Suite

audio-video interaction for 8 self-observing audio files

Stereo rendering at: https://youtu.be/BNzoDeourno

Suite is a formal self-regulating method of composition.
Eight stereo files are processed, mixed and spatialised in real-time.
The resulting audio output feeds a sound analysis module upon which, in a ring fashion, the algorithmic processor creating the sound-video composition is based.
The internal analysis-synthesis loop is deterministic but the video-sonic result is never the same, although showing asymmetric recursions.
The module continuously analysing the overall sound output is called **"Observer"**.

The Observer interprets and slices the sound signal regressed as a stream of features representing:
1) the overall amplitude, fundamental frequency, brightness, roughness, noisiness, onset time-point,
2) the local amplitudes of the 8 lowest Bark bands (the energy content detected inside the range 20-920 Hz, segmented as 8 perceptually relevant frequency bands).
-The overall sound qualities (1) are tracked in real-time;
-The Bark amplitudes (2) are instead analysed in their behavioural flow, taking in account their individual stability, prominence and gait with respect to their own short-time "history".

This streaming and complex analysis vector is assigned to a dense net of algorithmic decisions and nuances operating upon the 8 stereo files, forming in this way the final audio.
A similar net processes 8 fix images producing the parallel video rendering.
The audio output thus obtained is in turn sent to the analyser (the Observer), circularly feeding the algorithmic "decision-making" process.
In this way this structurally closed system is based on the absolute coincidence of input and output.

Even if the compositional internals are strictly shaped, the result shows an organisationally open behaviour, enhanced in addition by the fact that even the most detailed analysis cannot apparently avoid reductions and distortions, being a conceptual representation. Besides the metaphor of a virtual embodied knowledge of the sound upon itself, this dynamic system shows different levels of resistance between its input sound matter and its human-organised reading, computationally re-assembled inside the compositional machine. More than exploring the new-cybernetic idea of a self-aware system able to grow and behave, this work is intended as **a study upon the concept of instrument,** at the edges of its interesting boundary state of a complete input-output conjunction.

The crucial composition algorithms of *Suite* mostly rely on the Bark-bands energy behaviours (agencies) with respect to their recent time evolution (spanning up to 2 seconds).
Obviously the length of 2 seconds is not enough to speak of "history", but it is a sufficient feedback time for disengaging the mechanism from a straight real-time dimension. In this way the compositional decision-tree acts as the consequence of a virtual short-term action-reaction domain, in this way taking into account the mediating dimension of temporal expressivity, which could be defined as "the time of the performance".

Since these 8 audio files globally self-process themselves through audio analysis and automations, we could say that the musical result is a sound output and an instrument at the same time. A similar loop grounds the concept of the hyper-instruments, inside which the performative actions (of the living instrumentalist) feed the output sound and the methods of machine processing in one take. In this sense the compositional *Suite* ecology could be viewed as a virtual hyper-performer.

The affordance of this interactive feedback loop in the context of a hyper-instrument can be often motivated by the aim to extend the compositional tasks of the performer, or to dynamically ground complex self-emergent structures to perceptible human performative gestures.
This kind of interaction puts the performative gestural dimension of a living sound at a same level of information trade as the compositional dimension of abstract structural choices.
In other words it allows direct synergies between the low levels of signal analysis and control (i.e. sonic parameters, raw energies, physical interactions and modulations) with the mid levels of structural decisions (i.e. musical patterns, regularities, directions, densities, repetitions, formal interactions) which are considered as linguistic-compositional tasks.

The choice of exploring the interactive potentials of fixed sound files inside an automatic and autonomous performative time domain, is linked to the operational necessity to aesthetically test the consistency of infrastructures which elicit non obvious paradigms of composing by listening, and symbolisms emerging from energetic perception/action gaps.

HISTORY OF THE COMPOSITION
The 8 sound files were originally composed as short electroacoustic commentaries for the play "Il Padre de li Santi"[1] (The Father of the Saints).
The concrete, satiric and psychoanalytic contents of the show, and in addition the theatrical requirement of short occasional sound commentaries, suggested me the idea of brief time-expiring automatic systems: "found" sound files sampling and processing themselves through the data obtained by chains of self-automatic sound analysis. The resulting audio files were devised in order to comment precise contents happening during the course of the show: my choice was to exploit anecdotic music quotations, whose recognisable content and deconstruction was intended to convey the required satire, thus each quotation was self transforming, self processing and self expiring through a built-in strategy of automatic audio analysis.

After the play I decided to collect the separated short excerpts as an autonomous work.
How could I make these satiric, quotational and self-reflexive short sound entities dance and unify?
My decision was to extend their means of composition (automatic self-analysis) as a further global meta-level: a process of automatic decision-making about the occurrence, fragmentation, mixing and spatial movement of the thus obtained assemblage. At every new occurrence, each file is exposed to a different pitch transposition, speed of reproduction, point of departure and

---

[1] by Luigi Lunari, performed at the *Teatro dell'Orologio* in Rome from the 29 October until the 3rd of November 2013

"life duration"; in addition each triggered file follows a trajectory of internal shifts between the states of normal, slightly granulated, and dynamically equalised reproduction.

The individual files are called:

1-Ingresso (entrance/ouverture)

2-Campane (bells)

3-Folla (crowd)

4-Wanda (recalling Wanda Osiris)

5-Esorcista (exorcist)

6-Tamburi (drums)

7-Maggiordomo (butler)

8-Uccello di Fuoco (Firebird)

Fig.1 The pre-sampled sound materials

The only explicit historic quotation originally requested by the actors team was Strawinsky's *Firebird*. The consequent idea of structuring a formal system based on quotations and stolen sounds, taking Strawinsky's music as a structural model appeared to be a quite rational solution. Just as the music is a formal reshuffling of "stolen musics", the video in parallel works on "stolen scores" intensively remixed by exactly the same sound analysis methods manipulating the audio.

Fig.2 Screenshots from the video

3

ANALYSIS

The total output sound is sent to the analyser (*analyzer~*, and *roughness~* MAX objects)

The following features of the 8 lowest Bark bands are tracked in real-time:

-their amplitudes in dB

-their amplitude difference (positive or negative) inside a time window of 2"

-their standard deviation inside a time window of 500 ms.

-their amplitude derivatives with respect to the previous sample (delta distance)

-selection of the currently most increasing and decreasing bark bands (inside their last 2" window)

-the sorted index (and current delta value) of the 8 bands with respect to their energy of change

METHODS

Layer A

Each of these 8 bands references a different output audio file (all the files are in loop mode):

-the positive and negative amplitude maxima determine the file occurrence (start/stop)

-the standard deviation influences the output amplitudes (volumes)

-the delta distances influence the spatial movements (speaker assignment)

-the absolute individual amplitudes determine the starting point of the files ("seek" function)

Further mappings connect the bark amplitudes to the parameters of granulation, equalisation and delay, and determine the final video rendering. The onset attacks influence some step-by-step processing modules (in opposition to further continuous effects).

Amplitude, frequency, noisiness, roughness and brightness vectors distribute their effecting details and internal shapes inside the audio-video processing machine.

Layer B

A second collateral process is active: the quasi-random probabilistic selection of very short excerpts taken from the original Strawinski's Firebird, extremely fragmented and strongly filtered. It acts as a nested sound skeleton of the overall music, able to offer a point of departure for the analysis-composition loop, and an opportune background region filling the unavoidable intervals of stillness occurring inside the "layer A" procedure.

Part of this fragment selection, and also the whole filtering, are also determined through mappings by the global sound analysis engine (the Observer).

MOTIVATION

The automatic system was built with the aim of conveying formal abstract associations through its internal parts, as a study on the form-bearing potentials of sound gesturally conceived and segmented. The audio analysis treatments collect the energy behaviours of perceptually relevant frequency bands (Barks), formally framed by long term segmentation through onset detection.

By associating the selection of each sound file to recurrent global timbre qualities the result is to drive musical form through functional extensions of timbre.

The internal tensions and natural energy articulations of the analysed (output) sound add, through global mappings, a further layer of pulses, selections, dynamic movements, cyclic ornaments and contours to the principal semantic layout of the 8 original stereo files.

PERFORMANCE

The performance coincides with the real-time automatic audio-video rendering.
Every new performance will be different, since the system is a meta-composition. Potentially the system could maintain its behaviour of output -> self-analysis -> self-structuring, as an eternal circular loop; that's why the final performance is conceived in the style of a sound-installation, and does not require the audience to be positioned in a concert-like fashion. The audience could be moving and walking around the central video projection and inside the space of the 8 speakers ideally arranged as an external circle.

Every speaker is intended as a singular instrumental-like sound source, therefore site-specific unconventional different speaker deployments are advocated.

The demo version lasts 10', and it should be be differently or indefinitely extended in case of a gallery-style performance. The overall duration (in minutes, i.e. 3 hours = 180 minutes) has to be manually set in advance inside the application.

The rendering is completely automatic, after having set the analogue system and a few software parameters, as it is described inside the main checking-list of the application.

The performance interaction accounts for the minimal gesture of just pressing the start button. As shown inside the main MAX application:
1) Check the sound card (audio settings at the bottom)
2) Select a different diffusion option if 8 speakers are not available
3) Press start (yellow button on the left)
   Optionally:
- enable/disable the video rendering (by default performed by a second laptop).
-select different means of sound analysis (the self-observing state sets the "no-file" message)
 otherwise you will define 1 vs. multiple different internal sound analysis sources.
-set the duration time

Default: self-observing, 8-speakers, video-off, 10' demo-length.
It is not recommended to perform audio and video on a same laptop.

Laptop_1 loads the application *Suite* for the audio rendering.
Laptop_2 loads the application *Image* for the video rendering.

HARDWARE EQUIPMENT

 2 laptops in network (minimum OS X 10.8, 2 Ghz)
1 Ethernet cable (1000Mbit/s)
1 projector
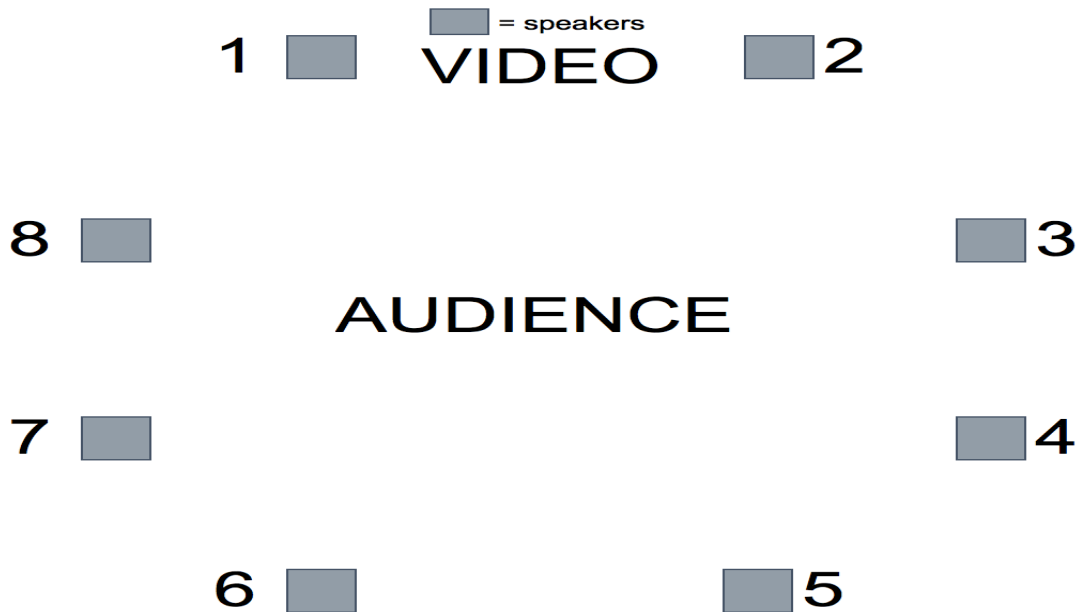1 sound interface (possibly RME or MOTU), minimum 8 outputs
8 speakers

Fig.3 Speaker arrangement

SOFTWARE

LAPTOP_1 LIST OF EXTERNALS AND ABSTRACTIONS
ambiencode~, ambidecode~, ambimonitor (Jan Schacher)
http://trondlossius.no/articles/743-ambisonics-externals-for-maxmsp-and-pd

analyzer~ (Tristan Jehan)
http://web.media.mit.edu/~tristan/maxmsp.html

dot.smooth, dot.std (Joseph Malloch et al.)
http://idmil.org/software/digital_orchestra_toolbox

ej.line (Emmanuel Jourdan)
http://www.e--j.com

f0 distance, f0.fold, f0.round, (Fredrik Olofsson)
http://www.fredrikolofsson.com/pages/code-max.html

fiddle~ (Millar Puckette et al.)
http://vud.org/max/

ftm, ftm.list, ftm.mess, ftm.object,
mnm.list2row, mnm.minmax, mnm.onepole, mnm.sum, mnm.winfilter
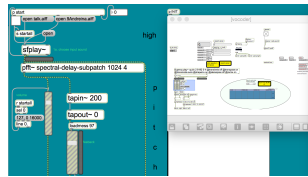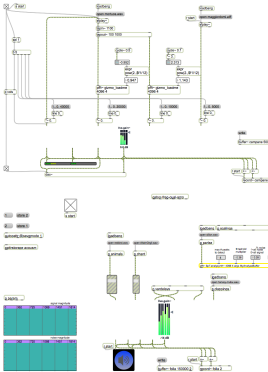FTM library (Frederic Bevilacqua et al.)
http://ftm.ircam.fr/index.php/Download

jg.granulate~, jg.spectdelay~ (John Gibson)
http://pages.iu.edu/~johgibso/software.htm

multiconvolve~ (Alex Harker and Pierre Alexandre Tremblay)
http://www.thehiss.org/

roughness (John MacCallum), o.route (Adrian Freed)
http://cnmat.berkeley.edu/downloads

LAPTOP_2 LIST OF EXTERNALS AND ABSTRACTIONS
dot.smooth, dot.std (Joseph Malloch et al.)
http://idmil.org/software/digital_orchestra_toolbox

f0 distance, f0.fold (Fredrik Olofsson)
http://www.fredrikolofsson.com/pages/code-max.html

ftm, ftm.list, ftm.mess, ftm.object,
mnm.list2row, mnm.minmax, mnm.onepole, mnm.sum, mnm.winfilter
FTM library (Frederic Bevilacqua et al.)
http://ftm.ircam.fr/index.php/Download

Jitter
https://cycling74.com/

o.route (Adrian Freed)
http://www.cnmat.berkeley.edu/MAX

Table 1. The 8 original external self-observing systems, from which the stero files were created.

1-Ingresso (entrance/ouverture)

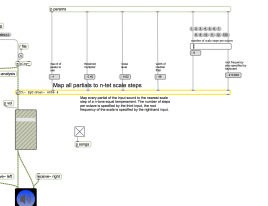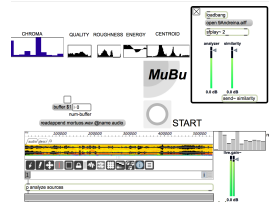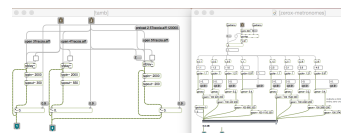2-Campane (bells)

3-Folla (crowd)
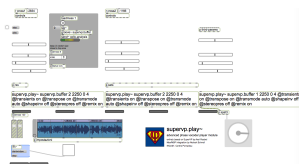

4-Wanda (recalling Wanda Osiris)
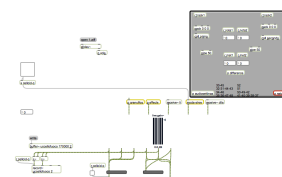

5-Esorcista (exorcist)

6-Tamburi (drums)

7-Maggiordomo (butler)

8-Uccello di Fuoco (Firebird)

INDEX