ARS_CMI: AUDIOVISUAL REAL-TIME SYSTEM FOR CONTEMPORARY MUSIC IMPROVISATION

Andreas Weixler
a.weixler (at) bruckneruni.at

Se-Lien Chuang
cse-lien (at) sime.com

Atelier Avant Austria
http://avant.mur.at

ABSTRACT

This is a concept of an composition/improvisation with the audiovisual interactive computer system created by Andreas Weixler and Se-Lien Chuang in the graphical programming environment Max Msp Jitter. Instruments are played in the style of contemporary composition/improvisation. The computer system consists of a pitch and dynamic detection, real time visual processes and a live multi channel granular synthesis with advanced controlling and performing methods. All computing devices, the audio detection, the visual and the audio processing are linked via a wireless Lan to reciprocally influence each other.

We developed a road proofed audiovisual interactive system in a work in progress of many performances all over the world in different arrangements and achived so a practical training in bringing this artwork to a live stage.

The sounds of live instruments serve as an interface in an audiovisual interactive concert that merges acoustic instrumental sound and real-time computing in an improvisation. In the combination of intuitiv improvisation and real time computing, we want to create an synaestheical artwork in which all audio and visual parts are equaly contributing. While visual images and processes are being generated during the concert, a multi-channel granular synthesis fits together minute tonal particles that make up the instrumental sounds into a constantly changing acoustic stream made up of different pitches, duration and positions in the electro-acoustic space. The musical and visual components interact and reciprocally influence each other in order to blend into a unique, synaesthetic, improvisational work of art.

1. INTRODUCTION

Our artistic creation is a work in progress. Every performance of our interactive audiovisual works, even of the same title, is unique not only because of the inherent concept of improvisation, but because of the computer system and the programming are further developed for every event. Also all other works like our contemporary instrumental compositions, algorithmic composition, electroacoustic music, videos and computer animations belong to a flow of creation, therefore each presentation can be seen as a window, in which you can experience the current status of our art. In our live performances we use a computer system for audiovisual interaction, which we developed in Max Mps Jitter, described as follows:

The computer system and Max/Msp/Jitter

Three modules of our computer system on Macintosh G4 Powerbooks are interacting with the performers i.e. an acoustic musical instrument and the electroacoustic hyper instrument, analyzing the sounds, exchanging data between analysis, audio and video computing to create an unique, synaesthetic, improvisational work of art.

Module 1 The audio analysis

To turn a musical instrument into an interface we use the information of pitch and dynamics to trigger events in the Max MSP Jitter patch [1].
To control the flow of triggering events of the audio analysis we implemented a dynamic threshold system, which avoids multiple triggering especially by sounds with a long sustain, like the bass recorder has.
The fiddle~ object by Miller Puckette [2] serves as a pitch tracker, like any other, yin~ by Norbert Schnell [3] at IRCAM or pitch~ by Tristan Jehan, [4] could do.

Module 2 The multichannel granular sound synthesis

The live sounds of the musical instruments are fed in a multichannel granular sound synthese [5] [6], which we understand and perform as an musical instrument of its own with virtuoso and complex potential.
This electronic instrument is highly depending on its input from the acoustic musical instrument.
The peculiarity of our sound system is not only the real time process and its multi channel features, more the intuitiv and artistic control of the parameters, in a sense of setting limited randoms as meta parameters which makes the system acting like flock of fish, in which you do not focus on a single fish but how the whole flock behaves like one beiing.

Module 3 the interactive visuals

A main emphasis of our interactivity is to achive a synaesthetical effect, which leads us to visuals which directly react to the parameter of sounds and to the exchange and integration of the various data of audio and visual data mutual depending on the live action of the both performers, of the acoustic and the electronic instruments, which per defintion is our audiovisual interactive computersystem.

Acoustic musical instruments

Se-Lien Chuang is controlling the interactive visuals system and also often performs an acoustic musical instrument as an interface to the system such as arco piano, Yang qin and bass recorder.
In addition we performed with other musicians performing modern western instruments as well as Japanese, Korean and Chinese traditonal instruments.
The acoustic musical instrument is not only acting as a musical instrument, but furthermore as an interface for the computing system. The musical instrument is controlling the creation of the visuals in real time and the instrumental sound will feed the granular synthesis distributed on 8 channels.

1.1. Arco Piano

Arco-piano is a special contemporary technique of performing the piano with hairs of bows, which gives a very sophisticated sound of the piano. We know Henry Cowell with the cluster sound, Colon Nancarrow with the “Studies for Player Piano“and John Cage with the “Prepared Piano Music” [7]. Se-Lien Chuang developed arco-piano technique to the artistic point in a manner of exploring the studies of bowing strings of the piano more technically and aesthetically, where she can control a polyphonic sound of the bowed piano strings and their upcoming overtones.

We use the arco-piano and our audiovisual interactive computer-system in the improvisational performance *Interlude for Pi and for Io*.

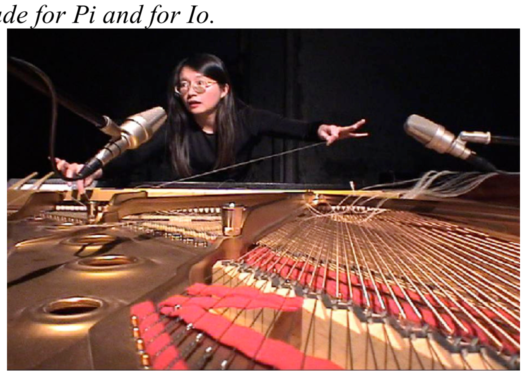


Figure 1. Arco piano in *Interlude for Pi and for Io*

1.2. Yang-Qin

Yang-qin is a traditional Chinese hammered dulcimer [8] with a near-squared soundboard, together with our computer-system it represents the performance of *Erinnerung in jedem Laut*.

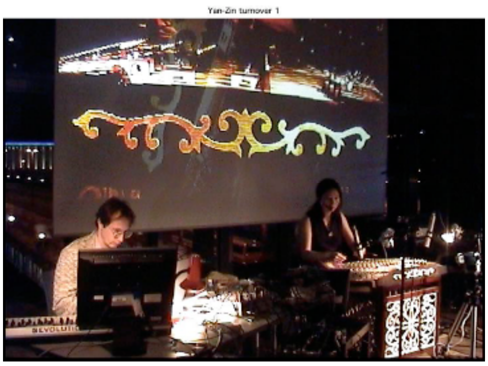


Figure 2. Yang qin in *Erinnerung in jedem Laut*

1.3. Other Instruments Used

We used various instruments as audio input and controlling devices in various performances worldwide, in some previous audiovisual real-time computing concepts there were 2 prepared guitars (*Waon* at the ISEA 02, Nagoya, Japan), Chinese mouth organ sheng and Chinese two-string violin erhu (*Waon* in Taiwan, 03), traditional Korean instrument haegeum, cheolhyungeum, geomungo and percussion (*Seoul Lightingale*, South-Korea, SICMF 03), traditional Japanese biwa (*The Story of Heike,* Nagoya, Japan, 2001), western instruments like violoncello and clarinet (*Waon* in NYC 2000), bass clarinet, saxophone and electric guitar (*Waon* at VNM 01, Graz, Austria), jazz bass and piano (*Lightingale* at VNM 03, Graz, Austria), experimental violoncello and violin (*Living Cave* at artport Mediaselect 01 and *running-figure-lantern* at the ISEA 02 in Nagoya, Japan) and also human voices (*Living Cave* at artport Mediaselect 01 Nagoya, Japan and at Laval Virtual, France 2001) and most recntly an succesfully a modern flute (*Das Kichern der silbernen Flöte*, electronic access @ Goldsmiths, London, 2006), a bass recorder (*The Colours of A Wooden flute*, electronic access @ Goldsmiths, London and experimental intermedia, New York, 2006) and a classical bassoon played by specialist of old music (*quod erat…,* vnm Festival, Graz 2007)

Prevoius and special works

We created a predetermined accompaniment of visuals with a Japanese biwa master, telling the traditional narrative *the story of Heike*, this was performed in Nagoya Japan in 2001 during a research at the Nagoya City University [9]. The intonation of the voice and the articulation of the instrument controlled the flow and processing of the picture associated with the content of the story. At the same time we created the audiovisual interactive installation *living cave*, which became performed in exhibitions in Japan and France and immediately let to the audiovisual instrumental theatre *running-figure-lantern* which was selected for the ISEA - International Symposium on Electronic Art, in Nagoya, Japan 2002 [10].

1.4. History of Audiovisual Instruments

Historically seen, it seems to be an age-old dream of artist to make music visible and visuals audible.

Musical instruments as an interface for a audiovisual system have been created ages before including the perspective-lute of Giuseppe Arcimboldi (1527-1593), the Clavecin oculaire of Louis Bertrand Castel (1688-1757), the ocular harpsichord (1757) of A. Morley, the Clavilux (1922) of Thomas Wilfred (1889-1968), the Optophon (1915) which composer Alexsander Skrjabin (1871-1915) used in his orchestra piece *Prometheus* (1910/11), the optophonic piano (1917) of Vladimir Baranoff-Rossiné (1888-1942), the >Reflektorische Lichtspiele< (1923) of Ludwig Hirschfeld/Mack and Kurt Schwerdfeger at Bauhaus, the Sonchromatoskop (1925) of Alexsander László (1895-1970) [11] and the Optophon (1927) which Dadaist poet Raoul Hausmann (1886-1971) experimented

with the optical-acoustic montage process to project the kaleidoscopic pictures. [12] [13]
The instruments we mentioned above are manual, machine, mechanical. Until 1960 Livinus va de Bundt and Nicolas Schöffer made the first light organs constructed with electronic control which later are computer-navigated keyboards with sensors and colorful projection lamps or laser technology.
There was always a desire to stimulate the brain by a combination of audio and visual effects: for example the light of colored glass in the strong acoustic of churches, light organs, simply hooked up by frequency response, in the disco music area up to today´s high-tech audiovisual shows of popular music events, and countless more examples of contemporary artwork. [14] [15] [16]

Technical concept

1**. Multichannel Granular Synthesis**
By processing the instrumental sound in real time the computer becomes a hyper instrument played in an improvisation along with the bass recorder, spreading its sound on an eight channel sound system, controlling flocks of grains, rearranging those in terms of pitch, time, order and position within the acoustic environment.We use a granular synthesis based on the rgrain3~ object by Richard Dudas [17] and have usual controls of transposition, position in the audio buffer, grain-length and panning. We record a live signal into a 22 second audio buffer and read all data out from there. The buffer will be overwritten all the time, so we have access to nearly real-time data up to 22 seconds back in time.

1.1 granular synthesis in multi layer structure
Unlike the usual granular patch we use a four layer granular synthesis, which makes it an eight channel granular synthesis. We added a new interface to it, which results in a very artistic access different to stochastic and chaotic algorithms described for example by C. Roads, N. Schnell, K. Essl among others. . [18] [19] [20]

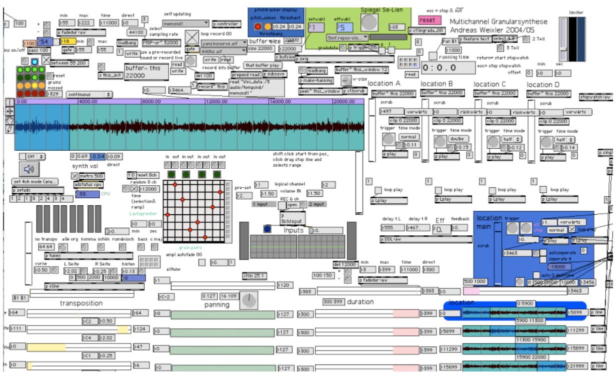


Figure 1. The main patch of multi-channel

1.1.1 Parameter Handling
In our multi-channel granular synthesis the parameter of transposition, panning, duration and location exist four times, i.e. in each layer A, B, C and D. Besides these we added some extended access to those by creating our own parameter handling techniques such as an algorithmic spatialisation and sequencing a granular window.In general the parameters of the four layers can be either linked or individually controlled in the domains of time, pitch and location.Transposition is the pitch shifting in one layer of granular synthesis.Panning is the amount of random panning, which can be limited to any range between left and right, in this case between the two audio channels, which belong to this layer. We have eight of this internal audio channels, each has a pair dedicated to a certain layer.

Algorithmic Spatialisation
All of these eight audio channels have to go trough an automated matrix, which keeps those fading over all eight loudspeakers, so we have a constant movement of the four stereo pairs of sound. By doing this the program takes care that every audio channel starts from one speaker and fades to another speaker in that way that each of the eight audio signals always has its own and unique position in the electro-acoustic room. The position in the room enhances the perception of a multi-channel granular synthesis. Duration is the length of the grains, which can be from 1 ms up to 400 ms and also have a controlled randomizer in within the chosen range. Location is the position in the audio buffer, which is replayed by the specific layer. It can be a definitive exact point or a window in which the grains are played randomly.

1.1.1.1 Time Handling
Multi Speed
The four layers of granular synthesis either can be linked in time to create a certain chord, within this each of the granular synthesis can have their own pitch, grain length and panning using the same audio data or we can choose its own direction and speed for each layer! In this case we can control each granular synthesis independently, each layer can have its own values for all parameters including the time factor, which is the sequential read out of the audio buffer. The writing position and the re-play locations are independent in their direction, forward or backward and in tempo. Playback time can be original time, half or double time, even any floating point or integer multiplication factor of time, or use a freeze effect to create a standing chord which not only consists of four different pitches, but even of four different position within the available audio data.
Oncoming Traffic
The playback is backward, and therefore has the opposite direction as the recording. The live signal is recorded into a 22 second audio buffer and one or more of the four

virtual play heads play the audio buffer in reverse, i.e. it goes back in time until it hits the moment of the actual recording, this acts like a threshold in time between the recording which is 22 seconds ago and the live signal. The audio playback goes back in time from there until it hits the actual time again and so forth. This give astonishing moments in a rehearsed live improvisation of manipulation of older materials up to sounds in real-time in its original context, just by manipulating the time.

Overtaking

The recording and the re-play are performed in the same direction, in this case forward. Depending on the playback speed we can have two forms of handling: either the recording takes over the re-play position time by time, when the re-play tempo is less then 1, meaning slower than the original tempo, or otherwise, by a time factor greater then 1, the play heads are faster then the recording of the live signal, and take over the actual recording moment, before it will be refilled by the live input. This means sometime the result is a real-time processing (0 sec delay) for a moment and then suddenly slips into the 22 seconds old sound material and from there again plays the audio buffer faster then real-time till it again is in sync with the actual recording and so on.

Memory

In addition to this we also have memory functions to save the temporal content of the audio buffer and to bring back materials to give an impression of a composed performance.

1.1.1.2 Pitch Shifting

As granular synthesis offers the treatment of time compression/ expansion independent from the pitch manipulation, we are dealing with three types of controlling the pitch:

Original Randomizers

Like in Richard Dudas granular synthesis patch transposition, panning, duration and location are controlled due to limited random functions. All of these parameters have possibilities to select a range for setting a random movement in-between or choose a fixed value. In the original patch using rsliders the transposition is limited in the range of 1 octave higher and 1 octave lower. In our multi-channel granular synthesis every layer has its own slider to set the range of the transposition separately.

Floating Point Multiplication of the Frequency /Spectrum

To achieve even very low or very high pitch shifting we programmed an override function into the rgrain3~, which allows any floating point or integer number to be set as a fixed multiplication factor. This is very useful to create certain tunings and chords, which even can be microtonal.

Chord Playing, Multiple Transposition of the Whole Sample

In addition to the point above we create a keyboard like pitch shifting, where transpositions can be played as piano chords, which then will create four kinds of pitch shifting at the same time, can be changed in real-time, and can even follow some kind of chord progression. Of course the outcoming sound depends on the source materials as its spectrum becomes pitch shifted in a multiple way, simple chords turn into tensions or altered chords then.

1.1.1.3 Selection of the Location in the Audio Buffer Sequencing a Granular Window

For the read-out of the audio data we created various methods:First the location of the four layers can be either linked or act independently. We have one main audio buffer display, which is linked to four smaller displays, one for each layer. A selection in the waveform~ object represents the window of the actual granular synthesis. This window, in which the random read-out of the granular synthesis is happening, can be moved forward or backward in any tempo by a simple sequencer and the window can have any size inside the audio buffer.When the layers are linked, they play the same position of the audio buffer, standing still or in movement. This is useful for creating chords and echoing a melody when this is driven by a timeline.

A click in the recording buffer sets the starting point of the movement, it enables the audio performer to select a special note out of the visual display of the audio wave. This gives the joyful possibility to click in the real-time display of the recording buffer and select a certain point of the recording and even catch the actually recording to make it a real-time effect, which is following the live audio signal. Then there is the „seperate 4" function, every click in the display of the recording buffer sets a new position to only one layer, the next click sets a new position to the next layer and so forth. This is creating a chord or ambient including freely selected sound bits, which can be chosen visually by the performer in the display of the audio buffer.The „auto separate" function is a certain window follower, which hands over the position from one layer to the other by freezing the last position. By this operation each layer re-plays a certain still standing spectrum of the sound but the oldest layer always updates its position to the actual position of the newest recording. So you hear slightly changing frozen sounds of the recording, which in some kind of double windowing re-play four windows, altogether also moving one by one like a bigger window through the audio buffer. (see figure 2 „auto separate function")

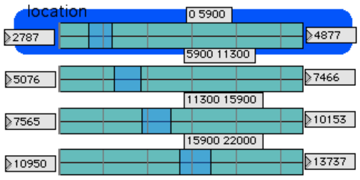


Figure 4. Auto separate function

2. Visual Processing

We developed a patch in Max Jitter called *ModularFilter*. It either takes a live video input, for example the performer, or some prepared pictures and movies as a source of visuals. Sophisticated video effects will be exchanged from part to part in the performance and their parameters are controlled simultaneously by the sound of the instrument and the linked parameter of the live granular synthesis as well.

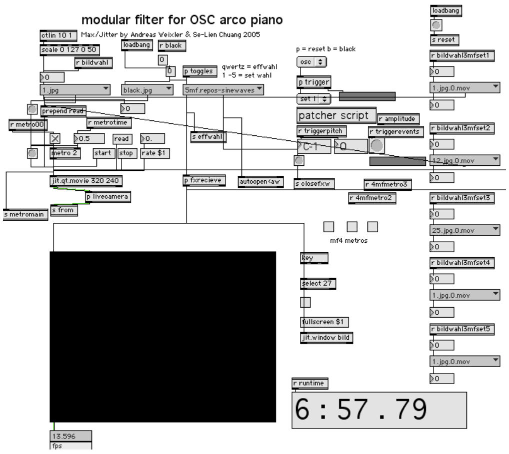
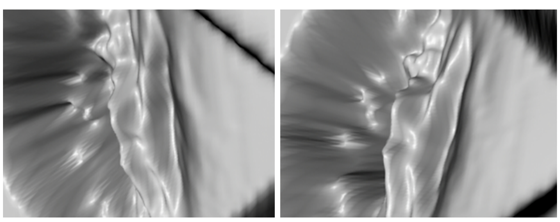


Figure 5. The main patch of *ModularFilter* in Jitter

There are three subjects about the initial materials for the source of visuals: moire, perspective and symmetry. For all of the subjects we use a live camera input and a selection of different short video clips of infinity, an art video made in 2004 by Se-Lien Chuang. Since the subjects are determined we can access the video processing with three criteria of manipulation: aliasing, zooming and mirroring. The module of effects we are using for example is: Rutt/Etra Scan Processor, inspired by the work of Woody Vasulka in the mid-70s, this patch displaces a 3D plane based on the luminance of a video signal.



a) b)

Figure. 6 *Variation* a) and b)

.

Other effects we are using for example are: Spatial Map -> SineWaves, Spatial Map -> LowPass Noise with FFT rotation and OpenGL [21] transformation etc.

Figure 7. The OpenGL transformation of live input of yang- qin in *Erinnerung in jedem Laut*

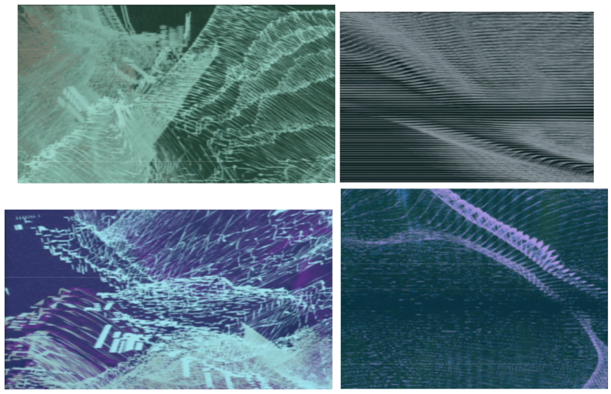


Figure 11. visuals of Colours of A Wooden Flute

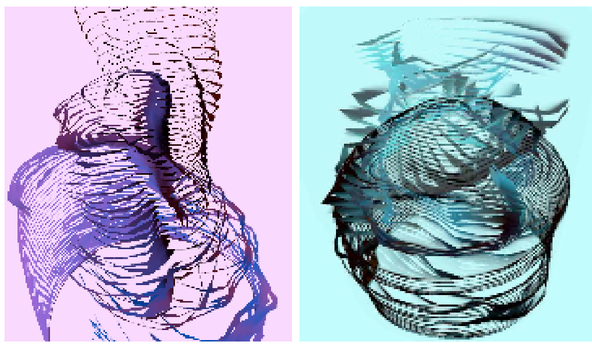
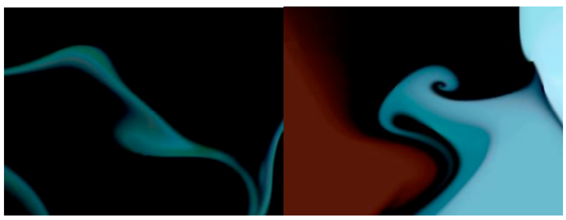


Figure 8. visuals of *for reynolds*



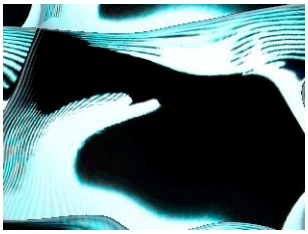


Figure 9. visuals of *Interlude for Pi and for Io*

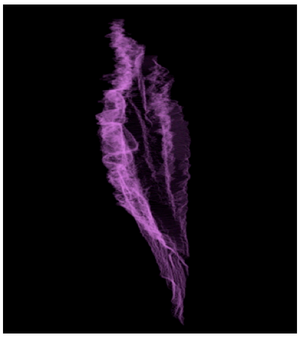


Figure 10. visual of *Das Kichern der silbernen Flöte*

3. Interactivity

3.1 Manual/Automatic Mode

The modules of the visual processing vary during the performance to attract the audience's attention and to have an stimulating effect. The change of effects can be done either manually, so the user of the visual computer becomes a performer, or automatically, then the computer follows a predefined time line. This necessary specially in performances where the musician (e.g. the arco piano player) might feel arduous to handle a computer device at the same time when he/she concentrates his/her mind on instrument, like in the performance of *Interlude for Pi and for Io*.

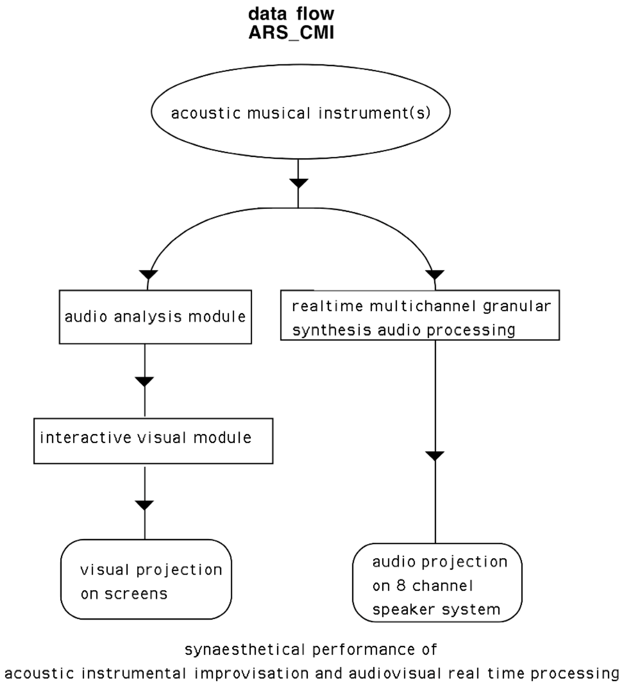


Figure 4. The data flow diagram of audiovisual real time computer system

3.3 Interaction

Acoustic musical instrument to Visual Computer

The audio analysis of the live instrument provides information about the pitch and the dynamics. This data is transmitted to the visual computer, where it triggers the visual event relating to the sound.

Audio Computer to Visual Computer

The parameters of the visual processes are influenced by selected parameters of the granular synthesis.

Visual Computer to Audio Computer

If the program in the visual computer changes the effects, it sends information to the computer of the granular synthesis to indicate the audio performer which parameters of the granular synthesis are now linked to visual parameters. Active parameters are shown as blue lighted in the granular patch as shown in figure 2.

Musician and Granular Synthesis Performer

As usual in a musical improvisation the musician and granular synthesis performer react to each other´s sounds and both to the computer aided creation of the visuals.

3.4 Network

We use OSC - Open Sound Control [19] - via LAN and wireless LAN. This gives us the opportunity to send a big punch of parameters, even a time clock and opens sufficient possibilities, but also challenges for data mapping.

4. Mapping

To achieve a synaesthetical effect where events in the audio are always related to events in the visuals we not only use a common trigger like the pitch information and dynamic attack of the notes played by the live instrument. We also linked the parameter of the granular synthesis and the visual processes and alter this links during the performance in order to keep it attractive for us and the audience. The mapping of these data of course is an important decision made by the artists.

Organization Levels

In our data mapping we distinguish between different organization levels: parameter, module, script

Parameter of Granular Synthesis

On the audio side we have the following parameter to be linked to the visual processes:Transposition, panning, duration and location, all parameters exist four times in each of the layers A, B, C and D.

Parameter of Visual Processing

On the video computer we link parameters like zoom, rotate, scale, saturation, fade etc.

Module

This is a central idea of ModularFilter. A main patch handles the visual source, i.e. a live camera or the initial materials, the exchange of the filters, effects and scripts in a modular system.(Which module we are using is discussed in chapter 2. Visual Processing).

Script

Is related to the data mapping, in that way it is the artistic decicion what will happen similar to a story teller.

It contains for example the reaction of the visual system according to a trigger from the audio anaysis module.

Figure 8. will give a first idea of the flow of data in our system.

2. CONCLUSIONS AND OUTLOOK

The goal of this project in computer music is to create an interface for visual and music computing for an associated audiovisual composition and to create a performance of equal participation of sounds and visuals.

To achive this we combine improvised music and real time computer processes to an audiovisual relatime composition, learning from each other, technical wise and as in an artistic way.

Currently we are working on extending the system for a chamber ensemble or a group of instruments. We plan to separate the granular synthesis as well as the analysis and link the parameters of individual instruments to form a complex visual response to a live audio event. We also hope to develop a system in which the visual events cause reasonable audio responses, to achieve equality of both, the audio and the video domain [24].

3. REFERENCES

[1] http://www.cycling74.com/

[2] http://crca.ucsd.edu/~tapel/software.html

[3] http://recherche.ircam.fr/equipes/temps-reel/maxmsp/anasynth.html

[4] http://web.media.mit.edu/~tristan/

[5] Ruschkowski, A. *Elektronische Klänge und musikalische Entdeckungen*, Philipp Reclam, Stuttgart 1998., pp 314-319

[6] Gabor, D. »Acoustical quanta and the theory of hearing«. In: *Nature 159*, London 1947

[7] Cage, J. Prepared Piano Music. Volume 1& 2, 1940-1947, No. 67886 a & 67886b, Edition Peters 2000.

[8] Gifford, Paul M. *The Hammered Dulcimer: A History.* Lanham, Maryland: Scarecrow Press, Inc. 2001.

[9] Mizuno, M., Weixler, A., Chuang S. L., „Interactive Musical Art in the Second Half of the Twenthieth Century and the NCU Project 2000“ in *The Yearbook of Nagoya City University-School of Design and Architecture 6, Introduction to design and architecture 6*, 2002, pp 197 – 225.

[10] http://vision.mdg.human.nagoya-u.ac.jp/isea/

[11] László, A. Einf. in die F., Leipzig 1926; M. Peternak, *Licht und Klang aus Ungarn* (Hg. H. Gaßner), *Wechselwirkungen. Ungar. Avantgarde in der Weimarer Republ.*, Kat., Kassel, Bochum, Marburg 1986.

[12] Caduff, C. Fantom Farbenklavier. Das Farbe-Ton-Verhältnis im 18. Jahrhundert oder Vom Einspruch gegen das clavecin oculaire und seinen ästhetischen Folgen. In: *Zeitschrift für deutsche Philologie 121*, 2002.

[13] Wellek, A. Farbenharmonie und Farbenklavier. Ihre Entstehungsgeschichte im 18. Jahrhundert. In: *Archiv für die gesamte Psychologie 94*, 1935.

[14] Centre Pompidou. *Sons & Lumierés*, Une historie du son dans lárt du XXe siècle, cataloque of the exhibition im Centre Pompidou, 2004.

[15] http://www.aec.at/de/archives/prix_archive/prix_projekt.asp?iProjectID=2494

[16] http://dips.dacreation.com/

[17] Dudas, R. Max/Msp, Jitter Nightschool: http://www.cnmat.berkeley.edu/Max_MSP2004.html

[18] Roads, C. *Microsound*, MIT press, 2001. pp 85-118

[19] Schnell, N. Beiträge zur Elektronischen Musik 4 - *GRAINY - Granularsynthese in Echtzeit - Eine Objekt/Sammlung für MAX auf der ISPW*, IEM Graz 1995.

[20] Essl, K. Beiträge zur Elektronischen Musik 6 - *Strukturgeneratoren - Algorithmische Komposition in Echtzeit* , IEM Graz 1996..

[21] http://www.opengl.org/

[22] Open Sound Control. http://www.cnmat.berkeley.edu

[23] http://vision.mdg.human.nagoya-u.ac.jp/isea/

[24] Atelier Avant Austria, http://avant.mur.at