines Einfachechos auf die Hörsamkeit von Sprache*, PhD Thesis, University of Gottingen, Gottingen, Germany, 1951. Also published in Acustica 1, 49-58 (1951).