SITUATED INTERACTIVE MUSIC SYSTEM: CONNECTING MIND AND BODY THROUGH MUSICAL INTERACTION

Sylvain Le Groux

Pompeu Fabra University
SPECS, IUA

Paul F.M.J. Verschure

Pompeu Fabra University, ICREA
SPECS, IUA

ABSTRACT

It is now widely acknowledged that music can deeply affect humans' emotional, cerebral and physiological states (20). Yet, the relationship between music features, emotional and physiological responses is still not clear. In this context, we introduce SiMS, a Situated Interactive Music System designed for the artistic and scientific exploration of the relationships between brain activity, physiology, emotion and musical features. SiMS provides the user with tools for the acquisition and processing of electroencephalogram (EEG) and heart-rate (HR) signal. We propose an original perceptually-grounded synthesis model, and the possibility to generate interesting musical structures in real-time as well as a mapping scheme based on psychophysiological evidence for the generation of affective music based on physiology input.

1. INTRODUCTION

Recent advances in human computer interaction have provided researchers and musicians with easy access to physiology sensing technologies. Although the idea is not new (7, 16), the past few years have witnessed a growing interest from the computer music community in using physiological data to generate or transform sound and music (1). In the literature, we distinguish three main trends: the use of physiology to modulate pre-recorded samples (2), to directly map physiological data to synthesis parameters (1), or to control higher level musical structures with parameters extracted from the physiology (6). In most of the cases, the relation between the emotional information, the physiology and the composer intention is not clear. Here we propose a software framework, based on solid psychophysiological and psychoacoustic principles, which facilitates the different stages of data analysis and synthesis and paves the way for the study of the interaction between brain activity, physiology and music perceptual features.

2. SYSTEM OVERVIEW

SiMS is grounded on previous works dealing with large-scale interactive multimedia systems and on the psychophysiology of emotional sonic expression (21, 5). The overall 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 and affective musical structures can be generated (11).

SiMS architecture is modelled after the three component information processing model of human behaviour, which has physiological input (sense organs), processing and decision-making (central nervous system), and motor output (muscle systems and glands). This architecture is also referred to as sensing, processing and response model (17). We tested SiMS within different sensing environments such as physiology sensors (heart-rate, electroencephalogram), or virtual and mixed-reality sensors (camera, gazers, lasers, pressure sensitive floors, ...) and focus here on physiology.

SiMS software is implemented as a set of MaxMSP abstractions and C externals (23). We constantly refined its design and functionalities to adapt to different contexts of use. Each MaxMSP module has a GUI as well as an OSC namespace that allows networked control over its parameters in real-time.

3. SENSING: PHYSIOLOGY

Our main objective with SiMS is to promote and investigate the interaction between perceptually meaningful acoustic cues, bodily and brain responses and musical emotions. With the advent of new human-computer interaction (HCI) technologies, it is now possible to derive emotion-related information from physiological data and use it as an input to our interactive music systems. The first step is to get clean and meaningful signals.

For Electroencephalogram (EEG) and Heart Rate (HR) acquisition, we used the ENOBIO system from Starlab (18); a wearable, wireless, 4-channel, all-digital electrophysiology recording system that has been optimised for dry electrodes. The acquisition and analysis software was designed as a set

of MaxMSP abstractions and C-based externals for acquiring and decoding the stream of bytes from the ENOBIO server, calibrating and filtering the EEG spectrum into its different wave bands, and computing the alpha band. Similarly, for the HR signal, we devised a set of MaxMSP tools to calibrate, square, bandpass filter, threshold the signal and finally detect heart rate, defined as the interval between successive heart beats. The next step involves the extraction and use of emotion-related relevant information from the physiology.

4. PROCESSING: MAPPINGS

The mapping process incorporates two distinct scenari. The first scenario is one where the output is the translation of some emotionally meaningful properties extracted from the brain activity and HR signal into the musical domain. The second scenario is one where the generation of music's goal is to try to influence the physiological response. In the first case SiMS "simply" follows the "mood" of the listener (illustrative feedback) whereas in the second case SiMS tries to induce a specific affective state in the listener (emotion induction).

In both cases the mappings between physiology, emotions and musical parameters are based on previous works on the psychophysiology of emotion. (5) To represent emotions, we use the well established bi-polar dimensional theory of emotions where musical emotion is a point on the axes of hedonic valence (or pleasantness) and intensity of activation (or arousal) (19).

4.1. The Physiology of Emotions

Heart activity measurement has a long tradition in emotion and media research, where it has been shown to be a valid real-time measure for attention and arousal (8). Attention evokes short-term (phasic component) deceleration of heart rate, while arousing stimuli accelerates heart rate in longer term (tonic component). Heart rate change has been also shown to reflect stimuli valence. While the heart rate drops initially after presentation of the stimuli due to attention shift, the negative stimuli result in a larger decrease of a longer duration (3).

Similarly, the study of brainwaves has a rich history, and different brainwave activities have been shown to correlated with different states. We choose to focus on the alpha waves which typically correlates with states of relaxation (13).

4.2. Correlations Between Musical Features, Affect and Physiology

Several studies have already investigated the impact of music structural and sonic features on the induction of emotional state (9, 5). Those studies can be summarized as follows: an increase in tempo, volume or register positively correlates with an increase in valence and arousal. Shortening of articulations correlates with increase of valence. Increase of brightness correlates with increase of arousal, increase of noisiness correlates with increase of arousal, while harmonicity negatively correlates with valence.

From those correspondances between physiology, affective state and musical parameters, it becomes possible to envision a physiologically-based affective music generator.

For instance, in order to induce a calm and relaxed state in the participant (high valence, low arousal), SiMS will have to decrease the global volume, increase harmonicity, lower the brightness, lengthen the articulations, and decrease the tempo of the interactive music piece, until the alpha power from the EEG, characteristic of a relaxed state, and the HR rate reach significant values.

One of the main difficulty of this task is to find the appropriate thresholds to fine-tune SiMS behavior. At this stage, we just explored the parameter space heuristically and can't give any quantitative statistical data analysis yet. This will be the subject of future studies. We can envision machine learning mapping strategies such as reinforcement learning, where SiMS tunes its musical parameters to optimize the reward expressed as a high alpha wave power (relaxed) and low HR, for instance.

We have presented a set of well-studied plausible relationships between physiology and musical features, and will now introduce the music engine, in charge of generating an interactive musical piece from physiological i nput.

5. RESPONSE: THE AFFECTIVE MUSIC ENGINE

The music generation system is composed of a hierarchy of perceptually and musically meaningful agents interacting and communicating via the OSC protocol (22). Each agent influences the induction of emotion in the listener. We introduce those agents starting from the top of the hierarchy (macro structure) and ending at the bottom (micro structure).

5.1. Musical Performance

Inspired by previous works on musical performance modeling (4), we implemented a set of simplified modules that controls the expressivity of music generation by modulating some parameters lower in the hierarchy.

Phrasing: Crescendo/decrescendo envelopes are applied to dynamics and tempo parameters to create a more animated musical structure.

Articulation: Since articulation is crucial in the overall character of a piece (4), we can vary the amount of legato or staccato with wich the note will be played.

Performance Noise: Human musicians are not completely accurate when producing and perceiving small timing variations. To mimic this factor, we introduce some performance noise using white noise added to each tone onset time and tone sound level (4).

5.2. Musical Material

While conceptually fairly simple, the music material generator has been designed to keep the balance between predictability and surprise. The basic idea is to have the external world act as a input that modulates prepared musical cells. We thus devised a set of modules in MaxMSP allowing the scoring of basic midi musical material such as Rhythm, Pitch, Velocity and Duration via a text editor or step-sequencer interface.

The generation of musical elements from base material relies on the extended dodecaphonism paradigm. Once a set of predefined series is chosen for every parameter, the "serialist" modules can use different principles to selects specific elements from the material and generate music (e.g choose a random element from the series, choose all the elements successively in the order of the score like a sequencer, choose all the elements in reverse order, or play all the elements once without repetition...).

Additionally, we built a set of "stochastic" modules that produces musical elements in a simple probabilistic manner: each musical dimension is generated from a random-walk process characterized by its range of action and brownian noise factor.

5.3. Perceptual Sound Generation

We wanted to have a synthesizer allowing the modulation of subtle timbral features that are perceptually relevant in the context of music and emotion research. We implemented a polyphonic synthesizer in Max MSP which relies on the tristimulus model of timbre (14, 15). Here, the sound is basically represented as a sum of weighted sinusoids at frequencies multiple of the fundamental frequency (12) plus a filtered noise component. The tristimulus analysis of timbre proposes to quantify timbre in terms of three coordinates (x, y, z) associated with band-loudness values. Inspired from the tristimulus theory of colour perception, it associates high values of x to dominant high-frequencies $f_{5...N}(n)$, high values of y to dominant mid-frequency components $f_{2...4}(n)$ and high values of z to dominant fundamental frequency $f_0(n)$.

The noisy part of the sound was generated following the subtractive synthesis paradigm. We filtered a random noise generator with a bank of three passband filters centered at f_0, $3f_0$ and $9f_0$ respectively as suggested in (15) so that the noisy portion of the sound follows a tristimulus spectral distribution.

This design allows control over perceptually relevant sound parameters which variation correlates with different emotional states (5). Noisiness is defined as the relative amplitude of the filtered noise generator. Inharmonicity relates to the factor by which successive partials deviate from the harmonic spectrum. Finally, we define the even partial attenuation factor as the relative amplitude level of even partials in the spectrum of the signal

6. IMPLICATIONS / APPLICATIONS

The SiMS framework is the first step towards the development of applications and experiments around interaction, emotion and physiology.

One of our central interest is to study the emotional response to specific sound features to design efficient tools for interactive music therapy. Some previous studies show that a significant correlation exists between electrodermal activity, heart rate, heart rate variability and the subjective evaluation of well-defined musical parameters (5, 9). We plan to examine in more detail the nature of this correlation and will extend the study to other musical parameters in order to build an interesting interactive system.

We informally started to explore the conscious learning of relaxation through sonic feedback or auditory neurofeedback where the interactive music feedback system helps the listener keeping track of and modulate her agitation state as measured by EEG (10). We believe an interactive music system which design is grounded on strong psychophysiology principles will prove useful for musicotherapy applications.

On a more artistic note, we recently premiered a piece for the "Multimodal Brain Orchestra [1]" using SiMS for real-time modulation of the musical material. It involved the use of a multi-person orchestra to control a virtual string quartet through brain states, the integration of multi-person physiological states in an evolving synthetic audio-visual composition and the real-time control of the affective expression of a performance by the physiological state of a human or "emotional conductor".

7. CONCLUSION

With the goal to explore the interaction between emotion, physiology, audio features and musical interaction, presented a unified framework and a set of software modules for acquiring and processing physiology data, a sound synthesis algorithm based on perceptual principles, as well as a musical structure generator. Additionally, we proposed a possible mapping model based on the review of correlational studies from physiology domain to induced emotion and

---

[1] http://specs.upf.edu/?q=installation/2025

musical feature. However, the whole system needs to be robustly evaluated. In the near future we plan a series of controlled experiments and statistical analysis to assess the validity of the mapping hypothesis, and to fine-tune the different parameters of the system.

8. REFERENCES

[1] B. Arslan, A. Brouse, J. Castet, R. Lehembre, C. Simon, J. J. Filatriau, and Q. Noirhomme, "A real time music synthesis environment driven with biological signals," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 2, 2006, pp. II–II.

[2] G. Bieber and H. Diener, "Stepman - a new kind of music interaction," 2005.

[3] M. M. Bradley and P. J. Lang, "Affective reactions to acoustic stimuli." *Psychophysiology*, vol. 37, no. 2, pp. 204–215, March 2000.

[4] A. Friberg, R. Bresin, and J. Sundberg, "Overview of the kth rule system for musical performance," *Advances in Cognitive Psychology, Special Issue on Music Performance*, vol. 2, no. 2-3, pp. 145–161, 2006.

[5] A. Gabrielsson and E. Lindström, *Music and Emotion - Theory and Research*, ser. Series in Affective Science. New York: Oxford University Press, 2001, ch. The Influence of Musical Structure on Emotional Expression.

[6] R. Hamilton, "Bioinformatic feedback: performer bio-data as a driver for real-time composition," in *NIME '06: Proceedings of the 6th international conference on New interfaces for musical expression*, 2006.

[7] B. R. Knapp and H. S. Lusted, "A bioelectric controller for computer music applications," *Computer Music Journal*, vol. 14, no. 1, pp. 42–47, 1990.

[8] A. Lang, "Involuntary attention and physiological arousal evoked by structural features and emotional content in TV commercials," in *Communication Research*, vol. 17, no. 3, 1990, pp. 275–299.

[9] S. Le Groux, A. Valjamae, J. Manzolli, and P. F. Verschure, "Implicit physiological interaction for the generation of affective music," in *Proceedings of the International Computer Music Conference*, 2008.

[10] S. Le Groux and P. F. Verschure, "Neuromuse: Training your brain through musical interaction," in *Proceedings of the International Conference on Auditory Display*, Copenhagen, Denmark, May 18-22 2009.

[11] J. Manzolli and P. F. M. J. Verschure, "Roboser: A real-world composition system," *Comput.Music J.*, vol. 29, no. 3, pp. 55–74, 2005.

[12] R. McAulay and T. Quatieri, "Speech analysis/synthesis based on a sinusoidal representation," *IEEE Trans. on Acoust., Speech and Signal Proc.*, vol. 34, pp. 744–754, 1986.

[13] P. L. Nunez, *Electric Fields of the Brain: The Neurophysics of Eeg*. Oxford Univ Pr (Txt), February 2005.

[14] H. Pollard and E. Jansson, "A tristimulus method for the specification of musical timbre," in *Acustica*, vol. 51, 1982.

[15] A. Riley and D. Howard, "Real-time tristimulus timbre synthesizer," University of York, Tech. Rep., 2004.

[16] D. Rosenboom, "Biofeedback and the arts: Results of early experiments," in *Computer Music Journal*, vol. 13, no. 4, 1989, pp. 86–88.

[17] R. Rowe, *Interactive music systems: machine listening and composing*. Cambridge, MA, USA: MIT Press, 1992.

[18] G. Ruffini, S. Dunne, L. Fuentemilla, C. Grau, E. Farres, J. Marco-Pallares, P. C. P. Watts, and S. R. P. Silva, "First human trials of a dry electrophysiology sensor using a carbon nanotube array interface," 2007.

[19] J. A. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology*, vol. 39, pp. 345–356, 1980.

[20] O. W. Sacks, *Musicophilia : tales of music and the brain / Oliver Sacks*. Alfred A. Knopf, New York :, 2008.

[21] K. C. Wassermann, K. Eng, and P. F. Verschure, "Live soundscape composition based on synthetic emotions," *IEEE MultiMedia*, vol. 10, no. 4, pp. 82–90, 2003.

[22] M. Wright, "Open sound control: an enabling technology for musical networking," *Org. Sound*, vol. 10, no. 3, pp. 193–200, 2005.

[23] D. Zicarelli, "How I learned to love a program that does nothing," *Computer Music Journal*, no. 26, pp. 44–51, 2002.