AN INTERACTIVE FRAMEWORK FOR MULTILEVEL SONIFICATION

Nuno Diniz, Michiel Demey and Marc Leman

IPEM, Ghent University
Ghent, Belgium
nuno.diniz@ugent.be

ABSTRACT

In this paper, a conceptual framework for interactive sonification is introduced. It is argued that electroacoustic composition techniques can provide a methodology for structuring and presenting multivariable data through sound. Furthermore, an embodied music cognition driven interface is applied in order to provide an interactive exploration of the generated output. The motivation and theoretical foundation for this work, the framework's implementation and a exploratory use case are presented.

1. INTRODUCTION

The development and application of processes that allow the transmission of information using sound has always been a main concern of music composition practice. Particularly in the 20th century, several theories have been suggested for establishing a meaningful and coherent binding of individual sound streams or events. However diverse these approaches might be, they all address the same problem: how to establish a unified context between hierarchical levels of communication that are exposed simultaneously through time. As in music, it is relevant to take this problem under consideration when presenting multivariable data through sound. For illustration purposes, consider the situation where three variables are sonified at a given moment with the C, E and G musical pitches. The presence of a higher level of meaning (a major chord) as well as the intermediate ones (such as the intervals formed by the combination of the individual elements in the pitch set) should be taken into account with the same degree of importance as the individual pitches. The work presented here is focused on the exploration of a framework which provides a simultaneous consideration and encoding of these interrelated planes as this process is a key element in the definition of structures that convey the forming of contexts in sound data presentation. In the following section, an brief overview of the compositional views of Pierre Schaeffer and Karlheinz Stockhausen is presented in order to establish a relation between musical composition practice and multilevel sound communication. Afterwards, we present the motivation underlining the use of embodied music cognition theory's concepts as interface paradigms for interactive sonification. Then, the interface design, the technological aspects and user evaluation of a exploratory use case are addressed. Finally, a discussion of the present work is provided.

2. MUSICAL COMPOSITION AND MULTILEVEL SOUND COMMUNICATION

The application of musical enabled processes in non speech sound communication has been present in the auditory display research since the early stages of this discipline [1] [2] [3]. However, our goal here is to extend the scope of previous investigation with a unified top down/bottom up approach as "the human approach combines a bottom-up approach with a top-down approach" and has a "tendency for organizing event structures in coherent sections" [4].

In the two main initial trends in electroacoustic music, the french Musique Concrete and the Electronic Music from Cologne, the search of ways for establishing relations between material and form is present in the theoretical and compositional production of their leading advocates, Pierre Schaeffer and Karlheinz Stockhausen. According to Michel Chion's Guide to Sound Objects [5], the sound object, as defined by Schaeffer, is "perceived as an object only in a context, a structure, which includes it". This dependency relationship between individual and group is further develop in the sense that "every object of perception is at the same time an object in so far as it is perceived as a unit locatable in a context, and a structure in so far as it is itself composed of several objects". One can extract from such postulates that the dialogue condition that is imposed to the sound object and the structure holds a dynamic perspective shift that reassures the relationship between these two concepts. From his part, Stockhausen's concept of unity concerning the possibility to trace all musical parameters to a single compositional principle [6] envisioned the unified control of the musical structures in a given work through the establishment of inherent relationships between the micro and the macro level of the musical discourse. Although possibly initially driven by the aims of integral serialism, his search for such mechanisms of scope transposition continued throughout his career. Of such techniques, one can highlight moment form, a structuring paradigm based on a non linear distribution of gestalts known as moments, or the formula based composition in which all aspects of a given work derive, either directly or indirectly, from a initial short composition. As an example, his over twenty-nine hours long opera cycle "Licht" is based on a three part, eighteen bar only score formula. Although, as argued by Vickers [7], "The difference (...) between sonication and musical composition is largely one of perspective", it is surely arguable that these concepts can be fully applied outside the art and music realm. Nevertheless, they encapsulate a set of guidelines that can be of service in functional sound based communication, as defined in [8]. As Delalande pointed out [9], there is a communality of processes in electroacoustic composition practice that concern the relationship between singularity and regularity of events which underlines structural dependencies between singular entities used in the musical discourse. As such, the aim of this work is to transpose the above mentioned compositional concepts to the interactive sonification domain and apply the relationships between material and form to the micro and macro sound levels of data presentation. As a result, functional context definitions are generated by data dependent hierarchical levels that nevertheless preserve their informational identity and significance.

3. INTERACTIVE SONIFICATION AND EMBODIED MUSIC COGNITION

As defined by Hermann and Hunt, interactive sonification is "the use of sound within a tightly closed humancomputer interface where the auditory signal provides information about data under analysis" [10]. Given the initial premisses described in the last sections, the proposed interface for interacting with the framework's musically structured output follows an embodied music cognition perspective [11]. With the objective of promoting a fruitfully dialogue between the user and the data, an approach based on the expansion of the mediating role of the body through the manipulation of virtual entities within an immersive environment is considered (in relation to [12]). First, virtual objects can act as mediators representing multilevel mapping layers that conform with the premiss of a hierarchical object oriented decomposition of sound entities. Second, given its inherent multimodal nature, a virtual reality based framework presents itself as an appropriated setting for the investigation and development of interfaces between body and music in which the natural communication tools are covered through the immersion of the actors involved. By enabling a configurable location and form representation of the data in space, this methodology invites the user to a physically based approach to the inspection process through a shared space of multilevel interaction. As such, an embodied cognition approach is expected to further enable a perceptual link between the data under inspection and the semantic high level representations of the user (see Figure 1).

The framework's concrete implementation and an exploratory use case are the subject of the following sections.

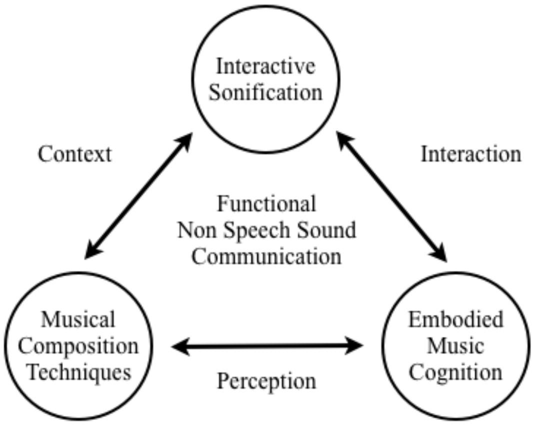

Figure 1: *Focused relations in this project.*

4. FRAMEWORK

4.1. Architecture overview

The design is based on a functional division of the multimodal spectrum into individual branches around a virtual environment state representation. Following a top-down approach, a first level is composed of abstract managing cores and their respective elements per modality - visual, auditory and human interface. A second level is then obtained by concrete implementations through the use of external libraries. So, as a result of this encapsulation, the concrete implementations of the virtual worlds, their visual and auditory representations and the human interfaces that enable the manipulation of the virtual objects can be either refined or substituted according to the desired performance, access or functional needs of the intended use cases. The user configured binding between the elements in play follows the observer design pattern. It is provided through the implementation of custom tracker objects that read and update the relevant entities through event triggering or user defined refresh rates. Furthermore, this modular design allows both static and realtime processing of data as well as physical model based interaction.

To further illustrate the framework's design follows a concise description of the sonification package structure.

- Core/Element - Both core and elements implement generic interfaces concerning the frameworks kernel (ISoundCore; ISoundElement) and the external library used in the implementation (Ex. ISoundCoreSC3). It is segmented per library and functional task and contains the implementation of the synthesis controller. Ex. SonificationIntervalSC3 class.
- Sonification - Implementation of the sonification levels. These provides the triggering algorithm for the synthesis controller instances. Ex. SonificationLevel0 class.
- Model - Provides in real time the data for sonification. Defines the specified model for data conversion and source connectors. Ex. WiiPitchValueToFreqConverter class.

4.2. Java Technology

For portability and scalability purposes, the framework's kernel was implemented using Java technology. The primary reasons for this choice are Java's object oriented paradigm, cross-platform support, a wide range of modular freely available open source libraries and a robust interconnecting framework with virtually every IT application area. As particularly relevant, we can underline databases connectors, mobile and data mining framework, web service based access, web start deployment technology and support for various functional and/or interpreted languages (Ex. Python). In the case of specific performance and/or compatibility demands, it is possible to make use of C/C++ code via the Java Native Interfaces through component wrapping. Finally, a strong argument in favor of the implementation of real-time software in Java is the continuous evolution in the Real-Time Specification for Java's implementations (RTSJ).

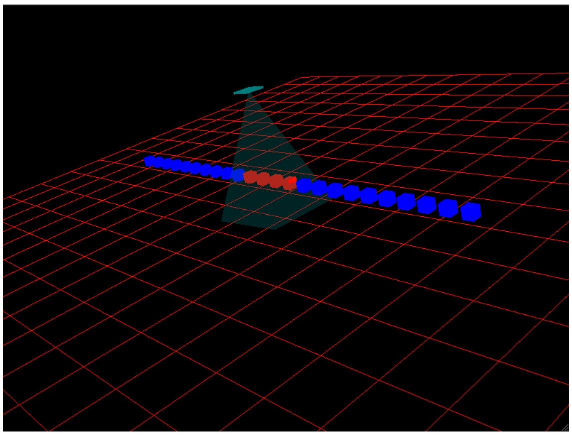

Figure 2: *Visual feedback of the described use case showing the inspection window and the inspection tool controlled by the user's hand.*

5. USE CASE

5.1. Description

The presented use case consists of the interactive exploration of one dimension dataset through sound (See Figure 2). The main goal was to present the test subjects with a simple use case in order to extract preliminary issues concerning the framework.

Both the users hand and the frequency stream are mapped as virtual objects in a 3D space. The latter represents an inspection window composed, in this case, of 24 independent geometrical objects. Each one of them constitutes an access point to a single variable's frequency values calculated by the provided model. By using his hand as a sonic magnifying glass or a virtual microphone, the user can zoom in and out in order to investigate either one elements output or its relationship with other members of the set. This interaction mode was strongly inspired by Stockhausen's Mikrophonie I composition where the active use of microphones is a base concept in the performance of the piece. As Stockhausen said concerning this approach, "normally inaudible vibrations . . . are made audible by an active process of sound detection (comparable to the auscultation of a body by a physician); the microphone is used actively as a musical instrument, in contrast to its former passive function of reproducing sounds as faithfully as possible" [13]. Each independent virtual sound source is activated through collision detection when the inspection tool's volume intersects the virtual objects. At this point, the activated items are feed into the sonification levels responsible for calculating the respective sonic outputs according to their specific implementation.

At this point, the virtual objects, their structure and their relationship with the sonification layers are addressed. In this example, only the individual virtual elements that compose the array set of the inspection window are subject to sonification procedure. However, following the previously referenced theoretical guidance of Schaeffer, several sonification layers can be defined in order to map this hierarchical definition of the virtual entities. As mentioned, the inspection window's representing array (the parent object) is composed of 24 cubes (the child objects). Here, the manipulation of the parameters involved in the sonification of the individual cubes comes into play. As one gets the inspection tool closer to the activated elements, the distance between them has an affect on the amplitude and depth of the reverberation. The shorter this distance is, higher the loudness and smaller the depth of the reverberation will be. Although this behavior is assigned to the individual elements, it conveys information about the activated set as a whole. It stimulates a perceptual interpolation between the relative whole and the individual nodes. Furthermore, this approach implements the philosophy that the difference between a sound object's constitution being either parameters or a set of subordinated sound objects is mainly one of perspective.

5.2. Sonification Levels

Three independent sonification levels were define in which the data mining processes are driven by musical relations present in the data.

- Level 0 - This level manages the sound output concerning the individual entities in the scene. It updates and triggers the assigned pitch of the activated items. This level was implemented through individual sine wave oscillators for each activated element.
- Level 1 - This level is responsible for detecting and sonically activating musical intervals between two virtual entities under inspection. These relations are defined as a ratio between two given frequencies and used used to highlight degrees of variation of the data. For example, a perfect fifth interval can be used for detecting a relation of 3/2 between two elements within the array. This level was implemented through the use of a resonant filter bank per interval. Its application consisted in a percussive type activation each time a given interval was detected.
- Level 2 - This level establishes a relation between several elements and their frequencies in the inspection scope. The presence of a music chord is calculated through the detention of N ratios or intervals from a base frequency. For example, a C major chord is detected through the simultaneous presence of three frequencies: the base F0 and two other that, in relation to F0, respect the 3/2 and 5/4 ratios conditions. By defining and sonically highlighting these relations, further information is provided through a wider view of the data's progression. This level was implemented through the use of a set of delayed sine wave oscillators per chord detection.

5.3. Technology

This preliminary use case several external libraries and additional technologies. They are presented by modality in the following items.

- Visual - Java 3D Library was selected for the visual engine for its high level scene graph based implementation, well structured overall design and functionalities.
- Sound - The sound engine has been implemented using Supercollider 3 through JCollider, a Java based SCLang implementation using the NetUtil OSC Java library [14].
- Human interface - The NaturaPoints OptiTrack motion capturing system provided the tridimensional position and orientation tracking through an OSC custom client.

5.4. User evaluation

The following preliminary evaluation consisted of measurement of the users's performance, while conducting predefined tasks and collecting their personal appraisal regarding the interface.

The proposed tasks were comprised of exploring a predefined dataset using different combinations of the sonification levels. For example, a user would be asked to find a certain relation present in the proposed dataset with only the lowest level of sonification. Then, this user would be asked to repeat the task using the same level combined with one of higher degree (Ex. Level1). After exploration of the interface and performance of the set tasks, participants were required to evaluate the human interface used in terms of performance, maneuverability and precision. Moreover, participants were asked to comment on the visual output completeness (in order to find and interact with the virtual array) and on the sonification output in terms of distinguishability, information carrying potential and aesthetic design. The feedback provided by the subjects pointed out some problems concerning the virtual objects interaction (i.e. the need to dynamically change the morphology of the inspection tool), the visual output (i.e. the need for a second view that conveys depth perception) and to interactively change the dimensions of the inspection window. Concerning the sonification output, the users reported being able to perceive all levels and discern the information conveyed to them. However, it was generally noticed that the activation of such levels should be interactive and made available during task performance. Nevertheless, performance considerably improved by the use of multiple levels of sonification.

6. FUTURE WORK

The future development of this project will progress in several ways. Firstly, we will focus on the expansion of the sonification levels and their intercommunication in order to progressively incorporate higher levels of representation. These will developed not only as a function of the simultaneous data streams at a certain point (a "vertical" score analysis) but a time based analysis ("horizontal" score analysis) in which the result of the sonification process takes into account previously examined samples. Such development will contribute to a more global, musical form inspired perspective of the data's inner relationships by sonically placing local behaviors within a broader context. Other modes of interaction with the sonification levels will be explored. For example, besides the regulation of the amplitude and reverberation parameters, the relative distance of the virtual microphone and the object(s) under inspection could also be used for the activation and mixing of the sonification levels. Furthermore, in order to the morphology and sonic feedback of the virtual elements to reflect the data's behavior, further investigation in incorporating physical model based interaction will be carried out. As Stockhausen commented about Mikrophonie I, "Someone said, must it be a tam-tam? I said no, I can imagine the score being used to examine an old Volkswagen musically, to go inside the old thing and bang it and scratch it and do all sorts of things to it, and play Mikrophonie I, using the microphone" [15]. Secondly, concerning the interface, future testing will include the real time configuration by the user for positioning the inspection window in the dataset, adjusting the inspection tool dimension parameters and the activating the sonification's levels. Third and finally, all of these features will be subject of a more comprehensive usability study in order to validate the present and future modes of user interaction in new inspection scenarios (Ex. the simultaneous inspection of N variables datasets).

7. CONCLUSIONS

The presented article aims to established further relationships between the interaction sonification field and musical composition practices. Although the present development is still in an initial stage, preliminary testing has shown that the progressive inclusion of the previously mentioned concepts and its related techniques can contribute to close the semantic gap between the user and data through sound.

8. ACKNOWLEDGMENTS

The authors would like to give a special thanks to Alexander Deweppe for conducting the user's feedback preliminary study and Pieter Coussement for the guidance concerning the sound design. This research has the support of the EmcoMetecca project.

9. REFERENCES

[1] J. Alty, "Can We Use Music in Computer- Human Communication?" in *proceedings of HCI '95*, Huddersfield, August 1995.

[2] M. Blattner, D. Sumikawa, R. Greenberg, "Earcons and Icons: Their Structure and Common Design Principles" in *Human-Computer Interaction*, pp. 11-44, 1989.

[3] J. Hankinson, A. Edwards, "Musical Phrase-Structured Audio Communication" in *Proceedings of the 6th International Conference on Auditory Display*, Atlanta, GA, USA, 2000.

[4] O. Kuhl, K. Jensen, "Retrieving and Recreating Musical Form" in *Lecture notes in computer science*, Springer, 2008.

[5] M. Chion,"Guide To Sound Objects. Pierre Schaeffer and Musical Research", (Trans. John Dack and Christine North), http://www.ears.dmu.ac.uk/.

[6] M. Clarke, "Extending Contacts: The Concept of Unity in Computer Music" in *Perspectives of New Music*, Vol. 36, No. 1, pp. 221-246, Winter 1998.

[7] P. Vickers, B. Hogg, "Sonification Abstraite/Sonification Concrete: An ' Aesthetic Perspective Space For Classifying Auditory Displays In The Ars Musica Domain' " in *Proceedings of the 12th International Conference on Auditory Display*, London, UK June 20 - 23, 2006.

[8] T. Hermann, "Taxonomy And Definitions For Sonification And Auditory Display " in *Proceedings of the 14th International Conference on Auditory Display*, Paris, France June 24 - 27, 2008

[9] F. Delalande, "Towards an Analysis of Compositional Strategies" in *Circuit : musiques contemporaines*, vol. 17, n 1, p. 11-26, 2007.

[10] T. Hermann, A. Hunt, (eds.) "Special Issue on Interactive Sonification" in *IEEE Multimedia*, April-June, Vol. 12, No. 2, 2005.

[11] M. Leman, "Embodied Music Cognition and Mediation Technology". Cambridge, MA: MIT Press, 2008.

[12] A. Mulder, S. Fels, K. Mase, " Mapping virtual object manipulation to sound variation" in *IPSJ Sig Notes*, 1997.

[13] K. Stockhausen, " Mikrophonie I (1965), fr Tamtam, 2 Mikrophone, 2 Filter und Regler." in *Stockhausen, Texte zur Musik 3*, Cologne: Verlag M. DuMont Schauberg, p. 5765.

[14] H. Rutz, JCollider and Netutil Java library. https://www.sciss.de/.

[15] K. Stockhausen, "Stockhausen on Music. Lectures and Interviews compiled by Robin Maconie." London and New York: Marion Boyars, 1989.