Crossroads: Interactive Music Systems Transforming Performance, Production and Listening

Mathieu Barthet
Queen Mary University of London, Centre for Digital Music
Mile End Road, London E14NS, UK
m.barthet@qmul.ac.uk

Mark Sandler
Queen Mary University of London, Centre for Digital Music
Mile End Road, London E14NS, UK
mark.sandler@qmul.ac.uk

Florian Thalmann
Queen Mary University of London, Centre for Digital Music
Mile End Road, London E14NS, UK
f.thalmann@qmul.ac.uk

Geraint Wiggins
Queen Mary University of London, Centre for Digital Music
Mile End Road, London E14NS, UK
geraint.wiggins@qmul.ac.uk

György Fazekas
Queen Mary University of London, Centre for Digital Music
Mile End Road, London E14NS, UK
g.fazekas@qmul.ac.uk

Abstract
We discuss several state-of-the-art systems that propose new paradigms and user workflows for music composition, production, performance, and listening. We focus on a selection of systems that exploit recent advances in semantic and affective computing, music information retrieval (MIR) and semantic web, as well as insights from fields such as mobile computing and information visualisation. These systems offer the potential to provide transformative experiences for users, which is manifested in creativity, engagement, efficiency, discovery and affect.

Author Keywords
audience, performer, production, mixing, mood, adaptive player, semantic audio, semantic web, visualisation

ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous; J.5 [Arts and humanities]

Introduction
A popular interpretation of *Crossroads*, a blues song by Robert Johnson, is that it refers to the place where the artist supposedly sold his soul to the devil in exchange for his musical talents. It often seems as though we have to face similar decisions when working with music and computer technology. In this paper we illustrate that recent advances

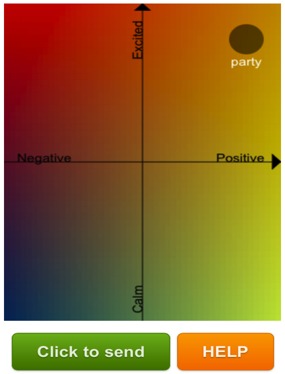

Figure 1: Mood Conductor app.

Figure 2: Mood Conductor visual feedback.

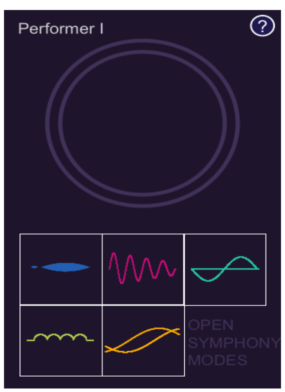

Figure 3: Open Symphony app.

in semantic and affective computing and the semantic web can be useful in humanising our interaction schemes. Ontologies, for instance, allow us to elegantly map between human concepts and technical parameters. The crossroads between human-computer interaction and music [5] thus become a fertile ground for transforming the way we create and experience music. We discuss several state-of-the-art interactive music systems devised for musical activities belonging to different parts of the musical chain: (i) composition, (ii) production, (iii) performance, and (iv) listening.

Revisiting the Shared Values Between Audience and Performers

In traditional art, popular and folk music performances, audiences are not usually engaged in the musical creation process and act as receiver. There are however many examples of audience participation in arts in which the role of the receiver-spectator is extended by means of active creative decisions influencing the artistic text; for instance, in literature, Steve Jackson and Ian Livingstone's fantasy roleplay gamebooks "in which you are the hero!" invite the reader to determine the progression of the narrative based on interactions with a dice; in improvisational theatre and stand-up comedy, it is not unusual for performers to incite the audience to provide cues for their sketches, orally. Until recently such creative participation was handled with low technology (e.g. dice, gestures, written notes, etc.) restricting the creative interactions to a limited number of possibilities, and preventing scalability to large audiences. The advent of mobile, web and information visualisation technologies has opened new possibilities for the design of computer-supported collaborative music making systems.

Two examples of such systems are Mood Conductor [3] and Open Symphony [4] which let audience members conduct performers and take part in the musical composition process over the course of live performances. They were developed using HCI-based methodologies, by user-centred design and iterative evaluations of prototypes with both performers and audiences, and subsequent improvements. Both systems rely on (i) a client/server web infrastructure allowing the exchange of creative data between audience and performers, and (ii) visualisation to provide explicit feedback and musical directions. The audience's creative interactions are expressed through a voting system. Votes can be made from a client device (PC, tablet, smartphone) running a smartphone-friendly and platform-independent web application which send or pull data to a central server. A visual client accesses the server to generate visuals responding to audience inputs. One of their advantages compared to previous audience participation forms using low technologies is that they are scalable to large audiences.

Mood Conductor was designed based on results from psychological research on human emotion modelling [7]. The app (see GUI on Figure 1) allows users to select desired moods in a two-dimensional space representing arousal (or excitation) and valence (or pleasantness). To facilitate the characterisation of different moods the user interface also displays colours and tags related to affect in the AV space. Performers are guided by the audience's mood votes using time-varying visualisations projected on a screen (Figure 2). The moods "elected" by the audience following a proportional representation system serve as expressive indications to guide musical interpretations of either existing written music pieces or spontaneous improvisations.

By reconfiguring the chain of communication and redistributing the skills of musicianship, the Open Symphony system overrides the unidirectional creative relationship between audience and performer, establishing a co-leading partnership for music-making. With the Open Symphony

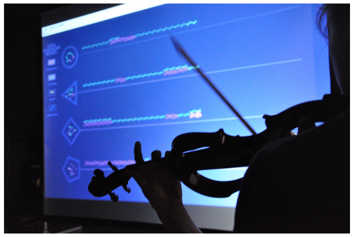

Figure 4: Open Symphony visual feedback with audience-generated symbolic score.

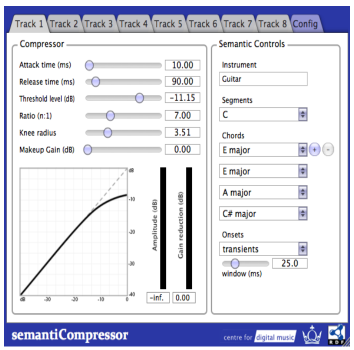

Figure 5: GUI of the intelligent semantic compressor.

Figure 6: GUI of the Semantic Audio Parametric EQ.

app (see GUI on Figure 3) audience members can choose amongst various musical playing modes (e.g. drone, motif, improvisation) which are then interpreted by performers from a dynamically generated symbolic score (Figure 4). Mood Conductor and Open Symphony hence propose a new creative balance between audiences and performers that do not require audience members to have expert musical skills. The shared roles of participants create an environment which draws upon the improvisational and performance expertise of the musician and the active listening and reflection position of the audience.

Transforming the Music Production Workflow: Intelligent Audio Plugins

Digital audio effects (DAFx) are essential tools in contemporary music production as they can be used to enhance the perceived quality of sonic elements or for sound design. Current DAFx plugin interfaces for digital audio workstations (DAWs) require expert knowledge from users as they generally rely on multiple low-level control parameters affecting one or several perceptual sound attributes, such as dynamics, timing and timbre. Mixing engineers typically use different parameter settings for different sections (e.g. chorus, verse) to reinforce the arrangement of a piece and create variations between sections. The tuning of DAFx plugins for a multitrack project can hence be a time-consuming and complex process. Semantic web technologies and music information retrieval (MIR) offer promising prospects to transform the music production workflow and make it more efficient by revisiting the way to control plugins and automate their application when judged appropriate. In [10] the authors propose an intelligent semantic audio compressor (Figure 5) which can adapt its parameters based on features such as instrumentation, structural segments, chords, and note onsets. [8] presents a suite of DAFx plugins that automatically map low level control parameters (e.g. filter cutoff frequency) to high level controls more easily understood by users (e.g. bright, warm). The plugins can load and save semantic profiles stored on a server which gather user descriptions of the sound, plugin parameters and extracted audio features (Figure 6). The systems described above have the potential to ease the audio mixing task, making it more accessible and efficient.

Transforming the Listening Experience: Interactive and Adaptive Music Player

The aural side of the listening experience of commercial music recordings provided by mainstream technologies has not changed much from how it was with a cassette Walkman in the early eighties; except in rare cases, tracks are played back in linear and inflexible ways and their selection remains based on bibliographic (e.g. artist, title) rather than high level creative metadata (e.g. mood). The Semantic Music Player [9] is a platform built to investigate new music playback paradigms on mobile devices based on context and user interactions. It builds on an abstract representation of the musical structure and semantic information (e.g. analytical features extracted from the audio or manually annotated metadata) which can be queried and navigated in configurable ways using semantic web technologies. While the player is designed to hide the complexity from the listener, the framework enables technologists and composers/producers to define arbitrary functional mappings that modify the music dynamically based on (i) mobile sensor data, (ii) user interface controls, (iii) contextual information, (iv) musical metadata pulled from online resources, as well as (v) semantic information queried from the abstract representation of the music. With the appropriate configurations, the player can for instance adapt music playback to the geographical situation of the listener and react to sensor inputs. Song durations can be altered using similarity information between sections at various hierarchical levels, and

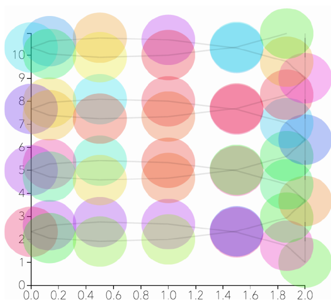

Figure 7: Note-based 3D spatial rendering of recordings based on Drobisch-Shepard's pitch helix. *x: chroma, y: octaves, size: note duration (random colors).*

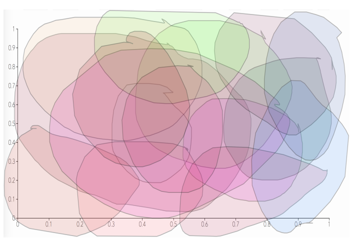

Figure 8: Complex mappings defined by shapes.

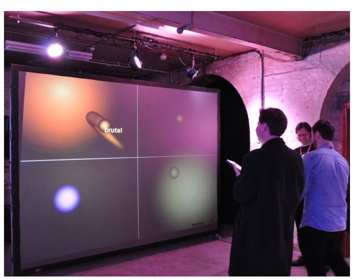

Figure 9: Moodplay interactive music player installation.

automatic transitions can be made using automatic beat-matching, time-stretching and cross-fading algorithms. The player can also be configured to re-render a monaural audio recording decomposed into separate events [2] and tune spatial position in the mix as a function of analytical information (Figure 7). Listeners can experience the music by moving around within the musical space using mobile sensors. More complex mappings can be created using continuous interlocking shapes in a multidimensional space each of which can control musical parameters such as tempo, volume, and effects (Figure 8). This space can then be navigated by directly mapping input values to spatial position.

User studies highlighted the inclination of listeners to select and discover music based on mood (e.g. uplifting) [6]. Moodplay [1] is an example of an interactive music player allowing users to collaboratively determine which songs are played based on the songs' mood characteristics. The player is controlled through the web app presented in the first section (Figure 1). The system uses crowd-sourced tag statistics to determine the location of songs in the arousal valence space and semantic web technologies to query songs by mood coordinate values. Visualisations and adaptive lighting effects can be added to enrich user experience (Figure 9). A rich set of application contexts were suggested by users for Moodplay, ranging from party and home usage, to therapy and fitness [1].

Conclusion

We discussed various interactive systems along the musical value chain focusing on simple and intuitive high-level controls. These examples show that the use of music informatics, affective computing and semantic technologies can not only help transform the way we interact with such systems, but the musical experiences themselves.

Acknowledgements

This work has been partly supported by EPSRC Grant EP/L019981/1, Fusing Audio and Semantic Technologies for Intelligent Music Production and Consumption.

References

[1] M. Barthet, G. Fazekas, A. Allik, and M. Sandler. 2015. Moodplay: An Interactive Mood-based Musical Experience. In *Proc. of Audio Mostly*. ACM, 3:1–3:8.

[2] S. Ewert, B. Pardo, M. Müller, and M. D. Plumbley. 2014. Score-informed source separation for musical audio recordings: An overview. *IEEE Signal Processing Magazine* 31, 3 (2014), 116–124.

[3] G. Fazekas, M. Barthet, and M. Sandler. 2013. The Mood Conductor System: Audience and Performer Interaction using Mobile Technology and Emotion Cues. In *Proc. of the Computer Music Multidisciplinary Research conference*.

[4] K. Hayes, M. Barthet, Y. Wu, L. Zhang, and N. Bryan-Kinns. 2016. A Participatory Live Music Performance with the Open Symphony System. In *In Proc Int. Conference on Computer Human Interaction (Interactivity)*.

[5] S. Holland, K. Wilkie, P. Mulholland, and A. Seago. 2013. *Music and Human Computer Interaction*. Springer, Chapter Music Interaction: Understanding Music and HCI.

[6] J. A. Lee and J. S. Downie. 2004. Survey of music information needs, uses, and seeking behaviors: preliminary findings. In *Proc. ISMIR*.

[7] J. A. Russell. 1980. A circumplex model of affect. *J. of Personality and Social Psychology* 39, 6 (1980), 1161–1178.

[8] R. Stables, S. Enderby, B. De Man, G. Fazekas, and J. D. Reiss. 2014. SAFE: A system for the extraction and retrieval of semantic audio descriptors. In *Proc. ISMIR*.

[9] F. Thalmann, A. C. Perez, G. Fazekas, G. A. Wiggins, and M. Sandler. 2016. The Semantic Music Player: A Smart Mobile Player Based on Ontological Structures and Analytical Feature Metadata. In *Proc. of the 2nd Web Audio Conference*.

[10] T. Wilmering, G. Fazekas, and M. Sandler. 2012. High-Level Semantic Metadata for the Control of Multitrack Adaptive Digital Audio Effects. In *Proc. of the AES 133rd Convention*.