Expressive Music Interaction: tools, prototypes and hackathons

Jordi Janer
MTG-Universitat Pompeu Fabra
Barcelona
jordi.janer@upf.edu

Frédéric Bevilacqua
STMS lab IRCAM-CNRS-UPMC
Paris
Frederic.Bevilacqua@ircam.fr

Frederic Font
Sebastián Mealla
Sergi Jordà
MTG-Universitat Pompeu Fabra
Barcelona
first.last@upf.edu

Joseph Larralde
Emmanuel Fléty
Gaël Dubus
Norbert Schnell
STMS lab IRCAM-CNRS-UPMC
Paris
first.last@ircam.fr

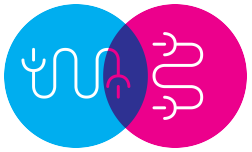
Abstract
This position paper aims to introduce to the HCI community two recent initiatives that bring music research tools to developers, hackers, digital makers and creative industries. By setting up a compendium of both hardware and software tools and APIs, we developed prototypes and took part in events such as workshops and hackathons. These events allowed digital makers to access to various building blocks and prototype new interactive music systems. The experience we gained by making available interoperable tools to large communities and building prototyping opens a large set of methodological questions that are valuable to discuss within the HCI community.

Author Keywords
Music; Audio Processing; Gesture Analysis; Sensors; Repositories; API; Communities

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

Introduction
Taking advantage of a large set of tools available, from tangible and embedded interfaces to music processing software and web repositories, new paradigms for Music Interaction can be designed.

Two Initiatives:



MusicBricks aims to promote the use of state-of-the-art music research tools from different institutions by creative communities. Partners are: Stromatolite, Sigma-Orionis, Fraunhofer-IDMT, TU Wien, IRCAM and Universitat Pompeu Fabra.
http://musicbricks.net

rapidmix

RapidMix aims to accelerate the production of the next generation of Multimodal Interactive eXpressive (MIX) technologies. Partners are: Goldsmiths,Universitat Pompeu Fabra, IRCAM,Plux, Orbe, Reactable, Roli Audio-gaming
http://rapidmix.goldsmithsdigital.com

Nevertheless, a major challenge today resides in effectively and rapidly integrating multiple technologies, allowing creative users and makers to build prototypes and evaluating their potential applications.

In this paper we report on two projects that explored methodologies for prototyping music interaction in the context of a large 'ecosystem' including academic partners, hackers/makers, and the industry. We believe that this experience should bring key elements, at both technological and methodological levels, for discussing how HCI and Music practice could enrich each other.

First, we briefly describe the basic technologies that we developed. This leads to a first issue on how to make the latest research tools available for creative users, and how to inter-operate with various open music technologies. As illustration, we briefly present two prototypes that were built using these technologies.

Second, we describe the different events such as workshops and hackathons that were set to give makers, designers and musicians the opportunity to experiments with our technological tools.

Finally, we provide a list of issues and open questions that should stimulate discussion and debates.

Tools and APIs for Makers

Our research labs have developed a number of technology tools for building Music Interaction systems. We deployed these technologies in the form of sensors, integrated software tools and APIs for an easy integration by creative users. Free and/or open source software are seen as crucial aspects for the widespread and fast adoption of these tools. We describe below the three types of tools and APIs:

1. Motion Sensors

   Ircam's wireless motion interface (called R-IoT) embeds a 9 DOF (Degree of Freedom) inertial measurement unit sensor, with embedded processing capabilities[1]. It allows to acquire 3D acceleration, 3-axis angular velocity and absolute orientation at a frame rate of 200 Hz and 16 bits over a WiFi connection, using the OpenSoundControl protocol.

2. Audio Processing Libraries
   Two audio analysis and processing libraries are made available (essentia of MTG-UPF[1] and MuBu & friends of Ircam[3]). They include a number of audio analysis algorithms (e.g. pitch and beat detection, note transcription, key estimation, etc.); and also a set of high quality audio processing and synthesis algorithms, allowing the transformation of audio samples in the context of audio effects or concatenative synthesis. These libraries are documented with tutorials to rapidly learn the algorithms features in a variety of use-cases.

3. Online Content Repositories
   We provide an online collaborative sound database called Freesound.org, where people from different disciplines share recorded sound clips under Creative Commons licenses. It currently hosts more than 290,000 sounds and has more than 4 million registered users (as of January 2016). This database was started in 2005, and is still developed and maintained by MTG-UPF[2].

Prototype Examples

Following we provide a short description of the two prototypes that were built combining the aforementioned tools.

[1] https://github.com/Ircam-RnD/RIoT

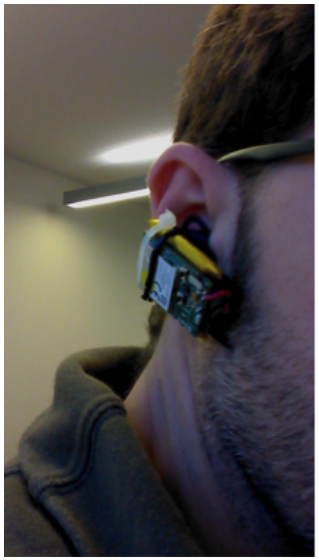


Figure 1: HandsFreeSoundMachine. Picture of the mounted speech and gesture interfaces. The user interacts with voice commands and head movements.

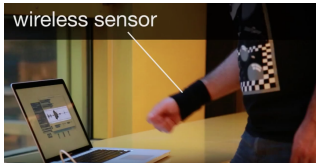


Figure 2: FreeMix prototype. The wireless sensor is placed in the wrist of the user through an elastic band. Visual feedback of the played sound is displayed on screen.

These prototypes introduce novel music interaction systems, while acting as example applications for creative users.

HandsFreeSoundMachine
Our first prototype uses voice and gesture inputs to control timing and timbre of a drum machine. Voice input is used to query sounds from an online repository, while gesture analysis is used to set the rhythm pattern of a 4-track step sequencer. Voice input can have two functions: (i) a speech recognition input, and (ii) a vocal imitation input to control the acoustic characteristics of the retrieved sound.

This system[2] combines hardware and software tools together with online web services. It integrates: a commercial bluetooth headset, a wireless motion sensing platform (R-IoT), an online sound repository API (Freesound) and a publicly available Speech Recognition online service API (Google Speech API). This hands-free interaction system combines voice and head gesture, hence allowing a musician to play another instrument simultaneously (e.g. a guitar).

FreeMix
As for the *HandsFreeSoundMachine*, this second prototype also uses Freesound and R-IoT along with additional movement and sound processing available in the Max environment[3]. This prototype illustrates how one can build a collaborative live performance using sound web repositories as initial material for a collective musical performance.

The sound synthesis is based on descriptors based concatenative synthesis: selecting and replaying small bits of sound samples using sound descriptors. Sound segmentation and analysis is thus central of this approach.

The real-time audio processing can be controlled using different inputs: motion, sounds and physiological signals. Different types of sound control and interaction modes can be combined, such as pre-determined 'playing techniques' as well as interactive machine learning techniques.

Importantly, the prototype is aimed to ultimately offer a collaborative experience. New sounds can be recorded on-the-fly and uploaded to a database so that users can remix sounds produced by others.

Interacting with Makers/Designers/Musicians

An important aspect of the two projects reported in this paper is the methodology applied to present the tools to the expert users or makers. We carried out both guided workshops and hackathons to collect more experimental ideas.

Workshops
The workshops were ideal scenarios to present the APIs to makers through hands-on sessions that directly engaged participants. Under the umbrella name of *Music Processing*, each session addressed one music processing topic, giving hints about the use of these technologies in practical applications.

Hackathons
Besides the workshops, we offered our tools and APIs in hackathons. This setting provided hackers with a good creative environment to experiment with the tools and develop an early prototype, proof-of-concept or idea. Hackathons took place in several international locations during 2015 (MusicTechFest Umeä, MusicTechFest Ljubljana, Music Hack Day Barcelona).

[2] Source code available:https://github.com/MTG/hands-free-sound-machine
[3] Mubu http://forumnet.ircam.fr/product/mubu-en/

Hackathons provided valuable user feedback on the offered tools. As they are being continuously extended with recent research outcomes, hackathons represent an excellent opportunity for applying user-centered design.

Discussions

The field of digital music has been historically at the origin of several communication standards that greatly facilitated the modular use of tools, hardware and software. The MIDI standard and later the OpenSoundControl protocols are relevant concrete examples. The increased complexity of dealing simultaneously with multimodal data, processing/synthesis software and tangible interfaces opens new challenges for prototyping. The projects we reported here represent an attempt to answer such challenges by first providing makers and hackers a large set of open tools with documented APIs. The two prototypes we shortly described represent original examples of possible applications of our tools that we believe to be at the cutting-edge of music technology research.

Still the appropriation of such tools, from ideas to functional prototypes, remains difficult. Inspired from an HCI approach based on user-centered design, we conducted various events from workshops to hackathons.

As a methodology for ideation, design and evaluation of the technology readiness, hackathons represents an appropriate approach we would like to discuss in the context of a larger HCI view, using our concrete experience.

First, at hackathons, the burst of energy, ideas and possible unexpected collaboration between person of various background offer a fertile ground for 'out of the box' thinking. Yet, the access to technological tools, through predefined APIs and examples might also limit or 'constrain' for the imagination.

The balance between short-term experimental events and long-term iterative design process should thus be discussed in light of other HCI practices. The specificity of the music field resides in the generally high expertise of the designers/makers/musicians which represent unique opportunities to build new methodologies for interaction design.

Acknowledgments

This work has been supported by two H2020 Research and Innovation grants funded by the European Commission : MusicBricks (Grant Agreement Nr. 644871) and RapidMix (Grant Agreement Nr. 644862) . We also thank our collaborators from the other partner institutions working in these two projects.

References

[1] D. Bogdanov, Nicolas Wack, Emilia Gómez, Sankalp Gulati, P. Herrera, O. Mayor, Gerard Roma, J. Salamon, J. Zapata, and Xavier Serra. 2013. ESSENTIA: an Audio Analysis Library for Music Information Retrieval. In *International Society for Music Information Retrieval Conference (ISMIR'13)*. Curitiba, Brazil, 493–498.

[2] Frederic Font, Gerard Roma, and Xavier Serra. 2013. Freesound technical demo. In *Proceedings of the 21st ACM international conference on Multimedia - MM '13*. 411–412. DOI: http://dx.doi.org/10.1145/2502081.2502245

[3] Norbert Schnell, Axel Röbel, Diemo Schwarz, Geoffroy Peeters, and Riccardo Borghesi. 2009. MuBu & Friends - Assembling Tools for Content Based Real-Time Interactive Audio Processing in Max/MSP. In *Proceedings of the International Computer Music Conference (ICMC)*.