Audio Engineering Society

Convention e-Brief 344

Presented at the 142nd Convention
2017 May 20–23 Berlin, Germany

Design of an algorithm for VST Audio mixing based on Gibson diagrams

Belman Jahir Rodríguez Niño[1], Marcelo Herrera Martínez[2]

[1] *Universidad de San Buenaventura sede Bogotá, Cra 8H No. 170-20, Bogotá, Colombia*

[2] *University of Iceland, Taeknigardur, Dunhagi 5, Reykjavík, Iceland*

Correspondence should be addressed to Marcelo Herrera Martínez (marcelo@hi.is)

ABSTRACT

This project consists on the creation of a plugin on the Ableton Live platform with the aim of providing visually the audio mixing process in real-time. The software programming is developed on Max for Live – a program to establish the link between Max Msp and Ableton Live. The plugin is assigned for each channel, with the aim of visualizing the correspondent sound to a "sphere" object on a 3D window and there to observe the variations in real time of loudness, panning and frequency analysis, based on the David Gibson´s interpretation, on his book "The Art of Mixing".

1 Introduction

The use of interactive algorithms and programs has been on development in the musical production industry [1], [2]. On a daily basis, a vast number of programs are being optimized, in order to be used for end-users and firms dedicated to audio software commercialization, as well as persons dedicated to the creation of free software. At the moment, there are high quality audio interactive audio software applications. Advantage of audio manipulation is to handle signal characteristics as amplitude, phase and frequency, along with other options, in this case visual through algorithms in order to achieve real world simulations. Audio software count with extensions or subprograms, so-called ¨plug-ins¨, developed to optimize the mixing process; they give to the user graphic interfaces which simulate control and fader elements which originally belong to the real audio hardware. Mixing process is one of the key aspects difficult to perform in the production process, due to the fact that it depends on the expertise and experience while performing such a process. We interrelate with sound on two manners: We feel (and hear) the physical sound waves which are loudspeaker outputs, and we imagine the correspondent coloration. The problem is that people rarely differentiate among the individual parts of the pieces which form a complex recorded musical piece. For this reason, the development of an algorithm which enables the visualisation of the mixing process becomes necessary.

2 Development of the algorithm

The algorithms were developed in Max Msp for (Demo Version); this is a joint between the audio production software Live 8 from Ableton and the programming environment Max Msp 5.1. The logic of the algorithm for the project is observed in the following scheme. This algorithm will be applied independently to each channel. Each channel represents a different sound according to their correspondent spatial location. Audio signal –inside the Live 8 environment- enters through a patch to the Max Msp programming platform. There, the signal is processed according to the Gibson´s visualisation paradigm.

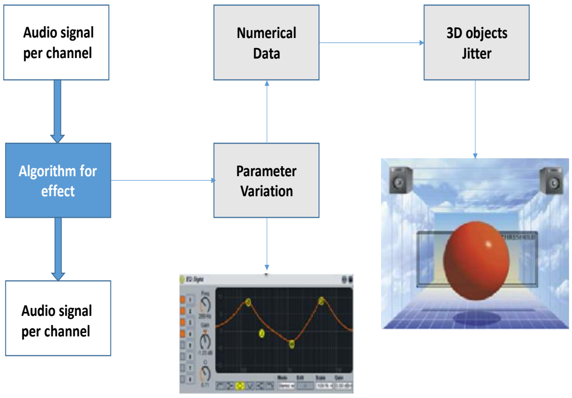

Figure 1. General algorithm.

Inside the Max Msp for Live, audio signal may be transformed to numerical data, in order to be handled according to the user´s needs. Additionally, Max Msp for Live counts with Jitter, a subprogram which handles all the visual environment. Finally, after this algorithm is performed, the audio signal comes back to the Live environment in order to be manipulated from there, without the necessity to enter to the Max Msp environment. At the initial performance, while linking Max Msp to Live, only the last versions from Live 8 and forth have to be available and Max Msp 5.1. When these programs are installed, while opening Live, it will recognise the Plug-Ins developed in Max Msp and rename it into a newer extension ".amxd", which only will work inside the Ableton platform Live 8, and forth. The logic behind the algorithm designed inside Max Msp is explained further. The audio signal processing individual parts are: Input Equalizer algorithm, FFT algorithm, Panning algorithm and Volume algorithm. Finally, the creation of Jitter Visualizer process is explained, which contains every previous algorithm. Stereo signal which comes to the Max Msp programming environment through Plug-In object. This signal enters an input equalizing algorithm. This one has the function to enable o the user the signal characterization through 1-band equalizer. For the sake of illustration, a bass guitar must be processed according to the physical nature of the instrument. First algorithm is the equalization, which will change the type of the signal for subsequent processes.

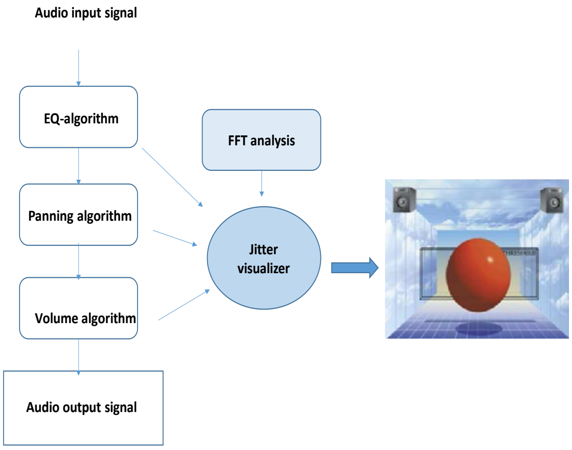

Figure 2. Flux diagram of the general algorithm.

A FFT frequency analyser receives the monophonic signal, already equalized, and assigns to it a definite colour depending to the sound nature of the channel. This in order to characterize or differentiate the sound from each channel, according to Gibson´s theory, some frequencies are associated to some definite colours, as it is shown in Figure 3. An algorithm for the interpretation of the sound Frequency Response, coming from the correspondent channel in order to associate it to a definite colour in the "sphere" objects is designed.

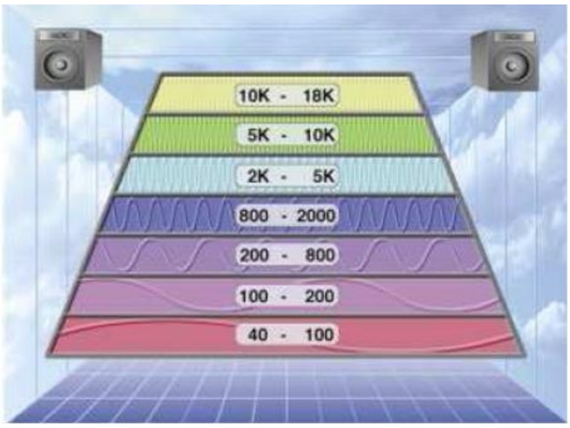

Figure 3. Colour classification according to the frequency.

In the 3D visualisation window, for instance, if in the channel where the Plug-In has been assigned, there is a bass, its colour must be represented to a

frequency between 100 and 800 Hz, which is the violet. Besides, the input equalizer, according to its configuration will help to define the colour and the sound for each channel.
After analysing the frequencies characteristics of the channel, the amplitude value summation is received (from 20 Hz – 320 Hz) and this averaged summation is sent to an object which assigns the colour to "sphere" objects. If the sound characteristic of the channel contains a register in low frequencies, then a reddish colour will be noted. The middle range frequencies register is taken from 640 Hz till 2 kHz, making to it an assignation as blue colour, and finally the high frequencies will be from 4 kHz to 20 kHz with a green colour –everything according to Gibson´s theory. Afterwards the stereo signal passes to the panning algorithm, which enables to receive the signal for further modification by the user in the stereophonic space. The output signal of this algorithm will have equalization and panning. The algorithm varies the channel level or volume, in a stereophonic manner, in order to take back the processed audio signal to the Live environment. All these algorithms imply changes to objects in Jitter, which are the ones in charge to convert the audio information into video.

3 Results

The creation of a virtual VST processing system was achieved; its purpose is to achieve a signal transformation of the sound signal by a graphic interface which facilitates the audio mix spatialization VST, through 5 different processes per channel.

4 Analysis of colours by pure tones

By pure tone insertion per channel, and the assignation of the project plugin, there may be compared the results of colour frequency with the theory exposed in Gibson´s book. The used frequencies in this test are: 50 Hz, 200 Hz, 500 Hz, 1000 Hz, 5000 Hz, 8000 Hz and 15000 Hz. The vertical disposition is according to the frequencies that were used in this test, for this reason the objects are not symmetrically separated. Colour theory according to Gibson´s theory may be compared to the results obtained in the Live session, inside the visualisation window of the Plugin. The colour is determined by the FFT analyser, which assigns a colour depending on the frequency.

5 Volume and panning analysis

By the same session with pure tones, a depth distribution is performed through the knob gain, in order to observe the volume levels and their position in the visualisation window. This analysis was done for the whole range for the volume in the plugin.
Volume levels are distributed symmetrically; it is noticed the objects´ depth – by their radius -. The less their volume, the less the sphere radius. Some objects keep active the option "edge" which acts as transparency. As for instance, the closest sphere keeps a different render mechanism with respect to the next, so to be distinguishable. The range of panning values covers the whole stereophonic space; in the horizontal extremes of the visualisation window, 50L for the left-limit and 50R for the right-limit belong. The frequency values are the same for the previous tests, ordered from less frequencies at the left to higher frequencies at the right. Continuing with the results analysis for the panning parameter, we may confirm that the values which control the sphere radius are dependent of the Knob Gain of the input equalizer, due to the fact that this parameter increases the frequency level that has been determined. In an implicit way, this value increases the object spatial amplitude. The spatial disposition of this analysis was randomized in order to show the visualisation possibilities of the objects. When they are sharing the same physical space, then the sphere radius ranges from larger to smaller values; in other words, from higher gains to less gains at the input equalizer. Also there were used pure tones from previous sessions, keeping the low frequencies from left to right, as it was seen previously in the volume analysis. The radius value is only controlled by the input equalizer gain.

6 Control analysis through camera

Final results are tested; vertical and horizontal movement control of objects through a camera: the detection process of movement for this algorithm recognizes a determined colour by the user.

Figure 4. Volume analysis of the plugin.

In the upper left window of the figure 5, click is done in the desired colour for movement capture.
While this action is in progress, at the left bottom window all the colours get filtrated and at the right window, the movement detection is observed through a central orange point; in this point camera is focused for detection and parameter variation which will actuate the parts of the general algorithm exposed in the development process. In the upper left part, there is a button which activates and deactivates the plugin control by the camera.

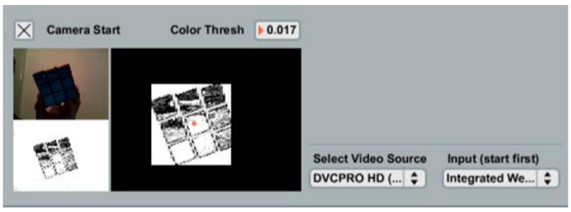

Figure 5. Control by movement.

Additionally, there is a numerical control "Colour Thresh" which varies the capture radius of the screen movement. In other words, the larger the number, the smaller will be the ability to recognize the desired colour for the movement; the smaller the number, the stricter will be the colour filter, which will enable a better control for the plugin. It is necessary to keep the plugin section performance under an environment where no big colour contrasts will be experienced; this means in a definite background, and a colour different from the background.

7 Conclusions

The design of plugin virtual devices is not limited to sound quality improvement or recreation of traditional instruments. The key motivation for this task is to enable the use of a definite software which could use these plugins in all working stations, in order to spread out information and ideas in the virtual recreation of audio mixes. The presented plugin may be manipulated in Ableton Live working stations. It is possible to recreate virtually the sound spatial disposition of the mixing process, making use of the advantages of Max Msp Software, which creates panning, processes and automations in real time. These virtual sound architectures describe and recreate the environments that are rendered from the visual side, generating a correspondence between what is being seen and what is been heard. The GUI (Graphical User Interface) format of Max for Live enables to generate an intuitive user interface, which is appropriate for the Live environment and easy to use by multiple users, due to the fact that these virtual systems may be distributed through massive digital formats as Internet or compact discs.

References

[1] M. Herrera, "Compresor perceptual basado en la transformada Wavelet-Daubechies Tipo IV con dos niveles de descomposición". In: Congreso Iberoamericano de Acústica, FIA, Valdivia, Chile, 2014.

[2] A. Cruz, M. Herrera, "Frequency band displacement for optimizing acoustic boxes above the natural frequency of the loudspeaker". In: XIX Simposio Internacional de Tratamiento de Señales, Imágenes y Visión Artificial - STSIVA 2014.

[3] D. Gibson, *The art of mixing – A visual guide to recording, engineering and production.* 1997.

[4] K. G. Pohlmann. *Principles of Audio Digital.* 4th edition, McGraw-Hill, 2000.