TOWARDS THE INTERNET OF MUSICAL THINGS

Luca Turchet
Department of Network
and Systems Engineering
KTH Royal Institute of Technology
turchet@kth.se

Carlo Fischione
Department of Network
and Systems Engineering
KTH Royal Institute of Technology
carlofi@kth.se

Mathieu Barthet
Centre for Digital Music
Queen Mary University of London
m.barthet@qmul.ac.uk

ABSTRACT

In this paper we propose to extend the concept of the Internet of Things to the musical domain leading to a subfield coined as the Internet of Musical Things (IoMUT). IoMUT refers to the network of computing devices embedded in physical objects (Musical Things) dedicated to the production and/or reception of musical content. Musical Things, such as smart musical instruments or smart devices, are connected by an infrastructure that enables multidirectional communication, both locally and remotely. The IoMUT digital ecosystem gathers interoperable devices and services that connect performers and audiences to support performer-performer and audience-performers interactions, not possible beforehand. The paper presents the main concepts of IoMUT and discusses the related implications and challenges.

1. INTRODUCTION

Recent years have seen a substantial increase in smart devices and appliances in the home, office and other environments that connect wirelessly through local networks and the Internet. This is the manifestation of the so-called *Internet of Things* (IoT), an umbrella term encompassing the augmentation of everyday physical objects using information and communication technologies. In the Internet of Things, *Things* refer to embedded systems that are connected to the Internet, which are able to interact with each other and cooperate with their neighbours to reach common goals [1]. The core technology enabling the IoT consists of wireless sensors networks (WSNs) [2]. These are networks of tiny autonomous sensor and actuator nodes that can be embedded in any physical object for control and monitoring via wireless transmission. To date, however, the application of IoT technologies in musical contexts has received little attention compared to other domains such as consumer electronics, healthcare, smart cities, and geospatial analysis [3].

In this position paper we propose to extend the concept of IoT to the musical domain leading to a subfield that we coin as the *Internet of Musical Things* [1] (IoMUT). Through its technological infrastructure, the IoMUT enables an ecosystem of interoperable devices connecting performers and audiences, to support novel performer-performer, audience-performers and audience-audience interactions.

Section 2 examines works and technologies related to the envisioned IoMUT. In Section 3, we argue for the prospect of a holistic integration of these technologies to create the envisioned IoMUT. Sections 4 and 5 discuss the implications and the current major challenges to establish the IoMUT, and Section 6 concludes the paper.

2. RELATED WORK

In this section we review key related works on which our IoMUT vision is founded.

2.1 IoT technologies

Wireless sensors networks (WSNs) design has been the object of much research, both in academia and industry (e.g., [5]). This has resulted in the definition of new communication protocols for WSNs, especially for low data rate and low power consumption, such as IEEE [6], Zigbee [2], ROLL [3]. Very recently, researchers are investigating the integration of WSNs with future wireless cellular networks, the so called 5G networks. The state-of-the art of the activity is the narrow-band IoT (NBIoT) [7].

Unfortunately, most of these protocols are not favourable for the networking of musical instruments. While the most cutting-edge networks (e.g., the fifth generation of cellular wireless networks) will deliver very high data rates, they will also provide communication delays of the order of 25ms [7]. The interconnection of musical instruments poses stringent requirements in terms of end-to-end latency to transmit and receive messages (which will have to be of the order of milliseconds [8,9]), and the reliability or probability of successful message receptions (which will have to be of the oder of 10^{-10} bit error probability). However, such requirements are not only for the application of IoT to the networked music domain, but also for emerging classes of services, such as telepresence, virtual reality, and

[1] The term "Internet of Musical Things" (or "Internet of Music Things") [4] has previously been employed in the context of specific semantic audio applications as part of the EPSRC FAST-IMPACt project (www.semanticaudio.ac.uk), or as challenge to develop wearable instruments in hack sessions. However, to the best of our knowledge, what it entails has not been formalised in the wider context of audience/performer interactions, which is the aim of the initiative described in this work.

[2] www.zigbee.org

[3] www.ietf.org/dyn/wg/charter/roll-charter.html

mission-critical control. This is pushing the development of the emerging paradigm of "Tactile Internet" [10] and millimeter Waves (mmWaves) communications [11, 12].

Tactile Internet refers to an Internet network where the communication delay between a transmitter and a receiver would be so low that even information associated to human touch, vision, and audition could be transmitted back and forth in real-time, making remote interaction experiences, virtual or human, effective, high-quality and realistic. This novel area focuses on the problem of designing communication networks, both wireless and wired, capable to ensure ultra-low latency communications, with end-to-end delays of the oder or few milliseconds. The technological vision of Tactile Internet is that the round-trip time to send information from a source to a destination and back experiences latencies below 5ms. An important aspect of the Tactile Internet vision is the reliability and low latency for wireless communications. One of the emerging technology for the wireless communication is mmWaves. Such a technology uses wireless frequencies within the range of 10 to 300 GHz and offers data rates of giga bits per seconds over short distances. These frequencies make possible the design of small antennas that can be easily embedded in musical instruments (as proposed in the Smart Instruments concept [13]). Moreover, the high data rates will enable the transmission of multimodal content in high resolution.

2.2 Digital ecosystems

Digital ecosystems are the result of recent developments of digital network infrastructure inheriting from principles of ecological systems [14]. They are collaborative environments where species/agents form a coalition to reach specific goals. Value is created by making connections through collective ("swarm") intelligence and by promoting collaboration. The concepts underlying digital ecosystems match well the proposed goals of the Internet of Musical Things where multiple actors, smart devices and intelligent services interact to enrich musical experiences.

2.3 Networked music performance systems

Networked music performance (NMP) systems were proposed to enable collaborative music creation over a computer network and have been the object of scientific and artistic investigations [9, 15–17]. A notable example is the ReacTable [18], a tangible interface consisting of a table capable of tracking objects that are moved on its surface to control the sonic output. The ReacTable allows multiple performers to simultaneously interact with objects either placed on the same table or on several networked tables in geographically remote locations.

Networked collaborative music creations can occur over a Wide Area Network (WAN), or a Local Area Network (LAN) and in particular over a wireless one (WLAN) [17], and different methods have been proposed for each of these configurations. In [9], the authors provide a comprehensive overview of hardware and software technologies enabling NMP, including low-latency codecs, frameworks, protocols, as well as perceptually relevant aspects.

2.4 Participatory live music performance systems

Within interactive arts, participatory live music performance (PLMP) systems capitalising on information and communication technologies have emerged to actively engage audiences in the music creation process [19]. These systems disrupt the traditional unidirectional chain of musical communication from performers to audience members who are "passive" from the creative point of view (see e.g., [19–22]). Interaction techniques for technology-mediated audience participation have been proposed exploiting a wide range of media and sensors, from mobile devices [19–21, 23, 24] to tangible interfaces [25] such as light sticks [26] (see [19] for a review and classification framework).

Most PLMP systems require the audience to use a single type of device and application. Nevertheless, different types of devices could be exploited simultaneously to enrich interaction possibilities. To date, audience creative participation has mainly been based on manual controls or gestures using smartphones (e.g., screen touch, tilt). Expressive modalities could be increased by tracking physiological parameters (e.g., electrodermal activity, heart rate) [27] [28] at the individual and collective levels using devices specifically designed for this purpose, or by tracking more complex audience behaviors and body gestures. Furthermore, means of interaction in current PLMP systems typically rely on the auditory or visual modalities, while the sense of touch has scarcely been explored to create more engaging musical experiences.

2.5 Smart Instruments

Recently, a new class of musical instruments has been proposed, the *Smart Instruments* [13]. In addition to sensor and actuator enhancements provided in the so-called *augmented instruments* [29, 30], Smart Instruments are characterised by embedded computational intelligence, a sound processing and synthesis engine, bidirectional wireless connectivity, an embedded sound delivery system, and a system for feedback to the player. Smart Instruments bring together separate strands of augmented instruments, networked music and Internet of Things technology, offering direct point-to-point communication between each other and other portable sensor-enabled devices, without need for a central mediator such as a laptop. Interoperability is a key feature of Smart Instruments, which are capable of directly exchanging musically relevant information with one another and communicating with a diverse network of external devices, including wearable technology, mobile phones, virtual reality headsets and large-scale concert hall audio and lighting systems.

The company MIND Music Labs [4] has recently developed the *Sensus Smart Guitar* [13, 31] which, to the best of our knowledge, is the first musical instrument to encompass all of the above features of a Smart Instrument. Such an instrument is based on a conventional electroacoustic guitar that is augmented with IoT technologies. It involves several sensors embedded in various parts of the instrument, which allow for the tracking of a variety of

[4] `www.mindmusiclabs.com`

gestures of the performer. These are used to modulate the instrument's sound thanks to an embedded platform for digital audio effects. Acoustical sounds are produced by the instrument itself by means of an actuation system that transforms the instrument's resonating wooden body into a loudspeaker. Furthermore, the *Sensus Smart Guitar* is equipped with bidirectional wireless connectivity, which makes possible the transmission and reception of different types of data from the instrument to a variety of smart devices and vice versa.

2.6 Smart wearables

The last decade has witnessed a substantial increase in the prevalence of wearable sensor systems, including electronic wristbands, watches and small sensor tokens clipped to a belt or held in a pocket. Many of such devices (e.g., Fitbit [5]) target the personal health and fitness sectors. They typically include inertial measurement units (IMUs) for capturing body movement and sensors for physiological data (e.g., body temperature, galvanic skin response, heart rate). Such devices, here referred to as *smart wearables*, include wireless communication options to link to mobile phones or computers. In some cases, a small display, speaker or tactile actuator may be included. A distinguishing characteristic of wearable devices is their unobtrusiveness: they are designed to be worn during everyday activity and to passively collect data without regular intervention by the user. Such features make these devices suitable to track and collect body movements and physiological responses of audience members during live concerts. However, to date, this challenge has been scarcely addressed.

Moreover, to date, the use of the wearable devices exploiting the tactile channel in musical applications has been rather limited. Noticeable exceptions are *Rhytm'n'shoes*, a wearable shoe-based audio-tactile interface equipped with bidirectional wireless transmission [32], and *Mood Glove*, a glove designed to amplify the emotions expressed by music in film through haptic sensations [33].

2.7 Virtual reality, augmented reality, and 360° videos

The last two decades have seen an increase of both academic and industrial research in the fields of virtual reality (VR) and augmented reality (AR) for musical applications. Several virtual musical instruments have been developed (for a recent review see [34]), while musical instruments augmented with sensors, such as the Sensus Smart Guitar, have been used to interactively control virtual reality scenarios displayed on head-mounted-displays (HMD) [31].

AR has been used to enhance performance stages for augmented concert experiences, as well as for participatory performances applications. Mazzanti et al. proposed the *augmented stage* [35] an interactive space for both performers and audience members, where AR techniques are used to superimpose a performance stage with a virtual environment, populated with interactive elements. Spectators contribute to the visual and sonic outcome of the performance by manipulating virtual objects via their mobile phones. Berthaut et al. proposed *Reflets*, a mixed-reality environment that allows one to display virtual content on stage, such as 3D virtual musical interfaces or visual augmentations of instruments and performers [36]. Poupyrev et al. proposed the *augmented groove*, a musical interface for collaborative jamming where AR, 3D interfaces, as well as physical, tangible interaction are used for conducting multimedia musical performance [37].

Immersive virtual environments have been proposed as a means to provide new forms of musical interactions. For instance, Berthaut et al. proposed the *3D reactive widgets*, graphical elements that enable efficient and simultaneous control and visualisation of musical processes, along with *Piivert*, an input device developed to manipulate such widgets, and several techniques for 3D musical interaction [38, 39].

The growing availability of 360° videos has recently opened new opportunities for the entertainment industry, so that musical content that can be delivered through VR devices offering experiences unprecedented in terms of immersion and presence. Recent examples include Orchestra VR, a 360° 3D performance featuring the opening of Beethoven's Fifth Symphony performed by the Los Angeles Philharmonic Orchestra, accessible via an app for various VR headsets [6], Paul McCartney's 360 cinematic concert experience app allowing the experience of recorded concerts with 360° video and 3D audio using Google's Cardboard HMD, Los Angeles Radio Station KCRW, which launched a VR App for "intimate and immersive musical performances" [7] and FOVE's Eye Play The Piano project, which allows disabled children to play a real acoustic piano using eye tracking technologies embedded in HMDs [8].

3. THE INTERNET OF MUSICAL THINGS

The proposed Internet of Musical Things (IoMUT) relates to the network of physical objects (*Musical Things*) dedicated to the production, interaction with or experience of musical content. Musical Things embed electronics, sensors, data forwarding and processing software, and network connectivity enabling the collection and exchange of data for musical purpose. A Musical Thing can take the form of a Smart Instrument, a Smart Wearable, or any other smart device utilised to control, generate, or track responses to music content. For instance, a Smart Wearable can track simple movements, complex gestures, as well as physiological parameters, but can also provide feedback leveraging the senses of audition, touch, and vision.

The IoMUT arises from (but is not limited to) the holistic integration of the current and future technologies mentioned in Section 2. The IoMUT is based on a technological infrastructure that supports multidirectional wireless communication between Musical Things, both locally and remotely. Within the IoMUT, different types of devices for performers and audience are exploited simultaneously to enrich interaction possibilities. This multiplies affordances and ways to track performers' and audience mem-

[5] www.fitbit.com

[6] www.laphil.com/orchestravr

[7] www.kcrw.com/vr

[8] www.eyeplaythepiano.com/en

bers' creative controls or responses. The technological infrastructure of the IoMUT consists of hardware and software (such as sensors, actuators, devices, networks, protocols, APIs, platforms, clouds, services), but differently from the IoT, these are specific to the musical case. In particular, for the most typical case of real-time applications, the infrastructure ensures communications with low latency, high reliability, high quality, and synchronization between connected devices.

Such an infrastructure enables an ecosystem of interoperable devices connecting performers with each other, as well as with audiences. Figure 1 shows a conceptual diagram of the different components that are interconnected in our vision of the IoMUT ecosystem. As it can be seen in the diagram, the interactions between the human actors (performers and audience members) are mediated by Musical Things. Such interactions can be both co-located (see blue arrows), when the human actors are in the same physical space (e.g., concert hall, public space), or remote, when they take place in different physical spaces that are connected by a network (see black arrows).

Regarding co-located interactions, these can be based on point to point communications between a Musical Thing in possession of the performer and a Musical Thing in possession of a audience member (see the blue dashed arrows), but also between one or more Musical Things of the performers towards one or more Musical Things for the audience as a whole, and vice versa (see the blue solid line arrow). An example of the latter case could be that of one or more Smart Instruments affecting the lighting system of a concert all. Regarding remote interactions, these can occur not only between audience members/performers present at the concert venue and remote audience members/performers (see the solid black arrows), but also between remote audience members/performers (see the black dashed arrows).

The communication between Musical Things is achieved through APIs (application programming interfaces, indicated in Figure 1 with the small red rectangles), which we propose could be based on a unified API specification (the IoMUT API specification). The interactions mentioned above, based on the exchange of multimodal creative content, are made possible thanks to Services (indicated with the green areas). For instance, these can be services for creative content analysis (such as multi-sensor data fusion [40], music information retrieval [41]), services for creative content mapping (between analysis and devices), or services for creative content synchronization (between devices). In particular, the implementation of novel forms of interactions that leverage different sensory modalities makes the definition of *Multimodal Mapping Strategies* necessary. These strategies consist of the process of transforming, in real-time, the sensed data into control data for perceptual feedback (haptic, auditory, visual).

4. IMPLICATIONS

Thanks to the IoMUT it is possible to reimagine the live music performance art and music teaching by providing a technological ecosystem that multiplies possibilities of interaction between audiences, performers, students, teachers, as well as their instruments and machines. This has the potential to revolutionise the way to experience, compose, and learn music, as well as even record it by adding other modalities to audio. In particular, IoMUT has the potential to make NMP and PLMP more engaging and more expressive, because it uses a radically novel approach to address the fundamental obstacles of state-of-the-art methods, which hinder efficient, meaningful and expressive interactions between performers and between performers and audiences.

The IoMUT ecosystem can support new performers-performers and audience-performers interactions, not possible beforehand. Examples of use cases include novel forms of: jamming (e.g., using apps running on smartphones to control the sound engine of a smart instrument); enhanced concert experiences (e.g., audience members in possession of haptic feedback smart wearables "feel" the vibrato of a smart violin or the rhythm of a smart drum; the emotional response of audience is used to control the timbre of Smart Instruments or the behavior of stage equipment such as projections, smoke machines, lights); remote rehearsals (point-to-point audio streaming between Smart Instruments).

To date, no human computer interaction system for musical applications enables the many interaction pathways and mapping strategies we envision in the IoMUT: one-to-one, one-to-many, many-to-one, many-to-many, in both co-located and remote situations. In contrast to traditional acoustic instruments, the IoMUT framework allows to establish "composed electronic instruments" where the control interface and the process of sound production is decoupled [42]. New situations of "performative agency" [42] can be envisioned by letting audience members and the intelligence derived within the IoMUT digital ecosystem influence the outcome of specific musical performances.

By applying the IoT to music we envision to go from the traditional musical chain (i.e., composers writing musical content for performers, who deliver it to a unique and "creatively passive" audience) to a musical mesh where possibilities of interactions are countless. We envision both common (co-located participatory music performance in a concert hall) and extreme scenarios (massive open online music performance gathering thousands or hundreds of thousands of participants in a virtual environment).

Combining such IoMUT musical mesh model with VR/AR applications it is possible to enable new forms of music learning, co-creation, and immersive and augmented concert experiences. For instance, a violin player could see through AR head-mounted-displays semantic and visual information about what another performer is playing, or be able to follow the score without having to look at a music stand. An audience member could virtually experience to walk on stage or feel in the "skin" of a performer. A whole audience could affect the lighting effects on stage based on physiological responses sensed with wireless smart wristbands. Such smart wristbands could also be used to understand audiences' affective responses. An audience could engage in the music creation process at specific times in a performance as prepared in a composer's score. A concert-

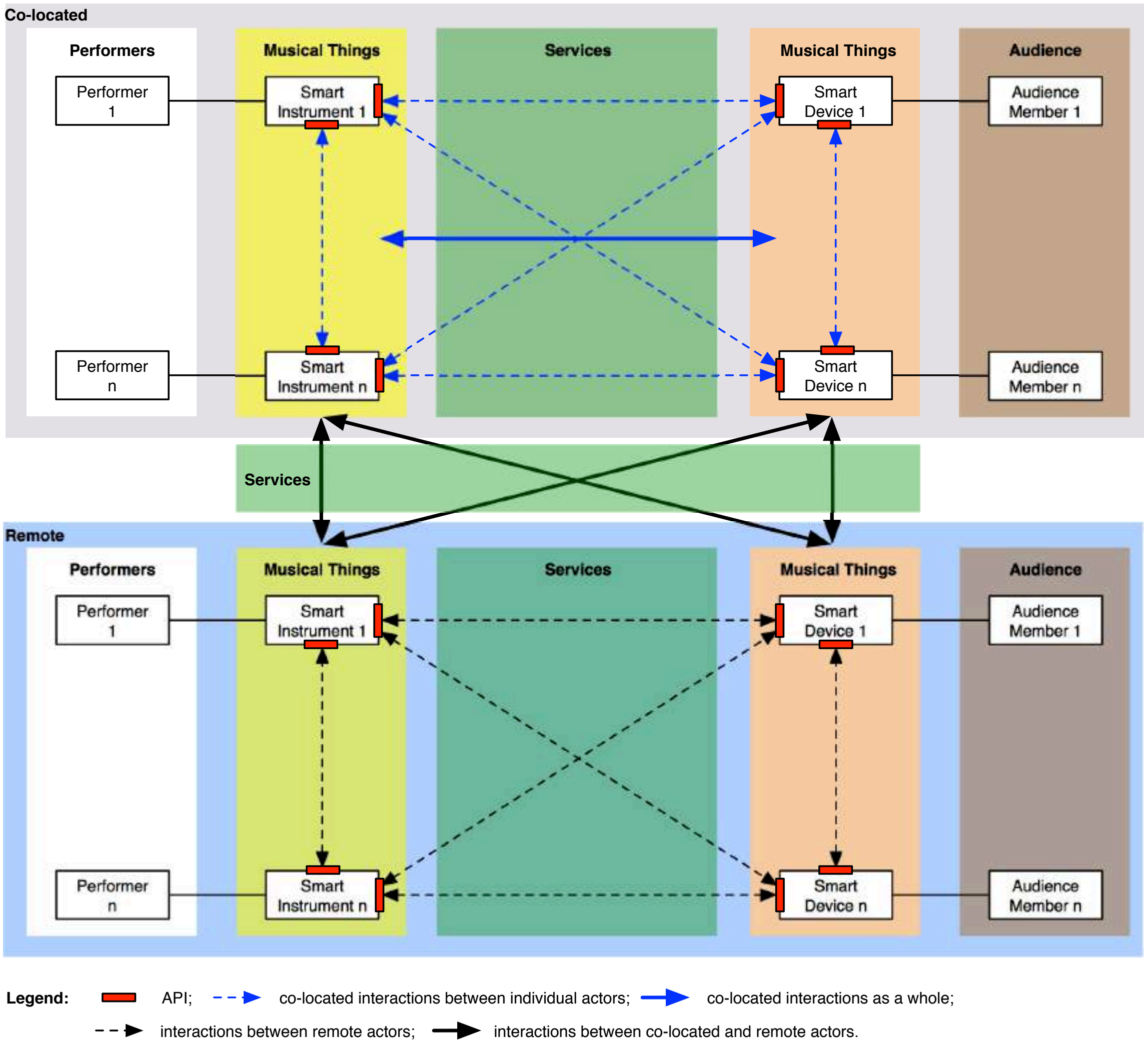

Figure 1. Block diagram of the IoMUT ecosystem.

goer could have access to "augmented programme notes" guiding and preparing them prior to the concert experience by learning more about historical and compositional aspects and listening to different renderings interactively, and letting them see the evolution in the score as the music is being played or additional information about the soloist during the concert.

In a different vein, the IoMUT has the potential to generate new business models that could exploit the information collected by Musical Things in many ways. For instance, such information could be used to understand customer behaviour, to deliver specific services, to improve products and concert experiences, and to identify and intercept the so-called "business moments" (defined by Gartner Inc. [9]).

5. CURRENT CHALLENGES

The envisioned IoMUT poses both technological and artistic or pedagogical challenges. Regarding the technological challenges, it is necessary that the connectivity features of the envisioned Musical Things go far beyond the state-of-the-art technologies available today for the music domain.

[9] www.gartner.com/newsroom/id/2602820

Between a sensor that acquires a measurement of a specific auditory, physiological, or gestural phenomenon, and the receiver that reacts to that reading over a network, there is a chain of networking and information processing components, which must be appropriately addressed in order to enable acceptable musical interactions over the network. Nevertheless, current NMP systems suffer from transmission issues of latency, jitter, synchronization, and audio quality: these hinder real-time interactions that are essential to collaborative music creation [9]. It is also important to notice that for optimal NMP experiences, several aspects of musical interactions must be taken into account beside the efficient transmission of audio content. Indeed, during co-located musical interactions musicians rely on several modalities in addition to the sounds generated by their instruments, which include for instance the visual feedback from gestures of other performers, related tactile sensations, or the sound reverberation of the space [43]. However, providing realistic performance conditions over a network represents a significant engineering challenge due to the extremely strict requirements in terms of network latency and multimodal content quality, which are needed to achieve a high-quality interaction experience [44].

The most important specific technological challenges of IoMUT against the more general ones from the IoT are the requirements of very short latency, high reliability, high audio/multimodal content quality, and synchronization to be ensured for musical communication. This implies the creation of a technological infrastructure that is capable of transmitting multimodal content, and in particular audio, from one musician to another/others not only in hi-fi quality, but also with a negligible amount of latency, which enable performers to play in synchronous ways. Current IoT scientific methods and technologies do not satisfy these tight constraints needed for the real-time transmission of audio/multimodal content both at short and at large distances [9, 17]. The envisioned Tactile Internet [10] is expected to solve at least in part such issues.

The establishing of the Tactile Internet vision will require, however, the redesign of networking protocols, from the physical layer to the transport layer. For example, at the physical layer, messages will have to be inevitably short to ensure the desired low latencies, and this will pose restrictions to the data rates. At the routing layer, the protocols will have to be optimized for low delays rather than for high throughput of information. Possible avenues for future research include: the identification of all the components of a high-performance network infrastructure that is physically capable to deliver low-latency across the radio as well as in the core network segments; the proposition of new distributed routing decision methods capable to minimize the delay by design; the investigation of new dynamic control and management techniques based on optimization theory capable to configure and allocate the proper data plane resources across different domains to build low-latency end-to-end services.

The IoMUT digital ecosystems can benefit from ongoing work in the semantic web. Metadata related to multimodal creative content can be represented using vocabulary defined in ontologies with associated properties. Such ontologies enable the retrieval of linked data within the ecosystem driven by the the needs of specific creative music services (e.g., a service providing vibrato of notes played by a guitar player to enact events in the Musical Thing of an audience member). Alignment and mapping/translation techniques can be developed to enable "semantic information integration" [14] within the IoMUT ecosystem.

An important aspect of the IoMUT regards the interconnection of different types of devices. Such devices target performers or audiences (both co-located and remote), and are used to generate, track, and/or interpret multimodal musical content. This poses several other technological challenges. These include the need for ad-hoc protocols and interchange formats for musically relevant information that have to be common to the different Musical Things, as well as the definition of common APIs specifically designed for IoMUT applications.

Wearable systems present many opportunities for novel forms of musical interaction, especially involving multiple sensory modalities. Related design challenges concern the optimization for musical qualities of the sensor and actuators capabilities of these devices (e.g., temporal precision, low latency, synchronicity of audio, visual, and tactile modalities). Another related challenge is how to effectively use multiple sensory modalities in PLMP systems. In particular the haptic one leveraged by Smart Wearables could have a high impact potential on the musical experience of audience.

A major challenge to the approach of *Multimodal Mapping Strategies*, consists of how to determine mappings flexible enough to allow for musical participation expressive and meaningful to both experts and novices. These mappings could be based on features extracted in real-time from sensors data and musical audio analysis. *Multi-sensor data fusion techniques* [40] could be exploited for this purpose, which explicitly account for the diversity in acquired data, (e.g., in relation to sampling rates, dimensionality, range, and origin).

Moreover, the IoMUT demands new analytic approaches: new analytic tools and algorithms are needed to process large amounts of music-related data in order to retrieve information useful to understand and exploit user behaviours.

Regarding the non-technological challenges posed by the IoMUT, from the artistic perspective, previous attempts to integrate audiences into performances have not completely released the barriers inhibiting musical interactivity between performers and audiences. For a performance to be truly interactive, each member of the audience should be as individually empowered as the performers on stage. To achieve this, it is necessary to reimagine musical performances and invent new compositional paradigms that can catalyse, encompass and incorporate multiple and diverse contributions into a coherent and engaging whole. This poses several challenges: how can we compose music that gives individual freedom to several (potentially hundreds or thousands) of performers so that they feel empowered without compromising the quality of the performance? How do we manage all of these inputs from individuals who have different backgrounds, sensibilities and skills and leverage them so that the result is satisfying for all of them and in which each person can still recognize his or her own individual contributions? How do musicians compose and rehearse for a concert where they do not fully control the end result? In short, how do we use technology to integrate each individual expression into an evolving whole that binds people together?

A framework such as the IoMUT and what it entails for artistic and pedagogical agendas will require to be assessed. This could pave the way for novel research on audience reception, interactive arts, education and aesthetics. Such research would for instance help to reflect on the roles of constraints, agency and identities in the participatory arts. Finally, issues related to security and privacy of information should also be addressed, especially if such system was to be deployed for the masses.

6. CONCLUSIONS

IoMUT relates to wireless networks of *Musical Things* that allow for interconnection of and interaction between performers, audiences, and their smart devices. Using IoMUT we proposed a model transforming the traditional linear composer-performer-audience musical chain into a musical mesh interconnecting co-located and/or remote performers and audiences. Many opportunities are enabled by such a model that multiplies possibilities of interaction and communication between audiences, performers, their instruments and machines. On the other hand, the IoMUT poses both technological and non-technological challenges that we expect will be faced in upcoming years by both academic and industrial research.

Acknowledgments

The work of Luca Turchet is supported by MSCA Individual fellowship of European Union's Horizon 2020 research and innovation programme, under grant agreement No 749561. The work of Carlo Fischione is supported by the TNG SRA ICT project "TouCHES" - Tactile Cyberphysical Networks, funded by the Swedish Research Council. Mathieu Barthet acknowledges support from the EU H2020 Audio Commons grant (688382).

7. REFERENCES

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

[2] W. Dargie and C. Poellabauer, *Fundamentals of wireless sensor networks: theory and practice*. John Wiley & Sons, 2010.

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

[4] A. Hazzard, S. Benford, A. Chamberlain, C. Greenhalgh, and H. Kwon, "Musical intersections across the digital and physical," in *Digital Music Research Network Abstracts (DMRN+9)*, 2014.

[5] A. Willig, "Recent and emerging topics in wireless industrial communication," *IEEE Transactions on Industrial Informatics*, vol. 4, no. 2, pp. 102–124, 2008.

[6] *IEEE Std 802.15.4-2996, September, Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs)*, IEEE, 2006.

[7] S. Landström, J. Bergström, E. Westerberg, and D. Hammarwall, "Nb-iot: a sustainable technology for connecting billions of devices," *Ericsson Technology Review*, vol. 93, no. 3, pp. 1–12, 2016.

[8] C. Rottondi, M. Buccoli, M. Zanoni, D. Garao, G. Verticale, and A. Sarti, "Feature-based analysis of the effects of packet delay on networked musical interactions," *Journal of the Audio Engineering Society*, vol. 63, no. 11, pp. 864–875, 2015.

[9] C. Rottondi, C. Chafe, C. Allocchio, and A. Sarti, "An overview on networked music performance technologies," *IEEE Access*, 2016.

[10] G. P. Fettweis, "The Tactile Internet: applications and challenges," *IEEE Vehicular Technology Magazine*, vol. 9, no. 1, pp. 64–70, 2014.

[11] H. Shokri-Ghadikolaei, C. Fischione, G. Fodor, P. Popovski, and M. Zorzi, "Millimeter wave cellular networks: A MAC layer perspective," *IEEE Transactions on Communications*, vol. 63, no. 10, pp. 3437–3458, 2015.

[12] H. Shokri-Ghadikolaei, C. Fischione, P. Popovski, and M. Zorzi, "Design aspects of short-range millimeter-wave networks: A MAC layer perspective," *IEEE Network*, vol. 30, no. 3, pp. 88–96, 2016.

[13] L. Turchet, A. McPherson, and C. Fischione, "Smart Instruments: Towards an Ecosystem of Interoperable Devices Connecting Performers and Audiences," in *Proceedings of the Sound and Music Computing Conference*, 2016, pp. 498–505.

[14] H. Boley and E. Chang, "Digital ecosystems: Principles and semantics," in *Inaugural IEEE International Conference on Digital Ecosystems and Technologies*, 2007.

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

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

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

[18] S. Jordà, G. Geiger, M. Alonso, and M. Kaltenbrunner, "The reactable: exploring the synergy between live music performance and tabletop tangible interfaces," in *Proceedings of the 1st international conference on Tangible and embedded interaction*, 2007, pp. 139–146.

[19] Y. Wu, L. Zhang, N. Bryan-Kinns, and M. Barthet, "Open symphony: Creative participation for audiences of live music performances," *IEEE Multimedia Magazine*, no. 48-62, 2017.

[20] G. Fazekas, M. Barthet, and M. B. Sandler, *Novel Methods in Facilitating Audience and Performer Interaction Using the Mood Conductor Framework*, ser. Lecture Notes in Computer Science. Springer-Verlag, 2013, vol. 8905, pp. 122–147.

[21] 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 international conference on New Interfaces for Musical Expression*, 2016, pp. 170–175.

[22] 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 International Conference on Live Interfaces*, 2016.

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

[24] N. Weitzner, J. Freeman, S. Garrett, and Y.-L. Chen, "massmobile-an audience participation framework." in *Proceedings of the international conference on New Interfaces for Musical Expression*, 2012, pp. 21–23.

[25] B. Bengler and N. Bryan-Kinns, "Designing collaborative musical experiences for broad audiences," in *Proceedings of the 9th ACM Conference on Creativity & Cognition*. ACM, 2013, pp. 234–242.

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

[27] D. J. Baker and D. Müllensiefen, "Hearing wagner: Physiological responses to richard wagner's der ring des nibelungen," in *Proc. Int. Conference on Music Perception and Cognition*, 2014.

[28] A. Tanaka, "Musical performance practice on sensor-based instruments," *Trends in Gestural Control of Music*, vol. 13, pp. 389–405, 2000.

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

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

[31] L. Turchet, M. Benincaso, and C. Fischione, "Examples of use cases with smart instruments," in *Proceedings of AudioMostly Conference*, 2017 (In press).

[32] S. Papetti, M. Civolani, and F. Fontana, "Rhythm'n'shoes: a wearable foot tapping interface with audio-tactile feedback," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2011, pp. 473–476.

[33] A. Mazzoni and N. Bryan-Kinns, "Mood glove: A haptic wearable prototype system to enhance mood music in film," *Entertainment Computing*, vol. 17, pp. 9–17, 2016.

[34] S. Serafin, C. Erkut, J. Kojs, N. Nilsson, and R. Nordahl, "Virtual reality musical instruments: State of the art, design principles, and future directions," *Computer Music Journal*, vol. 40, no. 3, pp. 22–40, 2016.

[35] D. Mazzanti, V. Zappi, D. Caldwell, and A. Brogni, "Augmented stage for participatory performances." in *Proceedings of the international conference on New Interfaces for Musical Expression*, 2014, pp. 29–34.

[36] F. Berthaut, D. Plasencia, M. Hachet, and S. Subramanian, "Reflets: Combining and revealing spaces for musical performances," in *Proceedings of the international conference on New Interfaces for Musical Expression*, 2015.

[37] I. Poupyrev, R. Berry, J. Kurumisawa, K. Nakao, M. Billinghurst, C. Airola, H. Kato, T. Yonezawa, and L. Baldwin, "Augmented groove: Collaborative jamming in augmented reality," in *ACM SIGGRAPH 2000 Conference Abstracts and Applications*, 2000, p. 77.

[38] D. Berthaut, M. Desainte-Catherine, and M. Hachet, "Interacting with 3d reactive widgets for musical performance," *Journal of New Music Research*, vol. 40, no. 3, pp. 253–263, 2011.

[39] F. Berthaut and M. Hachet, "Spatial interfaces and interactive 3d environments for immersive musical performances," *IEEE Computer Graphics and Applications*, vol. 36, no. 5, pp. 82–87, 2016.

[40] D. Hall and J. Llinas, "Multisensor data fusion," in *Multisensor Data Fusion*. CRC press, 2001.

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

[42] O. Bown, A. Eldridge, and J. McCormack, "Understanding interaction in contemporary digital music: from instruments to behavioural objects," *Organised Sound*, vol. 14, no. 2, pp. 188–196, 2009.

[43] W. Woszczyk, J. Cooperstock, J. Roston, and W. Martens, "Shake, rattle, and roll: Getting immersed in multisensory, interactive music via broadband networks," *Journal of the Audio Engineering Society*, vol. 53, no. 4, pp. 336–344, 2005.

[44] M. Slater, "Grand challenges in virtual environments," *Frontiers in Robotics and AI*, vol. 1, p. 3, 2014.