

Interactive Sonification of the Spatial Behavior of Human and Synthetic Characters in a Mixed-Reality Environment

Sylvain Le Groux, Jonatas Manzolli, and Paul F.M.J. Verschure
 Pompeu Fabra University, University of Campinas, Pompeu Fabra University & ICREA
 {slegroux@iua.upf.edu, jonatas@nics.unicamp.br, paul.verschure@iua.upf.edu }

Abstract

It is generally admitted that music is a powerful carrier of emotions [4, 21], and that audition can play an important role in enhancing the sensation of presence in Virtual Environments [5, 22]. In mixed-reality environments and interactive multi-media systems such as Massively Multiplayer Online Games (MMORPG), the improvement of the user's perception of immersion is crucial. Nonetheless, the sonification of those environments is often reduced to its simplest expression, namely a set of prerecorded sound tracks. Background music many times relies on repetitive, predetermined and somewhat predictable musical material. Hence, there is a need for a sonification scheme that can generate context sensitive, adaptive, rich and consistent music in real-time. In this paper we introduce a framework for the sonification of spatial behavior of multiple human and synthetic characters in a Mixed-Reality environment. Previously we have used RoBoser [1] to sonify different interactive installation including the interaction between humans and a large-scale accessible space called Ada [2] Here we are investigating the applicability of the RoBoser framework to the sonification of the continuous and dynamic interaction between individuals populating a mixed-reality space. We propose a semantic layer that maps sensor data into intuitive parameters for the control of music generation, and show that the musical events are directly influenced by the spatial behavior of human and synthetic characters in the space, thus creating a behavior-dependant sonification that enhance the user's perception of immersion.

1. Introduction

Sonification is a relatively new field of research that has developed rapidly in the recent years. It has been used for various purposes, and some standard techniques and applications have emerged [3]. For instance, Auditory Icons and Earcons models are commonly used to signal specific events to the users. Audification models convert data series into sound samples for scientific data mining [3]. Furthermore, model-based sonification allows a better understanding of the structure of data by using plausible physical sound synthesis algorithms driven by the data itself.

Our work relates to yet another type of sonification, called *behavioral mapping* where data from the environment

and the behaviour of agents (human and/or machines) are mapped to acoustic and musical attributes. If scientific sonification of abstract data is usually intended to give an insight in the structure of the data for analysis purposes, our objective is somewhat different. We aim at improving the social interaction in the environment and create a musical world that has to be both informative and aesthetically appealing. We want to enhance the comprehension of the environment, embed the user in a rich and consistent audio-visual environment while producing a musical experience.

It is widely acknowledged that music is a powerful carrier of emotions, and that auditory cues play an important role in enhancing the sensation of presence in virtual environments [4, 5, 6, 7]. Mixed-Reality spaces are becoming more and more common and various labs are building their own environments (e.g. the Allosphere at UCSB, the intelligent House at MIT, the Nanohouse at UTS and the Sentient Lab in the Faculty of Architecture, University of Sydney...).

Until now music generation and sonification for large-scale mixed-reality environments has not been extensively studied. However, we can extract basic design guidelines from previous studies. For instance, [24] showed that multichannel systems enhance the sense of presence. It is also assumed that moving sound sources increase the dynamicity of the environment, and that interactive sonification makes it more interesting for the user [23]. In addition, repetitive loops are likely to be detected by the user and perceived as artificial [23], which means our system should be able to produce a reasonable variety of sounds.

In this paper we present an original framework for adaptive sonification of the spatial behavior of entities in a mixed-reality space based on the aforementioned guidelines. In contrast with many data sonification approaches, we do not sonify directly the sensory data, but we rely on higher-level behavioral information that is extracted from sensor data. We propose to extract this high-level spatial/social behavior from sensor data via a neuromorphic processes designed with a real-time neural network simulator [3]. This spatial behavioral information is then used to modulate a real-time real-world composition engine. The final correlation between auditory representation and understanding of the space is based on the study of perceptual effects of musical and sound generation parameters. As such, applications of our system could be also

related to very recent virtual reality therapy systems [9] and we are studying effects of behavioral mapping in sonic therapies.

2. The mixed reality space

To better understand the context of our work, we give a brief description of the mixed-reality space we have constructed. This mixed-reality environment is the continuation of the Ada project, an interactive space exhibited for the Swiss national Expo in 2002. [10]. It consists of two main parts. On the one hand we have a physical interactive space of 7x7 m. called XIM (eXperience Induction Machine) and on the other hand the virtual environment called PVC (for Persistent Virtual Community). The PVC is divided into four nested regions: the Clubhouse is the virtual counterpart of the XIM physical space, and is itself surrounded by the P-Garden and the Avatar-Heaven, a retreat area for avatars (see Figure 1).

The community is interfaced to this mixed-reality environment via several communication channels. For instance, interactions can take place between visitors of the physical space, and/or between users that connect over the Internet (Remote Users), but it can also take place between fully synthetic characters or the virtual environment itself.

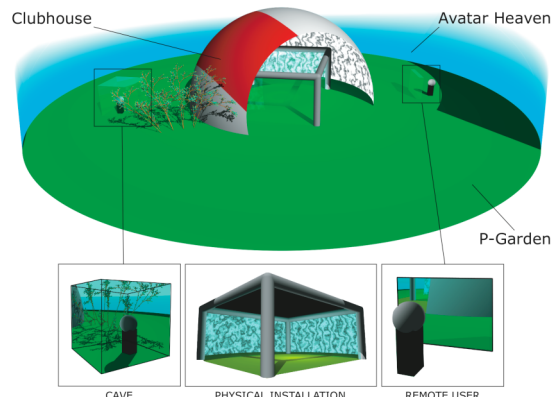


Figure 1. The Persistent Virtual Community consists of the physical installation (the Experience Induction Machine), the ClubHouse, the P-Garden and Avatar Heaven. Users can visit the physical installation XIM, connect over the Internet (Remote Users) or visit a CAVE installation.

From the general structure of the PVC, we distinguish four main areas where sound should contribute to the overall interactive experience. First, we would like to generate sound in the physical space itself (the XIM) to accompany, illustrate, transform and influence the interaction of physical and virtual users within the physical space. Secondly we wish to find a coherent sonification scheme where each avatar's personal contribution is taken into account and added to the global sonic output. Finally, the virtual environment at both

the client and server side has a role to play in the global sonification scheme, involving remote control, sound streaming, client/server sound synthesis. The variety of the possible modes of interaction and communication paradigms is apparent, and the sonification of such a complex mixed-reality system is an intricate task. Hence, there are three main tasks for the sonification process: 1) soundscapes: here the concept is to embed the physical and virtual environment within a dynamic soundscape. The sonic result can be related to a naturalist approach in which the sounds are used to correlate natural or biological-like structures with sonic events. Concepts such as Sonic Ecology can be related to this approach, 2) synthetic voices: here the concept is to bring to the acoustic space the agent's voices and from that to construct structures such as dialogues and narratives and 3) musical realm: here the concept is to bring a development of the Roboser [1] to create a dynamic musical discourse in which the agent's behavior is related to expression and emotional experience.

In this paper we will more specifically focus on the sonification of the physical space XIM in relation to the behavior of humans and avatars that populate this mixed-reality environment.

3. System architecture

The real-time/real-world system we designed is made up of three main interconnected modules (C.f. Figure 2.) that can be described in terms of a sensor interface, a central-processing stage, and a sound synthesis stage [11]. This is also sometimes referred to as the sensing, processing, and response paradigm [12].

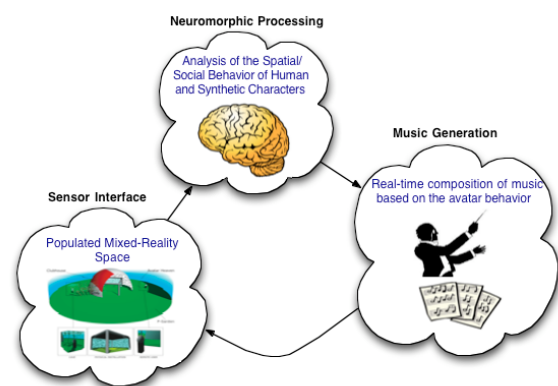


Figure 2. The architecture of our real-time interactive system follows the sensing/processing/response paradigm.

The first module consists of the mixed-reality environment populated by “behaving” human and synthetic characters. The second module consists of a neural model that processes the sensor data and interprets it as behavioral data. Finally, the third module deals with the interactive composition of musical structure and the generation of

sounds. The musical output is modulated by the behavior of the human and synthetic users.

The different modules are all interconnected via TCP/IP connections via Ethernet on a local network.

4. Overview

In this section, we present the strategies underlying the implementation of the different modules that compose our interactive real-time sonification system.

4.1. Sensing

Accurate and robust tracking of movement and people interacting with the XIM is a complex task carried on by a team of colleagues in our lab. Their approach focuses on multimodal information fusion to improve tracking, automatic assessment of sensor data quality and analysis of methods for intelligent recruitment of sensors to resolve conflicting situations.

XIM is equipped with a number of sensors that are used to determine its internal states and to map the physical XIM onto its virtual representation. In addition, XIM can use its effectors to influence the behavior of its as well as to regulate human-avatar interactions and to provide the best operating conditions for all sub-systems. The sensors that are currently used are 3 overhead cameras and frame grabbers, 4 gazers, 3 microphones, and the pressure sensitive floor. They provide information for the accurate tracking of the visitors and can influence the emotions and behaviors of XIM. The effectors are 8 speakers, sound synthesizers, 6 video projectors, 8 lightfingers and the floor (see Figure 3).

With our technology, we can offer a reliable tracking of movement and people in the XIM, which, combined with the trajectories of the avatars, form an accurate map of each individual trajectory.

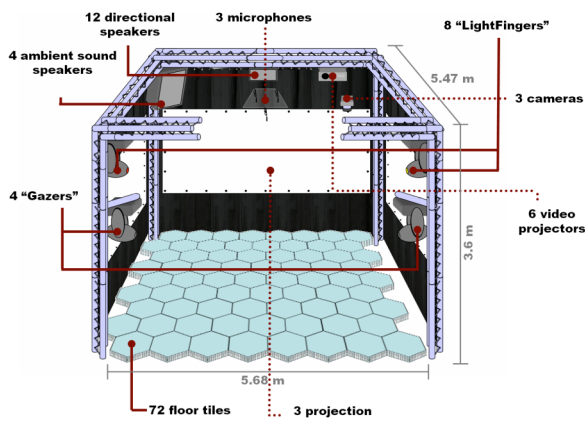


Figure 3. XIM is provided with a set of sensor and effectors. Sensors allows for the tracking of people in the space. Effectors are used to influence the emotions and behavior of the visitors.

4.2. Neuromorphic analysis of user behavior

A simulation environment named iQR [8] (iqr.sourceforge.org) handles the real-time processing of the tracking system outputs. iQR is a flexible tool for creating and running simulations of large-scale neural models and provides a user-friendly graphical interface to design multilevel neural models (see Figure 4).

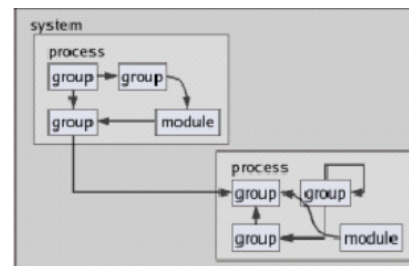


Figure 4. This example illustrates an iQR module where two processes are interconnected. Each process contain a neuronal circuit made up of three neuronal groups. (courtesy U. Bernardet)

In this project, we used iQR to transform the tracking data from the real and synthetic characters into higher-level relevant control parameters for an interactive real-world music generator. iQR allows us to model a biologically plausible neural network for spatial/behavioral analysis.

The spatial behavior of each individual is mapped to a neuron group called Floor Layout of size 56 * 53 neurons representing the spatial topology of XIM (Figure 5). When individuals move from one point in space to another, the corresponding neurons are activated. With this representation, we can design neuromorphic processes performing an analysis of the behavior of the population in the space.

For instance, motion is defined as the difference of neural activity in the cell group representing people's position at two time intervals. We also defined a neuronal group that counts the number of people populating the space by summing up the total activity of the occupancy group. We

then virtually divided the space in four main activity areas (up/down/left/right corners), and created the corresponding four neuronal groups. The synapse connectivity from the floor layout to those four groups can be defined so that each group receives the activity of only a subset of the Floor Layout group, thus allowing us to know which of the four areas of the space is the most active. The different modes of interaction of XIM are based on such analysis (if one individual is in the space, the XIM will behave differently than if there are two people, or none, etc. See Figure 5.)

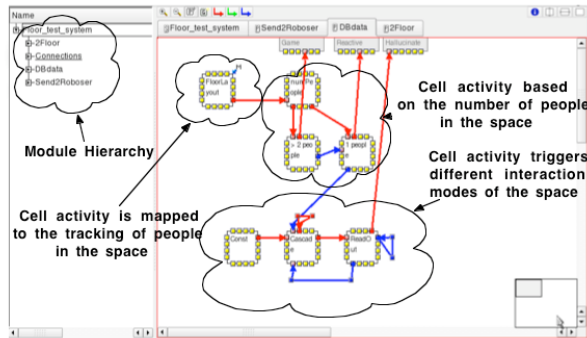


Figure 5. With iQR, we graphically designed a module that extracts behavioral information from the tracking system. Cell activity of the different interconnected neuronal groups is related to the activity of people in the environment.

We now have a high-level representation of the spatial behavior of people in the space and are able to define musical parameter and specific neuronal groups, which activity will control the musical output generation.

4.3. Real-World Composition Engine

The music generation, interaction and composition tools are based on our work on the synthetic composition engine RoBoser [1]. The paradigm for musical interaction in XIM, called Real-World composition, is grounded in our work on large-scale interactive multi-media systems [10].

Our aim is to integrate sensory data from the environment in real time and interface this interpreted sensor data to a composition engine. In this way unique emergent musical structures can be generated. In our previous work on Roboser, we have shown how the dynamics of a real-world system induces novelty in the micro-fluctuations of sound control parameters [6]. Here our goal is to use the Roboser framework to transform the social and spatial behaviors of human and avatars in the XIM environment into new musical structures.

5. Music generation

In order to synchronize musical composition to the behavior of real and synthetic characters in the mixed-reality environment, we represent the music as a parameterized structure, which values are dynamically controlled over time. We classify the musical parameters into two main categories. On the one hand are the macro-level structural parameters such as tonality, rhythm, tempo, etc. On the other hand are the micro-level parameters related to the synthesis of sound (instrumentation, brightness, harmonicity, spectral complexity...)

5.1 Music Parameterization and Generation

We chose a set of standard musical parameters for controlling the generation of music based on the requirement that their modulation should have a clear perceptual effect. We kept a list of parameters that has been extensively studied, and which effect on emotional expression is widely acknowledged, as described in [13]. Those parameters are tempo, mode, volume register, tonality and rhythm for musical structure and articulation, brightness and harmonicity for the sound generation. With this set of parameters we are trying to study how the macro and micro levels can be used to express an effective sonification.

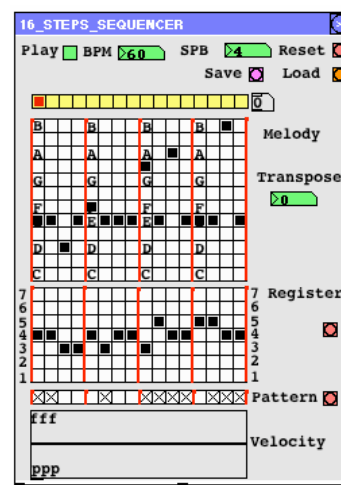


Figure 6. This is our GUI-based abstraction in Pure Data that allows the creation of a basic cell with musical material. Musical parameters such as register, pitch, rhythmic pattern and tempo can be defined, saved and loaded at anytime. All those parameters can also be modulated in real-time via external Open Sound Control messages (OSC protocol).

The generation of music is based on the real-world composition paradigm (c.f. Section 5.3.), where prepared musical material is dynamically modulated as the users interact with the mixed-reality space. We designed a GUI interface in Pure Data, a real-time graphical dataflow programming environment for audio [4] that allows writing, loading and saving musical structure material by hand. The parameters corresponding to different settings are saved in a text file or Style File. This structure defines a basic musical cell (Figure 6). It is possible to recall different musical cells from outside the abstraction by sending appropriate Open Sound Control messages (OSC is a protocol for communication among computers, sound synthesizers, and other multimedia devices optimized for modern networking technology) to the right inlet of the abstraction [16]. This will load the parameters saved in the corresponding text file.

A group of various related cells defines a certain style or atmosphere and correspond to a specific mode of interaction in the XIM (see Figure 7). When the interaction between people and XIM takes place, these basic musical events are dynamically modified. The musical material is modulated following some rule of interaction (see Section 6). The initial musical material is amplified, transformed, nuanced, as the interaction between the XIM and the users evolves.

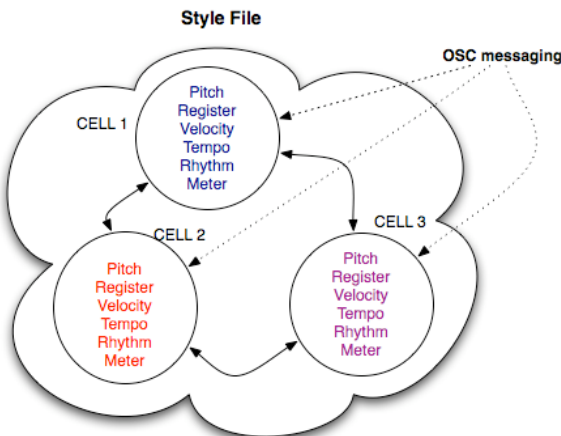


Figure 7. A group of related musical cells defines a certain style file which specific structural and musical parameters are associated with a mode of interaction of the XIM. When a style file is enabled, we cycle through the different cells whose properties are modulated in real-time via OSC control messages.

5.2. Sound Generation

For the generation of the sound itself, we designed a Pure Data module that interfaces MIDI-based control messages to wavetable synthesizers. Figure 8 shows the Graphical User Interface of our wavetable synthesizer where

the user can choose an instrument, a midi channel and control parameters such as volume, note duration, modulation, pitch bend,... MIDI, or Musical Instrument Digital Interface is the precursor of OSC and the industry standard protocol that enables electronic musical instruments to control and synchronize with each other. One major drawback of midi samplers is that they usually lack variety in their controllable parameters. The wavetable synthesizer provides us with limited control over the generation of sound at the micro-level properties of timbre. However, it is generally admitted that the modulation of subtle timbral features is also perceptually relevant [13].

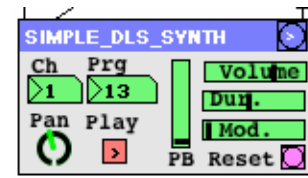


Figure 8. This abstraction allows Pure Data to control MIDI-based wavetable synthesizers in real-time.

To provide our system with fine timbral control over the generation of sound, we implemented a simple FM synthesizer. We chose the FM synthesis algorithm because of its simplicity, efficiency, and ability to synthesize a wide range of timbre with only a few control parameters [15].

We used the MIDI instruments for most of the melodic “discrete” material while the FM synthesizer added continuous layers of sound with slowly varying timbre characteristics.

Finally, to make the overall musical result more lively and interesting, we designed a soundscape generator that can trigger samples of environmental natural sounds (such as forest, sea, footsteps,...) chosen from the Freesound database (<http://www.freesound.iaa.upf.edu>).

5.3. Spatialization

Spatialization is a crucial parameter to take into account for improving the sensation of presence in a VR environment [7]. It creates the perceptual illusion of a physically realistic soundfield. We wanted to be able to virtually place various sound sources in the environment and move them around the listener in real-time in XIM. For this purpose, we used the VBAP technique, a flexible and efficient method for positioning virtual sources using multiple speakers [17].

We have built a set of musical tools, based on the real-world composition paradigm, which allows us to generate musical material in a parameterized way. The next decisive step is to define a strategy for controlling the musical output in function of the spatial/social behavior of users in the mixed-environment space.

6. Mappings

In this project we had to deal with many levels of transformation of the data. Here we present the different semantic and technical mappings we found were necessary to achieve a reasonable correlation between auditory representation and understanding of the space (Figure 9).

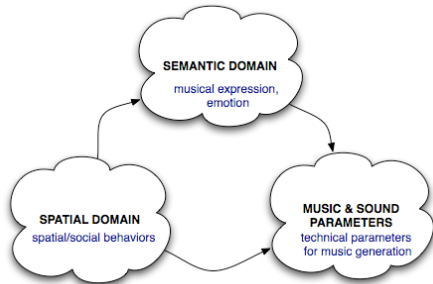


Figure 9. This diagram shows the different type of data we are dealing with during the sonification process.

First, we defined a set of relations between musical structural parameters and musical expression based on a review of literature on music and emotion. [4,5,13] (See Table 1).

Musical Parameter	Level	Semantics of Musical Expression
Tempo	Slow	Sadness, Calmness, Dignity, Boredom
	Fast	Happiness, Activity, Surprise, Anger
Mode	Minor	Sadness, Dreamy, Anger
	Major	Happiness, Grace, Serenity
Volume	Loud	Joy, Intensity, Power, Anger
	Soft	Sadness, Tenderness, Solemnity, Fear
Register	High	Happiness, Grace, Excitement, Anger, Activity
	Low	Sadness, Dignity, Solemnity, Boredom
Tonality	Tonal	Joyful, Dull
	Atonal	Angry
Rhythm	Regular	Happiness, Dignity, Peace
	Irregular	Amusement, Uneasiness, Anger

Table 1. This table displays the relation between standard musical parameters and semantics of musical expression.

In the second place, we defined a table that allows us to map from the perceptual domain to FM synthesis technical parameters [15].

Sound Parameter	Synthesizer Parameter Level	Semantics of Musical Expression
Articulation	Staccato (short)	Gaiety, energy, fear, anger
	Legato (long)	Sadness, tenderness, solemnity
Brightness	High Modulation Index	Potency, Anger, Disgust, Fear, Activity, Surprise
	Low	Tenderness, Sadness
Harmonicity	Harmonicity ratio is non integer	Calmness, tenderness, Light
	Harmonicity ratio is integer	Energy, Anger

Table 2. This table shows the relation between sound synthesis parameters and perceptual parameters.

Finally, in order to design the interaction between spatial behavior and music generation, we made use of the semantic mappings previously introduced. We described the musical effects we wanted to express, and associated them with a corresponding set of music and sound parameters using tables 1 and 2. The choice of the set of musical parameters in relation to behavioral/spatial information is very much the result of a trial and error design process and reflects some subjective and artistic decisions.

Spatial Parameter	Semantics	Associated Sound and Music Parameters
Spatial Location	Activity Surprise	Trigger specific samples. Change brightness of FM layer.
Motion	Happiness, Activity, Excitement	Tempo, Velocity, Register
Number of Users	Activity, Excitement	Density of events. Rhythmic regularities, Harmonicity, Articulation
Interaction Scenario	Surprise, Change, Anger	Change of Style Files, Scale Mode

Table 3. Table illustrating the relationship between spatial behaviors and musical parameters.

7. Data

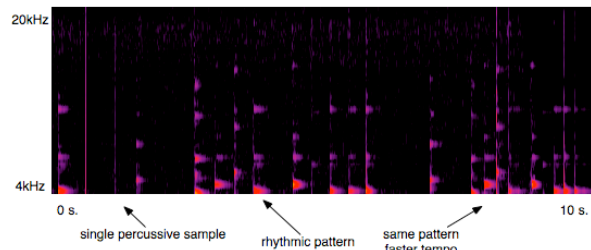


Figure 10. This spectrogram (x is time in seconds and y frequency in kHz) illustrates the sonic result of a simple interaction scheme. One user is in the space and moves from one corner to the other. When approaching the corner she triggers a percussive sample. She then moves away from the corner. As the motion of the user is accelerating the rhythmic pattern of the current cell is played at a faster tempo. The duration of each note is rather long, and most of the energy is in the low frequencies.

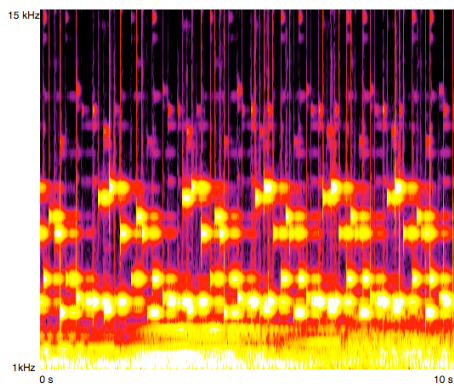


Figure 11. This spectrogram (x is time in seconds and y frequency in kHz) illustrates an interactive session where six users are moving around in the space. We observe that the density of event is high. There is also a large amount of energy in the high frequency spectrum.

We found that our music generation system offered a good representation of the spatial behavior of people in the space. The combination of soundscapes, event-based samples, midi melodies, and slowly varying FM synthesis layers provided a rich and lively sonic environment. Different set of parameters produced significantly distinct sounding results. Figure 10 and 11 illustrate the spectral content of the audio signal generated by the system for two different interaction scenarios where augmentation of tempo and of density of events were correlated to the sensation of increased activity in the space. The users perceived this auditory feedback as natural and playful. They would often start playing with the space by running around, moving slowly, staying still, thus trying to provoke various musical outputs. In the future we wish to perform a series of

experiments to quantitatively study the relation between the emotional state of the users and the different parameters of our system.

Conclusions and future work

In this paper, we have presented a complete system that can generate a complex spatialized sonification of the behavior of multiple humans and avatars in a Mixed-Reality environment based on their spatial/social behavior. We proposed a system architecture that relies on interaction in real-time between a mixed reality space, a neural network simulator for neuromorphic sensor processing and a sound and music generator. We introduced a semantic mapping layer to help the interaction design and ease the choice of relevant musical and sonic parameters.

In the future, we would like to operate on a higher level of integration between the neural control and the musical structure. To achieve this we want to use a more elaborate and adaptive neuromorphic model of sensor processing combined with a more advanced cognitive architecture called Distributed Adaptive Control, [18] for musical learning and decision-making generating a synthetic behavioral music composition system. DAC distinguishes three levels of control in a cognitive behaving system: reactive, adaptive and contextual. In the current system, the mappings are made directly from the spatial sensor data. By relying on the DAC paradigm to infer more refined behavioral and emotional information, we can use this higher-level information as a composition parameter as shown in [6].

In the sound synthesis domain, we used basic mappings from perceptual properties to FM synthesis parameters; this could be extended to other synthesis algorithm and parameters. For this purpose, a complete and detailed study of the effect of sound timbre properties on the emotional state of a person is necessary. We also plan to conduct experiments involving self-reported as well as physiological measurements of users' emotional response to different set of parameters of our system.

We limited the scope of this paper to the sonification of the XIM, and didn't address the problem of sonification at the client side for remote users; we will have to tackle this issue in the near future.

Acknowledgements

The authors wish to thank all the members of the SPECS team for their valuable help and expertise, especially Ulysses Bernardet and Sergi Bermudez for their support in developing and designing the different parts of the system. This project is supported by PRESENCCIA (FP6 IST-FET 27731).

References

- [1] J. Manzolli, and V. PFMJ, Roboser: a Real-World Composition System. *Computer Music Journal* 29 (2005) 55-74.
- [2] K. Eng, D. Klein, A. Bähler, U. Bernardet, M. Blanchard, M. Costa, T. Delbrück, R.J. Douglas, K. Hepp, J. Manzolli, M. Mintz, F. Roth, U. Rutishauser, K. Wassermann, A.M. Whatley, A. Wittmann, R. Wyss, and P.F.M.J. Verschure, Design for a brain revisited: the neuromorphic design and functionality of the interactive space 'Ada'. *Rev Neurosci* 14 (2003) 145--180.
- [3] T. Hermann and A. Hunt, An Introduction to Interactive Sonification. *IEEE MultiMedia* 5 (2005) 20--24
- [4] A. Gabrielsson, and P. Juslin, Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of Music* 24 (1996) 68-91.
- [5] P.N. Juslin, and P. Laukka, Communication of emotions in vocal expression and music performance: different channels, same code? *Psychol Bull.* 129 (*Psychol Bull.*) 770-814.
- [6] K. Wasserman, J. Manzolli, K. Eng, and P.F.M.J. Verschure, Live soundscape composition based on synthetic emotions: Using music to communicate between an interactive exhibition and its visitors. *IEEE MultiMedia* 10 (2003) 82-90.
- [7] R.H. Gilkey, and J.M. Weisenberger, The Sense of Presence for the Suddenly-Deafened Adult: Implications for Virtual Environments. *Presence* 4 (1995) 357-363.
- [8] U. Bernardet, M. Blanchard, and P. Verschure, Iqr: a distributed system for real-time real-world neuronal simulation. *Neurocomputing* 44-46 (2002) 1043--1048.
- [9] K. Eng, Siekierka, E., Cameirao, M.S., Zimmerli, L., Pyk, P., Duff, A., Erol, F., Schuster, C., Basseti, C., Kiper, D., and Verschure, P.F.M.J., Cognitive Virtual-Reality Based Stroke Rehabilitation., *World Congress on Medical Physics and Biomedical Engineering* 2006, Seoul, Korea, 2006.
- [10] K. Eng, A. Baebler, U. Bernadet, M. Blanchard, A. Briska, M. Costa, T. Delbruck, R. Douglas, K. Hepp, D. Klein, J. Manzolli, M. Mintz, T. Netter, F. Roth, K. Wassermann, A. Whatley, A. Wittmann, and P.F.M.J. Verschure, Ada: Constructing a Synthetic Organism. in: R. Siegwart, (Ed.), *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.
- [11] M.M. Wanderley, N. Schnell, and J.B. Rovin, ESCHER - Modeling and Performing Composed Instruments in real-time, *IEEE Int. Conference on Systems Man and Cybernetics*, San Diego - CA, USA.
- [12] R. Rowe, *Interactive music systems: machine listening and composing*, MIT Press, Cambridge, MA, USA, 1992.
- [13] The influence of musical structure on emotional expression. In P. N. Juslin, & J. A. Sloboda (Eds.), *Music and emotion: Theory and Research* (pp. 223-248).
- [14] M. Puckette, *Pure Data*, International Computer Music Conference, International Computer Music Association, San Francisco, 1996, pp. 269-272.
- [15] Chowning, John; Bristow, David (1986). *FM Theory & Applications - By Musicians For Musicians*. Tokyo: Yamaha.
- [16] Freed and Wright, "Open SoundControl: A New Protocol for Communicating with Sound Synthesizers", *ICMC* 1997.
- [17] V. Pulkki "Virtual sound source positioning using vector base amplitude panning," *Journal of the Audio Engineering Society*, vol. 45, num. (1997) 6, pp. 456-66.
- [18] R. Wyss, P. König, and P.F.M.J. Verschure, A model of the ventral visual system based on temporal stability and local memory. *Public Library of Science Biology* 4 (2006) 836-43
- [19] W.R. Sherman, A.B. Craig. *Understanding Virtual Reality: Interface, Application and Design*. Morgan Kaufmann. 2002.
- [20] J.A. Lozano, M. Alcañiz, C. Monserrat, M.C. Juan, J.A. Gil, H. Varvaró. Virtual food in virtual environments for the treatment of eating disorders. In *Proceedings MMVR 2002-Medicine Meets Virtual Reality 02/10*, 268-273. January 2002.
- [21] L.B. Meyer, *Emotion and Meaning in Music*, University Of Chicago Press, 1961.
- [22] R.H. Gilkey, and J.M. Weisenberger, The Sense of Presence for the Suddenly-Deafened Adult: Implications for Virtual Environments. *Presence* 4 (1995) 357-363.
- [23] Stefania Serafin, Giovanni Serafin: Sound Design to Enhance Presence in Photorealistic Virtual Reality. *ICAD 2004*
- [24] Freeman, J., & Lessiter, J (2001). *Hear There & Everywhere: The Effects of Multi-channel Audio on Presence*. Presented at *ICAD 2001 - The Seventh International Conference on Audio Display*, Helsinki University of Technology, July 29th - August 1st 2001.