CompoVOX: REAL-TIME SONIFICATION OF VOICE

Daniel Molina Villota
University of Málaga, Spain
Jean Monnet University, France
danielmolinavillota@gmail.com

Isabel Barbancho, Antonio Jurado Navas
University of Málaga, Spain
ibp@ic.uma.es, navas@ic.uma.es

ABSTRACT

It has been developed an interactive application that allows sonify human voice and visualize a graphic interface in relation to the sounds produced. This program has been developed in MAX MSP, and it takes the spoken voice signal, and from its treatment, it allows to generate an automatic and tonal musical composition.

1. INTRODUCTION

The main objective of any interactive application, is to strongly move the viewer to play with that application and to show him novel aspects. For that reason, using voice as a generator of tonal musical sequences, can be interesting both for composers and amateur users, since they can explore aspects of the voice that are not usually threated for sound generation.

Voice is the most important tool of human communication, but also the most important musical instrument, therefore, the non-daily use of voice (whether communicative or musical) based on a synesthetic treatment as well as not synesthetic but with both at same time can be quite interactive.

CompoVOX has been developed in MAX MSP using the tools of data processing, generation of visual effects and sound effects offered by this programming method.

2. TALKING ABOUT INTERACTIVITY

In the last two decades, there has been a great technological advance in processing capacity, therefore use of technologies applicable to the numerical arts has changed very much. Here they are mentioned some interesting works related to this work.

The importance of interactivity (directly or indirectly) can be seen in projects like Opto-isolator [1], who induces the viewer to a high interaction, it's proposed that it would be the work itself who observes the viewer. Otherwise there are project called Re: MARK, who employs voice to create image employing voice's analysis (identification of phonemes), and the movement of the participants to produce real-time animations. The importance of an interface who allows viewer to play actively is very important. projects like WIP [4] show the relevance of this fact. WIP is a project who takes curve and the amplitude of sound to allow in real time generate multiple visual combinations, projection of the geometric shapes and appearance modification.

This work has seemed inspired by the fact of being very attractive visually, that is achievable using real time synesthetic and no synesthetic procedures to generate sound and image, and at same time creating a tonal music sequence from voice and taking to generate visual forms related to the sound performed on the screen shown in the installation.

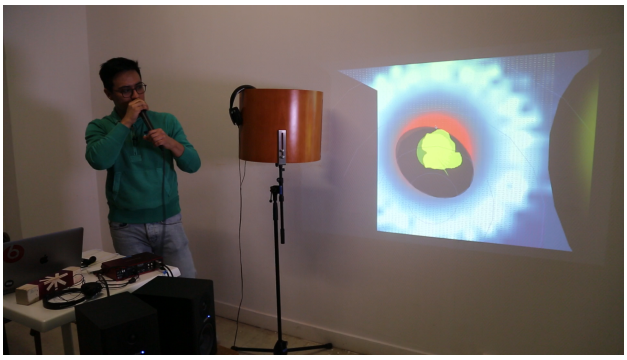

Figure 1. Doing a demo of CompoVOX

3. SOUND SYNTHESIS

To carry out the generation of sound, this project has made use of the acquisition of voice through a dynamic microphone. This microphone allows to isolate the environment and in this way, it is possible to focus only on the capture of voice, also an audio card is used, then the signal is captured by the microphone and transferred to the computer through the audio card.

This project has been developed through a program built under the structure of MAX MSP, the program is responsible for taking the audio signal and through different stages of treatment, obtain a signal that can be translated into midi values and for so in a tonal musical sequence.

At first, this program uses the fast Fourier transform, in this way, the region of interest for this project is filtered, that is the spectrum of the human voice that ranges between 50 and 600 Hz. Once this process has been carried out different Fourier filters are performed in parallel, which focus on taking only the signal of certain regions of the spectrum. To define such regions of interest, several tests have been made by differentiating the levels reached

in each region and that level is responsible for activating different areas of the musical scale, usually the low notes activate lower sounds and the high notes will activate higher sounds.

It is evident that the levels of each one of the signals that are obtained will vary rapidly, since they come directly from the voice signal, therefore to smooth the changes that are in the signal an averaging function is used, this averaging avoids sudden variations in the control signals, thus eliminating intrusions of sound from the environment and at the same time a signal that varies more slowly facilitates the control in real time of musical parameters that must be audible to the user.

The entire system is controlled by the same clock, but each synthesis stage is activated only when an appropriate level range is reached in the indicated region of the spectrum and in that case, there will be a mapping (by means of a scaling between frequency and level) of midi notes, all sound generation systems are passed through a tonal filter, which forces the system to have some regularity and stay in the same key.

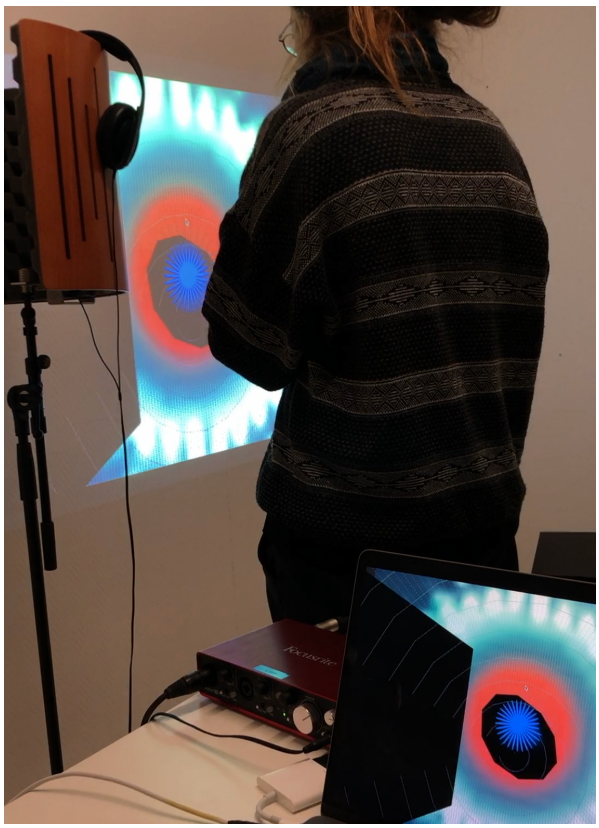

Figure 2. Student playing with device

The different parts of the signal are used both to control the type of note and to control the attack and the temporal length of the note. A control parameter is assigned to each signal. On the other hand, sound synthesis is also done by taking the central frequency of the voice and its variations to generate a sequence of notes that is passed through a tonal filter, and subsequently the frequency value serves as a tone control parameter of a synthesizer.

For the graphic interface, several aspects are used. A background treated by MAX MSP that is responsible for reproducing a space environment and whose edges come from noise, a conical gang that covers the space and that varies in sharpness and size proportionally to the central frequency of the voice, a flat circular object located in the center it takes different forms correspondingly to the waveform of the voice, and objects that change shape according to the level in a region of low frequencies of the spectrum. The whole system moves proportionally to the level of sound, that is, the louder the voice, the closer the objects are, and the fainter they are, the farther they go. The frequency is also responsible for changing the color of the objects.

4. PERFORMING THIS DEMO

During the presentation, the sounds that the person makes with the microphone will be captured. In real time, the participant's voice will be used to filter and give different colors to the sampled sounds. The voice will also be used as a mean of controlling and generating patterns / sequences that will be reproduced by the sequencer. This filtering and sequence will generate in a synesthetic manner a visual analogue that will show the details of the spectrum of the participant's voice. Additionally, the participant will be able to play with a rhythmic base.

5. CONCLUSIONS

Sonification of voice signal for the generation of tonal musical sequences accompanied by a graphic interface that relates some aspects of the voice with sounds and images, is highly interactive and allows exploring aspects of sound generation that are not traditionally used in the voice treatment. CompoVox allows to work voice signal in a different way from traditional use in music.

6. REFERENCES

[1] G. Levin, “OptoIsolator”. Project at M.I.T Massachusetes technological Institute. 2007.

[2] Z. Lieberman, G. Levin, “Re: MARK”. Project for FutureLab of Ars Electronics (Siemens). 2002.

[3] N. Nordmann, J.R.Valentin, “Variationen für Tenorion”. Project presented at Kunstvolkslauf Hannover, Zinnober in 2013.

[4] B.Guesnon, “W.I.P Work.In.Processing”. Project by himself begun at 2013