DIGITAL WATERMARKING BY DILATED AUDIO SAMPLES

***J UAN ANTONIO RIOS CHAVEZ, **CARLOS AQUINO RUIZ,**

***** CELEDONIO ENRIQUE AGUILAR MEZA**

*Master of Science in Engineering in Microelectronics

**Engineer in Communications and Electronics

*** Engineer in Communications and Electronics

E-mail: [1]jriosc0500@ipn.mx [2]caquino@ipn.mx [3]caguilarm@ipn.mx

ABSTRACT

It is currently experiencing a growth in multimedia files, piracy, that is the reason why it is necessary to protect these files that takes a major boom in society. To solve this problem, algorithms have developed watermarking which consists in hiding data (binary sequence known as a watermark) within the multimedia file to be protected and which must meet certain requirements. In this article we present a detailed method that consists of marking sound files dilating the original samples depending on the binary series number that you want to hide and the evaluation of this methodology.

Keywords: *Audio, Dilation, Watermark, Samples, Protection.*

1. INTRODUCTION

The digitization of multimedia has become a daily activity due to handling, transmission and reception of these files. These activities cause a problem, distribution and copying files without permission request from the original author. As shown in the graph (see Figure 1), musical piracy has grown in the past year significantly in Latin America (illegal downloads, illegal copies, distribution).

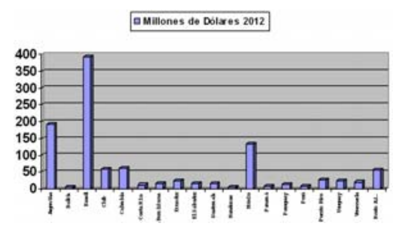

Fig 1. Million losses due to piracy in Latin America.

The watermarks have been proposed as a solution to protect or frame these files. This technique is to embed data within the original file in order to mark, label or protect the file and get better control of legal copies, frame the property or simply avoid plagiarism. We call multimedia files to audio, images and video. For the last two there are several algorithms that perform the action of watermarking but relatively few people are oriented to audio due, to the complexity of the human ear. The watermark embedded in the audio should not alter the audio quality of it, and must support various digital processing that will also do the same natural processes to which the signal undergoes (noise added by the channel of transmission, filtering, re-sampling).

The watermark is a binary string and the marking process is sectioned into audio blocks that you wish to dial, and delay a number of samples to embed a '1 'binary and dilate other different numbers of samples prior to a '0' bit. The organization of the article is as follows: Section II shows the development of the theoretical algorithm, divided into insertion process, marking and evaluation of file detection or the extraction process. Section III presents the implementation of the method, the initial parameters necessary and a thorough comparison of the original audio with the audio to demonstrate the visual and audio quality which remains almost exactly identical. The conclusions are in Section IV and finally Section V, shows the references consulted for the project.

II DEVELOPMENT

As it is well known, an algorithm of digital watermarking has three parts: the process of embedding (encoder), the intermediate process (simulated attacks and / or unexpected processes) and the extraction process and / or generation of hidden message (decoder) (see Fig.2)

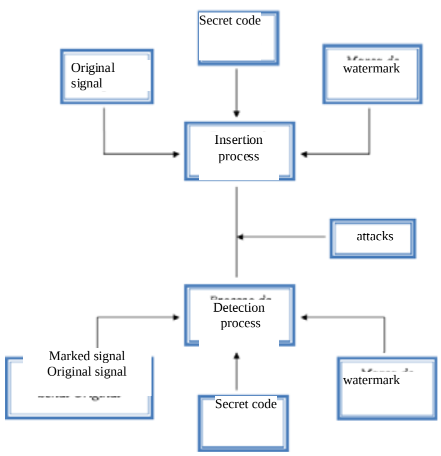

Fig 2. Generic diagram of digital watermarking.

The proposed method is based on the traditional method of using echo data hiding, that is, a delay in the audio samples. This paper aims to embed a '1 'or '0' bit in the original audio by a dilation of the samples of the same, delaying a number of them being careful not affect the audio quality of the same. For this it is necessary to use the theory of post-masking, which is not to exceed a limit of samples in the delay so that it can not be perceptible to the human ear. This delay should not be greater than 50 ms mark so as the post-masking theory which states that sample l delayed less than or equal to 50 ms is approvable for preventing hearing there of by humans.

For the insertion process consider the following: first the original audio signal is sampled at 44.1 kHz and according to post-masking, a delay of no more than 441 samples could be used to encrypt the watermarking bits. The insertion methodology is summarized in the following steps:

a) The audio is divided into blocks of size N. This size is determined by dividing the total number of audio samples between the number of bits of predetermined watermark, since each bit will be inserted into each audio block.

b) Two vectors are generated: one is delayed to 441 samples of the original version and the other is delayed 220 samples of the original version, the first denotes the inclusion of a '1 'and the second denotes the insertion of a '0'. mathematically:

$$y(n) = \begin{cases} x(n) + 0.1 * x(n - d_1) \\ x(n) + 0.1 * x(n - d_2) \end{cases} \quad (1)$$

where;

and (n) = marked block
x (n) = original block
x (n-d), x (n-d) = delayed block

For the extraction process it requires to take the following considerations: the detection system and / or extraction of the watermark is blind, that is, no need to know the most of the insertion parameters in order to perform optimum detection of the watermark and so to recover the hidden data.

To recover the hidden message the cepstrum theory was used, which traces its origins to detect abnormal vibrations around the close areas of volcanoes and so prevent an eruption or earthquake; in our case, we used the cepstrum to detect if there are delays in each block added audio if any, defining the magnitude of these to see if what was hidden was a bit '0 'or bit '1'. All this without knowing the original watermark that was embedded. The following equation shows the detection by cepstrum:

$$C(x_n) = F^{-1}(\log|F(x(n))|^2) \quad (2)$$

To evaluate the inaudibility existing in the algorithm, one hundred music files were marked and divided into five different categories which were tested by objective and subjective tests. Objective tests were carried out through established standards, for this case bases established by the ITU (International Telecommunication Union) such as ODG (Objective Difference Grade) and PEAQ (Perceptual Audio Quality Evaluation). These standards are based on the imperfection of the human auditory system. Below are the standard tables:

Deterioration description	ODG
Imperceptible	0.0
Perceptible, but not disturbing	-1.0
Lightly disturbing	-2.0
disturbing	-3.0
Very disturbing	-4.0

Table 1. ODG Range Values

Subjective tests are those that are made by way of a vote on a scale of degrees as a MOS proof (Mean Opinion Score).

To measure the robustness of algorithm simulation it was made by adding white Gaussian noise, such noise has zero variance and a flat spectrum, it was tested with different intensities in dB. The next section shows the results for various tests.

III RESULTS

The watermark embedded in audio files is as follows: W = (1 0 1 0 1 0 1 0 1 0 1 0) and proceeded as follows: first, a comparison was made between the input signals (original) and the blazes. Below is a comparison between an audio clip over time and with frequency (see Fig 3 and Fig 4)

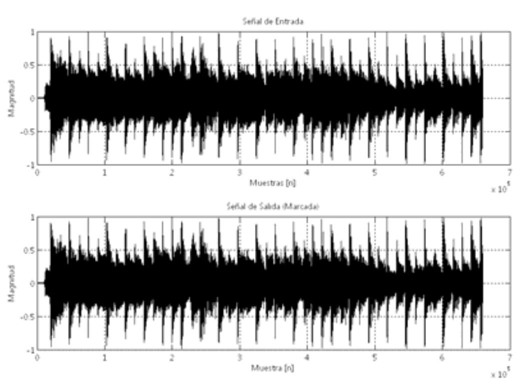

Fig 3. Graph of an audio clip, original and marked.

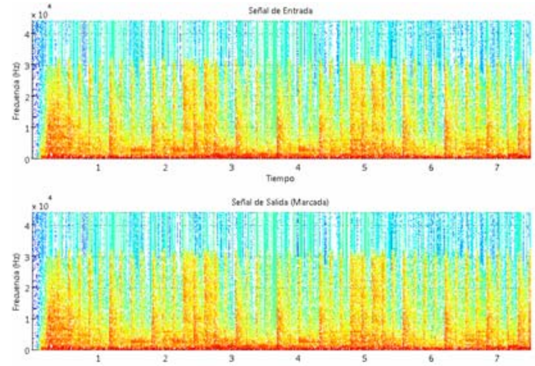

Fig. 3 Frequency spectrum of an audio clip original and marked

To detect the bit that was embedded it was necessary to calculate and display the cepstrum (see Equation 2) and show how it detects a delay within a signal to an audio clip and a delay was added and samples were obtained as the graph below:

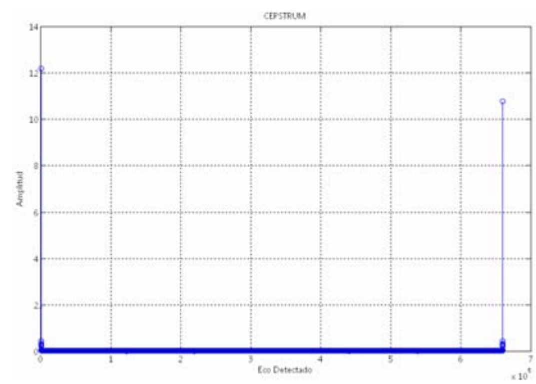

Fig 4. Graph showing the detection of delayed samples in an audio clip.

It is noteworthy that the initial peak representing the original samples and the last peak represents the delayed samples (dilated).

As mentioned in the previous section, to verify the inaudibility of the embedded watermark it was necessary the ODG measurement and the results obtained for each of the five categories listed in the following table:

Category	ODG
Band	-0.122
Pop	-0.134
Electronic	-1.872
Rock	-0.462
Balad	-0.281

Table 2. ODG of audio clips marked.

For the inaudibility subjective tests were surveyed with an audience of 72 people and in the following graph we show the results (see Figure 4):

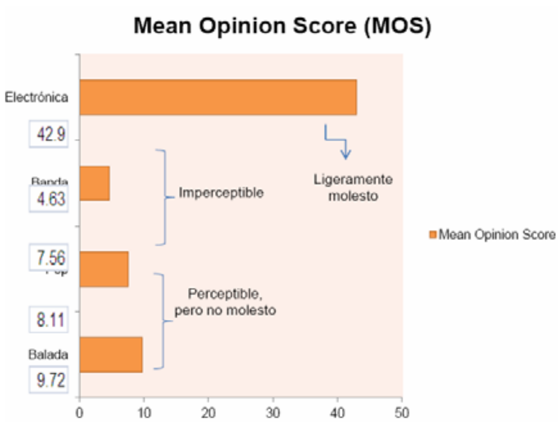

Fig 4. Graph showing the inaudibility MOS test.

To evaluate the robustness of the algorithm, the marked files were tested by adding white Gaussian noise (AWGN, Additive White Gaussian Noise) and the result is shown in Table 3 where we can see that band music is the least affected by the addition of noise tones because the entails are too loud and prevent any presence of noise.

Categoría / Prueba	Pop			Banda			Electrónica			Rock			Balada		
SNR de Ruido Blanco (dB)	20	40	60	20	40	60	20	40	60	20	40	60	20	40	60
Bits erróneos / Total de bits incrustados	6/12	4/12	1/12	3/12	1/12	0/12	7/12	6/12	2/12	3/12	1/12	0/12	5/12	2/12	1/12
% BER	50.0	33.33	8.33	25.0	8.33	0.0	58.33	50.0	16.66	25.0	8.33	0.0	41.66	16.66	8.33

Table 3. BER (Bit Error Rate) for each intensity of AWGN.

IV DISCUSSION

The watermarks are to be a binary string insertion through a methodology that may be useful to label multimedia files to thereby protect or simply control works that have copyright and also necessarily have to satisfy a number of requirements depending on the application that is implanted as the authors note in [1, 2]. In [3] the author speaks of a methodology in which to generate binary insertion multiple echoes are added to the signal in order to differentiate the bit that has been embedded, it is feasible to extract a bit due to the amplitudes of the marked signals (both '1 'to '0') appear to be the same.

V CONCLUSION

The data hiding algorithm for expansion samples, according to the results, is highly reliable for the authentication applications due to robustness tests s yield strength of the insert watermark becomes fragile, since mostly the bit error detection in the process proved to be about half of which those that were embedded. Some applications where the algorithm may be used are as follows:

a) As an identifier: as a way to show the name of the owner of a musical work.
b) Broadcasting: songs or label information of radio emission.
c) As evidence in court, and that if any part of recording illegal evidence and it is modified in favor of any party, the fragile watermark will be supportive to determine when the recording was processed.

If the methodology, at some point, would liked to be used for security applications the following recommendations may be followed:

a) Insert multiple delays to the signal to insert '1 'and '0', respecting the post-masking margin.

b) Modulate the amplitude of the marked blocks and increase the robustness of the same, that is, they will be less likely to disclose which bit they were embedded.

c) Optimize the detection process by a purely blind methodology, that is, that it is not necessary to have prior knowledge of what was embedded, where in the audio was embedded and under what criteria, sampling rate, etc.

d) Check the possible transmission of the file marked in different channels to verify the real robustness.

As future work the algorithm can be implemented on general purpose hardware to validate its performance in real-time processing and establish a possible generation of a device able to perform this action.

VI REFERENCES

[1] M. Hrncar, J. Krajcovic, "Principles of Audio Watermarking", Bytcicka Zilina, Slovak Republic, 2005, pp. 1-3.

[2] Nedeljko, C., T. Seppanen, "Digital Audio Watermarking Techniques and Technologies: Applications and Benchmarks", Information Science Reference, Hershey New York, 2007.

[3] Foo Say Wei, Dong Qi, "Audio Watermarking of Stereo Signals Based on Echo Hiding Method", School of Electrical & Electronic Engeneering, Nanyang Technological University, Singapore, 2009.

[4] Recommendation ITU-RBS.1387-1 Method for Objective Measurements of Perceived Audio Quality 2001.

AUTHOR PROFILES:

Master´s Degree Juan Antonio Rios Chavez. Born in Mexico DF on May 7, 1988, holds a degree in Communications and Electronics from the National Polytechnic Institute in 2009 and earned his Master of Science in Engineering with specialization in Microelectronics Digital Signal Processing by the National Polytechnic Institute in 2011, his research interests include digital audio processing for robotic applications and the hardware implementation of general purpose DSP (Digital Signal Processor) and FPGA (Field Programmable Gate Array)

Ing. Carlos Aquino Ruiz Born in Mexico on April 2, 1973 is in Communications and Electronics Engineering from the National Polytechnic Institute, a research professor and areas of development are applied computing, data networking and security.

Ing. Celedonio Enrique Meza Aguilar Born in Mexico on March 3, 1959 and holds a degree in Communications and Electronics from the National Polytechnic Institute, is currently research professor and development areas are the processing and handling, coding and compression, safety and efficient transmission.