

# An approach to visualization of complex event data for generating sonic structures

Sinan BOKESOY\*, Jean Baptiste THIEBAUT†

\*CICM, University of ParisVIII  
sinan@sonic-disorder.com

†IMC Department of Computer Science  
Queen Mary, University of London,  
jbt@tcs.qmul.ac.uk

## Abstract

*This paper presents a visualization strategy adapted to a complex event generation system. For this purpose the Cosmos application has been addressed which organizes sounds in real-time on different timescales and layers within a self similar structure. The visualization aims at providing a feedback of the distribution of sound in time. It incorporates the representation of event parameters like onset, duration and modulation sources which define the emergent structure and also the visualization of the extracted high level parameters with STFT analysis tools. Finally different experiments are presented and interpreted.*

## 1 Introduction

Visualization of musical information is an important process in compositional processes besides using the ears as a first hand guide. We assess that the visualisation of sonic parameters in real time is a useful feedback for the composer, the performer and the audience. According to A. Miller [Miller 1996], visualisation is an abstraction of the phenomena witnessed in the world of perceptions. These abstractions are useful to understand complex structures and to communicate them to others. For instance, the structure of an atom or the motion of planets is easily understandable through their visual representation, although their theory is less accessible.

The process of composing the sound and composing with the sound needs visual feedback about the evaluated data, in order to project the parameter space efficiently to our sensory systems as a perceptual model. The visualisation of complex organisation of sound is useful for the composer or the performer as operational properties like the representation of time may emerge from the display.

In traditional notation systems, such as within classical music, where the instruments have been already defined, any musical event is defined with its starting and duration on the timeline. Around its symbol, one can add other

features of articulation such as playing techniques, dynamics etc. All of them encapsulate a sound object delivered by the acoustical instrument and the notation of the particular event represents an interval of time and a region in the timbre space of this instrument.

In the 20<sup>th</sup> century, the tools that analyzed the sound object did give the opportunity to dive onto the micro-time scale of sonic events and observe the spectral evolution, which is practically not possible to notate within traditional systems. The ability to reach the micro-timescale has opened new horizons in composing the sound material itself. Short Time Fourier Transform (STFT) has been a popular tool to realize harmonic analysis on the sound material. The encapsulation of time does exist, but it is measured now with the analysis window size of STFT rather than the long durations of traditional note lengths. Inside this micro time-span we freeze the sonic evolution and represent the partial analysis. The spectral development from the analysis data occurs within the overlapping of these discrete analysis windows.

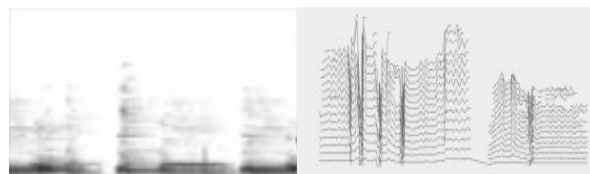


Figure 1. STFT analysis and partial tracking analysis within AudioSculpt<sup>1</sup>. Depending on the purpose the image on the right is more comprehensible and editable.

High level musical parameter extraction is necessary on the STFT analysis of complex sound material, because of the immense data flow.

For this purpose, we need the help of the software tools, which deliver the analysis and extracted data visually.

---

<sup>1</sup> AudioSculpt analysis/resynthesis software developed by IRCAM

Visual control of sound transformation has been explored as a new way of interaction for musical purposes. In the past it has taken several forms, such as AudioSculpt (Figure 1), which displays an editable spectrogram, and allows interacting with the amplitude and frequencies of individual partials. The extracted high level parameters like fundamental pitch, centroid, noisiness, spectral envelope etc. help to evaluate the data musically better and makes it applicable for compositional processes. Hence, the process of linking the visual data in this case with the perceived sonic output has been always important for intuitive interaction.

## 1.1 Basic properties of the event generation system

In our case, the analysis material we aim to visualize will be generated algorithmically on another computer application. For instance, we can capture the generated event space before it has been rendered as a global sound output by establishing a link between our visualization tool and the event generator. Such event generation systems apply a bottom-up approach to synthesize complex sound structures with being involved particularly with the granular representation of sound, where each grain of sound is defined with a certain duration and synthesis parameters in the timbre space. (Figure 2) Works such as *Achorripsis*, *Analogique A/B*<sup>2</sup> by I. Xenakis, and applications like *Cloud Generator*<sup>3</sup> by C. Roads are early examples using models, where the sonic elements structure themselves to formalize the sonic evolution.

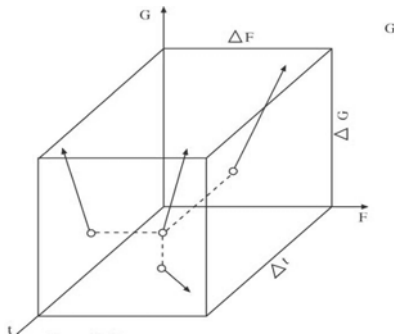


Figure 2. Grain distribution inside a division of time in the *Analogique* model of I. Xenakis. The vectors show the glissandi as a transformation on the sonic entities. The timbre space is represented by frequency  $F$  and intensity  $G$  as being evolved on time  $t$ .

<sup>2</sup> *Achorripsis* (1956-57), *Analogique* (1958-59) are first examples which use stochastic models with a bottom-up approach in composition

<sup>3</sup> Granular Synthesis Application for distributing events in time and showing them as point clouds

According to Bregman [Bregman 1990], the granular representations are especially useful in the analysis and perceptual modeling of dynamical events of peculiar complexity. As a methodological decision it has been preferred to utilize the organization of minimal units to introduce secondary sonorities, where the composer makes the decisions that apply at the smallest scales of time, and where things happen that result in the large scale (emergent properties), perceivable attributes we might call timbre and texture [2].

Since the design and the definition of this kind of sound space involve the calculation of properties of each individual entity, they can be easily represented by graphical tools and interpreted in different ways. The direct visual mapping of grain parameters or incorporating them within the sonic experience goes back to 70's to Xenakis' *Diatope* installations consisting of spatialized sound, moving lasers, pinpoint lights through the space. Today there are a number of researchers and artists collaborating for artistic visualization of music. For instance Roads' recent audio/visual works interests us in the manner he evaluates the visualization of structured micro sounds as a formal expression of parameters building his music. An amount of data with certain entropy is being delivered within the limits of visual perception.

We explore in this paper a mapping from the sonic event data and a mapping from the audio signal analysis. The sonic event data representation is used to give cues on the structure of the generated macro sound, while the audio signal analysis retrieves information on morphological aspects, such as harmonics, noisiness, amplitude or pitch. For this purpose, we benefit from STFT analysis by using its high level interpretations and the event representation visually. It is not just the analysis of the whole being similar to the STFT technique but a representation of each element responsible in creating this complex structure. For the experiments in this case we have chosen the *Cosmos* application [Bokesoy 2005], which is a complex event generation system for synthesizing emergent sonic structures on multiple time-scales with using stochastic mechanisms. After a small introduction to *Cosmos*, we will continue to explain the mapping system and the structure of the visualization process in greater detail.

## 2 *Cosmos* as an complex event generation system

The *Cosmos* application is a real-time dynamic event distribution system, which generates sonic entities as a result of the organization on multiple time scales.

The discrete events of certain density are distributed in a time space with their onset time and duration parameter, which are calculated with stochastic functions. (Figure 3) Each event in macro space defines a time region of meso space equal to its duration, and the sub events are distributed inside this time duration. The same organization is also true for the events of meso space and micro space.

As new spaces are created, new events and therefore new sub spaces are distributed. The micro events are assigned to sonic entities with specific waveforms rendering a multidimensional timbre space using sound parameters like pitch, intensity, spatial distribution and signal processing tools being modulated with continuous curve functions.

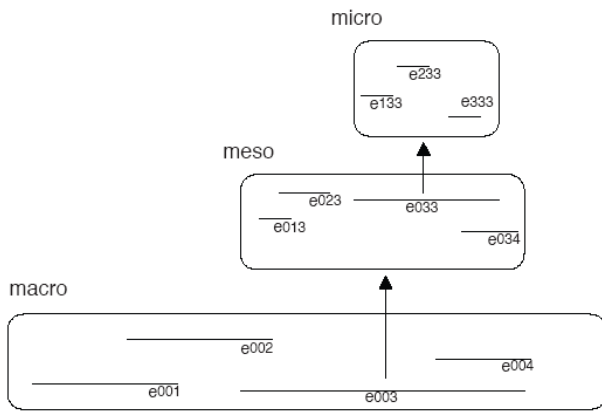


Figure 3. Self-similarity in the structure of the *Cosmos* model. Each event in the event space opens another space in the lower level timescale with the same distribution mechanism.

The modulation generators in ‘Cosmos’ are the stochastic modulation sources addressing the synthesis parameters, which are assigned to sonic entities defined within the micro event space. There are four different curve generators for each of the macro, meso and micro space events. Finally, we can combine together different modulation sources belonging to three different time scales, from macroscopic to microscopic, to obtain higher complexity. Each curve generator can be assigned to a unique destination, and if there are more than one modulation source for one destination, they will be superimposed as one modulation source with the complexity of representing the evolution of multiple time scales.

### 3 The mapping strategies

With its elaborated timescale operations, *Cosmos* fits very well as a mechanism to observe under our application.

On Figure 4, we can have a look to the possible data routing between *Cosmos* and the visual interface.

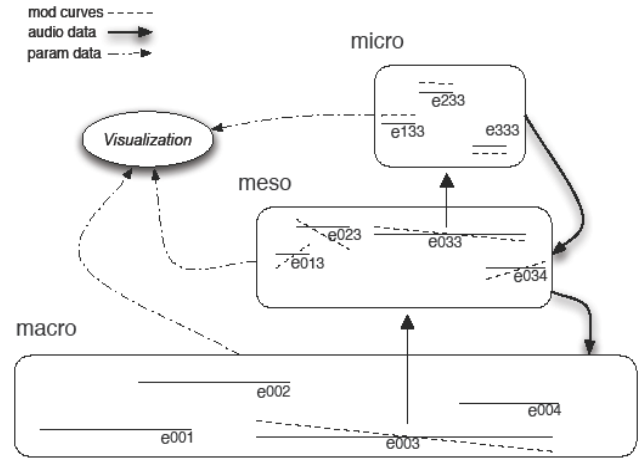


Figure 4. You can see the modulation data on each event and the audio data which is being routed exactly reverse to the event generation process. The micro-space audio sum is being routed to the relevant meso-event, etc. From each event space, the analysis and modulation data flows towards the visual interface.

The audio channels in *Cosmos* can be routed as the micro, meso and macro event audio data individually. The visual event data representation patch receives the data from *Cosmos* in the list form. The continuous curve generator data coming from *Cosmos* and the STFT interpretation of audio channels belonging to macro, meso and micro event spaces can be mapped as color information. Figure 5 illustrates the high level continuous parameter flow from *Cosmos* towards the visualization application.

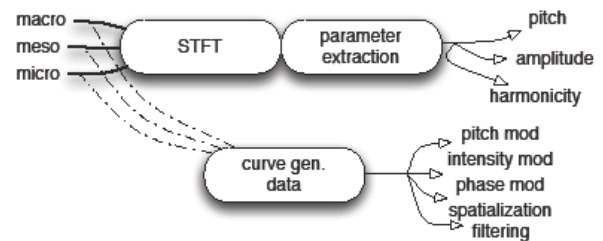


Figure 5. The macro, meso and micro space audio data enters the STFT and then the pitch, amplitude and noisiness parameters are extracted. The macro, meso and micro event modulation will be received also as audio rate data, in order to display the high level parameter modulation applied to sonic event spaces.

It is important to note that analyzing different time levels of *Cosmos* audio data allows for different zooming processes in time in comparison to what STFT proposes as an analysis on the whole macro sound form.

For instance, STFT analysis of the macro sound data may not be able to unveil the individual contribution of micro sonic events. We can shed new light on existing representations, exposing hidden transformations and revealing subtle emergent behavior by comparing these different analysis processes.

From a compositional point of view, for the bottom up construction of complex organized sound, the interpretation of different layers of the mechanism by using analysis tools is important to manipulate and evolve the compositional process. The representation of event data is respecting the WYSIWYG (what you see is what you get) the principle by supporting an intuitive visual interface for parameter access. For a better correspondence between the sonic material and the event data presented visually, a unique feature has been added to the *Cosmos* application, which lets the user to listen either individually or in combination to macro, meso or micro events including their modulation sources. This helps to synchronize the zoom in/out feature of *TeleScope* with the sonic output coming from *Cosmos*. The higher the visual density, the higher the sonic density.

## 4 Visual Representation

The traditional way of visualizing events on a timeline is first defining the timeline as the  $x$  axis and then displaying onset times and durations with referencing the timeline in a 2D display. (Figure 6) We depart from this approach for making possible to present the multiple timescale event information with a unique display.

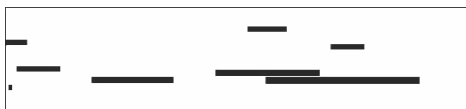


Figure 6. The event space in the *Achorripsis* model of I. Xenakis composed in 1956-57, calculated and visualized within Max/MSP<sup>4</sup>.

The visualization of the data happens within the following strategies;

### 4.1 Real-time event data representation

According to the onset time and duration information received from *Cosmos* as a form of list data,

we should observe on the screen the branching of macro, meso and micro events. The list data looks like;

```
1 1145 0 macro
1 1 127 0 meso
1 1 1 34 1205 micro
2 1 1 31 0 micro
3 1 1 31 31 micro
4 1 1 31 63 micro
2 1 127 127 meso
1 2 1 31 95 micro
.....
```

First column is the event number in the current event space. The second column in the meso and second and third column in the micro event data defines the inheritance. The last columns define the duration and onset time of the events. This data is being sent at every macro-space initialization of *Cosmos*. The time origin is represented by a circle. The macro events start drawing from this origin as soon as their onset time arrives. At that moment a line trajectory measured by the duration of the macro event is displayed. A blue ball representing the macro event moves on its trajectory. When it reaches the end of the trajectory, the ball and the line fades out. On a macro event trajectory, the meso event branches start to distribute according to their onset times but their trajectories are displayed thinner than the macro ones. Whenever a micro event onset time arrives, a violet square appears on the current meso branch and moves along its trajectory during its lifetime.

For example on Figure7, there are four snapshots presenting the sonic evolution in a macro-space of *Cosmos*. The branches of macro and meso events are shown instantly within the calculation of the macro-space. The balls representing the event time start moving along their trajectories on the branches. For instance we can see 4 macro event branches on the first top 3 snapshots, and then on the last snapshot two another has been added.

This is reminiscent of fractal organic structures in the nature, which fits very well with the self similar structure of the *Cosmos* application. In nature, the size of a tree, branches and the complexity certainly give us an idea about its age and other properties. Time is projected on our application display as the path, which the event balls have been following from their departure points along their lifetime. The event density per second in *Cosmos* can be calculated by multiplying the macro, meso and micro space densities and dividing this by the macro space length. This number gives us the amount of micro event balls per second on the screen. The density limit of *Cosmos* is 500events/sec and such fast updating of the displayed information is beyond the limits of visual perception.

<sup>4</sup> A programming environment, currently being developed by *Cycling74*

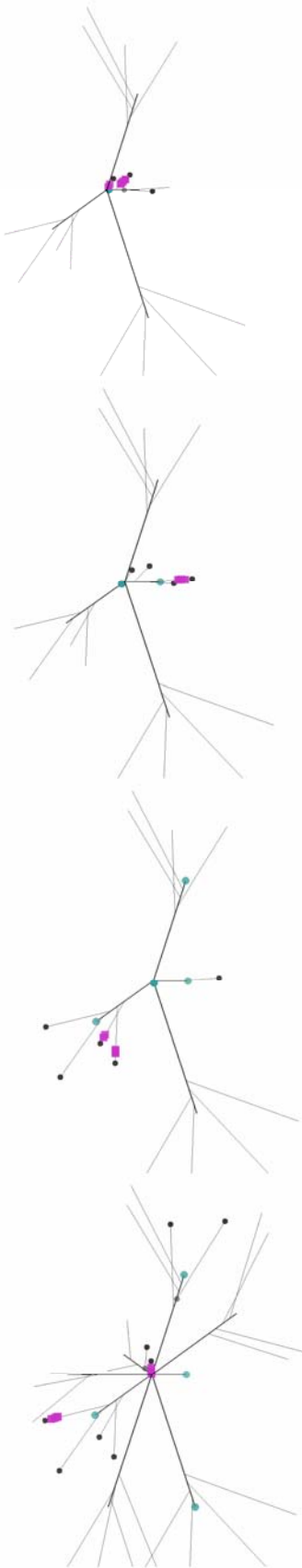


Figure 7. Representation of the event distribution process with 4 snapshots. On each macro space initialization the calculated macro and meso branches are displayed according to their onset times, while the balls start moving on those branches representing the event distribution.

## 4.2 Representation of the modulation data

The modulation data belonging to macro, meso and micro events are displayed on the trajectories or on a horizontal timeline axis using a mapping system between the modulation data and the color intensity of the trajectories. The mechanism displays one modulation source at a time on the event display we have described before. Our strategy here to display continuous data is similar to displaying the harmonic content in STFT, where the magnitude of each partial is given with color intensity. Visualization of the analysis parameters with color encoding has been useful for a clear output in many situations [Sedes, Courribet, Thiebaut 2004] such as in the work of Timbregrams [Tzanetakis, Cook 2000]. We can select among the modulation sources for pitch, intensity, spatialization, filter parameter and phase. Each of them is normalized in the range between 0 and 1, and we can map this range to color intensity value (Figure 8). The display of scanning multiple event modulation data with color bands is reminiscent to radio telescopes presenting the spectra emitted from distant stars to reveal the material content in their structure.

An interesting result appears when we combine the macro, meso and micro modulation sources under one destination. Since we introduce deeper level of timescale in the modulation source, the macro level modulation starts to fractalize on meso and micro levels, which are easily perceived on the meso and micro color bands as brightness change (Figure 8). The color bars on the figure 8 from top to down represent the 1. macro modulation, 1. meso mod. which belongs to the 1. macro mod., 1. and 2. micro mods which belong to the 1. meso mod, 2. meso mod., 1. and 2. micro modulations. Which belong to the 2. meso mod. The white color represents to highest modulation value, the dark the lowest. While the macro, meso and micro modulation do add up, the color transitions resolution also increase. The line graph on the Figure 8 represents the branching of the modulation from macro to micro as showing the 1. macro, meso and micro modulations. Finally the color mapping successfully provides a method to illustrate how the displayed parameter is varying over different timescales which exhibits a comprehensive formalism about the sonic structure.

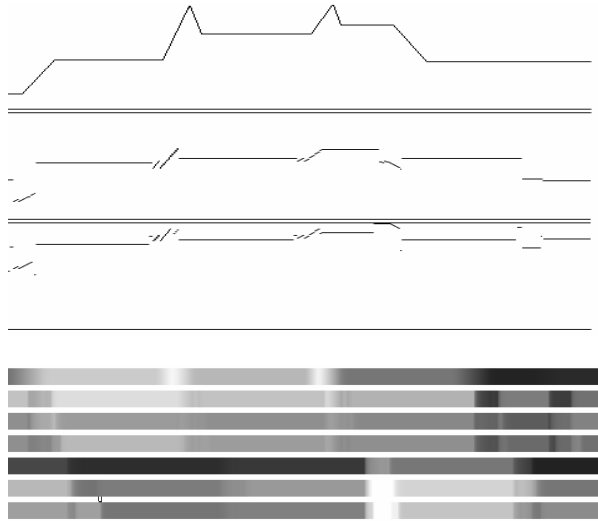


Figure 8. Time is represented from left to right by scanning the macro space with a time pointer. The line graph branches increase from macro to micro modulations on the above part of the figure, and the color transition resolution of the color bars increases from macro to micro event mod. in the lower part of the figure. Both represent the fractalization of the modulation curves clearly.

### 4.3 The extracted parameters derived from STFT analysis

Other than the modulation parameters, which we are receiving from *Cosmos*, we introduce now the analysis data extracted from STFT. This data is being obtained via the analysis tool of T. Jehan<sup>5</sup> made available for Max/Msp. The audio channel coming from each event space of *Cosmos* is analyzed with this object, which delivers continuous data of intensity and pitch analysis of the selected partials and noisiness parameters. This data can be interpreted like we have done for the modulation data of *Cosmos* and shown as color information.

At first sight, the intensity data from the analysis and the intensity modulation data are expected to have similar projections as color mapping since the effect of the intensity modulation applied by *Cosmos* is proportional to the intensity analysis values.

There is a lot to say in terms of psychoacoustic results about the pitch data. The representation of the perceived pitch and the pitch modulation data from *Cosmos* can be quite different depending on the sound source used for the micro-events.

There exists the rhythm to pitch phenomenon, which is introduced by high density event distribution inside event spaces of *Cosmos*. For instance, a small micro sound event distributed regularly with a density as much as 60/sec. introduces a secondary perceived pitch component. Where the analysis of the micro event spaces might become closer to the original pitch of the micro-event, the meso and macro event spaces could introduce different pitch analysis as an effect of higher density in those layers. If the source audio material for the micro events is pure harmonic and doesn't exhibit complex spectra, the interpretation of the pitch analysis becomes clearer.

On Figure 9a, we present an experiment where we have assigned an audio sample to micro events and kept all pitch modulation constant on all layers of *Cosmos*. The meso space densities have been controlled by the stochastic Poisson distribution. Macro event durations and the meso event onset time distribution have been controlled with the Exponential distribution. The macro and micro space densities have been kept constant at 4 events/space. Finally the macro space length is 1000 msec. These settings introduce a mechanism where local densities are being created on meso and macro spaces and are dynamically changing the analyzed pitch values on the meso and macro level analysis as shown on the Figure 9a. From top to down, the color bars show the analyzed pitch of the 1. macro event, 1. meso event and 2. meso event which belong to 1. macro event, 1. micro event (the white row) which belong to 1. meso event and finally the macro space. You can clearly see how the local densities change the pitch as being mapped to darker colors, which represent low pitch values.



Figure 9a. Pitch values derived from STFT analysis are mapped as to color intensity. 4. Row appears white, since the white color is assigned to 1452Hz, which is the analyzed micro event pitch being held constant by *Cosmos*.

Another experiment is presented on the Figure 9b. where the pitch of the micro events are controlled by the uniform stochastic function. This is the only difference from the experiment on Figure 9a. A modulation has been introduced on the micro level pitch by *Cosmos*. But the visualization gives us completely different projection of this modulation on macro and meso layers as a result of the non-linear effect of the local densities.

<sup>5</sup> analyzer~ object developed by T. Jehan included in his Max/Msp externals bundle

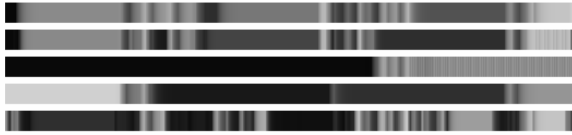


Figure 9b. The micro event pitch changes now stochastically as shown of the 4. row. But the local densities on meso and macro layers introduce different pitches then the modulation applied on the micro events.

The noisiness/harmonicity analysis will be also varying according to the elements described above. Although the source material may be inharmonic and show strong noise content on the level of micro-space, the secondary pitch introduced by the local densities in meso and macro spaces, could mask this property of the source audio material by introducing a harmonic characteristic leading to a change in the perceived macro structure. It becomes clear that the change introduced between the micro and meso levels, will be iterated further between meso and macro levels, which is clearly an emergent property of the system.

We have seen how useful it can be to utilize at the same time the event data representation and the STFT based analysis for high level sound parameters coming from Cosmos to reveal the structural facts happening on multiple layers of sonic event organization. The information delivered by the visualization application helps to realize a scene analysis of the macro-sound structure created by the *Cosmos* by revealing the layers of the complex sound.

## 5 Conclusions

The visualization experiments presented in this paper are examples of a practical representation of complex event organization. Operational parameters like time, pitch, intensity, onsets and duration of individual events need to be exposed during the immense data flow in real-time with distinct dimensions, as they are independent from each other and relevant to the perception. The visualization helps to provide an intuitive feedback to support the control of sound parameters. The *prototype* of the visual application will be developed further to represent 3D displays. The aim here is not to represent as much data as it can be on one screen but keeping the visual feedback on a level of hassles interpretation of the visual content. We are also very close to use now the visual processing techniques found on computer graphics applications for manipulating the sound parameter data backwards in a meaningful way for the synthesis application which is *Cosmos* in our case.

## 6 Acknowledgements

Thanks to Pat Healey from IMC, Department of Computer Science Queen Mary University of London for providing the residency of S. Bokesoy for the research collaboration.

## 7 References

- Miller, Arthur. I. 1996. *Insights of Genius*. Cambridge, Massachusetts: MIT Press.
- Bregmann, A. S. 1990. *Auditory Scene Analysis: The perceptual organization of sound*. Cambridge, Massachusetts: MIT Press.
- Bokesoy, S. 2005. The Cosmos Model: An event generation system for synthesizing emergent sonic structures. In *Proceedings of the International Computer Music Conference*, Barcelona: International Computer Music Association.
- Sedes, A., Courribet B., and Thiebaut J. B. 2004. Visualization of Sound as a control interface. In *Proceedings of the Digital Audio Effects Conference*, Naples.
- Tzanetakis G., Cook P. 2000. Audio Information Retrieval (AIR) Tools" In *Proceedings of International Symposium on Music Information Retrieval (ISMIR)*, Plymouth, Massachusetts.