Jamming with a Smart Mandolin and Freesound-based Accompaniment

Conference Paper · November 2018

DOI: 10.23919/FRUCT.2018.8588110

Jamming with a Smart Mandolin and Freesound-based Accompaniment

Luca Turchet and Mathieu Barthet
Center for Digital Music,
Queen Mary University of London
luca.turchet@qmul.ac.uk, m.barthet@qmul.ac.uk

***Abstract*—This paper presents an Internet of Musical Things ecosystem involving musicians and audiences interacting with a smart mandolin, smartphones, and the Audio Commons online repository Freesound. The ecosystem has been devised to support performer-instrument and performer-audience interactions through the generation of musical accompaniments exploiting crowd-sourced sounds. We present two use cases investigating how audio content retrieved from Freesound can be leveraged by performers or audiences to produce accompanying soundtracks for music performance with a smart mandolin. In the performer-instrument interaction use case, the performer can select content to be retrieved prior to performing through a set of keywords and structure it in order to create the desired accompaniment. In the performer-audience interaction use case, a group of audience members participates in the music creation by selecting and arranging Freesound audio content to create an accompaniment collaboratively. We discuss the advantages and limitations of the system with regard to music making and audience participation, along with its implications and challenges.**

I. Introduction

In recent years, research has started to investigate how to extend the Internet of Things (IoT) paradigm to the musical domain, proposing visions for the so-called *"Internet of Musical Things (IoMusT)"* [1], [2]. This emerging research field draws upon various lines of existing research such as IoT [3], [4], [5], new interfaces for musical expression [6], ubiquitous music [7], networked [8] and mobile music [9], [10], human-computer interaction [11], [12], participatory art [13], and artificial intelligence applied to musical contexts [14].

A definition for the IoMusT was proposed in [15] following a computer science perspective: *"the Internet of Musical Things refers to the ensemble of interfaces, protocols and representations of music-related information that enable services and applications serving a musical purpose based on interactions between humans and Musical Things or between Musical Things themselves, in physical and/or digital realms. Music-related information refers to data sensed and processed by a Musical Thing, and/or exchanged with a human or with another Musical Thing"*.

A Musical Thing was defined as *"a computing device capable of sensing, acquiring, processing, or actuating, and exchanging data serving a musical purpose"*. Musical Things, such as smart musical instruments [16], are connected by an infrastructure that enables multidirectional communication, both locally and remotely. An IoMusT ecosystem gathers interoperable devices and services that connect performers and audiences to support novel forms of performer-performer and audience-performers interactions. However, only a few use cases for these interactions have been implemented so far.

In a different vein, the growth of semantic web and cloud computing technologies enabled to store and provide timely access to vast amount of data online. A culture of sharing forged around new web services tailored for creative content (e.g., YouTube for music and video, Flickr for photos, etc.). Creative Commons (CC) [17] appeared as a legal framework enabling to license creative work and derivatives, as well as setting conditions for reuse and potential commercial exploitation. A recent endeavor, the Audio Commons Initiative [18] [19], aims to bridge the gap between audio content producers, providers and consumers through a web-based ecosystem. The approach combines techniques from music information retrieval (to extract creative metadata to automatically annotate audio content) and the semantic web (to structure knowledge and enable intelligent searches). Content aggregators part of the Audio Commons ecosystem, such as Freesound (see Section III), provide access to audio data through user-facing and application programming interfaces (APIs). The present work and our previous study reported in [20] are to the best of our knowledge the first initiatives to combine Internet of Musical Things and Audio Commons ecosystems to create new forms of musical interaction.

This paper presents an Internet of Musical Things ecosystem that combines a smart musical instrument (a traditional acoustic mandolin augmented with sensors, embedded intelligence and Internet connectivity [21]), smartphones, and the Audio Commons content provider Freesound [22]. Such ecosystem was devised to enable new forms of interactions between the performer and the instrument, as well as between the performer and audience members, through the generation of musical accompaniments leveraging online audio content.

The Freesound database used in our system provides a collection of several hundreds of thousands of crowd-sourced non musical and musical sounds licensed under Creative Commons. This collection is the source of audio for the web-based music making tool Playsound [23], which lets users search for and layer Freesound sounds to produce semantically-informed improvised music. A type of musical composition that can be generated using sounds from Freesound are soundscapes, an aural equivalent to landscapes, which convey a sense of place and emotional context to the listeners through the organisation over time of various (processed) acoustic scenes (see e.g., [24] for a discussion on soundscapes). In this work, we are interested in using Freesound content in a different way, by letting performers or audience members generate a musical accompaniment that can guide or complement what a musician is playing.

The remainder of this paper is organised as follows. Section II presents an overview of related works. Section III describes the developed IoMusT ecosystem, while Section IV presents a critical discussion of its advantages and disadvantages as well as implications and challenges. Finally, Section V provides summarising conclusions.

II. Related works

In this section we review key works on technologies related to the proposed architecture.

A. The Audio Commons initiative and its artistic use

Audio Commons provides an ecosystem through which musicians can access audio content e.g., with web browser-based interfaces (e.g., [22], [25]), plugins (AudioGaming's AudioTexture [26], Waves Audio's SampleSurfer), or live coding tools [27]. Traditionally, digital audio workstations (DAWs) and digital musical interfaces were designed to operate with local audio content, for example personal recordings gathered by the musician. Audio Commons proposes an architecture that can disrupt linear media production providing access to distributed audio content in a user-friendly way. Several tools for music production have been developed leveraging this architecture. AudioGaming's AudioTexture is a plugin that enables users to generate unique textures [28] through transformations of Audio Commons content based on granular synthesis. Waves Audio's SampleSurfer is a plugin that acts as an audio content search engine informed by semantic metadata and audio content-based features. It enables quick access to hundreds of thousands of sounds from various online content providers according to requirements matching composers' needs. Playsound [29] [23] is a web-based tool designed for beginners or advanced musicians willing to explore music composition based on semantic ideation and spectrogram sound introspection. Over plugins aimed at DAW-literate musicians, Playsound appears as more inclusive and can be used in ubiquitous contexts [7]. However, it does not provide the versatility of a full DAW environment.

Although web browser-based interfaces have the advantages of not requiring additional software they do not integrate with the tools typically used by computer musicians, such as DAWs. In the context of computer-assisted music production, plugins that connect to Audio Commons directly within the DAW present some advantages avoiding a disruption of the workflow due to switching platform to find sounds that are relevant. Relatedly, enabling smart musical instruments users to access Audio Commons sounds appears as more straightforward if the search and retrieval can be made without having to switch to an external laptop or desktop computer. This workflow aspect has informed our design (see Section III) by exploiting smartphones that can easily be used in conjunction to smart musical instruments.

In a previous new interfaces for musical expression study [20], we proposed a sonic wearable interface letting users trigger and transform sounds downloaded from Freesound through body-based gestural interactions. This was achieved by combining a portable Musical Thing connected to Internet and e-textile sensors. As with this work, the system enabled to repurpose sounds from Freesound in a creative way.

B. Networked music performance

Networked music has been the object of much attention by the computer music community at least since the 80s. However, advances have not reached the level that web-based technologies provide for speech and video-based communication over networks. This is due to the stringent temporal constraints necessary for music [8] where delay as low as 20ms can make music performance unintelligible. Several reviews of networked music performance (NMP) systems can be found in [30], [31], [32], [8]. Dedicated streaming applications minimising latency and maximising audio quality were developed such as JackTrip [33] addressing the challenges of bi-directional performance using modern Wide Area Networks. IoMusT ecosystems [1], [15] can facilitate NMP and widen the range of applications e.g., including audience-performer interaction but share common challenges. Technical issues in network transmission of music and audio are discussed in the AES white paper [34] including latency, time stamping (synchronisation), quality of service, audio formats, number of channels and client-server issues.

C. Technology-mediated audience participation

Interactive arts has benefited from information and communication technologies providing novel communication pathways between producers and consumers of creative content. Technology-mediated audience participation (TMAP) is a burgeoning field which has been applied to theatre [35] and music (see [9], [36], [13], [37] for reviews). [13] proposed various dimensions to characterise participatory music performance systems including *participation level* (from partial to full), *motivation* (e.g., imitative, turn-taking, competitive, conducting), *agency distribution* (from individual to collective levels), *agency mediation* (e.g., direct production of sounds or indirect by communicating intentions to performers), *modalities and media* (e.g., sounds, visuals, tweets, etc.), *interfaces* (e.g., smartphones phones, dedicated tangible interfaces, etc.), and situation (e.g., co-located or remote). Several systems invite the audience to conduct the performers such as Mood Conductor [38] and Open Symphony [39], [13]. Others let the audience contribute to elements of a performance by generating musical/visual content with their mobile devices such as A.bel [40] or the present work. Systems where the performance is fully generated by the audience often exploit synthesis techniques on mobile devices [41], [36], [10], such as Web Audio [42], [43], [44].

To date, audience creative participation in music performance has mainly relied on ad-hoc tangible interfaces (see e.g., [45], [46]) or smartphones. The long evolution of lutherie and instrument crafting witnesses the complexity in providing meaningful expressive capabilities to instruments also ensuring their longevity. Typical controls afforded by smartphones (e.g., screen touch, sweep, tilt) and the shape and nature of the interface itself may fail to satisfy musicians expecting to find affordances from traditional musical instruments. Traditional musical instruments often provide intimate control and large degrees of freedom based on proprioception and kinaesthetics, systematic and reproducible gestures, subtle nuances of timbre

[47], [48], correct intonation, etc. There is scope for SMIs based on traditional instruments to be used in a participatory music performance context as they conform to what musicians are familiar with yet provide ubiquitous capabilities [49].

D. Smart Instruments

SMIs integrate various musical and network technologies initially developed for different purposes [16]. SMIs can be seen as an evolution of hyperinstruments [50] and augmented instruments that integrate sensors and/or actuators on conventional instruments [51], [52], [53], building up on Internet of Things [5] and artificial intelligence technologies. Examples of commercial smart instruments include the Sensus Smart Guitar developed by MIND Music Labs [16], an electro-acoustic guitar with on-board processing (ELK music operating system [54]), sensors and wireless communication, as well as the HyVibe guitar [55], an acoustic guitar with an active vibration control system enabling the production of various audio effects. SMI prototypes developed within academia include the smart cajón [56] developed using the Bela board for low-latency audio and sensor processing [57] applying sensor fusion and semantic audio techniques to detect strokes [58]. Gregorio and Kim [59] proposed augmented acoustic drums using electromagnetic actuation that can be controlled in a networked way to follow specific timbral configurations conferring them smart instrument properties.

III. Proposed IoMusT ecosystem

We developed an IoMusT ecosystem involving a smart mandolin, smartphones, and the Freesound repository with the goal of supporting two types of technologically-mediated interaction:

- Use case 1: performer-instrument interaction;
- Use case 2: performer-audience interaction.

Both use cases aimed to enable the creative use of content retrieved from Freesound as an accompaniment for music played on the smart mandolin. In the performer-instrument interaction use case, the performer can search for content through a set of keywords and can then arrange it in order to create a desired accompaniment. In the performer-audience interaction use case, the accompaniment is generated by a group of audience members who can retrieve and organize Freesound content collaboratively.

Fig. 1 shows a schematic representation of the implemented IoMusT ecosystem. The ecosystem comprised the following components:

Freesound repository. Freesound is a collaborative repository of Creative Commons-licensed audio samples developed at the Music Technology Group of Universitat Pompeu Fabra. At the time of writing, the Freesound database provides access to about 400k musical and non-musical sound samples. The available metadata information about the sounds depends on what has been provided by authors during uploads including tags, descriptions or file names. Freesound enables designers to create third-party applications exploiting its audio content in live applications by granting access to the database trough a REST API [60] [61].

Smart mandolin. The smart mandolin [21] (see Fig. 2) comprises a conventional acoustic mandolin enhanced with different types of sensors, a contact microphone (HotSpot by K&K Sound), a loudspeaker, wireless connectivity, embedded battery, and the Bela low-latency audio processing board [62]. The sensor interface consists of six Force Sensing Resistors of various sizes and types (Interlink Electronics squared FSR 406, three rounded FSR 402, one small-rounded FSR 400, and one strip FSR 408), one Soft Pot ribbon sensor (Spectra Symbol), an inertial measurement unit (IMU, Bosh BNO055), and an Infrared Proximity Sensor Short Range (Sharp GP2Y0A41SK0F). The ribbon sensor was attached, thanks to its adhesive film, on top of the strip pressure sensor in order to provide simultaneous information about position and pressure of the finger. The IMU used for tracking the instrument movements in the tridimensional space, was placed into a box containing the processing unit and wireless connectivity.

Wireless connectivity was achieved by means of a Wi-Fi USB dongle (A6100-100PES by NETGEAR, which supports the IEEE 802.11ac Wi-Fi standard). Wireless data reception and forwarding were achieved leveraging Open Sound Control (OSC) messages over the User Datagram Protocol.

The audio engine was coded in the Pure Data real-time audio processing environment and comprised a variety of ad-hoc sound effects modulating the instrument's string sounds, a library of sound samples to be triggered, as well as mapping strategies to control the sound production from the data gathered from the sensors as well as from the real-time extraction of features from the audio signal captured by the microphone. We extracted the note onset from the audio signal captured by the microphone, by leveraging the Pure Data object fiddle$\sim$ (which is described in [63]). From this low-level feature we then calculated the note density over 60 seconds (which was used to control the smartphones app's layout, as described in the following). Finally, the sound engine supported the playing and looping of backing tracks retrieved from Freesound.

Smartphones and apps. We developed a dedicated smartphone app serving both use cases (performer-instrument and performer-audience interactions). The app was created by using the TouchOSC [64] environment, enabling to rapidly build modular control surfaces for mobile applications leveraging OSC messages. The app was only used to send OSC messages to the smart mandolin, where all computations were performed (including the interaction with the cloud).

Fig. 3 illustrates the layout of the app for performer-instrument interaction case. The app allows for the selection of pre-defined keywords belonging to three categories: environmental, musical, and human sounds. The environmental sounds category comprises the following keywords: 'forest', 'storm', 'street', and 'sea'; the musical sounds category is composed of the keywords 'blues', 'folk', 'rock', and 'classical music'; the human sounds category is composed of the keywords 'voices', 'footsteps', 'singing', and 'applause'. The app comprises, for each category, a toggle to start and stop the playback of the retrieved sounds and a fader to control their volume.

In the performer-audience interaction case, three users were involved with each app displaying only one of the three categories (one per each smartphone user). The displayed

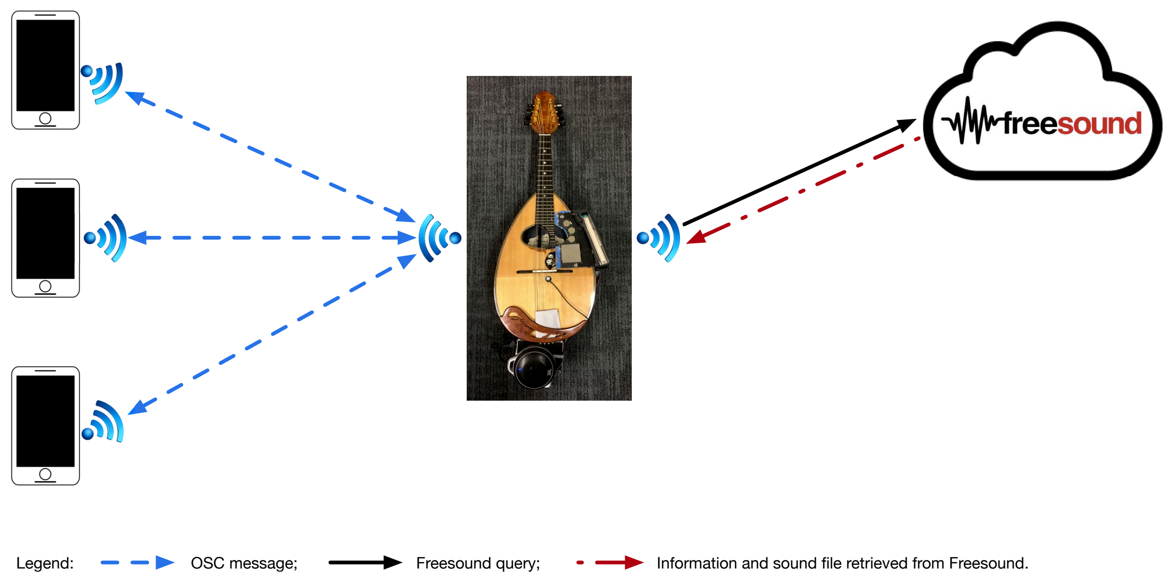

Fig. 1. Schematic representation of the IoMusT ecosystem.

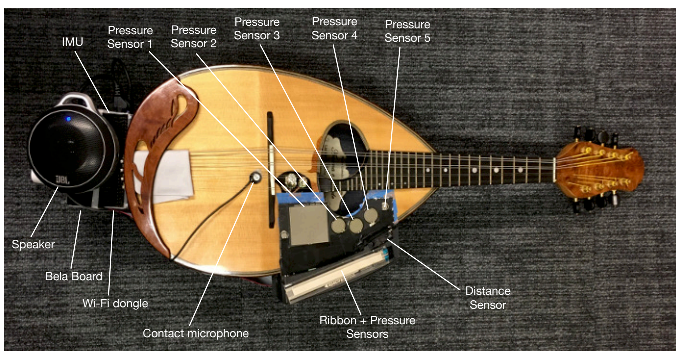

Fig. 2. The smart mandolin with the indication of its components.

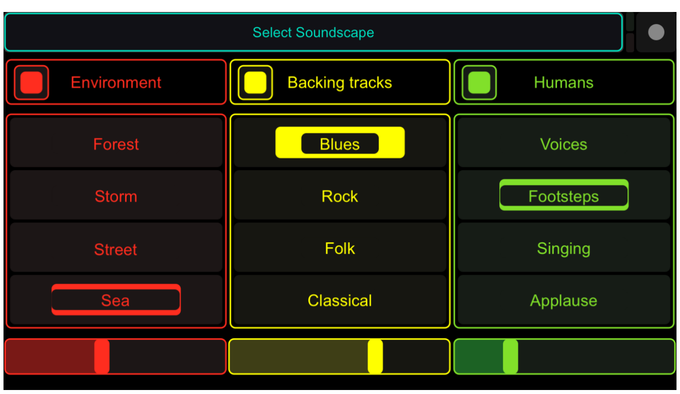

Fig. 3. The layout of the app.

category could change based on messages delivered by the smart mandolin to the app. Specifically, the layout change was based on the analysis of the density of the notes (which was performed and delivered every 60 seconds): the values of the density were divided into three ranges (low, medium, and high), which were mapped to a different configuration of the layout of the three smartphones (see Table I). The notes density was chosen as a control parameter because from the performer's standpoint is an expressive parameter easy to control in order to produce the designed changes in the app's layout.

TABLE I. MAPPING BETWEEN THE NOTE DENSITY COMPUTED FROM THE SMART MANDOLIN AUDIO SIGNAL AND THE CONFIGURATION OF THE LAYOUTS OF THE THREE SMARTPHONES.

Density [notes per minute]	Smartphone 1	Smartphone 2	Smartphone 3
Low [1, 30]	Environment	Human	Backing tracks
Medium [31, 150]	Backing tracks	Environment	Human
High [151, 500]	Human	Backing tracks	Environment

Network. All devices were connected using a Wi-Fi router (TP-Link TL-WR902AC), which features the IEEE 802.11ac standard over the 5GHz band. Following the recommendations reported in [65] to optimise the components of a Wi-Fi system for live performance scenarios to reduce latency and increase throughput, the router was configured in access point mode, security was disabled, and only the IEEE 802.11ac standard was supported. To enable Internet connectivity we used a 4G dongle, which was plugged into the dedicated USB port in the router.

Fig. 4 shows a picture of the system during the performer-audience interaction use case.

A. Software

The software responsible for retrieving the sounds from the Freesound repository runs on the smart mandolin's Bela board, and was implemented in Python. The script leveraged a Python client for the Freesound API (available at https://github.com/MTG/freesound-python) and the pyliblo library (available at http://das.nasophon.de/pyliblo) for handling OSC messages. An OSC server received the messages sent by the smartphones and, on the basis of the received keyword, sent a query to the Freesound repository. A list of a maximum of 20 filenames was retrieved after having sorted the sounds by

Fig. 4. Illustration of a jam session involving a smart mandolin player and three audience members using the proposed system.

descending duration. From that list one filename was randomly selected and then downloaded as an mp3 file (an mp3 file was downloaded instead of the (possible) .wav counterpart to reduce the download time and file size due to the live interaction). Since the Pure Data version running on Bela (i.e., libpd [66]) could only playback .wav files, the Python script converted each downloaded file from .mp3 into .wav thanks to the pysox library (available at https://github.com/rabitt/pysox) [67]. Finally, an OSC client instructed the Pure Data patch implementing the sound engine to start/stop playing the retrieved sound and tune its volume based on user interaction with the corresponding fader.

IV. Critical analysis

The two use cases (performer-instrument and performer-audience interactions) were preliminary tested with the goal of assessing the proof of concept of the system, mainly at technical level. The test of the performer-instrument interaction use case involved only the first author (who is a smart mandolin player). It consisted of the use of the system during two months. The test of the performer-audience interaction use case involved the first author and three audience members (1 female, 2 males, aged between 26 and 31 years old). It was performed in a room of Queen Mary University of London and lasted 20 minutes during which the first author played improvisations on the smart mandolin while the three audience members interacted with the apps on their smartphones to create the soundscape-based accompaniment. Future tests of both use case are planned involving other mandolin players and a larger number of audience members.

From the conducted preliminary evaluations it is possible to report the following critical analysis. Firstly, the app and the whole architecture allow one to create interesting evolving soundscapes as an accompaniment to the music played by the smart mandolin. For instance, as displayed in Fig. 3, one could create a soundscape mixing a blues track coupled with sounds of sea waves and of footsteps, as well as adjust their presence and volume in the mix using toggles and faders, respectively.

The need of having more choices in terms of sound categories emerged during both use cases. Both the performer and the three audience members deemed the three pre-defined keywords as too few, and felt the need to explore more sonorities especially after becoming familiar with the system. The app can be easily modified to involve other categories of sounds, which would lead to different resulting soundscapes. In addition, as far as the performer-instrument use case is concerned, the performer felt the need of controlling the soundscapes without interacting with the app using the hands. This usually was inconvenient and in some cases disrupted the flow of the music. The use of a pedalboard, having switches and volume controls, was suggested to lead to a better interaction with the sound content retrieved from Freesound, as the use of the feet would free the hands.

In terms of musical expressiveness, a strength of the system is the large variety of sounds that can be played related to the pre-defined keywords, which fosters the creation of a narrative. However, it is worth noticing that the system comes with an element of surprise: the retrieved sounds may not fully correspond to the keywords or to the expectations of the user. This is due to the noise often present in crowd-sourced data but also the breadth with which semantic terms can be understood and used. A limit of the system is that the retrieved sound samples can not be auditioned prior to being played. However, the volume fader can be used to generate slow fade ins, which limits the effect of unwanted sounds that could then be deactivated or substituted.

The developed system could scale to several smart instruments and smartphones. This poses a set of technological and artistic challenges. Technologically, one needs to extend the architecture here proposed to account for the increased number of Musical Things present in the ecosystem. This may lead to large increments in the bandwidth demand, since more files will be downloaded simultaneously which could create congestions of the network. Artistically, composers will have to face the challenge to devise structures of the compositions where several audience members can control categories of sounds that need to fit well together in order to create meaningful and appealing soundscapes and narratives.

V. Conclusion and future work

This paper presented an IoMusT ecosystem which linked an instance of a smart instrument with smartphones and an Audio Commons content repository. The ecosystem was shown to support interactions between the performer and the smart instrument as well as between the performer and audience members.

While we presented the architecture and a critical analysis of our system, this work has not formally assessed the user experience and creativity support. We plan to assess these aspects in future work by means of experimental procedures involving performers, composers, and concert goers. In addition, this study involved one smart instrument and three audience members using smartphones. However, the system could scale to several smart instruments and smartphones. This interesting and challenging scenario will be investigated in future work. Furthermore, a possible extension of the present work consists in developing a new app where the user can type any text that wish to retrieve (as is the case for Playsound.space [23]) rather than being constrained to a set of pre-defined keywords.

Acknowledgment

The authors acknowledge support from the EU H2020 Marie Curie Individual fellowship (No. 749561) and EU H2020 Audio Commons Initiative grant (No. 688382).

References

[1] L. Turchet, C. Fischione, and M. Barthet, "Towards the Internet of Musical Things," in *Proceedings of the Sound and Music Computing Conference*, 2017, pp. 13–20.

[2] D. Keller and V. Lazzarini, "Ecologically grounded creative practices in ubiquitous music," *Organised Sound*, vol. 22, no. 1, pp. 61–72, 2017.

[3] L. Atzori, A. Iera, and G. Morabito, "The internet of things: a survey," *Computer networks*, vol. 54, no. 15, pp. 2787–2805, 2010.

[4] D. Miorandi, S. Sicari, F. De Pellegrini, and I. Chlamtac, "Internet of things: Vision, applications and research challenges," *Ad Hoc Networks*, vol. 10, no. 7, pp. 1497–1516, 2012.

[5] E. Borgia, "The Internet of Things vision: Key features, applications and open issues," *Computer Communications*, vol. 54, pp. 1–31, 2014.

[6] A. Jensenius and M. Lyons, *A NIME Reader: Fifteen Years of New Interfaces for Musical Expression*. Springer, 2017.

[7] D. Keller, V. Lazzarini, and M. Pimenta, *Ubiquitous music*. Springer, 2014.

[8] C. Rottondi, C. Chafe, C. Allocchio, and A. Sarti, "An overview on networked music performance technologies," *IEEE Access*, vol. 4, pp. 8823–8843, 2016.

[9] A. Tanaka, "Mobile music making," in *Proceedings of the Conference on New Interfaces for Musical Expression*, 2004, pp. 154–156.

[10] G. Essl and S. W. Lee, "Mobile devices as musical instruments - state of the art and future prospects," in *Proceedings of the International Symposium on Computer Music Multidisciplinary Research*, 2017.

[11] Y. Rogers, H. Sharp, and J. Preece, *Interaction design: beyond human-computer interaction*. John Wiley & Sons, 2011.

[12] C. Rowland, E. Goodman, M. Charlier, A. Light, and A. Lui, *Designing connected products: UX for the consumer Internet of Things*. O'Reilly Media, Inc., 2015.

[13] Y. Wu, L. Zhang, N. Bryan-Kinns, and M. Barthet, "Open symphony: Creative participation for audiences of live music performances," *IEEE MultiMedia*, vol. 24, no. 1, pp. 48–62, 2017.

[14] J. Burgoyne, I. Fujinaga, and J. Downie, "Music information retrieval," *A New Companion to Digital Humanities*, pp. 213–228, 2016.

[15] L. Turchet, C. Fischione, G. Essl, D. Keller, and M. Barthet, "Internet of Musical Things: Vision and Challenges," *IEEE Access*, vol. 6, pp. 61 994–62 017, 2018.

[16] L. Turchet, "Smart Musical Instruments: vision, design principles, and future directions," *IEEE Access*, 2018.

[17] [Online]. Available: https://creativecommons.org

[18] [Online]. Available: http://www.audiocommons.org/

[19] F. Font, T. Brookes, G. Fazekas, M. Guerber, A. La Burthe, D. Plans, M. D. Plumbley, M. Shaashua, W. Wang, and X. Serra, "Audio commons: bringing creative commons audio content to the creative industries," in *Audio Engineering Society Conference: 61st International Conference: Audio for Games*. Audio Engineering Society, 2016.

[20] S. Skach, A. Xambó, L. Turchet, A. Stolfi, R. Stewart, and M. Barthet, "Embodied interactions with e-textiles and the internet of sounds for performing arts," in *Proceedings of the International Conference on Tangible, Embedded, and Embodied Interaction*. ACM, 2018, pp. 80–87.

[21] L. Turchet, "Smart Mandolin: autobiographical design, implementation, use cases, and lessons learned," in *Proceedings of Audio Mostly Conference*, 2018.

[22] [Online]. Available: https://freesound.org

[23] A. Stolfi, M. Ceriani, L. Turchet, and M. Barthet, "Playsound.space: Inclusive Free Music Improvisations Using Audio Commons," in *Proceedings of the Conference on New Interfaces for Musical Expression*, 2018, pp. 228–233.

[24] J. Pigrem and M. Barthet, "Datascaping: Data sonification as a narrative device in soundscape composition," in *Proceedings of the International Audio Mostly Conference*. ACM, 2017, p. 43.

[25] [Online]. Available: http://audiocommons.jamendo.com

[26] [Online]. Available: https://lesound.io/product/audiotexture/

[27] A. Xambó, G. Roma, A. Lerch, M. Barthet, and G. Fazekas, "Live repurposing of sounds: Mir explorations with personal and crowdsourced databases," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2018.

[28] L. Lu, L. Wenyin, and H. Zhang, "Audio textures: Theory and applications," *IEEE Transactions on Speech and Audio processing*, vol. 12, no. 2, pp. 156–167, 2004.

[29] [Online]. Available: http://playsound.space

[30] Á.. Barbosa, "Displaced soundscapes: A survey of network systems for music and sonic art creation," *Leonardo Music Journal*, vol. 13, pp. 53–59, 2003.

[31] G. Weinberg, "Interconnected musical networks: Toward a theoretical framework," *Computer Music Journal*, vol. 29, no. 2, pp. 23–39, 2005.

[32] L. Gabrielli and S. Squartini, *Wireless Networked Music Performance*. Springer, 2016.

[33] J. Cáceres and C. Chafe, "Jacktrip: Under the hood of an engine for network audio," *Journal of New Music Research*, vol. 39, no. 3, pp. 183–187, 2010.

[34] W. Woszczyk *et al.*, "Networking audio and music using internet2 and next-generation internet capabilities," *AESW P—1001 Technology Report TC-NAS 98*, vol. 1, 1998.

[35] T. Cerratto-Pargman, C. Rossitto, and L. Barkhuus, "Understanding audience participation in an interactive theater performance," in *Proceedings of the Nordic Conference on Human-Computer Interaction*. ACM, 2014, pp. 608–617.

[36] S. Lee, A. de Carvalho Junior, and G. Essl, "Understanding cloud support for the audience participation concert performance of crowd in c[loud]," in *Proceedings of the Conference on New Interfaces for Musical Expression*, 2016, pp. 176–181.

[37] F. Kayali, O. Hödl, G. Fitzpatrick, P. Purgathofer, A. Filipp, R. Mateus-Berr, U. Kühn, T. Wagensommerer, J. Kretz, and S. Kirchmayr, "Playful technology-mediated audience participation in a live music event," in *Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in Play*. ACM, 2017, pp. 437–443.

[38] G. Fazekas, M. Barthet, and M. Sandler, "Novel methods in facilitating audience and performer interaction using the mood conductor framework," in *Sound, Music, and Motion*, M. Aramaki, O. Derrien, R. Kronland-Martinet, and S. Ystad, Eds. Springer International Publishing, 2014, pp. 122–147.

[39] L. Zhang, Y. Wu, and M. Barthet, "A web application for audience participation in live music performance: The open symphony use case," in *Proceedings of the Conference on New Interfaces for Musical Expression*, 2016, pp. 170–175.

[40] A. Clément, F. Ribeiro, R. Rodrigues, and R. Penha, "Bridging the gap between performers and the audience using networked smartphones: the a.bel system," in *Proceedings of the International Conference on Live Interfaces*, 2016.

[41] N. Weitzner, J. Freeman, S. Garrett, and Y. Chen, "massMobile-an Audience Participation Framework," in *Proceedings of the Conference on New Interfaces for Musical Expression*, 2012, pp. 21–23.

[42] J. Lambert, S. Robaszkiewicz, and N. Schnell, "Synchronisation for distributed audio rendering over heterogeneous devices, in html5," in *Proceedings of the Web Audio Conference*, 2016.

[43] A. Stolfi, M. Barthet, F. Goródscy, A. Deusany, and F. Iazzetta, "Open band: Audience creative participation using web audio synthesis," in *Proceedings of the Web Audio Conference*, 2017.

[44] A. Stolfi, A. Milo, M. Ceriani, and M. Barthet, "Participatory musical improvisations with playsound.space," in *Proceedings of the Web Audio Conference*, 2018.

[45] B. Bengler and N. Bryan-Kinns, “Designing collaborative musical experiences for broad audiences,” in *Proceedings of the ACM Conference on Creativity & Cognition*. ACM, 2013, pp. 234–242.

[46] J. Freeman, “Large audience participation, technology, and orchestral performance,” in *Proceedings of the International Computer Music Conference*, 2005.

[47] M. Barthet, P. Depalle, R. Kronland-Martinet, and S. Ystad, “Acoustical correlates of timbre and expressiveness in clarinet performance,” *Music Perception: An Interdisciplinary Journal*, vol. 28, no. 2, pp. 135–154, 2010.

[48] ——, “Analysis-by-synthesis of timbre, timing, and dynamics in expressive clarinet performance,” *Music Perception: An Interdisciplinary Journal*, vol. 28, no. 3, pp. 265–278, 2011.

[49] L. Turchet and M. Barthet, “Ubiquitous musical activities with smart musical instruments,” in *Proc. of the Eighth Workshop on Ubiquitous Music*, 2018.

[50] T. Machover and J. Chung, “Hyperinstruments: Musically intelligent and interactive performance and creativity systems,” in *Proceedings of the International Computer Music Conference*, 1989.

[51] E. Miranda and M. Wanderley, *New digital musical instruments: control and interaction beyond the keyboard*. AR Editions, Inc, 2006, vol. 21.

[52] D. Overholt, E. Berdahl, and R. Hamilton, “Advancements in actuated musical instruments,” *Organised Sound*, vol. 16, no. 02, pp. 154–165, 2011.

[53] A. McPherson, “Buttons, handles, and keys: Advances in continuous-control keyboard instruments,” *Computer Music Journal*, vol. 39, no. 2, pp. 28–46, 2015.

[54] [Online]. Available: https://www.mindmusiclabs.com/ELK

[55] [Online]. Available: https://hyvibe.audio/

[56] L. Turchet, A. McPherson, and M. Barthet, “Co-design of a Smart Cajón,” *Journal of the Audio Engineering Society*, vol. 66, no. 4, pp. 220–230, 2018. [Online]. Available: https://doi.org/10.17743/jaes.2018.0007

[57] A. McPherson, R. Jack, and G. Moro, “Action-sound latency: Are our tools fast enough?” in *Proceedings of the Conference on New Interfaces for Musical Expression*, 2016.

[58] L. Turchet, A. McPherson, and M. Barthet, “Real-time hit classification in a Smart Cajón,” *Frontiers in ICT*, vol. 5, no. 16, 2018.

[59] J. Gregorio and Y. Kim, “Augmentation of acoustic drums using electromagnetic actuation and wireless control,” *Journal of the Audio Engineering Society*, vol. 66, no. 4, pp. 202–210, 2018.

[60] [Online]. Available: https://freesound.org/docs/api/

[61] V. Akkermans, F. Font, J. Funollet, B. de Jong, G. Roma, S. Togias, and X. Serra, “Freesound 2: An improved platform for sharing audio clips,” in *Proceedings of the International Society for Music Information Retrieval Conference*, 2011.

[62] A. McPherson and V. Zappi, “An environment for Submillisecond-Latency audio and sensor processing on BeagleBone black,” in *Audio Engineering Society Convention 138*. Audio Engineering Society, 2015. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=17755

[63] M. Puckette, T. Apel, and D. Ziccarelli, “Real-time audio analysis tools for pd and msp,” in *Proceedings of the International Computer Music Conference*, 1998.

[64] [Online]. Available: www.hexler.net/software/touchosc

[65] T. Mitchell, S. Madgwick, S. Rankine, G. Hilton, A. Freed, and A. Nix, “Making the most of wi-fi: Optimisations for robust wireless live music performance,” in *Proceedings of the Conference on New Interfaces for Musical Expression*, 2014, pp. 251–256.

[66] P. Brinkmann, P. Kirn, R. Lawler, C. McCormick, M. Roth, and H. Steiner, “Embedding pure data with libpd,” in *Proceedings of the Pure Data Convention*, vol. 291, 2011.

[67] R. Bittner, E. Humphrey, and J. Bello, “Pysox: Leveraging the audio signal processing power of sox in python,” in *Proceedings of the International Society for Music Information Retrieval Conference Late Breaking and Demo Papers*, 2016.