

Disembodied and Collaborative Musical Interaction in the Multimodal Brain Orchestra

Sylvain Le Groux
Laboratory for Synthetic
Perceptive Emotive and
Cognitive Systems (SPECS)
Universitat Pompeu Fabra
Barcelona, Spain
sylvain.legroux@upf.edu

Jonatas Manzoli
Interdisciplinary Nucleus of
Sound Studies (NICS)
UNICAMP
Campinas, Brazil
jonatas@nics.unicamp.br

Paul F.M.J. Verschure^{*}
Institutio Catalana de
Recerca i Estudis Avancats
(ICREA) and SPECS
Barcelona, Spain
paul.verschure@upf.edu

ABSTRACT

Most new digital musical interfaces have evolved upon the intuitive idea that there is a causality between sonic output and physical actions. Nevertheless, the advent of brain-computer interfaces (BCI) now allows us to directly access subjective mental states and express these in the physical world without bodily actions. In the context of an interactive and collaborative live performance, we propose to exploit novel brain-computer technologies to achieve unmediated brain control over music generation and expression. We introduce a general framework for the generation, synchronization and modulation of musical material from brain signal and describe its use in the realization of Xmotion, a multimodal performance for a “brain quartet”.

Keywords

Brain-computer Interface, Biosignals, Interactive Music System, Collaborative Musical Performance

1. INTRODUCTION

The Multimodal Brain Orchestra (MBO) demonstrates interactive, affect-based and self-generated musical content based on novel BCI technology. It is an exploration of the musical creative potential of a collection of unmediated brains directly interfaced to the world, bypassing their bodies.

One of the very first piece to use brainwave for generating music was “Music for solo performer” composed by Alvin Lucier in 1965 [28]. He used brainwaves as a generative source for the whole piece. In this piece, the electroencephalogram (EEG) signal from the performer was amplified and relayed to a set of loudspeakers coupled with percussion instruments. Some years later, the composer David Rosenboom started to use biofeedback devices (especially EEG) to allow performers to create sounds and music using their own brainwaves [25]. More recent research has attempted to create complex musical interaction between par-

ticular brainwaves and corresponding sound events where the listener EEG control a music generator imitating the style of a previously listened sample [20]. Data sonification in general and EEG sonification in particular has been the subject of various studies [13] showing the ability of the human auditory system to deal with and understand highly complex sonic representation of data.

Although there has been an renewed interest in brain-based music over the recent years, most projects are only based on direct mappings from the EEG spectral content to sound generators. They do not rely on explicit volitional control. The Multimodal Brain Orchestra (MBO) takes a different approach by integrating advanced BCI (Brain Computer Interface) technology that allows the performer complete volitional control over the command signals that are generated. MBO preserves the level of control of the instrumentalist by relying on classification of specific stimulus triggered events in the EEG. Another unique aspect of the MBO is that it allows for a multimodal and collaborative performance involving four brain orchestra members, a musical conductor and real-time visualization.

2. SYSTEM ARCHITECTURE

2.1 Overview: A client-server Architecture for Multimodal Interaction

The interactive music system of the Multimodal Brain Orchestra is based on a client-server modular architecture, where inter-module communication follows the Open Sound Control (OSC) protocol [30]. The MBO consists of three main components (Figure 1) namely the orchestra members, the multimodal interactive system, and the conductor. 1) The four members of the “brain quartet” are wired up to two different types of brain-computer interfaces: the P300 and the SSVEP (Steady-State Visual Evoked Potentials)(cf section 2.2). 2) The computer-based interactive multimedia system processes inputs from the conductor and the BCIs to generate music and visualization in real-time. This is the core of the system where most of the interaction design choices are made. The interactive multimedia component can itself be decomposed into three subsystems: the EEG signal processing module, the SiMS (Situating Interactive Music System) music server [17] and the real-time visualizer. Finally, the conductor uses a Wii-Baton (cf section 2.5) to modulate the tempo of the interactive music generation, trigger different sections of the piece, and cue the orchestra members (Figure 1).

2.2 Brain Computer Interface

The musicians of the orchestra are all connected to brain-computer interfaces that allow them to control sound events

^{*}c.f. Section 6 “Additional Authors” for the list of all additional authors

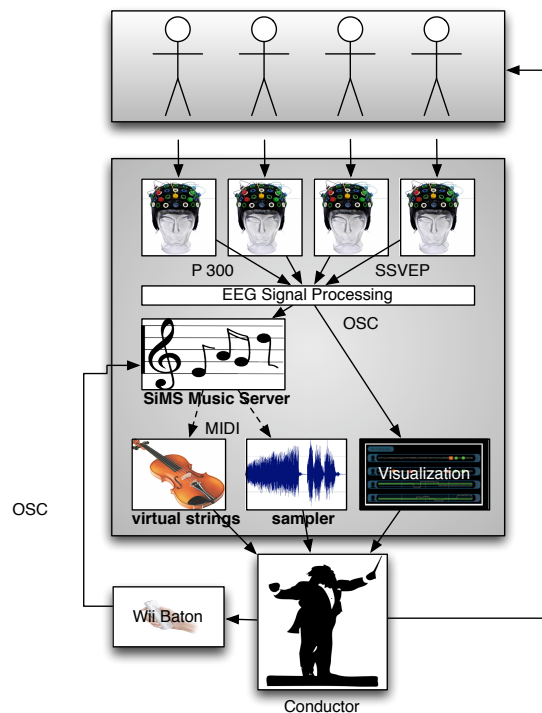


Figure 1: The Multimodal Brain Orchestra is a modular interactive system based on a client-server architecture using the OSC communication protocol. See text for further information.

and music expressiveness during the performance. These BCIs provide a new communication channel between a brain and a computer. These interfaces are based on the principle that mental activity can lead to observable changes of electrophysiological signals in the brain. These signals can be measured, processed, and later transformed into useful high level messages or commands [29, 9, 12].

The MBO is based on two different non-invasive BCI concepts which control the generation and modulation of music and soundscapes, namely the P300 and SSVEP. We worked with G.tec medical engineering GmbH products, providing BCI hardware devices (g.USBamp) and corresponding real-time processing software for MATLAB/Simulink¹. The control commands generated by the classification of the EEG using the, so called, P300 and SSVEP protocols were sent to the music server and visualization module via a simulink S-function implementing using the OSC protocol for Matlab².

2.2.1 The P300 Speller

The P300 is an event related potential (ERP) that can be measured with eight electrodes at a latency of approximately 300ms after an infrequent stimuli occurs. We used the P300 speller paradigm introduced by [8]. In our case, two orchestra members were using a 6 by 6 symbol matrix containing alpha-numeric characters (Figure 3) in which a row, column or single cell was randomly flashed on. The orchestra member has to focus on the cell containing the symbol to be communicated and to mentally count every time the cell flashes (this is to distinguish between common and rare stimuli). This elicits an attention dependent pos-

¹<http://www.mathworks.com>

²<http://andy.schmeder.net/software/>

itive deflection of the EEG about 200 msec after stimulus onset, the P300, that can be associated to the specific symbol by the system (Figure 2) [12]. We used this interface to trigger discrete sound events in real-time. Because it is difficult to control the exact time of occurrence of P300 signals, our music server SiMS (cf section 2.4) took care of beat-synchronizing the different P300 events with the rest of the composition.

A P300 interface is normally trained with 5-40 characters which corresponds to a training time of about 5-45 minutes. A group study with 100 people showed that after a training with 5 characters only, 72 % of the users could spell a 5 character word without any mistake [12]. This motivated the decision to limit the number of symbols used during the performance (Section 3.4).

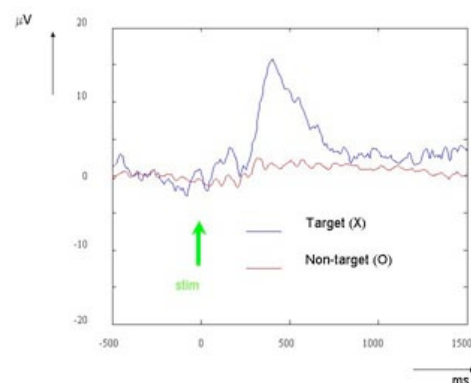


Figure 2: P300: a rare event triggers an ERP 300 ms after the onset of the event indicated by the green arrow.

A	B	C	D	E	F
G	H	I	J	K	L
M	N	O	P	Q	R
S	T	U	V	W	X
Y	Z	1	2	3	4
5	6	7	8	9	0

Figure 3: A 6 by 6 symbol matrix is presented to the P300 orchestra member who can potentially trigger 36 specific discrete sound events.

2.2.2 SSVEP

Another type of interface was provided by steady-state visually evoked potentials (SSVEP) triggered by flickering light. This method relies on the fact that when the retina is excited by a flickering light with a frequency ≈ 3.5 Hz, the brain generates activity at the same frequency [1, 2]. The interface is composed of four different light sources flickering at different frequencies and provides additional “step controllers” (Figure 4).

The SSVEP BCI interface is trained for about 5 minutes during which the user has to look several times at every flickering LED. Then, a user specific classifier is calculated that allows on-line control. In contrast to the P300, the SSVEP BCI gives a continuous control signal that switches

from one state to another within about 2-3 seconds. The SSVEP BCI solves also the zero-class problem. If the user is not looking at one of the LEDs then no decision is made[24].

SSVEP was used to control changes in articulation and dynamics of the music generated by SiMS.



Figure 4: Two members of the orchestra connected to their SSVEP-based BCI interfaces [24]

2.3 Visual Feedback

We designed a module that gave real-time visualization of the BCI output. More precisely, the different possible control messages detected by g.tec analysis software from the brain signal were sent to the visualizer via OSC and illustrated with simple color coded icons. From the two members of the orchestra using the P300 BCI interface we can receive 36 distinct control messages. Each of these 36 symbols was represented using a combination of six geometrical shapes and six different colors. The two members of the orchestra using the SSVEP BCI interfaces were able to trigger four possible events corresponding to the four different states (or in other words, four brain activity frequencies), but continuously. Each line in the display corresponded to a member of the orchestra: the first two using P300 and the last two SSVEP. When a P300 member triggered an event, the associated geometrical shape appeared in the left side and moved from left to right according to time. For the SSVEP events, the current state was shown in green and the past changes could be seen as they moved from left to right.

The real-time visualization played the role of a real time score. It provided feedback to the audience and was fundamental for the conductor to know when the requested events were actually triggered. The conductor could indicate to the orchestra member when to trigger an event (P300 or SSVEP) but the confirmation of its triggering was indicated by the real time visual score as well as by its musical consequences.

2.4 The Situated Interactive Music System (SiMS)

Once the signal is extracted from brain activity and transformed into high-level commands by g.tec software suite, a specific OSC message is sent to the SiMS music server [17] and to the visualization module. SSVEP and P300 interfaces provide us with a set of discrete commands we want to transform into musical parameters driving the SiMS server.

SiMS is an interactive music system inspired by Roboser, a midi-based composition system that has previously been applied to the sonification of robots and people's trajectories [7, 18]. SiMS is entirely based on a networked architecture. It implements various algorithmic composition tools (e.g: generation of tonal, brownian and serial series of pitches and rhythms) and a set of synthesis techniques val-

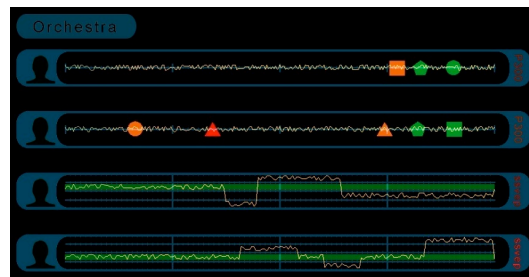


Figure 5: The real-time visualizer allows for real-time visualization of P300 system output (the two upper rows show combinations of shapes and colors) and SSVEP system output (the two lower rows)

idated by psychoacoustical tests [17, 15]. Inspired by previous works on musical performance modeling [10], SiMS allows to modulate the expressiveness of music generation by varying parameters such as phrasing, articulation and performance noise[17].

SiMS is implemented as a set of Max/MSP abstractions and C++ externals [31]. We have tested SiMS within different sensing environments such as biosignals (heart-rate, electroencephalogram) [16, 17, 15], or virtual and mixed-reality sensors (camera, gazers, lasers, pressure sensitive floors, ...) [3]. After constantly refining its design and functionalities to adapt to those different contexts of use, we opted for an architecture consisting of a hierarchy of perceptually and musically meaningful agents interacting and communicating via the OSC protocol [30] (Figure 6). For this project we focused on interfacing BCI to SiMS.

SiMS follows a biomimetic architecture that is multi-level and loosely distinguishes sensing (e.g electrodes attached to the scalp using a cap) from processing (musical mappings and processes) and actions (changes of musical parameters). It has to be emphasized though that we do not believe that these stages are discrete modules. Rather, they will share bi-directional interactions both internal to the architecture as through the environment itself. In this respect it is a further advance from the traditional separation of sensing, processing and response paradigm[26] which was at the core of traditional AI models.

2.5 Wii-mote Conductor Baton

We provided the orchestra conductor with additional control over the musical output using the Wii-mote (Nintendo) as a baton. Different sections of the quartet could be triggered by pressing a specific button, and the gestures of the conductor were recorded and analyzed. A processing module in SiMS (Figure 7) filtered the accelerometers and the time-varying accelerations were interpreted in terms of beat pulse and mapped to small tempo modulation in the SiMS player.

3. XMOTION: A BRAIN-BASED MUSICAL PERFORMANCE

3.1 Emotion, Cognition and Musical Composition

One of the original motivations of the MBO project was to explore the potential creativity of BCIs as they allow to access subjective mental states and express these in the physical world without bodily actions. The name XMotion designate those states that can be generated and experienced by the unmediated brain when it is both immersed

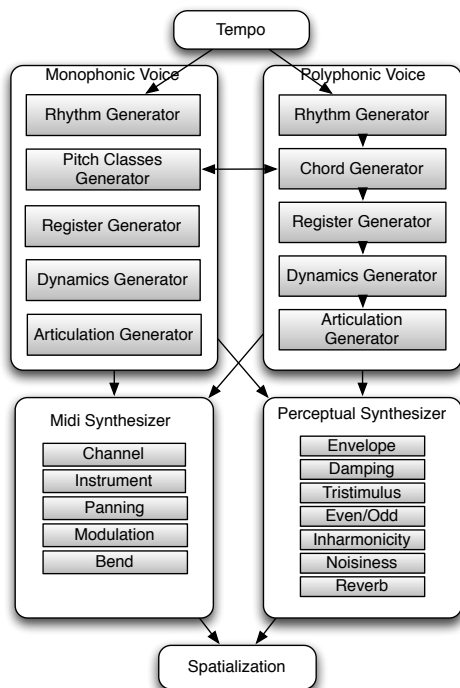


Figure 6: SiMS music server is built as a hierarchy of musical agents and can be integrated into various sensate environments. See text for further information.

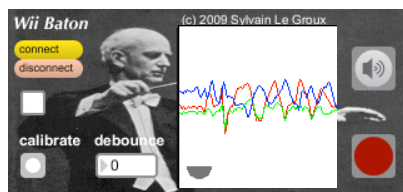


Figure 7: The wii-baton module analyzes 3D acceleration data from the wii-mote so the conductor can use it to modulate the tempo and to trigger specific sections of the piece.

and in charge of the multimodal experience in which it finds itself.

The XMotion performance is based on the assumption that mental states can be organized along the three-dimensional space of valence, arousal and representational content [21]. Usually emotion is described as decoupled from cognition in a low dimensional space such as the circumplex model of Russell [27]. This is a very effective description of emotional states in terms of their valence and arousal. However, these emotional dimensions are not independent of other dimensions such as the representational capacity of consciousness which allows us to evaluate and alter the emotional dimensions [14]. The musical piece composed for XMotion proposes to combine both models into a framework where the emotional dimensions of arousal and valence are expressed by the music, while the conductor evaluates its representational dimension.

Basing our ideas on previous emotion research studies [11, 15], we decided to control the modulation of music from Russell's bi-dimensional model of emotions [27]. The higher

the values of the dynamics, the higher the expressed arousal and similarly, the longer the articulation, the higher the valence. In addition, a database of sound samples was created where each sample was classified according to the Arousal and Valence taxonomy (Table 1).

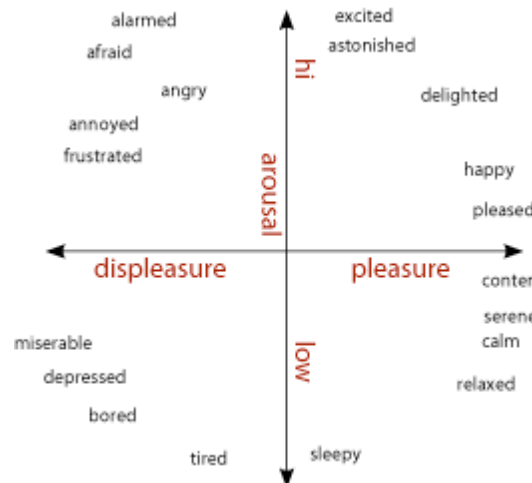


Figure 8: Russel's circumplex model of affect represents emotions on a 2D map of Valence and Arousal [27].

3.2 Musical Material

The musical composition by Jonas Manzolli consisted of three layers, namely the virtual string quartet, a fixed electroacoustic tape and live triggering of sound events. The four voices of a traditional string quartet setup up were precomposed offline and stored as MIDI events to be modulated (articulation and accentuation) by the MBO members connected to the SSVEP interfaces. The sound rendering was done using state of the art orchestral string sampling technology (using the London Symphony Orchestra library with Kontakt sampler³). The second layer consisted of a fixed audio tape soundtrack synchronized with the string quartet material with Live⁴ audio time stretching algorithms. Additionally, we used discrete sound events triggered by the P300 brain orchestra members. The orchestra members were coordinated by the musical conductor standing in front of them.

3.3 String Quartet

The basic composition strategy was to associate different melodic and rhythmic patterns of musical textures to variations in dynamics and articulation producing textural changes in the composition. The inspiration for this music architecture was the so called net-structure technique created by Ligeti using pattern-meccanico material [4]. The second aspect of the composition was to produce transposition of beats producing an effect of phase-shifting[5]. These two aspects produced a two-dimension gradual transformation in the string quartet textures. In one direction the melodic profile was gradually transformed by the articulation changes. On the other, the shift of accentuation and gradual tempo changes produced phase-shifts. In the first movement a chromatic pattern is repeated and legato increased the superposition of notes. The second and fourth

³<http://www.native-instruments.com/>

⁴<http://www.ableton.com>

movements worked with a constant chord modulation chain and the third with a canonical structure.

One member of the orchestra used the SSVEP to modulate the articulation of the string quartet (four levels from legato to staccato corresponding to the four light sources frequencies) while the other member modulated the accentuation (from piano to forte) of the quartet.

3.4 Soundscape

The soundscape was made of a fixed tape piece composition and discrete sound events triggered according to affective content. The sound events are driven by the conductor's cues and relate to the visual realm. The tape was created using four primitive sound qualities. The idea was to associate mental states with changes of sound material. "P300 performers" produced discrete events related to four letters: A (sharp strong), B (short percussive), C (water flow) and D (harmonic spectrum). On the conductor's cue, the performers concentrated on a specific column and row and triggered the desired sound. Two members of the orchestra were using P300 hundred and concentrated on 4 symbols each. Each symbol triggered a sound sample from the "emotional database" corresponding to the affective taxonomy associated with the symbol (for each symbol or sound quality we had a set of 4 possible sound samples).

Sound Quality	State	Arousal	Valence
Sharp Strong	A	High	Negative
Short Percussive	B	High	Negative
Water Flow	C	Low	Positive
Harmonic Spectrum	D	Low	Positive

Table 1: An affective taxonomy was used to classify the sound database

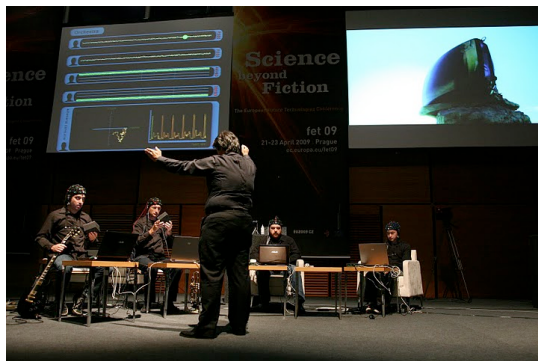


Figure 9: The MBO performance setup at FET European conference in Prague in July 2009.

4. CONCLUSIONS

We presented a disembodied interactive system designed for the generation and modulation of musical material from brain signal, and described XMotion, an interactive "brain quartet" piece based on novel brain computer interface technologies. The MBO shows how novel BCI technologies can be used in a multimodal collaborative context where the performers have volitional control over their mental state and the music generation process. Considering that the response time delays of the SSVEP and P300 interfaces are well above audio rate, we do not claim that these interfaces provide the level of subtlety and intimate control more traditional instruments can afford. Nevertheless, it is

a promising first step towards the exploration of the creative potential of collaborative brain-based interaction for audio-visual content generation. It is part of a larger effort to include physiological feedback in the interactive generation of music. We can envision many applications of this brain-based systems beyond the area of performance including music therapy (this system fosters musical collaboration and would allow disable people to play music together), neurofeedback [16, 6, 19] and motor rehabilitation (e.g. the use of musical feedback for neurofeedback training might be a good alternative to visual feedback for people with visual impairment)[22, 23]. We are further exploring both these artistic and practical applications of the MBO.

5. ACKNOWLEDGMENT

We would like to thank visual artist Behdad Rezazadeh, the brain orchestra members Encarni Marcos, Andre Luvizotto, Armin Duff, Enrique Martinez, An Hong and g.tec staff for their patience and dedication. The MBO is supported by Fundacio La Marato de TV3 and the European Commission ICT FP7 projects ReNaChip, Synthetic Forager, Rehabilitation Gaming System, and Presencia.

6. ADDITIONAL AUTHORS

Marti Sanchez*, Andre Luvizotto*, Anna Mura*, Aleksander Valjamae*, Christoph Guger†, Robert Prueckl†, Ulysses Bernardet*.

- * SPECS, Universitat Pompeu Fabra, Barcelona, Spain
- † g.tec Guger Technologies OEG, Herbersteinstrasse 60, 8020Graz, Austria

7. REFERENCES

- [1] B. Z. Allison, D. J. McFarland, G. Schalk, S. D. Zheng, M. M. Jackson, and J. R. Wolpaw. Towards an independent brain-computer interface using steady state visual evoked potentials. *Clin Neurophysiol*, 119(2):399–408, Feb 2008.
- [2] B. Z. Allison and J. A. Pineda. Erps evoked by different matrix sizes: implications for a brain computer interface (bci) system. *IEEE Trans Neural Syst Rehabil Eng*, 11(2):110–3, Jun 2003.
- [3] U. Bernardet, S. B. i Badia, A. Duff, M. Inderbitzin, S. L. Groux, J. Manzolli, Z. Mathews, A. Mura, A. Valjamae, and P. F. M. J. Verschure. *The eXperience Induction Machine: A New Paradigm for Mixed Reality Interaction Design and Psychological Experimentation*. Springer, 2009. [17 dec 2009] In Press.
- [4] J. Clendinning. The pattern-meccanico compositions of Gyorgy Ligeti. *Perspectives of New Music*, 31(1):192–234, 1993.
- [5] R. Cohn. Transpositional Combination of Beat-Class Sets in Steve Reich's Phase-Shifting Music. *Perspectives of New Music*, 30(2):146–177, 1992.
- [6] T. Egner and J. Gruzelier. Ecological validity of neurofeedback: modulation of slow wave EEG enhances musical performance. *Neuroreport*, 14(9):1221, 2003.
- [7] K. Eng, A. Babler, U. Bernardet, M. Blanchard, M. Costa, T. Delbruck, R. J. Douglas, K. Hepp, D. Klein, J. Manzolli, M. Mintz, F. Roth, U. Rutishauser, K. Wassermann, A. M. Whatley, A. Wittmann, R. Wyss, and P. F. M. J. Verschure. Ada - intelligent space: an artificial creature for the swissexpo.02. *Robotics and Automation*, 2003.

- Proceedings. ICRA '03. IEEE International Conference on*, 3:4154–4159 vol.3, 2003.
- [8] L. Farwell and E. Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and clinical Neurophysiology*, 70(6):510–523, 1988.
 - [9] E. A. Felton, J. A. Wilson, J. C. Williams, and P. C. Garell. Electroencephalographically controlled brain-computer interfaces using motor and sensory imagery in patients with temporary subdural electrode implants. report of four cases. *J Neurosurg*, 106(3):495–500, Mar 2007.
 - [10] 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*, 2(2-3):145–161, 2006.
 - [11] A. Gabrielsson and E. Lindström. *Music and Emotion - Theory and Research*, chapter The Influence of Musical Structure on Emotional Expression. Series in Affective Science. Oxford University Press, New York, 2001.
 - [12] C. Guger, S. Daban, E. Sellers, C. Holzner, G. Krausz, R. Carabalona, F. Gramatica, and G. Edlinger. How many people are able to control a P300-based brain-computer interface (BCI)? *Neuroscience letters*, 462(1):94–98, 2009.
 - [13] T. Hinterberger and G. Baier. Parametric orchestral sonification of eeg in real time. *IEEE MultiMedia*, 12(2):70–79, 2005.
 - [14] S. Laureys. The neural correlate of (un) awareness: lessons from the vegetative state. *Trends in cognitive sciences*, 9(12):556–559, 2005.
 - [15] S. Le Groux, A. Valjamae, J. Manzolli, and P. F. M. J. Verschure. Implicit physiological interaction for the generation of affective music. In *Proceedings of the International Computer Music Conference*, Belfast, UK, August 2008. Queens University Belfast.
 - [16] S. Le Groux and P. F. M. J. Verschure. Neuromuse: Training your brain through musical interaction. In *Proceedings of the International Conference on Auditory Display*, Copenhagen, Denmark, May 18-22 2009.
 - [17] S. Le Groux and P. F. M. J. Verschure. Situated interactive music system: Connecting mind and body through musical interaction. In *Proceedings of the International Computer Music Conference*, Montreal, Canada, August 2009. Mc Gill University.
 - [18] J. Manzolli and P. F. M. J. Verschure. Roboser: A real-world composition system. *Comput. Music J.*, 29(3):55–74, 2005.
 - [19] G. Mindlin and G. Rozelle. Brain music therapy: home neurofeedback for insomnia, depression, and anxiety. In *International Society for Neuronal Regulation 14-th Annual conference, Atlanta, Georgia*, pages 12–13, 2006.
 - [20] E. R. Miranda, K. Sharman, K. Kilborn, and A. Duncan. On harnessing the electroencephalogram for the musical braincap. *Comput. Music J.*, 27(2):80–102, 2003.
 - [21] A. Mura, J. Manzolli, B. Rezazadeh, S. L. Groux, M. Sanchez, A. Valjame, A. Luvizotto, C. Guger, U. Bernardet, and P. F. Verschure. The multimodal brain orchestra: art through technology. Technical report, SPECS, 2009.
 - [22] F. Nijboer, A. Furdea, I. Gunst, J. Mellinger, D. McFarland, N. Birbaumer, and A. K. ubler. An auditory brain-computer interface (BCI). *Journal of neuroscience methods*, 167(1):43–50, 2008.
 - [23] M. Pham, T. Hinterberger, N. Neumann, A. Kubler, N. Hofmayer, A. Grether, B. Wilhelm, J. Vatine, and N. Birbaumer. An auditory brain-computer interface based on the self-regulation of slow cortical potentials. *Neurorehabilitation and Neural Repair*, 19(3):206, 2005.
 - [24] R. Prueckl and C. Guger. A Brain-Computer Interface Based on Steady State Visual Evoked Potentials for Controlling a Robot. *Bio-Inspired Systems: Computational and Ambient Intelligence*, pages 690–697.
 - [25] D. Rosenboom. Biofeedback and the arts: Results of early experiments. In *Computer Music Journal*, volume 13, pages 86–88, 1989.
 - [26] R. Rowe. *Interactive music systems: machine listening and composing*. MIT Press, Cambridge, MA, USA, 1992.
 - [27] J. A. Russell. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39:345–356, 1980.
 - [28] Wikipedia. Alvin lucier — wikipedia, the free encyclopedia, 2008. [Online; accessed 28-January-2009].
 - [29] J. R. Wolpaw. Brain-computer interfaces as new brain output pathways. *J Physiol*, 579(Pt 3):613–9, Mar 2007.
 - [30] M. Wright. Open sound control: an enabling technology for musical networking. *Org. Sound*, 10(3):193–200, 2005.
 - [31] D. Zicarelli. How I learned to love a program that does nothing. *Computer Music Journal*, (26):44–51, 2002.