Audio Watermarking Based On The PSK Modulation

Wahid Barkouti
Physics Department
Faculty of Science, Manar University
Tunis, Tunisia
Wahid.Barkouti@fst.rnu.tn

Sihem Nasri
Physics Department
Faculty of Science, Manar University
Tunis, Tunisia
nasri_sihem@live.fr

Adnane Cherif
Physics Department
Faculty of Science, Manar University
Tunis, Tunisia
Adnane.cher@fst.rnu.tn

***Abstract*—Audio watermarking is a technique, which can be used to embed information into the digital representation of audio signals. The main challenge is to hide data representing some information without compromising the quality of the watermarked track and at the same time ensure that the embedded watermark is robust against removal attacks. Especially providing perfect audio quality combined with high robustness against a wide variety of attacks is not adequately addressed and evaluated in current watermarking systems. In this paper, we present a new phase modulation audio watermarking technique, which among other features provides evidence for high audio quality. PSK modulation has been proposed as an effective approach to watermarking.**

Keywords-PSK modulation,audio watermarking,copyright protection,data embeding.

I. Introduction

Data embedding in multimedia is a recent subject of research, which developed considerably in last years. The fast improvement of the Internet and the digital information revolution caused major changes in the overall culture [1]. Data embedding (also known as *data hiding*) is a multi-disciplinary area that combines signal processing with cryptography, communication theory, coding theory, information theory and the theory of human auditory and visual systems. The activity in the area is directed to both theoretical and practical aspects. Digital watermarking and information embedding systems have a number of important multimedia applications[2].These systems embed one signal, sometimes called an embedded signal or watermark, within another signal, called a “host signal”. The embedding must be done such that the embedded signal causes no serious degradation to its host.

Digital watermarking has become an active and important area of research, and development and commercialization of watermarking techniques is being deemed essential to help address some of the challenges faced by the rapid proliferation of digital content.

II. WATERMARKING ISSUES

A. *Perceptuel transparency*

The main requirement for watermarking is perceptual transparency. The embedding process should not introduce any perceptible artifacts, that is, the watermark should not affect the quality of the original signal. However, for robustness, the keeping perceptual artifacts as low as possible. Thus, there must be a trade-off between perceptual transparency and robustness. This problem can be solved by applying human perceptual modeling in the watermark embedding process [3]. The most important requirement is that the quality of the original signal has to be retained after the introduction of watermark. A watermark cannot be detected by listeners.

B. *Security*

The watermark must be strongly resistant to unauthorized detection. It is also desirable that watermarks be difficult for an unauthorized agent to forge. Watermark security addresses the secrecy of the embedded information. Where secrecy is necessary, a secret key has to be used for the embedding and extraction process.

For the optimal watermarking application, a trade-off has to be accepted between the above-mentioned criteria. Robustness means, for example, that many information of a watermark must be embedded that are, however, then in case of an attack, more visible or detectable [4]. Capacity refers to the amount of information that can be stored in a data source.

C. *Robustness*

Robustness means the resistance ability of the watermark information changes and modifications made to the original file. As modifications, resizing, file compression, rotation, and common operations will be particularly considered. Some data-embedding applications may take place in an error-free or lossless environment. For example, the embedded data may provide digital object identifiers for use in clean signals residing in a controlled data base. In these situations, robustness to signal degradations is not important. In many cases, however, lossy signal-processing operations may be present in the system [5].

These characteristics are partially exclusionary which means that other areas can be emphasized while deteriorating others. Trade-offs must be accepted for optimal performance. For example, a robust watermark cannot achieve both high capacity and imperceptibility [6].

III. AUDIO WATERMARKING

Digital audio watermarking involves the hiding of data within a digital audio file. Applications for this technology are numerous. Intellectual property protection is currently the main driving force behind research in this area. To combat online music piracy, a digital watermark could be added to all recording prior to release, signifying not only the author of the work, but the user who has purchased a legitimate copy [7].

A. *Principle*

There are three main stages in the watermarking process [4]:

- generation and embedding
- attacks
- detection

Generation of watermarks is an important stage of the process. Watermarks contain information that must be unique otherwise the owner cannot be uniquely identified. In embedding, an algorithm accepts the host and the data to be embedded and produces a watermarked signal. Various algorithms have been developed so far [1].

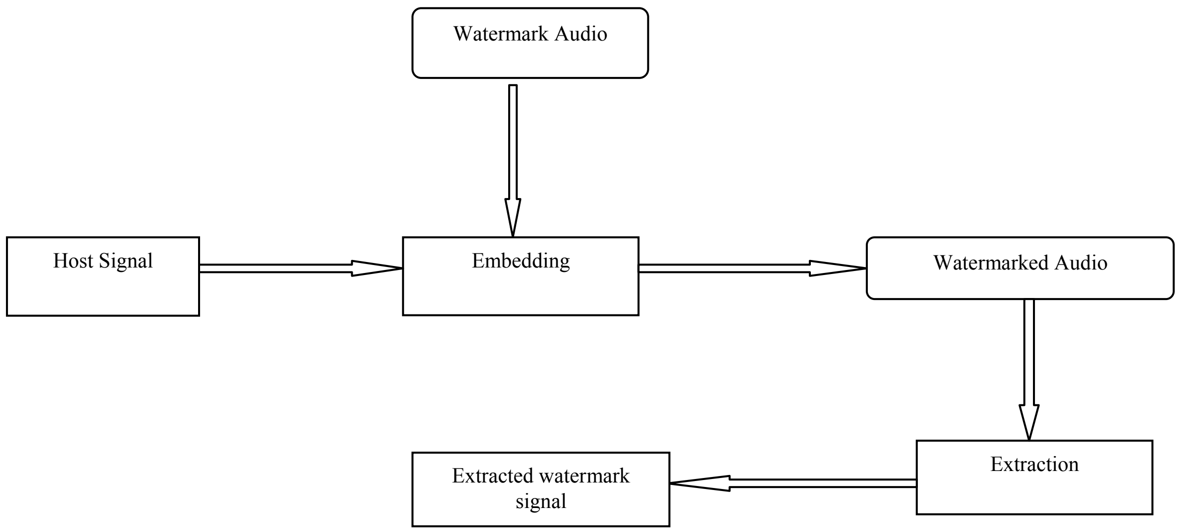

Figure 1. Block diagram of watermarking

The watermarked signal is then transmitted or stored, usually transmitted to another person. If this person makes a modification, this is called an attack. There are many possible attacks. Detection is an algorithm which is applied to the attacked signal to attempt to extract the watermark from it. If the signal is not modified during transmission, then the watermark is still present and it can be extracted. If the signal is copied, then the information is also carried in the copy. The embedding takes place by manipulating the contents of the digital data, which means the information is not embedded in the frame around the data, it is carried with the signal itself.

B. Human auditory system

Audio watermarking is quite challenging than image watermarking due to the dynamic supremacy of human auditory system (HAS) over human visual system (HVS) [8].

While the HAS has a large dynamic range, it often has a fairly small differential range. Consequently, loud sounds tend to mask out quieter sounds. Additionally, while the HAS has very low sensitivity to the amplitude and relative phase of a sound, it is difficult to perceive absolute phase. Finally there are some environmental distortions so common as to be ignored by the listener in most cases. These characteristics can be considered as positive factors to design watermark embedding and extracting schemes. HAS is insensitive to a constant relative phase. Shift in a stationary audio signal and some spectral distortions interpret as natural, perceptually non-annoying ones. [9].

C. Properties of the human auditory system

Two properties of the HAS dominantly used in watermarking algorithms are frequency (simultaneous) masking and temporal masking [9].

- **Frequency masking:** Frequency masking is a frequency domain phenomenon where low levels signal can be made inaudible by a simultaneously appearing stronger signal. A masking threshold can be found and is the level below which the audio signal is not audible. Thus, frequency domain is a good region to check for the possible areas that have imperceptibility [8].
- **Temporal masking:** In addition to frequency masking, two phenomena of the HAS in the time domain also play an important role in human auditory perception. Those are pre-masking and post- masking in time. The temporal masking effect appears before and after a masking signal has been switched on and off, respectively. The duration of the pre-masking is significantly less than one-tenth that of the post-

masking, which is in the interval of 50 to 200 milliseconds. Both pre- and post-masking have been exploited in the MPEG audio compression algorithm and several audio watermarking methods [9].

D. Problem and attaks on audio signals

The important requirements of an efficient watermarking technique are the robustness and inaudibility. There is a tradeoff between these two requirements; however, by testing the algorithm with the signal processing attacks that gap can be made minimal. Every application has its specific requirements, and provides an option to choose high robustness compensating with the quality of the signal and vice-versa. Without any transformations and attacks every watermarking technique performs efficiently [8].

IV. WATERMARKING INTO THE PHASE OF THE HOST SIGNAL

The important requirements of an efficient watermarking technique are the robustness and inaudibility. There is a tradeoff between these two requirements; however, by testing the algorithm with the signal processing attacks that gap can be made minimal. Every application has its specific requirements, and provides an option to choose high robustness compensating with the quality of the signal and vice-versa. Without any transformations and attacks every watermarking technique performs efficiently [8].

A. Phase coding

The basic **phase coding** method was presented in [9]. The basic idea is to split the original audio stream into blocks and embed the whole watermark data sequence into the phase spectrum of the first block. One drawback of the phase coding method is a considerably low payload because only the first block is used for watermark embedding. In addition, the watermark is not dispersed over the entire data set available, but is implicitly localized and can thus be removed easily by the cropping attack. It is a non-blind watermarking method (as the phase modulation algorithm) that limits the number of applications it is suitable for.

Approaches that embed the watermark into the phase of the original signal exploit the fact that the human auditory system has a low sensibility against relative phase changes, as stated in techniques. This method was presented in [8] and proposes to split the original audio stream into blocks and embed the whole watermark into the phase spectrum of the first block, as described in techniques.

One disadvantage of the phase coding approach is the low payload that can be achieved. Only the first block is used in embedding the watermark. Since the watermark is not distributed over the entire data set, but is implicitly localized, it can be removed easily if cropping is acceptable.

B. Phase modulation

Another form of embedding the watermark into the phase is by performing independent multiband phase modulation, [techniques] states. Inaudible phase modifications are exploited in this algorithm by controlled multiband phase alterations of the original signal.

Both phase embedding approaches use the psychoacoustic features of the human auditory system with regard to the just noticeable phase changes. They exploit the inaudibility of phase changes if the time envelope of the original signal is approximately preserved. Because of the phase alteration, embedding and detection of the watermark is done in the Fourier domain by processing the audio stream block wise. While the phase coding method is embedding the watermark in the phases of the first block, the phase modulation algorithm performs a long-term multiband phase modulation. Both algorithms are non-blind watermarking methods, since they require the original signal during the watermark retrieval, which of course limits their applicability. The watermark insertion in the phase modulation method is performed using an independent multiband phase modulation. Imperceptible phase modifications are exploited in this approach by the controlled phase alternation of the host audio [9].

V. PSYCHOACOUSTIC AUDITORY MODEL

An auditory model is an algorithm that tries to imitate the human hearing mechanism. It uses knowledge from several areas such biophysics and psychoacoustics. From the many phenomena that occur in the hearing process, the one that is the most important for this model is “simultaneous frequency masking.” The auditory model processes the audio information to produce information about the final masking threshold [10].

The first step is to translate the actual audio frame signal into the frequency domain using the Fast Fourier Transform. In the frequency domain the power spectrum, energy per critical band and the spread energy per critical band are calculated to estimate the masking threshold. This masking threshold is used to shape the “noise or watermark” signal to be imperceptible. Finally frequency domain output is translated into the time domain and the next frame is processed.

A. *Simultaneous masking*

Simultaneous masking of sound occurs when two sounds are played at the same time and one of them is masked or “hidden” because of the other. The formal definition says that masking occurs when a test tone or “maskee” is barely audible in the presence of a second tone or “masker.” The difference in sound pressure level between the masker and maskee is called the “masking level.

B. *Power spectra*

The first step in the frequency domain (linear, logarithmic or bark scales) is to calculate the power spectra of the incoming signal.

$$Sp(k) = |Sw(k)|^2 \qquad .1$$

The energy per critical band, *Spz* (*z*), is defined as:

$$Spz(z) = \sum_{k=k_{\min}(z)}^{k_{\max}(z)} Sp(k)$$

With z = 1, 2, 3… Z_t

The frequency indication k represents a frequency in linear scale (ladder). To calculate the threshold of masking, is needed a scale (ladder) in Barks. We note by kmin (z) and kmax (z) the frequency indications representing the lower and superior limits of the critical band (strip) z. The power spectrum *Sp (k)* and the energy per critical band *Spz (z)* are the base of the analysis in the frequency domain.

VI. WATERMARK EMBEDING PROCESS

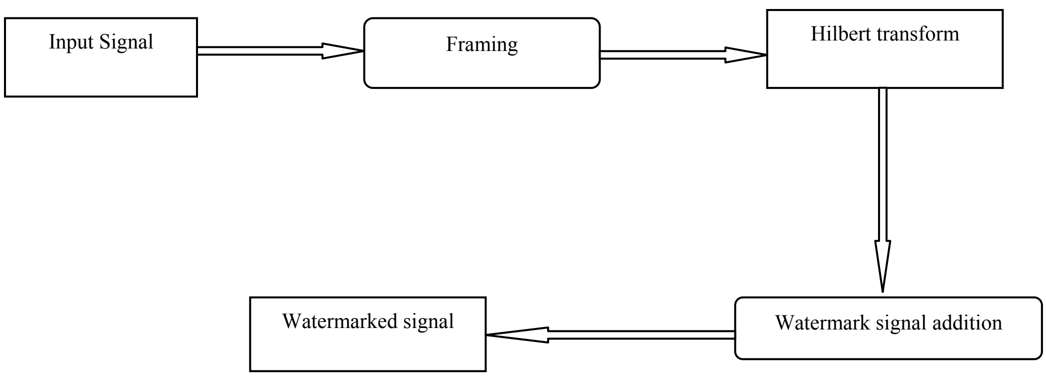

Figure 2. Watermark embeding

A. *fraimng*

The purpose of the segmentation of the signal lives (lies) in the fact of facilitating the insertion of the tattoo while respecting the constraint of inaudibilité hidden information. So we worked on wefts, each measures up (is considerable) 512. The size (cutting) 512 corresponds to a duration of 11,5 ms of every block.

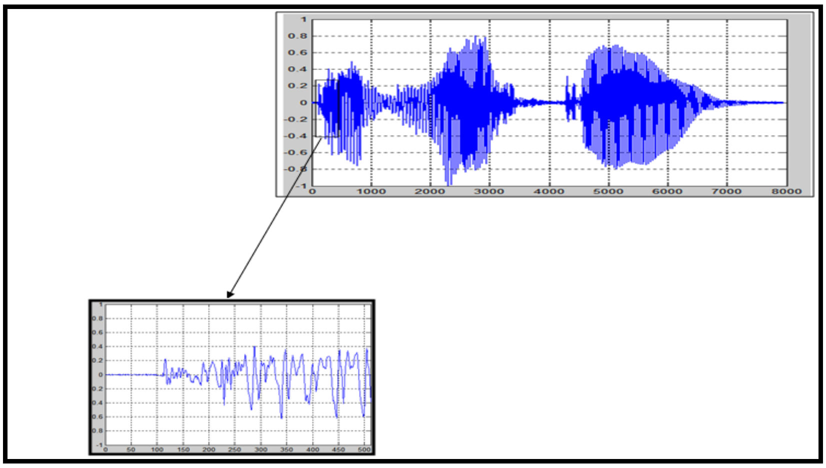

Figure 3. Fraiming of the audio file

B. Window of hamming

It is about a function applied to all the samples taken in the window before the calculation of the transformed of Fourier, so as to decrease side effects due to a rectangular window.

The use of a window of Hamming is preferable in that of the door. Indeed, if the truncation of the signal is realized in a not rectangular way but by making windows in the signal, the transitions in the signal will then be softer. The door has strong discontinuities, what pulls (entails) an important spectral flight (leak) that is we observe a flight (leak) of the power on frequencies lines. The more the chosen window will have a big temporal duration, the more it will be narrow in the frequency domain.

C. Spectral analysis

Subsequent to the framing of the unprocessed audio signal, spectral analysis is performed on the host audio signal, consisting of a fast Fourier transform (FFT), which allows us to calculate the low frequency components of each frame, as well as the overall frame power.

D. Hilbert transform

The Hilbert transform was originally defined for periodic function, or equivalently for functions on the circle, in which case it is given by convolution with the Hilbert kernel. More commonly, however, the Hilbert transform refers to a convolution with the Cauchy kernel, for functions defined on the real line. The Hilbert transform is closely related to the Paley wiener theorem, another result relating holomorphic functions in the upper half-plane and Fourier transform of functions on the real line.

It is the transformed which has a signal x (t) we associate the signal y (t) such as:

$$Y(f) = TH(f)X(f)$$

$$H(f) = \frac{Y(f)}{X(f)} = -j\,sign(f)$$

The transformed of Hilbert indicates the imaginary part of an analytical signal, the real part of which is the physical signal. The transfer function of the filter is:

$$HH(j2\pi f) = \Im\left(\frac{1}{\pi t}\right) = -j\,sign(f)$$

E. Embedding

The process of embedding a watermark into an audio file is divided into four main processes. An original audio file in wave format is fed into the system, where it is subsequently framed, analyzed, and processed, to attach the inaudible watermark to the output signal. Before being inserted, the mark must be coded then modulated to produce analogical information.

- **Coding:** the message containing the mark (brand) a digital information that is continuation (suite) of 0 and 1 consists in producing..
- **Digital phase modulation**, or phase shift keying (PSK), is very similar to frequency modulation. It involves changing the phase of the transmitted waveform instead of the frequency, and these finite phase changes represent digital data. In its simplest form, a phase-modulated waveform can be generated by using the digital data to switch between two signals of equal frequency but opposing phase. If the resultant waveform is multiplied by a sine wave of equal frequency, two components are generated: one cosine waveform of double the received frequency and one frequency-independent term whose amplitude is proportional to the cosine of the phase shift. Thus, filtering out the higher-frequency term yields the original modulating data prior to transmission.

  The modulated signal is:

$$m(t) = A\cos(\omega_0 t + \phi k)$$

$$\phi_k = \phi_0 + (2k+1)\frac{\pi}{M} \quad avec \quad 0 \le k \le M$$

$$M = 2^n$$

$$m(t) = A.\cos(\omega_0 t).\cos(\phi_k) - A.\sin(\omega_0 t).\sin(\phi_k)$$

F. Experimental results

The techniques of PSK modulation is the most adapted technique to the watermarking system. The PSK modulation permits to improve the performances of the system for transmission.

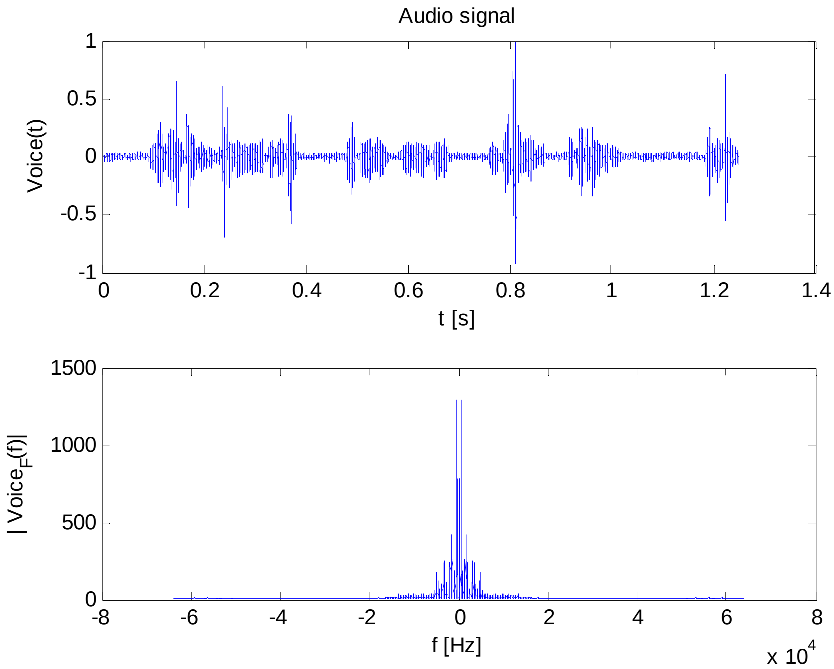

Figure 4. Original signal

Phase modulation, when it can be used, is one of the most effective modulation methods in terms of the signal-to perceived noise ratio. The modulated audio signal is given by the following figure:

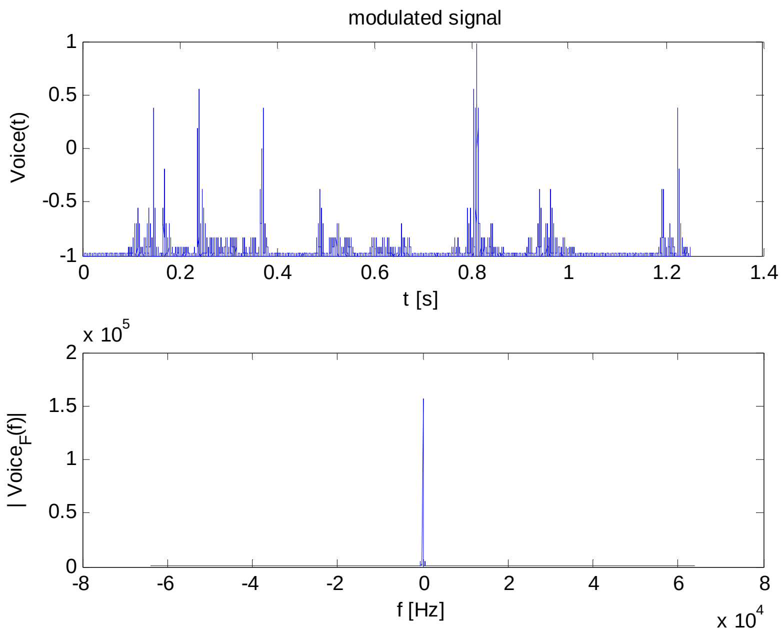

Figure 5. Modulated signal

Audio watermark is special signal embedded into digital audio. The signal is extracted by detection mechanisms and demodulated. The demodulated audio signal is given by the following figure:

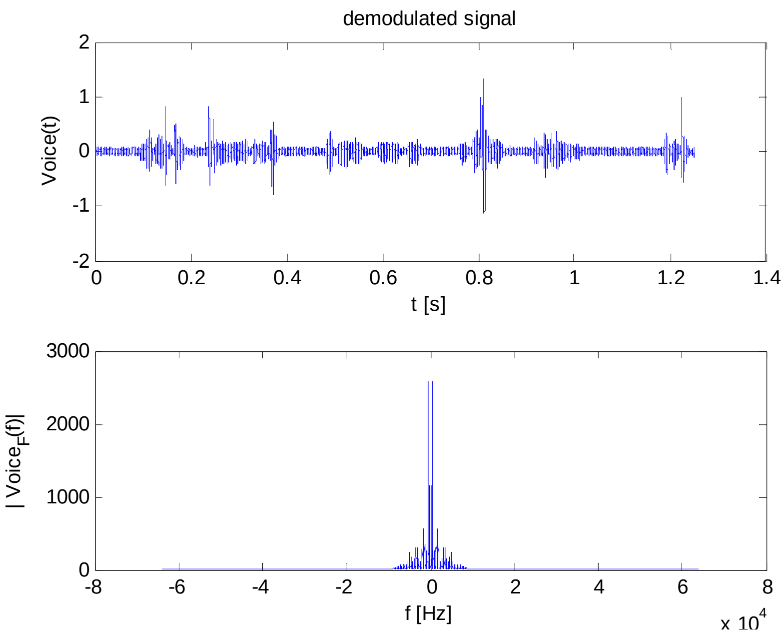

Figure 6. Demodulated signal

VII. WATERMAK EXTRACTION

The process of extracting the digital watermark from the audio file is similar to the technique for inserting the watermark. The host signal and the watermarked signal are given by the following figure:

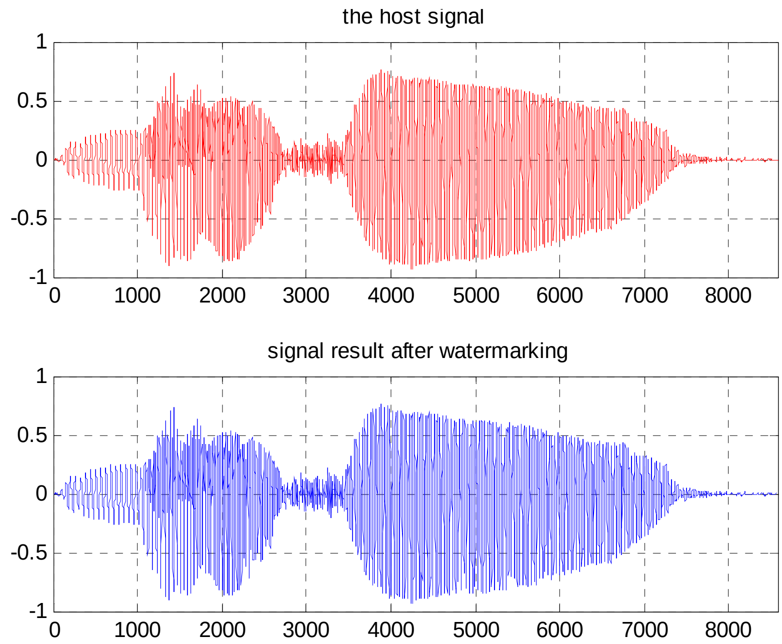

Figure 7. Host signal and the watermarked signal

We notice that there is no difference between the original signal and the watermarked signal. The resultant file undergoes no degradation in the quality (perception identical to the original) .The shape of wave is identical to the original shape.

VIII. CONCLUSION

The audio watermarking is a domain of very vast research. It can be exploited in various applications such is the case of our work. We used the audio watermarking as a tool of protection of copyright. The implementation of watermarking system means choosing a technique of adequate insertion, it is necessary to test the robustness of the system against the attacks managed at the level of the channel. In this paper, we used the technique of modulation of phase (PSK) which in fact does not exploit the models psychoacoustiques because the human ear is insensible in the variation of the phase of the sound signals. Throughout this work, we proved the robustness of our system so its performances.

ACKNOWLEDGMENT

We would like to acknowledge all the persons who have contributed to a great extent towards the initialization, development and success of this paper. With great respect and obedience,We would like to thank our parents who were the backbone behind our deeds.

REFERENCES

[1] P.Dutta, D. Bhattacharyya, and T. Kim, " Data hiding in audio signal " International Journal of Data base Theory and Application Vol. 2, No. 2, 2009.
[2] B. Chen and G.W. Wornell," Quantization index modulation methods for digital watermarking and information embedding of multimedia " Journal of VLSI Signal Processing 27,7-33,2001.
[3] J. Seok, J. Hong, and J.Kim, "A novel audio watermarking algorithm for copyright protection of digital audio," ETRI Journal,Volume 24, Number 3, 2002.

[4] J. Seitez, “Digital watermarking for digital media,”Information science publishing,2005 .
[5] M. D. Swanson, M. Kobayashi, andA.H. Tewfik “ Multimedia data-embedding and watermarking technologies ,” Proceeding of the IEEE, VOL. 86, NO. 6, 1998.
[6] M. Brockman, “Audio protection with removable watermarking for mobile distribution,”Masterr’s thesis oulu ,2009.
[7] M.Adya, “Audio watermarking resistant to mp3 compression ,” Masterr’s thesis ,India.
[8] R. Ravula “Audio watermarking using transformation techniques” Master’s Thesis, University of Louisiana State, 2010.
[9] C.Nedeljko, “Algorithms for audio watermarking and steganography,”, University of Oulu. ISBN 951-42-7383-4. 2004.
[10] R. A. Garcia, “Digital watermarking of audio signals using a psychoacoustic auditory model and spread spectrum Theory,”. The 107th Convention of Audio Engineering Society (AES), New York,USA, 1999.