Audiovisual Tool for understanding Audio concepts for being used in bachelor's degree programmes

Antoñanzas, Christian[a]; Gutierrez-Parera, Pablo[a]; Simarro, M. Ángeles[a] and Belloch , Jose A.[b]

[a]Instituto de Telecomunicaciones y Aplicaciones Multimedia, Universitat Politècnica de València, Spain . [b]Depto. de Ingeniería y Ciencia de Computadores, Universitat Jaume I, Spain.

Abstract

In the Audio Signal Processing field, there exists difficulties in order to explain different concepts such as compression, masking, quantization, sampling, among others. Further, most of these concepts require the use of audio laboratories and multiple practical session that must carry out students. Another issue is that there are students that are not able to internalize these concepts straightfordly and require more practical sessions. In order to address these problems, we have developed an audiovisual tool, designed with Matlab, that can be used for professors and students. This tool allows to analyze, test and apply the audio concepts to real audio signals. The developed tool has been successfully experienced by professors of the audio signal processing field that recommend its use in upcoming academic courses.

Keywords: *Audio Processing; Matlab; Software Interface; Teaching Tool.*

1. Introduction

Learning and teaching are two concepts that are highly related. However, the way we teach is changing day to day. Different works such as Shyr (2010), Bonwell (1991) and McKeachie (1994) indicate that students retain much more when they directly experiment at laboratory in contrast to the traditional classroom lectures where they only see or listen concepts. In agreement with the above mentioned perspectives, the field of education has been affected by Information and Communication Technologies (ICT's) which contribute to improve student learning and better teaching methods. Through ICT's, teachers can easily explain complex concepts and ensure student's comprehension. Furthermore, they are able to create interactive classes and make the lesson more enjoyable, which could improve the student attendance and concentration.

In this context, it is important to note that teaching audio concepts can be difficult for understanding if there is not a proper laboratory and all lectures are given in a theoretical way. Moreover, audio concepts are full of abstractions and signal theory that can tire and demotivate the students -see Christensen (2010), Hanrahan (1998), and Henning (2007)-. Practice lectures in audio processing are totally mandatory. However, there is occasionally schedule limitations that prevent to carry out the necessary practical sessions. As a consequence, multiple concepts are misunderstood by the students, which are not able to carry out a complete learning. Moreover, there exists universities that do not allow students to access to laboratories outside lectures schedule by themselves, which means an important restriction.

To solve this issue, we have developed a audiovisual tool that has mainly two applications: to show different audio concepts in order to help students to understand this concepts from a perceptual (visual and sound) point of view, and, at the same time, to help teachers in their teaching activity. One of the most advantages of this tool is that students are able to execute it in their personal computers and future versions will be available in mobile devices.

The paper is organized as follows. Section 2 describes the software and the design used for the development of this application. Section 3 explains in detail the four audio concepts that can be learnt with the proposed tool. Finally, conclusion remarks are given in Section 4.

2. Application features.

Currently, a significant number of signal processing researchers and engineers develop their algorithms under MATLAB (2015). MATLAB offers and excellent signal processing development environment because of the "DSP System Toolbox". Moreover MATLAB offers Object Oriented Programming (OOP) that simplify the development . The presented

tool provides the user with a friendly graphical interface (GUI). The fact of using MATLAB allows the tool to be run under different operating systems. Another important factor is that this tool does not requires previous knowledge about MATLAB programming.

The GUI represents the framework that allows the user interaction. As is appreciated in Figure 1, the GUI is divided in three blocks. At the left hand side, four applications that are linked to four different audio concepts are presented. When the user selects one of these concepts, it appears at the center of the GUI a description of this concept and different options that depends on the concept to evaluate. Normally these options are composed of different sounds such as noise or a signal generator, among others. In the following section, we will specify the characteristics for each implemented concept. At the right hand side, the user will appreciate three plots that will show the waveform in different domains: time-domain (top left), frequency-domain (top right), and spectrogram (bottom).
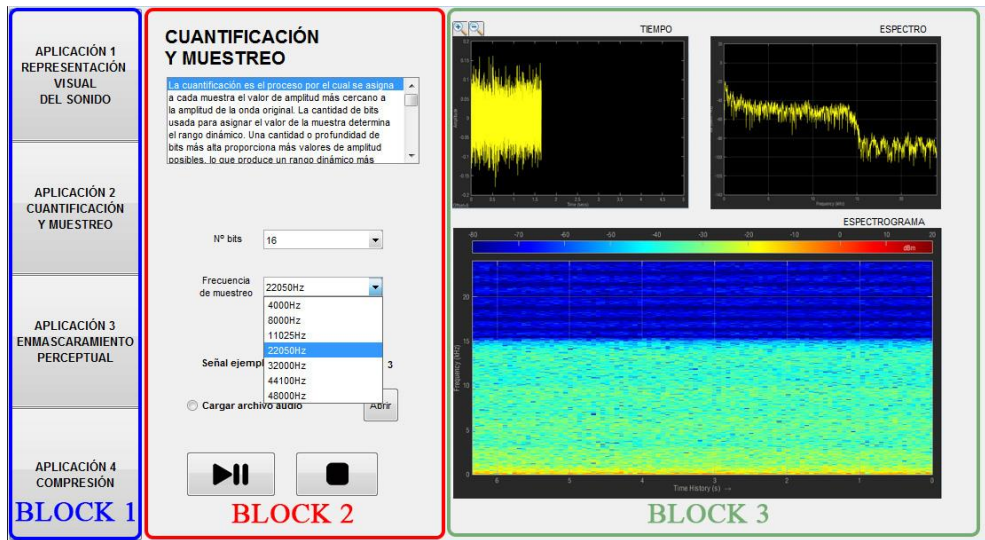

Figure 1. Main interface of the implemented tool.

3. Audio Concepts and Hands-on approach

During an audio course, the students must study multiple and different audio concepts. Among them, there exists audio concepts that are difficult to understand since they are plenty of mathematical formulations such as visualizing a signal in multiple domains, quantification and sampling, masking and compression. These four concepts are tackled in four different applications inside the tool from a perceptual (visual and sound) point of view as follows:

3.1. Visual representation of audio signals (time and frequency domain)

The temporal representation of a signal shows the time evolution with respect to an amplitude level and it is usually the most common way to see a waveform. But, in some cases, the frequency representation can help us to understand better some characteristics of audio signals since we perceive sounds from a tonal content point of view. The mathematical formula that transforms time domain functions to frequency domain functions is the Fourier Transform -Oppenheim (1975)-. A spectrogram is a useful visual representation where frequencies and amplitudes of audio signals are displayed at the same time as its time evolution. Figure 2 shows a capture of the first application of the tool that corresponds to the visualzation of signals in different domains. Spectrogram shows the amplitude at various frequencies (usually on the vertical axis) at a specific time (horizontal axis). The colors of the spectrogram indicate frequency power levels. Usually, bright warm colors indicate frequency content with higher power and dark cold colors indicate frequency content with very low power. A strong red horizontal line indicates the existence of a tone at a concrete frequency.

This application allows to generate sinusoid signals introducing its frequency and amplitude in an intuitive way. In addition, there is a noise generator that allows to select among different types of noises. Finally, this tool includes also the possibility of loading audio files in different audio formats.
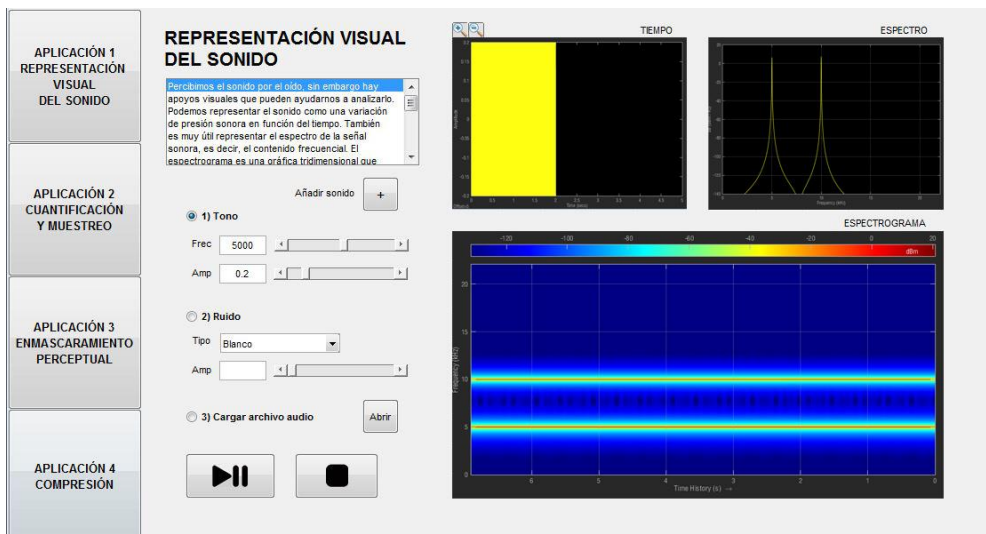

Figure 2. Capture of the first application of the tool that corresponds to the visualzation of signals in different domains.

3.2. Quantization and sampling

As described in the previous section, audio signals can be represented as a function that depends on amplitude and time, both continuous values in nature. However, to store audio signals into a digital device, as a computer, it is required to convert the continuous signal amplitude and time values into a discrete representation.

This occurs because digital devices store signals using finite numbers of bits. Thus, both amplitudes and time values can take discrete values. The process of digitizing the time is denoted sampling and the process of digitizing the amplitude is called quantization.

Sampling is the process of recording an analog signal at regular discrete moments of time. The sampling rate is the number of samples per second and the time interval between samples is called the sampling interval. Sampling allows us to convert continuous-time signals into discrete-time samples without loss information if the sampling rate is high enough. The Nyquist-Shannon sampling theorem states that the sampling rate for exact recovery of a signal is larger than twice the maximum frequency of the signal.

Quantization converts the amplitude of a signal in a finite set of values. Unlike sampling, it is generally irreversible and causes loss of information. As a consequence, quantization introduces distortion into the signal (that is denoted in the literature as error quantization). This distortion indicates the amplitude difference between the samples and their quantized versions.

The second application allows to select between quantization and sampling. The tool allows to select the number of bits for the quantization option and the sampled frequency for the sampling. In both cases, the user can select between different sample signals or load an own audio file.

Figure 1 shows a capture of the second application of the tool where it is illustrated the difference between the original signal and the processed signal in terms of time, frequency and spectrogram.

3.3. Auditory masking

The masking is a well-known and common phenomenon that frequently surrounds us. This phenomenon occurs when a louder sound (masker) occurs at the same time as a softer sound (masked). As a consequence, the louder sound prevents to hear the softer sound. This phenomenon is known as frequency masking. When the masker and the masked are not presented simultaneously, we refer to this phenomenon as temporal masking. In addition, masking can occur before the presence of the masker (pre-masking) and after the presence of the masker (post-masking). It is important to note that the masking is related with the hearing threshold that represents the lowest sound level that we can hear at a given frequency.

The third application allows to visualize and hear perceptual masking effects. The user can select amplitudes, both of noise and of a click signal and change them in real time. Figure 3 shows a capture of the third application related to the masking phenomenon.
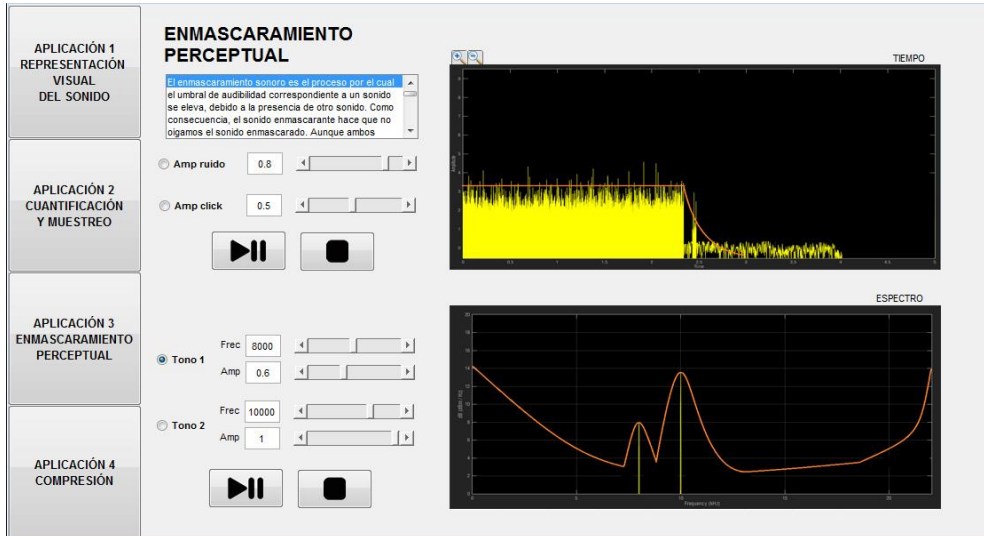

Figure 3. Capture of the third application of the tool where it is illustrated how the perceptual masking affects to a click signal and how a tone can perceptually mask another one.

3.4. Audio compression

The main objective of a perceptual audio compression system is to reduce the quantity of digital audio data by considering the manner in which humans hear. There exists two types of compression: lossless and lossy compression. In lossless audio compression, the decoded signal is exactly the same as the original one bit by bit. Examples of these audio formats are FLAC (Free Lossless Audio Coding) and WV (Wave Pack), among others. In lossy audio compression, the decoded signal is different as the original one. In fact, some data are removed. However, humans are not able to appreciate this effect. Examples of these audio formats are MP3, WMA (Windows Media Audio), among others.

It is interesting to highlight two important concepts in this fourth application: Compression ratio (the relation between the bits of the original signal and the bits of the compressed signal) and Bit rate (the number of bits that are processed per second). Note that a high number of bits involves more information and therefore more quality, but also more data to be stored in memory. Moreover, besides using different compression or bit rates, the design of the coders is based on different auditory phenomena. Figure 4 shows a capture of the

fourth application where the user is able to select among different types of encoders and bit rates. As in previous applications, the user can load its own audio files and the tool displays the differences between the original signal and the compressed signal in terms of time, frequency and spectrogram.
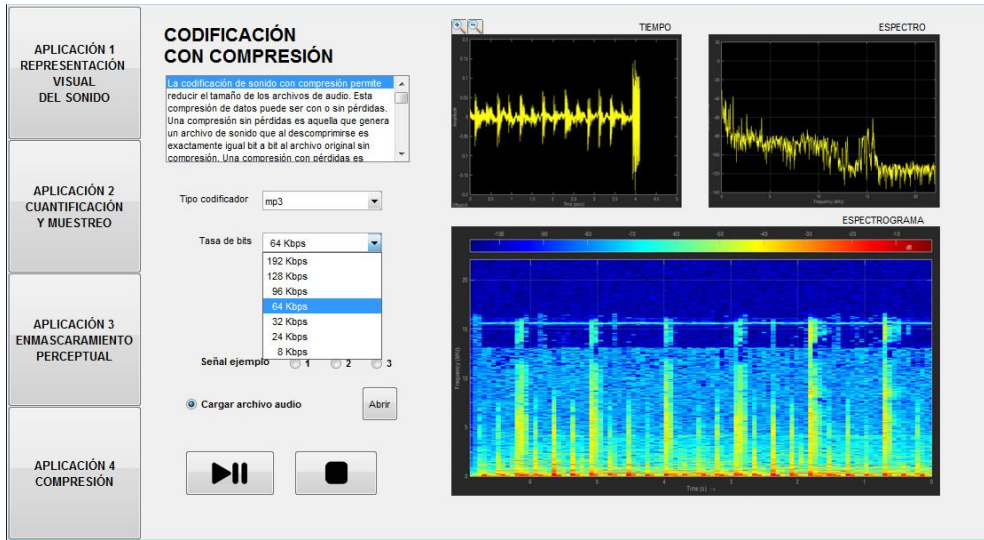

Figure 4. Capture of the fourth application where the user is able to select among different types of encoders and bit rates.

4. Conclusion

This work has presented an audiovisual tool composed of four applications that analyze and explain four audio concept such as visualizing a signal in multiple domains, quantification and sampling, masking and compression. The use of this application allows to teach these concepts more easily and help students to better grasp the concepts. As a future work, different improvements will be carried out. For example, this tool will be developed for being used also in mobile devices. Thus, the students may use their own mobile phones during class without being in a computer room. Other possibility consists in implementing others audio concepts such as filter design, delay-based effects (flanger and phaser), reverberation and spatial sound, among others. For this purpose we must consider if other theoretical audio concepts need a visual aid to understanding.

Acknowledgements

References

Bosi, M. & Goldberg, R.E. (2013). *Introduction to digital audio coding and standards.* Springer Science + Business Media.

Bonwell, C. C. & Eison, J.A. (1991). *Active learning: Creating excitement in the classroom*, Washington C. C. George Washington University

Christensen, C. M., Horn, M. B. & Johnson, C. W. (2010). *Rethinking student motivation: Why understanding the 'job' is crucial for improving education*, Boston: Innosight Institute.

Hanrahan, M. (1998). The Effect of Learning Environment Factors on Students' Motivation and Learning, *International Journal of Science Education*, 20, 737-753

Henning, M. (2007). *Student's Motivation to Learn, Academic Achievement, and Academic Advision* (PhD Thesis). New Zealand: AUT University

Matlab (2015). Online: http://se.mathworks.com/products/matlab/

Mckeachie, W. J. (1994). *Teaching Tips: Strategies, Research, and Theory for College and University Teachers, 9th ed*. Lexington, MA: Cengage learning.

Oppenheim, A. V. & Schafer, R.W. (1975). *Digital Signal Processing,* Prentice-Hall, Englewood Cliffs, New Jersey, USA.

Shyr, J. W. (2010). Integrating laboratory activity into junior high school classroom, *IEEE Transactions on Education*, 53, 32-37.