ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

PROGRAMMABLE LOGIC IMPLEMENTATION OF ECHO HIDING FOR AUDIO WATERMARKING

JUAN A. R. CHÁVEZ[1], CARLOS A. RUIZ[2], DOLORES Z. SAAVEDRA[3]

Telecommunication Academy, National Polytechnic Institute ESIME CU
Santa Ana Avenue #1000 Coyoacan, Mexico Federal District
E-mail:[1]jriosc0500@ipn.mx; [2]caquino@ipn.mx; [3]dzamorano@ipn.mx

ABSTRACT

Currently, distribution, and marketing copy music files are very common activities due to free internet connection that facilitates this. To solve this problem, digital watermarking is proposed. In this paper we propose an accurate and content based algorithm for embedding and detecting watermarks within audio signals. The algorithm is based on a new implementation of echo hiding technique for audio watermarking; this algorithm proposes embedding bits (watermark) within blocks of the original audio samples. This implementation was done into a FPGA (Field Programmable Gate Array) hardware; the watermarked audio clips were evaluated by calculating the ODG (Objective Difference Grade) established by the ITU (International Telecommunications Union). The ODG average obtained was -0.98, a result that indicates that the embedded watermark embedding using the proposed algorithm is imperceptible to the Human Auditory System (HAS). This will make the system highly accurate and robust against signal processing attacks in comparison with other methods, this tests was implemented with de Audiostirmark data base.

Keywords: *Audio Watermarking, Echo Modulation, Embeeding.*

1. INTRODUCTION

In recent years, the need to label or protect media has taken hold because of the distribution of those files imminently. In addition to copy and illegal distribution plays an important role because it causes an increase in worldwide piracy. The ease of connecting to the Internet access to illegal sites, and interactions in social networks, are other factors that influence the growth of these criminal activities. They do not have permission to be reproduced by the author and is violating their right to ownership. A recent example of this problem is the controversy regarding piracy of high-quality music across the Internet in MPEG Layer III best known as MP3 format.

To solve this problem, techniques for digital watermarking have been proposed. It should be noted that the problem of piracy is not only proper but also audible signals and video images, but relatively the research field of audio watermarking is complex because the human ear is the most imperfect organ in the human system, that is the focus of our study. The watermarking techniques consist of hiding data, known as watermarks, within media files. These data are binary strings that have the purpose of label and / or protect files with mention of the intellectual property rights therein.

A watermarking algorithm consists of three processes[1]: the embedding algorithm, the detection algorithm and consideration of the evaluation criteria. The first is the method of inserting data into the file that will house a hidden manner, the second is the process of decoding hidden data in order to be removed or modified only by the mastermind and the latter process of evaluating the robustness of the algorithm of embedding against alleged attacks (intentional or unintentional), signal processing for transmission, inaudibility and imperceptibility of hidden data and the playback quality of the watermarked file[2].

The algorithm used in this work is the concealment of data through modulation of echo, which is to add delays in the original audio signal to hide bits. These delays are based on post-masking theory which states that for a sample delay is inaudible (which is necessary in watermarking audio) must be up to 50ms.

To detect the watermark in the audio file processing there are two aspects: the first is called non-blind process, i.e. it is necessary to refer to the initial data used in the insertion process, such as the binary string which was hidden, the position of the embedded bits, the size of the watermark, etc. and the second is called blind process, i.e. it is not necessary to know any initial parameter to detect

Journal of Theoretical and Applied Information Technology
20th January 2013. Vol. 47 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the watermark, we simply need the watermarked file and submit it to the process. This paper proposes a blind detection system, which only requires knowledge of the audio file marked and evaluated in the cepstrum domain. Cepstrum technique was originally developed to detect vibrations of tectonic plates and prevent volcanic eruptions or earthquakes, applied to the audio, it detects when there is repetition of a block and by determining a threshold it will be detected if the bit was embedded as a binary'1 'or '0' [2]. Section I shows an introduction to the environment of digital watermarking. Section II describes the watermarking algorithm using echo modulation (embedding and detecting watermark), Section III shows the inaudibility and robustness tests in Section IV the conclusion is presented.

2. ECHO MODULATION ALGORITHM

1) Embedding Algorithm:

The traditional method of data hiding by insertion of echoes is by[3] adding a delayed version with d_1 samples of the original audio, this is to embed a binary '1' and to add a delayed version with d_1 samples of the original audio this is to embed a binary '0'. In the work we proposed an algorithm where this echoes are embedded and also an amplitude modulation is made. There is an amplitude modulation for the block where a binary '1' will be embedded and there is another amplitude modulation for the block where a '0' will be embedded but it must be necessarily different. Listed below are the steps to the methodology of embedding the watermark and the initial parameters considered to begin the process as shown in equation 1.

$$\hat{x} = \begin{cases} x(n) + x(n-d_1) & w_i = 0 \\ x(n) + x(n-d_2) & w_i = 1 \end{cases} \quad (1)$$

Where;

$\hat{x} = Watermarked\ Signal$

$x(n) = Block\ signal \qquad d_1, d_2 = Delays$

- The original audio is partitioned into blocks of size 'N', where 'N' is the number of bits of the watermark

- Each block is subjected to a comparison with the bit you want to hide, if it is '1 'is adhered to a delayed version of d_1 samples from the same block and with an amplitude α, but if that bit is a '0' is adhering a delayed version of d_2 samples of the same block and an amplitude β, as shown in Fig. 1.

2) Generation of watermarked audio

At the end watermarked blocks are joined by a time division multiplexer, the time which each particular temporal unit, allows the passage of each block which permits the union in a single channel of all blocks watermarked. It is necessary to adapt a number of "N" bits depending on the number of segmented blocks of the original audio and also a counter that determines the size of each block, as shown in Fig. 2.

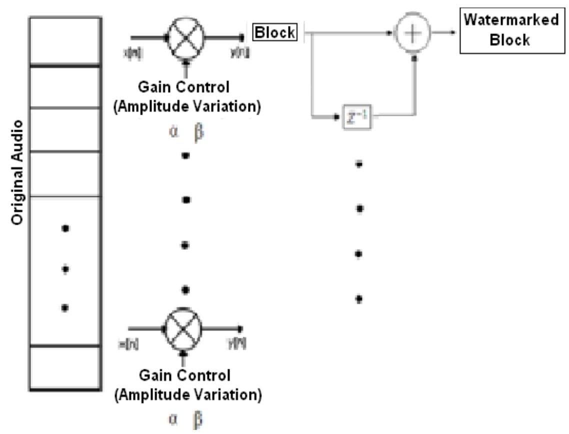

Fig. 1 Embedding Scheme For Echo Modulation Audio Watermarking

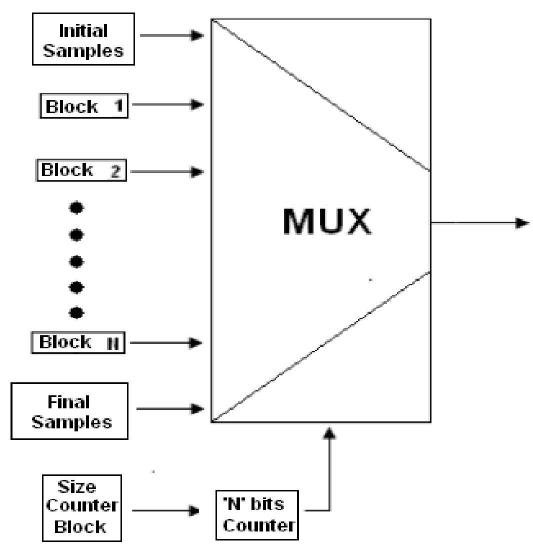

Fig. 2 Audio File Watermarked Generator

In Fig. 2 we can see a block called "initial samples" and another called "final samples", these two sets of samples are spaces that are used when they are silences in the audio and thus prevent the watermark as perceptible, detected or perceived audibly-

Journal of Theoretical and Applied Information Technology
20th January 2013. Vol. 47 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3) Detecting Algorithm:

The detection process is found to be optimal, if and only if it is performed blind, i.e., it is not necessary to reference data embedding, and the whole inlay by calculating the cepstrum, as shown as Fig. 3.

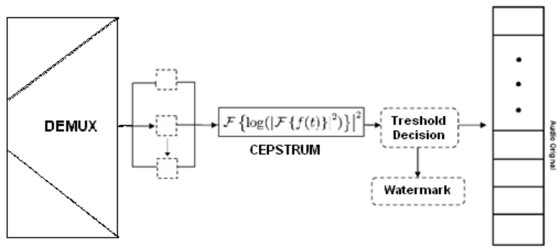

Fig. 3 Detecting Scheme For Echo Modulation Audio Watermarking

4) Evaluation Criteria

Four criteria were carefully selected as part of the evaluation framework. They were chosen to reflect the fact that watermarking is effectively a communications system. In addition, the criteria are simple to test, and may be applied to any type of watermarking system (audio, image, or video). It is important to note that the requirements of a practical watermarking system vary between applications, and so one criterion may be more important in some situations than in others. The criteria are described in the following paragraphs.

Signal to noise ratio (SNR): For transmission and signal processing is necessary to subject it to strong evidence that are likely to exist outside the file. The signal to noise ratio refers to the Additive White Gaussian intensity Noise (AWGN) that can be added to the signal for any of the above actions, the final calculated watermark bits can deteriorate. Bit rate refers to the amount of watermark data that may be reliably embedded within a host signal per unit of time or space, such as bits per second or bits per sample. A higher bit rate may be desirable in some applications in order to embed more copyright information. In this study, reliability was measured as the bit error rate (BER) of extracted watermark data. For embedded and extracted watermark sequences of length *B* bits, the BER (in percent) is given by the expression:

$$BER = \frac{Erroneous\ Bits}{Total\ Bits} \times 100 \qquad (2)$$

Resampling: The watermarked file may be subject to transmission and thus a possible scanning. Therefore, the watermark should have a high degree of robustness against resampling digital processing and, therefore, the test is to resample the watermarked signal with different sampling rates to check where it is able to bear the watermark

Inaudibility Test: As mentioned above, the hidden watermark has to be inaudible and it will be necessary to verify this using the Perceptual Evaluation of Audio Quality (PEAQ) established by the International Telecommunication Union (ITU) as shown in [6]and is to calculate the Objective Difference Grade (OGD) by the equation 3.

$$\hat{E}_s[k,n] = \propto [k]\hat{E}_s[k,n-1] + (1-\propto [k])E_c[k,n] \qquad (3)$$

The above equation represents the implementation of a first order filter in which two signals, the signal masked and masking which is in charge of covering the watermark signal to be imperceptible, also based on the imperfection of the Human Auditory System (HAS). With the result obtained it is compared to a standard established by the ITU, which is shown in Table 1.

Table 1. Range Of Values Of ODG

Impairment Description	ODG
Imperceptible	0.0
Perceptible, but not annoying	-1.0
Slightly annoying	-2.0
Annoying	-3.0
Very annoying	-4.0

5) FPGA Implementation

The hardware development is the Field Programmable Gate Array card (FPGA) Virtex-II Pro from Xilinx, is general purpose and uses parallel processing, i.e., that may implement two or more processes and work simultaneously in real time. FPGA's are semiconductor devices containing logic blocks whose connection and functionality can be programmed as often as the user / developer wants. Its scope is the same as that of the ASIC (Application Specific Integrated Circuit). Its advantages are versatility and low cost per unit and to develop systems, which today are widely used.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

In Fig. 4 shows the block structure of an FPGA:

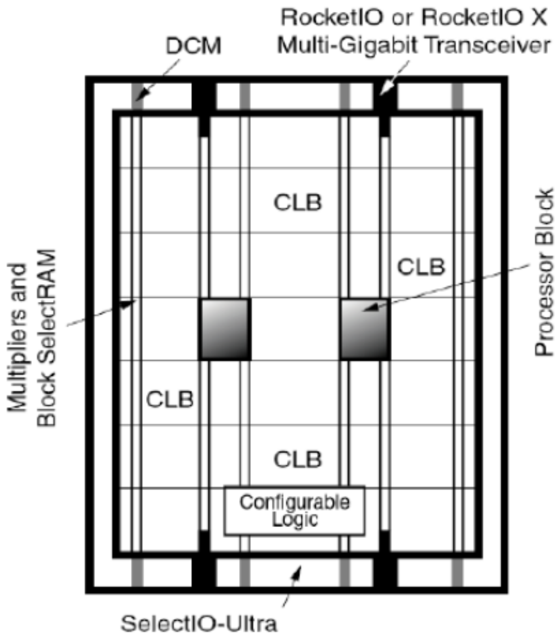

Fig. 4 FPGA Architecture Of Virtex-II Pro Of Xilinx

The next section discusses the results obtained according to initial conditions, testing and implementations that have been mentioned previously.

3. IMPLEMENTATION RESULTS

To implement the Xilinx FPGA algorithm embedding watermarks echo modulation is needed to design the schema using exclusive blocks of Xilinx system generator under Simulink platform. The requirements used in the audio clips are listed below:

- Format: WAV
- Sampling frequency: 44.1 kHz
- Average size: 20 MB
- Average length of audio: 10’000,000 samples
- Music Category: Pop, Rock, Band, Ballad, Electronic
- Watermark length: 120 bits
- Block size: 1024 samples

Fig. 5 and Fig. 6 show the diagrams of embedding watermark[4], we mention that to hide a '1' was used a delay of 50 ms and an amplitude modulation at 1/3 of the original and to hide a '0' only vacated the block unchanged.

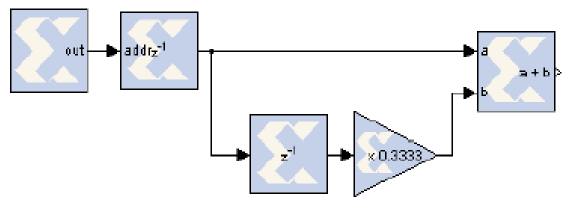

Fig. 5 Diagram Of Embedding A '1'

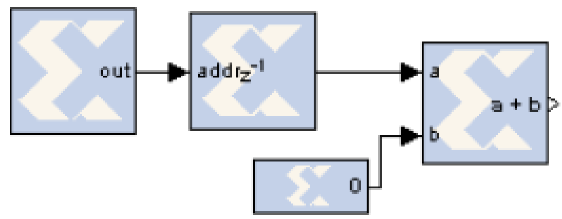

Fig. 5 Diagram Of Embedding A '0'

To verify that the proposed arrangement for embedding a "1" has an echo (delay) was displayed on an oscilloscope and the output signals are shown in Fig. 6.

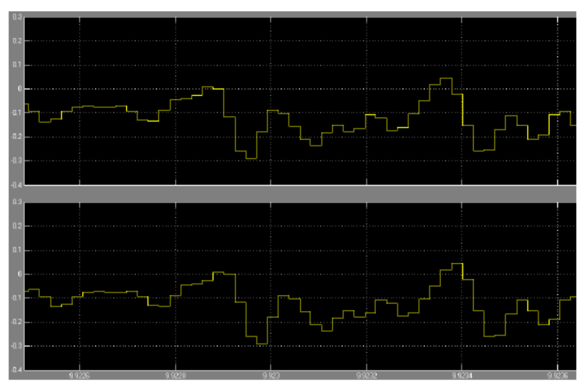

Fig. 6 Original And Delayed Signals

The first part shows a sample portion of an audio clip in normal play and the second part shows the same portion of sample of the same audio clip but delayed 50 ms.

The hidden watermark that was obtained by drawing a random binary sequence generator, the output which gives us a binary string will serve as the data to hide. Fig. 7 shows the implementation of the generator in blocks of Xilinx.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

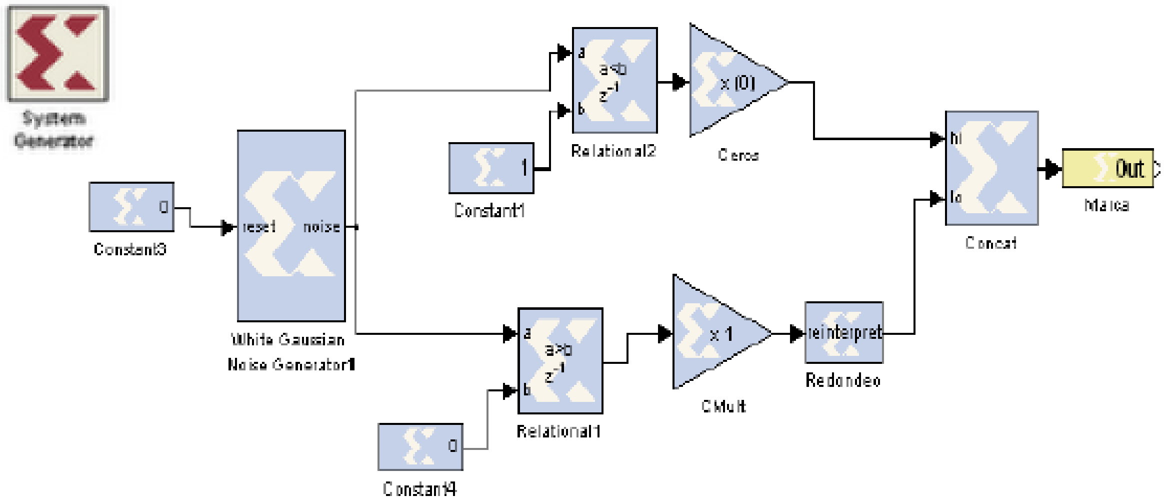

Fig. 7 Watermark Generator

The above diagram consists of a white Gaussian noise generator, which has the function of generating a random sequence which later through decision blocks, modules and concatenating gain control thereof, is modelled to obtain a pure stream of '1' and '0' binary.

Figure 8 shows the scheme of zero or an inlay for a block of samples that depends on the entry having either a "1" or a "0" implemented in blocks of system generator. It is noteworthy that the general outline of the inlay consists of 120 bits (which is not shown because the pattern is too large). It works by sectioning blocks of 1024 samples and submitting them to push rules, if a "1" there is a delay in adding it to the original block, otherwise, if a "0" the block is fully intact . The overall system output, will "switch" blocks through a multiplexer controlled by a router. To counter that, will be by putting the corresponding time to each one and generating the audio marking.

Hiding of a '1 'or '0' depends on the random sequence obtained and showed above by the diagram in Fig. 7.

In the detection process it was necessary to calculate the cepstrum in where we obtained a graph similar to Fig. 9

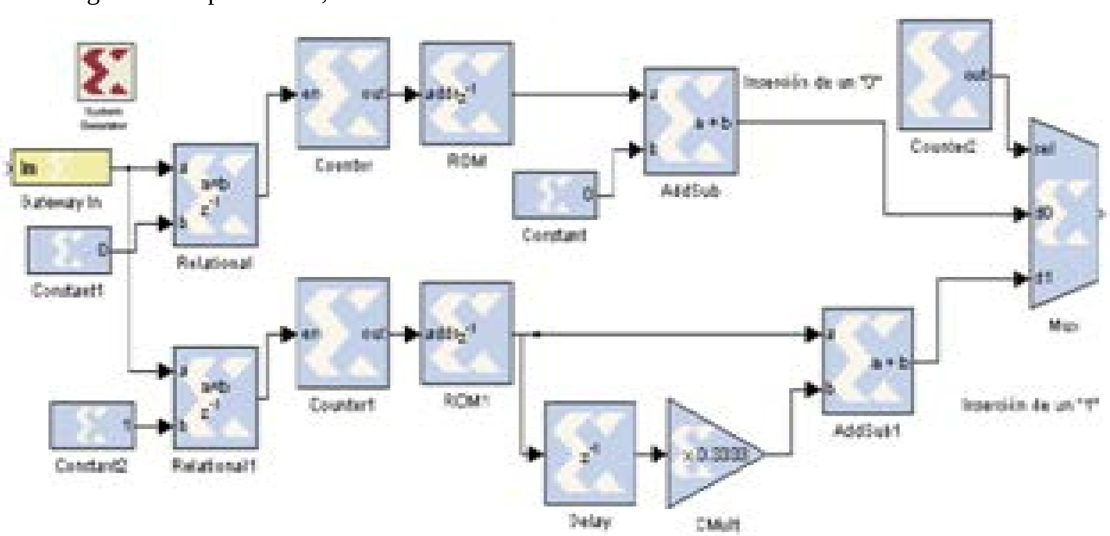

Fig. 8 Watermark Embedded By Echo Modulation

Journal of Theoretical and Applied Information Technology
20th January 2013. Vol. 47 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

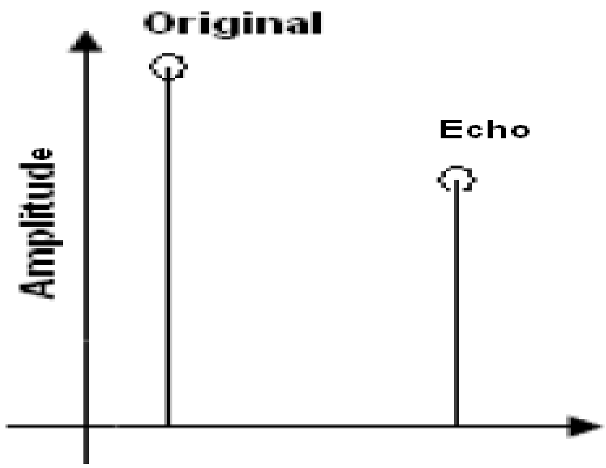

Fig. 9 Cepstrum Graphic (Original And Echo)

The figure above shows the marked audio but plotted in the cepstrum domain. This displays a peak for the original and a peak signal to the echo (if any) to decide whether it was hidden a '1 'or '0'. Fig. 10 shows the cepstrum implemented in blocks of Xilinx system generator.

To program the system in the FPGA it is necessary to establish the parameters required by the ISE software, it will help us build the file with *.bit that will be loaded to the FPGA. ISE has a graphical user interface GUI (Graphic User Interface) is called Project Navigator. For the start of a new project, you need the design stage. At this stage what is done is to select the type of design description. Designs can be entered in different formats. The most used are: the schematics, state diagrams and descriptions in VHDL or Verilog hardware.

Later in the synthesis and design its compiled net list and created, once compiled, it can be simulated in a functional level. After this simulation, using deployment tools such as translation, mapping and routing, it allows the specification of restrictions or directions, through the *.cfg, for an optimal implementation on FPGA. So you can make a simulation time, and estimate how the system will behave in the FPGA. The functional level simulation does not account for delays caused by the hardware level and simulates the design time considering the hardware configuration. Finally, create the programming file *.bit and downloaded on the device design.

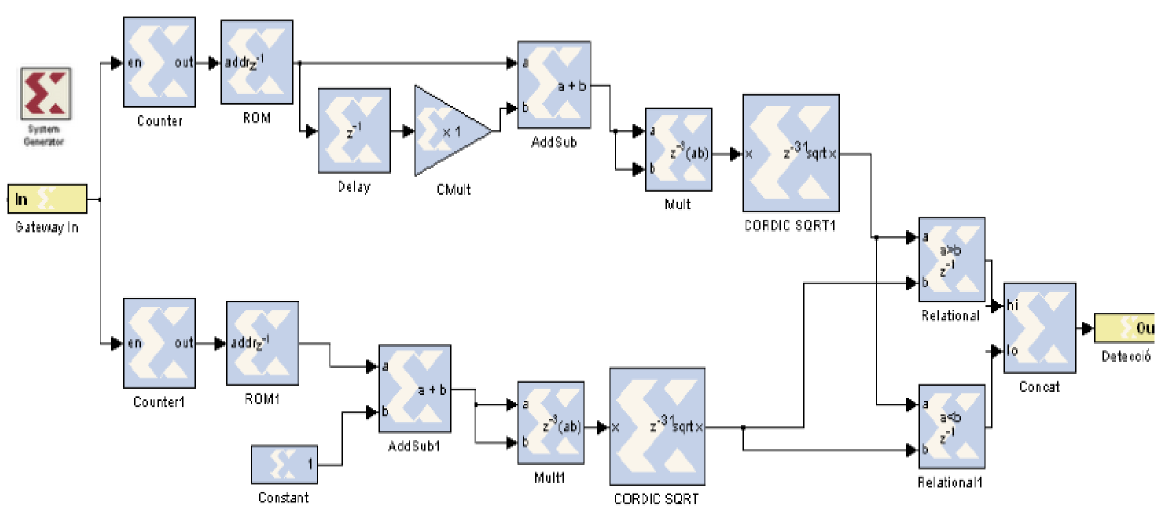

Fig. 10 Watermark Detector By Cepstrum Domain

Journal of Theoretical and Applied Information Technology
20th January 2013. Vol. 47 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Fig. 11 shows the Project Navigator interface. In point number 1 shows the toolbar. In item number 2 shows the window sources, which is the source files that make up the project. Point 3 shows the process window. These processes can be: synthesis, simulation and implementation among others. In Section 4, shows the work area, this area may be of schematic and VHDL, depending on the selected project description. Finally Section 5 is the transcript window, where you can view the status of transactions already processed and being processed.

In the implementation in FPGA embedded a binary string that is the phrase "JUAN ANTONIO RC" consisting of 120 bits and Fig 12 shows the watermark signal chosen for concealment and the watermark signal obtained in the recovery process under ideal conditions, i.e., that there is any alteration in the watermarked clip.

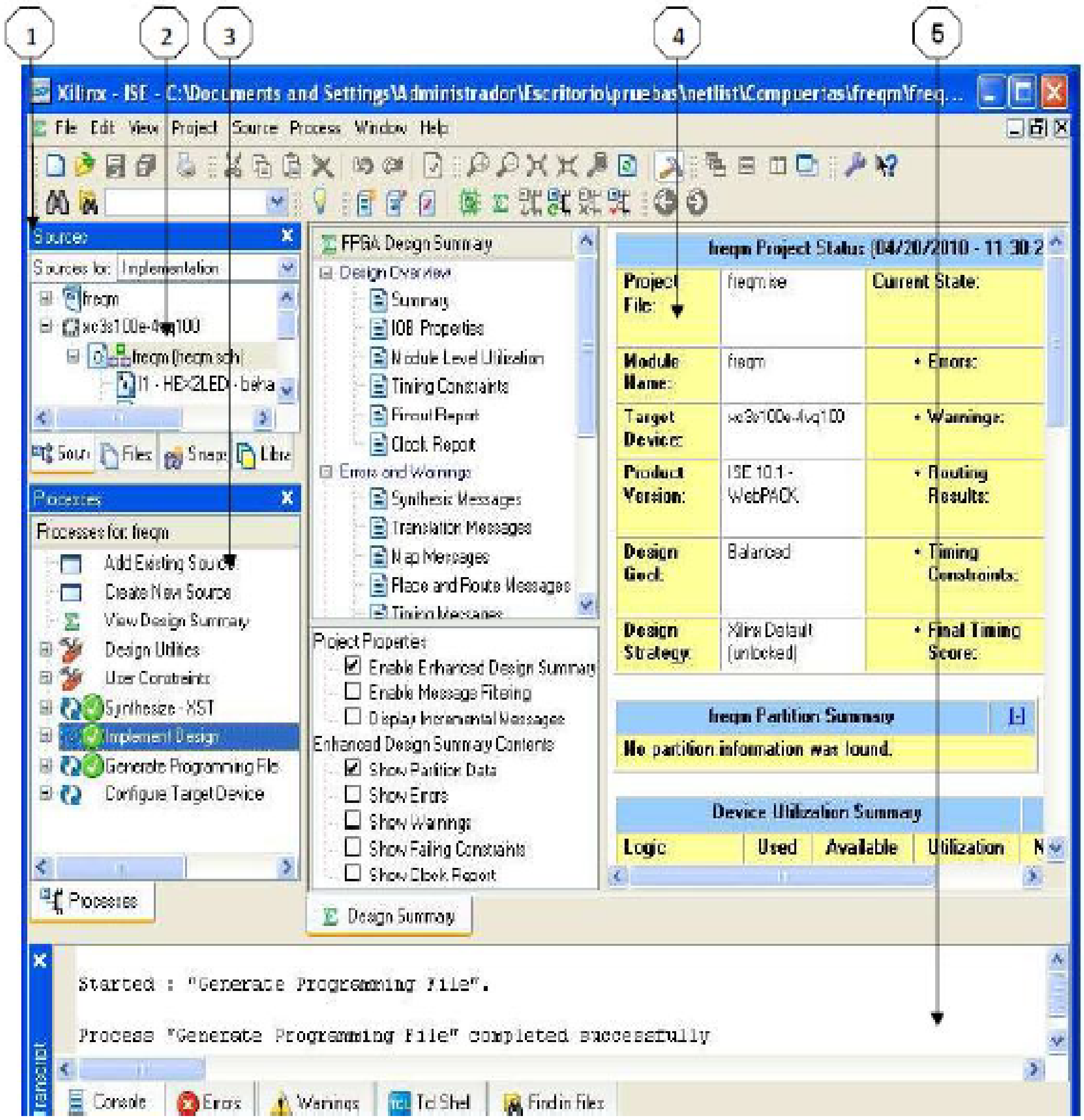

Fig. 11 Project Navigator

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

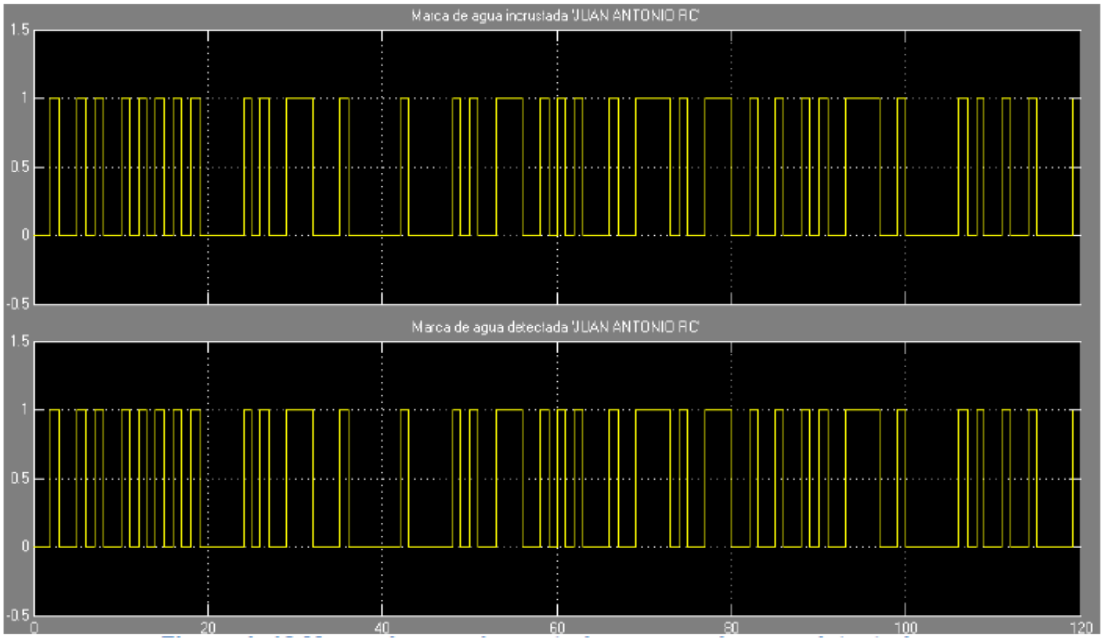

Fig. 12 Embedded Watermark And Watermark Recovered

As it can be generated watermark (120 bits embedded) is the same as the watermark signal obtained from the generator. In this case the optimum detection is 100% because there is no attack by half. Ideally the detection of the watermark should not suffer damage or modifications signal embedded watermark. At the exit of the insertion system, must be obtained, marked and reconstructed audio blocks that were previously processed, as shown in Fig 13.

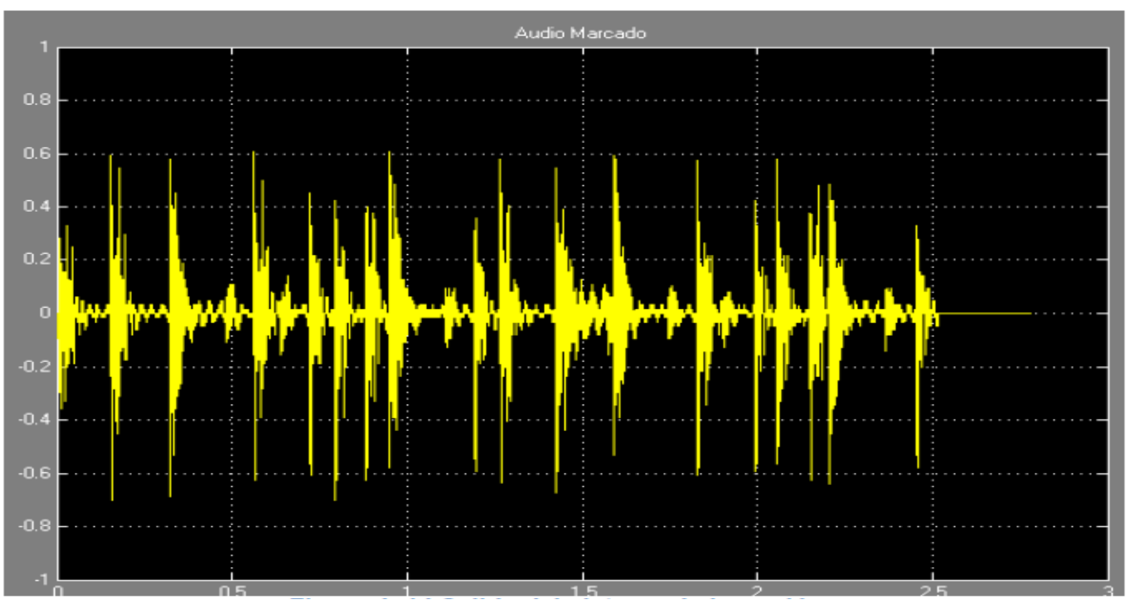

Fig. 13 Embedding System Output

Audiostirmark is an application designed exclusively to verify the robustness of a watermarking algorithm for digital audio. It is a database of various digital processes applicable to audible signals, and implements several scenarios that can be part of the audio clip watermarked. The algorithm was tested for robustness against attacks audiostirmark set in the software and Table 2 shows the results obtained.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 2. Attacks Supported And Not Supported By The Echo Hiding.

Eco Hiding	
Robustness	**Fragile**
Resampling	AWGN
Nothing	Add Dyn Noise
Amplify	Bit Changer
Extra Stereo	Copy Sample
Time Stretch	Cut Samples
Compressor	LSB Zero
Add Brumm	RC _High Pass
	RC _Low Pass
	Invert
	Replace Samples

To evaluate the inaudibility of the embedded watermark calculation was performed ODG (100 audio clips, 20 for each category) and the results obtained are shown in Table 3.

Table 3. Odg Of Musical Categories

Music Category	Average ODG
Pop	-0.695
Electronic	-1.876
Band	-0.178
Ballad	-0.293
Rock	-1.862

4. CONCLUSIONS

The techniques of digital watermarking can be analysed in the time domain or frequency hence depends on its mathematical complexity. The chosen algorithm works in time domain and its analysis is much more feasible.

The results of the evaluation of the robustness of the watermark embedded show that it is robust to some attacks or digital processing is submitted to the watermarked file. These results show us that the algorithm is generally fragile so the applications[5]that may have are:

- As an identifier: a way to show the name of the owner of a musical work.
- Broadcasting: tag songs or additional information of radio emission.
- As evidence in court and that if any recording is part of some illegal evidence and this is modified in favour of any party, the fragile watermark will be supportive to determine when the recording was processed.

The algorithm is a considerable fragility most digital attacks proposed by audiostirmark database which makes it feasible for authentication audio applications, for example, as evidence in court, security voice verification, etc.

As future work can increase the robustness of the algorithm. To increase the robustness of the watermark can opt to add multiple echoes in the insertion algorithm, but knowing that this would increase the complexity of the detection system and computational cost.

5. ACKNOWLEDGMENT

We thank the National Polytechnic Institute (IPN) for giving us the interest and knowledge required to perform this work in addition to the facilities granted to culminate it. Furthermore special thanks to the direct and indirect contributors that offered their views, suggestions and improvements to this work.

6. REFERENCES

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

[2] H.F. Eberhed, "*Psychoacoustics, facts and models*", Munchen, German, Ed. Springer, August 2006.

[3] Foo. S. W., Dong Q., "*Audio Watermarking of Stereo Signals Based on Echo-Hiding Method*", Nanyang Technological University, Singapore, IEEE, 2009.

[4] J. A Ríos Chávez, C. Aquino Ruiz, S. A. García, "*Audio Watermarking of WAV Files by Echo Modulation*", SEPI ESIME UC, México D.F., IEE Transactions, CONIELECOMP 2011

[5] Michael A."*Audio Watermarking: Features, applications and algorithms*", Department for Security Technology for Graphics and Communication Systems, Darmstadt, Germany, IEEE, 2000.

[6] Recommendation ITU-R BS.1387-1 Method for objective measurements of perceived audio quality, 2001.