Digital Audio Effects Based On Delay and Implementation on FPGA.

Jaime Andrés Laino Guerra
Faculty of Electronic Engineer
Universidad Pontificia Bolivariana
Medellín

RESUME **- This Article brings an analysis of some audio effects based on delay, as well as their diagram block on System Generator (hardware implementation). The resulting digital effect creates different sensations based on psychoacoustic perception.**

Keywords – Delay; Psychoacoustic; Simulink; System Generator; FPGA.

ABSTRACT **- This Article brings an analysis of some audio effects based on delay, as well as their diagram block on System Generator (hardware implementation). The resulting digital effect creates different sensations based on psychoacoustic perception.**

Keywords – Delay; Psychoacoustic; Simulink; System Generator; FPGA.

1. INTRODUCCIÓN

Now a days, as the study of music and electronics has grown over years and become possible to join them in one single field developing synthesizers and electronic instruments, a high amount of programs brings useful audio effects for music developers and engineers such as plug in's, which can work on different software right on your computer. There also lots of multi-effects pedal build in hardware that can process audio signals in real time and apply programmable parameters to pursue new tones and effects outside of conventional presets. The intent of this article is to show that some effects based on delay got a basis structure and his perception is based on psychoacoustic parameters selected for each one.

1. Acoustic and Psychoacoustic [2]:

First of all it needs to be clear about some concepts like the definition of acoustic, psychoacoustic and the basic knows of the hearing process.

Acoustic is "an interdisciplinary science that deals with the study of all mechanical waves in gases, liquids, and solids including vibration, sound, ultrasound and infrasound". Acoustic splits into multiple branches, one of them is psychoacoustic. This one is the scientific study of sound perception, its mean the psychological and physiological responses associated with sound.

As we know sound is a form of energy that moves through a media in waves of pressure, and although the ear is the sense organ that recognizes sound, it is the brain and central nervous system that "hears". Sound waves are perceived by the brain through the firing of nerve cells in the auditory portion of the central nervous system. The ear changes sound pressure waves from the outside world into a signal of nerve impulses sent to the brain, this is why the brain plays such an important role in this process, resuming the ear is a transduction with a specter analyzer that allows perceiving tiny different variation in frequency and energy.

2. Hass Effect:

This psychoacoustic phenomenon describes how human perceives the source of a sound by identifying the direction of the sound reflections by the time they take to arrive. This effect was work by Helmut Haas at the University of Gottingen, Gottingen, Germany. A simple experiment to a better understanding of this phenomenon is listening two equal signals using stereophonics headphones one of them lightly delay from the other an applied to each ear.

The stages of this experiment are illustrated next:

Note: the ears have more sensibility at the 1 KHz frequency so the experiment is illustrated at that frequency.

• Applying a single tone of audio signal in both speakers as seen in Figure 1: In this case the sound is perceived as one source in the direction of the angle form by both speakers (on headphones case, the source will be felt right in front).

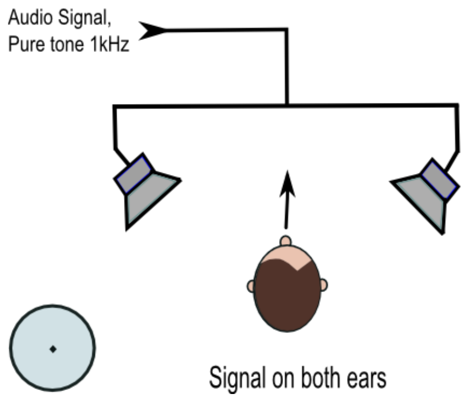


Figure 1. Pure tone applying to stereo speakers

• Applying a pure tone of 1 KHz audio signal in both speakers with one delay from 5 to 30 milliseconds from one sound to the other as seen in Figure 2: In this case is only perceived one source with twice of level in the direction of the first speaker (on headphones case, the source will be felt beside you).

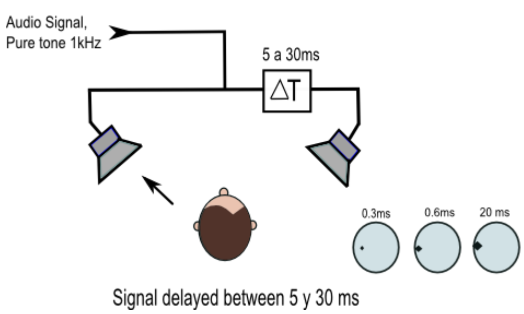


Figure 2. Pure tone applying to stereo speakers with delay

• Applying a pure tone of 1 KHz audio signal in both speakers with one delay over 35 milliseconds to the other as seen in Figure 3: In this case the sound delayed is perceived as an echo of the first, but from different sources (on headphones case, left and right ears felt two different sources).

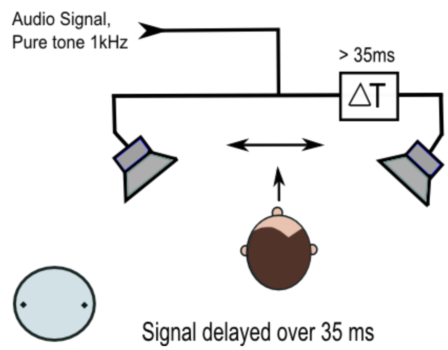


Figure 3. Pure tone applying to stereo speaker with delay over 35ms

This way the phenomenon characterized the delay experienced in acoustical spaces and give an idea of how the ear sum the sound energy's.

3. Delay :

Audio effects are communally use in recording, live sound and by musicians (and to describe sound environment for some enclosures). That's because they allow giving various interpretations and sensations with different sound textures. Digital effects can vary audio signal parameters such as loudness (amplitude), pitch (frequency), envelope, and others. It is possible to group effects depending on the change they made on the original signal. In this case we will refer to delay based effects.

3.1. Single Delay

Single delay is the foundation of several audio effects. Delays can be perceived in acoustical spaces. A sound wave reflected by any surface will be superimposed on the sound wave at the source. The distance of the surface will determinate the delay that is imposed to each reflected sound wave. Counterparts of these phenomena have been implemented by digital signal processing.

There are two representation of single delay but to better understanding of them in signal processing is necessary to know about comb filter.

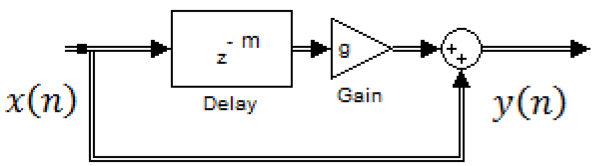
3.2. Comb Filter

Comb filter works adding a delay signal to the original, creating constructive and destructive interference. They are used in multiple signal processing applications, for example audio effects. In acoustics, comb filtering is presented, for example, when two loudspeakers are playing the same signal at differences distance from the listener. Because of the wavelength of the signal and the distance between speaker-listener, the sound can arrive with different phases creating constructive or destructive interferences. Basically when a reflected sound takes a longer path to arrive, a comb filter is created when the two are combined.

There are two forms of comb filter, feed-forward and feedback, or in digital case, FIR and IIR respectably.

3.3. FIR Comb Filter

This one corresponds to feed-forward and its diagram representation is in **Error! Reference source not found.**.



$$y(n) = x(n) + gx(n-m)$$

$m = t/fs$: fs is the sample frequency

$$H(z) = 1 + gZ^{-m} = {Z^m + g}/{Z^m}$$

FIR comb filter like acoustical delay present different sensations depending of the value of t, for larger values it's possible to separate the delayed signal from the original, for smaller values it isn't possible to segregate the events but it can be perceive the spectral effect.

3.4. IIR Comb Filter

The IIR comb filter produces an endless series of responses. That's because the input signal circulates in a delay line that is feed back to the input as can be seen in its diagram representation in Figure 5. Each time the signal goes through the delay line it is attenuated by g.

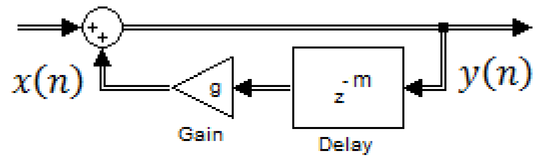


Figura 5. Block diagram IIR comb filter

That's because IIR comb filter is similar to FIR comb filter but the main differences is that the gain grows every time the input pass thru the delay line making the filter response infinite. The amplitude is g^p where p is the number of cycles the signal has gone through the delay line.

3.5. Universal Comb Filter

Since the ear perceives different the delay between 50 to 100ms or more and between 30ms or less, diverse sensation can be created playing only between those ranges that why with the same structure or block diagram, the effects discussed next can be created.

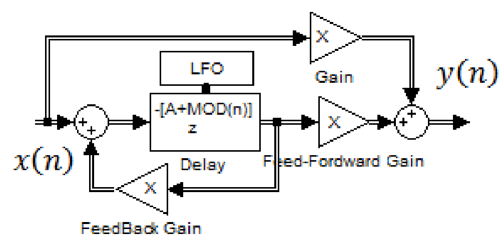


Figure 6. Universal comb filter with variable-length delay line

As can be seen in Figure 6 this block diagram represent a universal comb filter because it have the feed-forward loop of FIR comb filter and the feedback loop of the IIR comb filter also with a LFO block connected to the delay allowing to create the different effects according to the parameter set for each one.

4. ECHO

It is the result of a single reflection that returns to the source. The time it takes to go and come is at least 100ms (4410 samples in digital implementation), and that's why the minimum distance should be 17m. The equation for the delay time is:

$$t = \frac{2d}{c} \quad [=]s$$

Where c is the sound speed (345m/s) and d is the distance from the source and the reflection point en meters.

The equivalent to found the number of samples (Ns) for the delay is:

$$Ns = round(250 * d)$$

Figure 7 show the block diagram for Echo effect.

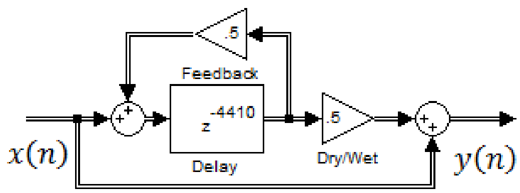


Figure 7. Echo block diagram

It is based on IIR comb filter, that's why the Feedback has to be less or equal than 1 to keep stability, also has to remember in digital implementation that m≥4410 samples taking in count the sampling frequency is 44100Hz

The parameters manipulable on echo system are:

• Feedback: is the amount of gain on the feedback loop. Range: 0:1.

• Dry/Wet: is the relation between direct signal (Dry) and delayed signal (Wet). Range 0:1 on the feed-forward loop and is represented by percent wet over dry

• Delay: is the delayed signal time in milliseconds. Range: 100ms to a maximum set by the programmer of the effect. The maximum time set should enable to distinguish between direct and delayed signal if that does not accomplish, the proper term is reverberation

5. VIBRATO

Appling the concept of modulation on single delay can create several effects, where the most simple is vibrato.

This effect consists on a periodical vibration of frequency sound, that creates a sensation of moving. A variable delay is applied around mean value by a low frequency oscillator (from 5 to 15Hz). Typical delay values are from 5 to 10ms. The low frequency oscillator should be set direct level or reference Delay of 7.5ms or 330 samples for digital implementation, which allow peak amplitude of 110 samples (2.5ms) that way the range for delay is set from 5 to 10ms. It can be seen in Figure 8.

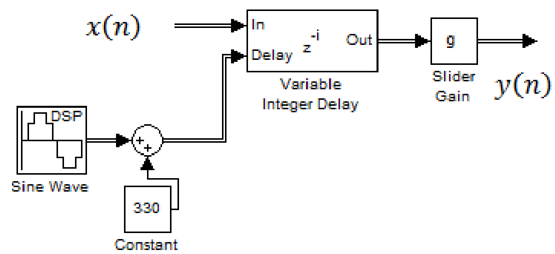


Figura 8. Vibrato block diagram

On this particular case the wave form of the LFO is sinusoidal, but it can be any form. Also is evident that the effect is created by modulation, but is worth mention that can be amplitude modulation or pulsatile effect, in this case is frequency modulation.

By delaying the signal will produce an apparent reduction in frequency. The parameters manipulable on Vibrato system are:

• Rate: Adjusts the rate of the Vibrato it means the frequency of the LFO. Range: 5 to 15Hz. perceptively it means how fast the sound vibrate.

• Depth: This adjusts the depth of the Vibrato. It means the wave amplitude of the LFO. Range: 0 to 2.5ms or 0:110 samples. Perceptively it means the amount of vibration.

• Gain: Is the direct gain of the effect. Range 0:1

6. FLANGER

It is a filter that changes the frequency content in a signal sampling at the entrance and shift phase modulation by a variable delay and adding the result to the bypass entrance. It has its origins in the tape-open recording. As can see in Figure 9 the block diagram is similar to vibrato except for the bypass signal add at the end.

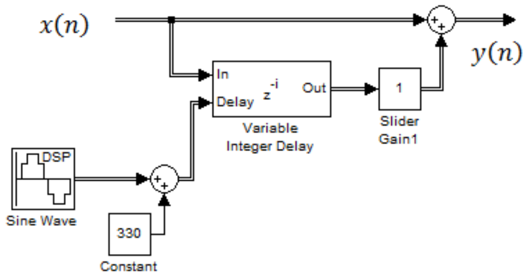


Figure 9. Flanger block diagram

This effect is achieved with variable delay oscillating periodically around certain value (less than 15ms) using LFO. The result signal will be affected by phase distortion period.

The LFO can be set with frequency 0.5 to 5Hz and waveform sinusoidal. To assure the right amount of delay the direct level of 7.5ms or 330 samples will be add; in this case the peak amplitude will be 320 samples, which guarantee delay range from 0.3 to 14.7ms.

There are 3 tuning parameters to set:

• Rate: Adjusts the rate of the frequency of the LFO. Range: 0.5 to 5Hz.

• Depth: This adjusts the depth of the Flanger. It means the wave amplitude of the LFO. Range: 0.3 to 14.7ms or 0 to 320 samples.

• Dry/Wet: is the relation between direct signal (Dry) and delayed signal (Wet). Range 0:1.

7. CHORUS

This effect simulates the presence of several sources play imperfect unison. It implemented by combining original signal to several signals delayed a variable number of samples. The typical delay is from 10 to 30ms and is affected by tiny low frequency variations from less than 3Hz for this case only 2.5ms up and down from the set delay.

For initial values the direct level is set at 20ms or 880 samples as can be seen in Figure 10 and the peak amplitude of the LFO is 110 samples at 1Hz.

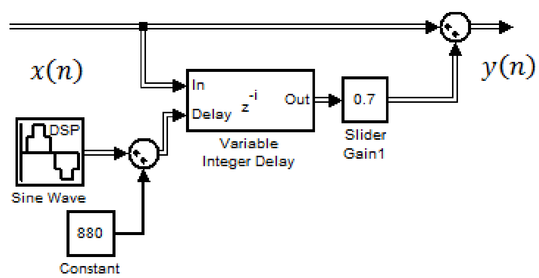


Figure 10. Chorus block diagram

The parameters for this effect are:

• Delay: controls the minimum delay or direct level that is used. If it is set to a very small value the chorus will act as a flanger. Optimal delay times range between 20 and 30 ms.

• Depth: control how much the total delay changes over time. The sum of the depth and delay parameters is the maximum delay used in processing the signal. Range 0 to 5ms or 0 to 220 samples.

• Rate: control the LFO frequency affecting the pitch modulation too. Range 0.2 to 2 Hz

• Dry/Wet: is the relation between direct signal (Dry) and delayed signal (Wet). Range 0:1.

8. Hardware simulation and implementation

Before continues it needs to be clear that in digital implementation all is control by clock time and samples, so the next table shows the ranges for the modifiable parameter for each effect given in number of samples for better understand in digital implementation. Remember the sample time is 44100 samples per second.

	Gain	FF G	FB G	Delay (samples)	Depth	Rate	MOD
Echo	1	0-1	0-1	0-13000	0	0 Hz	----
Vibrato	0	0-1	0	330	0-110	5-15Hz	Sine
Flanger	1	0-1	0	330	0-320	0.5-5Hz	Waveform
Chorus	1	0-1	0	880	0-220	0.2-2Hz	Waveform

For the digital implementation it was created a basic block representation that characterizes the Universal comb filter with variable-length delay line show in Figure 6 and based on it where added or delete block in order to use only what is needed. For all of this it was used System Generator a powerful tool that keep an abstraction level such as in traditional Simulink blocksets, permitting to model and create design with functional descriptions and automatically translate into hardware implementations. All of this host by the Simulink interface.

It was used a Spartan-3E 1600E that includes an on-board 50MHz clock oscillator. It is down sample by 1136 times to generate a 44014Hz clock. The ADC/DAC used work directly with 50MHz clock and give data at that clock rate, so this data also needs to be down sample, but the data update each 44014 clock signal. The ADC/DAC deliver 16-bits unsigned and required 12-bits unsigned, this is given by the device on the FPGA.

This FPGA counts with an SPI-compatible, four-channel, serial Digital-to-Analog Converter (DAC). The DAC device is a Linear Technology LTC2624 quad DAC with 12-bit unsigned resolution. This converter presents a typically performance characteristics shown on Figure 11.

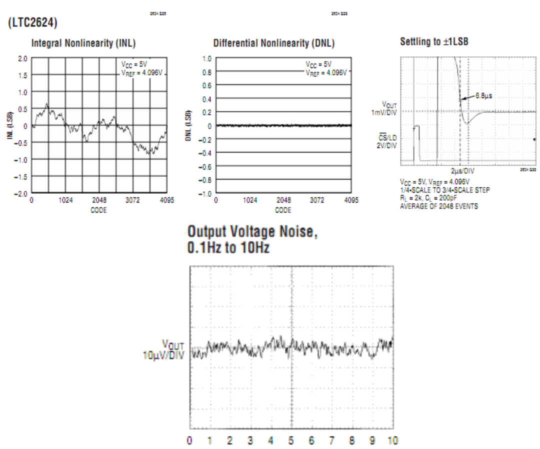


Figure 11. Typically performance characteristics LTC2624

This will present a background noise when there’s no signal at the input; it is cause by the least significant bit (LSB), this one gives the accuracy, but in the presence of zero it start to oscillate delivering a output voltage noise from 0.1 to 10Hz that be perceive as a white noise.

Having told the above, the final representation of the hardware on system generator is show next.

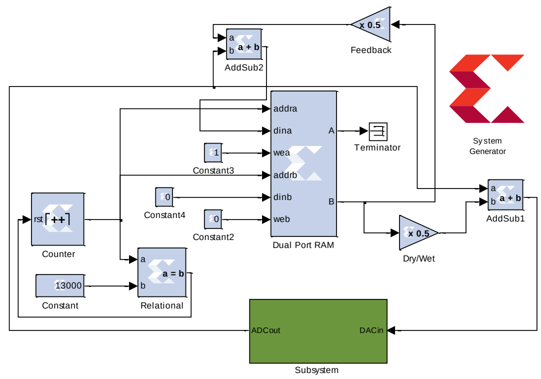


Figure 12. Echo Hardware implementation

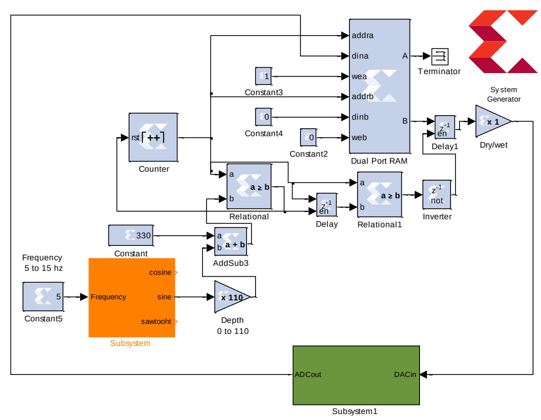


Figura 13. Vibrato Hardware implementation

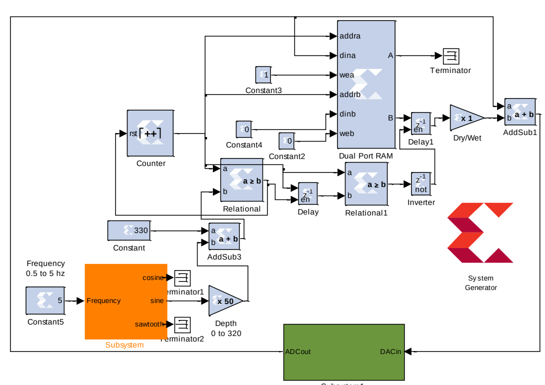


Figura 14. Flanger Hardware implementation

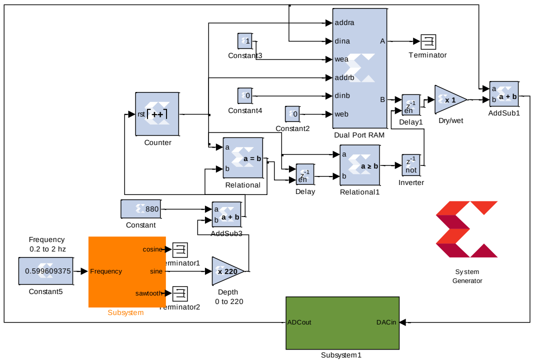


Figura 15. **Chorus Hardware implementation**

Analysis of latencies associated with both implementations (hardware and software). All the figures 7, 8, 9, 10 shows the software implementation on Simulink.

Delay-Based Effect	Simulink Implementation Latency	FPGA Implementation Latency
Echo	49171 Samples / 1.115 Sec	1 sample / 2.26e^{-5} Sec
Vibrato	52038 Samples / 1.18 Sec	2 samples /4.54 e^{-5} Sec
Flanger	65268 Samples / 1.48 Sec	2 samples /4.54 e^{-5} Sec
Chorus	48510 Samples / 1.10 Sec	2 samples /4.54 e^{-5} Sec

As seen in the table above, latency hardware implementations are tied to a flip-flop at the output of the Ram and the Ram itself that takes one clock cycle to update his output. While the latency in Simulink is linked to computation time required for each block. The time shown in the table corresponds to the latency present when using the "From Audio Device block" besides the already mentioned computation time of each block. This at the start of the simulation begins writing the input data to a buffer. When the buffer is full, it writes the contents of the buffer to the queue. As the audio device appends audio data to the bottom of the queue, it pulls data from the top of the queue to fill the Simulink frame.

When the Simulink models work directly with the "From Wave File Block" it already got
the wav file loaded on memory so it just need to star giving frames, for that case the latency on all the effects is 0.830 seconds or 36603 samples, this is the latency present from the moment is press the play bottom to the moment the sound start. In this case if the result expected is not the play sound but is the result record the effect is added and can be saved directly to a wav file without providing latency. This same work is archive with the hardware regardless of whether it is in real time or to a soundtrack.

Conclusions

• All the audio effects create with delays are psychoacoustic phenomenon base on the theory of the Haas effect, where the kind of effect expected depend of the time delayed because it will represent the sum of energy for any delayed signal and according to the time the sound will be perceive as one source, multiple source or just with twice the intensity.

• Given a universal architecture in the design effects, it is possible to simplify the resources and propose an implementation where the parameters can be manipulated in a straightforward manner and thus through the depth, frequency and constant delay generate all the desired auditory effects from a single architecture.

• Delays are commonly used in signal processing, in audio, they are used as basic building block to create audio effects, artificial reverberation, or to compensate sound reinforcement systems. Also delays in discrete time, as is in the digital world, are used as integrator allowing creating physical models such as spatial effects.

• Simulink is a friendly environment for simulation and model-based design providing an interactive graphical interface customizable from the set of block libraries that let create, simulate, implement and test time-varying systems such as signal processing.

• The hardware development kits have become in more than field programmable gate arrays consisting not only on look-up tables, register, multiplexers and distributed memory, but also dedicated circuitry for fast adders, multipliers, and I/O processing. Also the capability for implementing parallel arithmetic architectures makes the hardware suitable for implementing high performance DSP systems such as digital filtering, fast Fourier transforms, etc... The FPGA exceeds for far the clock rates at microprocessors or DSP processors running at two to ten times faster at the FPGA.

REFERENCIAS

[1] MÜLLER, Meinard. Information Retrieval for Music and Motion. Berlin: Springer, 2007. 31 p.

[2] GOLD, Ben and MORGAN, Nelson. Speech and audio signal processing. Nueva York: John Wiley & Sons, Inc, 2000. 1 p.

[3] ROSSING,Thomas D., MOORE,F. Richard y WHEELER, Paul A. The Science of Sound. 3th Edition. San Francisco: Addison wesley, 2002. 674 p.

[4] DAVIS, Don and DAVIS, Carolyn. Sound System Engineering 2nd Edition. Elveiser: 1997. 361-380, 423-468 p.

[5] DIAZ, Andrés Felipe and CADAVID, Jose Julian .Consola virtual de procesamientos de señales para instrumentos musicales con visualización musical por medio de imágenes SP Project. Bachelor dissertation, Engineering School, Electric and Electronic engineering, Universidad Pontifica Bolivariana 2008.

[6] CORREA, Juan Diego. Electrónica Musical. Bachelor dissertation, Engineering School, Electric and Electronic engineering, Universidad Pontifica Bolivariana 1999.

[7] STRELLIS, Chris. FPGA Synthesizer [Online]. Essex: Strellis.com [Access march 25, 2011]. Available at: http://www.strellis.com/fpga.shtml

[8] VEGA, Rafael, ARBELAEZ, Paulo, MEJIA, Carlos and VANEGAS, Mauricio. Módulo de reconocimiento de voz basado en los parámetros de predicción lineal [Online]. Montería: Vanegas, Mauricio [Access march 25, 2011]. Available at: www.upbmonteria.edu.co/electronica/gapa/descargas/Jornada/Memorias%20FPGA/Paper/CIDI_PAPER/Pper2.doc

[9] Rane Factory Store. Pro Audio Reference [Online]. Letter H [Access march 25, 2011] Available at: http://www.rane.com/par-h.html

[10] Memidex. Free online dictionary with encyclopedia definitions of acoustic [Online]. [Access march 25, 2011]. Available at: http://www.memidex.com/acoustics+physics

[11] WASSERTHAL, Christian and KUBERSKI, Bjorn. Phychoacoustics . OTTO-VON-GUERICKE-UNIVERSITÄT MAGDEBURG. November 2007.

[12] FASTL, H. The Psychoacoustics of Sound-Quality Evaluation. Institute of Man-Manchine-Communication, Technical University Munich Arcisstr. ACUSTICA – acta acustica. S.Hirzel Verlag – EAA vol. 83 :1997. 754-764 p.

[13] MACCALLUM, John, HUNT, Jeremy and EINBOND, Aaron. Timbre as a psychoacoustic parameter for harmonic analysis and composition [Online], Center for New Music and Audio Technology (CNMAT), University of California Berkeley [Access march 25, 2011]. Available at: http://www.music.mcgill.ca/~ich/research/misc/papers/cr1330.pdf

[14] ZOLZER, Udo. DAFX, Digital Audio Effects. John Wiley & Sons, Ltd. 2002. 63-73 p.

[15] J Comp Physiol A. The auditory cortex [Online]. Springer-Verlag 1997 [Access march 25, 2011] Available on: https://www.utdallas.edu/~kilgard/Ehret_AC_reviewJCompPhys.pdf

[16] Xilinx System Generator v2.1 For Simulink [Online]. User Guide, Xilinx Blockset Reference Guide [Access October 1, 2011]. Available at: https://safe.nrao.edu/wiki/pub/CICADA/WebHome/xilinx_ref_guide.pdf

[17] Xilinx. MicroBlaze Development Kit Spartan-3E 1600E Edition User Guide [Online]. [Access October 1, 2011]. Available at: http://www.xilinx.com/support/documentation/boards_and_kits/ug257.pdf

[18] Datasheet LTC2624 [Online]. Linear Technology [Access October 1, 2011]. Available at: http://cds.linear.com/docs/Datasheet/2604fd.pdf

[19] Fragment taken from the Matlab Help documents.

Resources Create for this work:

[20] Haas Phenomenon implemented on Pure Data

[21] FIR_Comb_Filter.mdl Example implemented on Simulink

[22] FIRcombfilter.m file show frequency and zero poles graph

[23] IIR_Comb_Filter.mdl Example implemented on Simulink

[24] IIRcombfilter.m file show frequency and zero poles graph

[25] Echo.mdl Example implemented on Simulink

[26] Vibrato.mdl Example implemented on Simulink

[27] Flanger.mdl Example implemented on Simulink

[28] Chorus.mdl Example implemented on Simulink

BIOGRAPHY

Jaime Andrés Laino Guerra.

Born in May 27 of 1988, in Barranquilla city (Atlántico). Bachelor from the school INSTITUTO JORGE ROBLEDO on 2006. Actually is ending his studies of Electronic Engineer at Universidad Pontificia Bolivariana from Medellín. E-Mail jlainog@gmail.com / jlainog@hotmail.com .