Interactive IoMusT-Based Concerts: Real-Time Pattern Recognition and Audience Experience

1st Nishal Silva
University of Trento
Trento, Italy
nishal.silva@unitn.it

2nd Alberto Boem
University of Trento
Trento, Italy
alberto.boem@unitn.it

3rd Luca Turchet
University of Trento
Trento, Italy
luca.turchet@unitn.it

***Abstract*—This paper explores the possibilities offered by the Internet of Musical Things paradigm to create new interactive performance experiences. We introduce a novel performance ecosystem where performers utilize smart musical instruments (SMIs) equipped with embedded Real-time Pattern Detection (RTPD) algorithms to trigger wirelessly connected devices. To validate our approach, the system was ecologically validated with two concerts, which involved two groups of performers and two groups of audience members. Each concert featured three SMIs (electric guitar, keyboard, and electronic drumkit) which were used to control three stage lights, one smoke machine, 3D graphics and haptic vibrations on Mixed Reality headsets and smartphones worn ad held by the audience members. We compared the use of RTPD-triggered effects against randomized triggering, and we assessed the experience of a composer, performers and audiences in interacting with the ecosystem using qualitative and quantitative assessments. Audience members preferred the condition in which the RTPD system directly triggered effects rather than random triggering. Overall, composers, performers and audience valued this new art format. The results indicate that combining smart musical instruments with pattern detection in an Internet of Musical Things ecosystem can open new avenues for artistic performance and audience engagement.**

***Index Terms*—Internet of Musical Things, Smart Musical Instruments, Real-time Pattern Detection, Audience Engagement.**

I. Introduction

The emergence of digital technology has brought about significant transformations in various domains, the realm of music being no exception. The Internet of Things [1] - network of physical devices embedded with sensors, software, and connectivity that enables them to collect and exchange data over the internet - has impacted the music domain, spurring the emergence of the realm of the Internet of Musical Things (IoMusT) [2]. The IoMusT refers to a network of musical instruments and other connected peripheral devices (i.e., stage lights, smoke machines, haptic devices, Mixed Reality (MR) headsets) - aptly named Musical Things, that exchange data over a wired or wireless network. Smart Musical Instruments (SMIs) play an integral role in the IoMusT paradigm [3]. These digital musical instruments possess computational capabilities, such Real-time Pattern Detection (RTPD), facilitated by embedded computing systems designed for real-time audio processing and networking tasks.

One of the central tenets of the IoMusT paradigm is enabling interconnections between Musical Things and multidirectional interactions between their users. This spurs the emergence of new performance ecosystems of humans and machines [4], [5], [6]. In recent years, various systems have been proposed to interconnect heterogeneous musical devices over co-located or remote networks [7], [8], [9], [10], [11], [12], [13]. The interconnection capabilities offered by the IoMusT paradigm have opened up a whole new realm of creative possibilities for composers and performers, allowing them to create connected, interactive and multi-sensory performances [14], [15], [16], [17], [18], [19]. However, to date only a handful of studies are available on such a topic and the potential of performances and live musical events enabled by the IoMusT is still largely unexplored. In particular, the capability of IoMusT technologies to support multi-sensory experiences for audiences has been scarcely addressed thus far.

To bridge this gap, we explore the creative possibilities of incorporating Musical Things into performances, enabled by the IoMusT paradigm by focusing on a class of SMIs that feature RTPD capabilities. These instruments can detect specific musical phrases, which have been previously defined by the musicians, as soon as they are played on the instrument. Such detections are then used to trigger external devices in real-time via wireless control messages, creating an interactive musical experience for the audience. To the best of the authors' knowledge, this is the first study investigating the integration of RTPD methods into a performance ecosystem involving Musical Things, controlled by SMIs, for the audience.

Let us consider a guitarist performing on a smart electric guitar with RTPD capabilities. The embedded computing device on the instrument is configured to detect a set of predefined musical patterns, and output commands to external devices. The SMI will analyze the music played on the instrument continuously, and if the musician plays one of the patterns, it will be detected immediately upon completion, and the command associated with that pattern will be sent wirelessly to trigger external devices such as lighting systems, smoke machines, visual projections, haptic devices or even other musical instruments.

To implement such an ecosystem there is the need of creating SMIs equipped with RTPD methods, as well as a mapping system that facilitates the composition of the unidirectional or multidirectional interactions between the patterns of each musical instrument and the parameters of the devices used by the audience members. First, we introduce such SMIs system which leverage pattern detection algorithms based on real-time embedded machine learning methods and ad-hoc created datasets. Second, we describe a mapping application that enables composers to interconnect musical instruments and multiple output devices. Finally, we present a validation of the proposed system through two live performances, focusing on the experience of performers and audience members. A video of the developed ecosystem utilized during the concert-experiments is available online[1].

II. Related works

Technologically-mediated multi-sensory music performances represent a rapidly evolving area of research and artistic exploration, with the goal of creating more immersive and engaging experiences for the audience and performers as well [19], [20]. These performances are not limited to traditional auditory elements, but also incorporate additional senses such as vision, touch and smell, aiming to enhance the emotional engagement of the audience. The majority of research in this space has focused on how to enhance the participation and immersion of audience in VR settings [21], especially involving audio and visual or audio and tactile modalities (e.g., [22], [23], [24], [25]).

The interplay between different senses in Augmented Reality (AR) performances was explored in [26]. The authors presented two systems designed to enhance music creation and playback through AR interfaces in 3D environments. The systems allow users to engage in collaborative music-making experiences that incorporate visual, auditory, and tactile modalities. In one system, users are able to interact with virtual bubbles containing musical segments which can be captured and arranged on a playback timeline. The other system allows users to manipulate virtual musical instruments through hand movements and speech commands. The authors also presented a user study which showed the systems' ability to improve user interactions significantly, compared to traditional interfaces.

Lindström et. al [27] evaluated three public multi-sensory performances that mapped expressive gestures and voice to real-time audiovisual outputs. Through questionnaires across different audience engagement levels (performers, active participants, and passive spectators), they found that direct participation and understanding of the technology significantly enhanced audience appreciation. Their findings suggest that successful performances require both robust technical implementation and clear narrative elements, with active engagement leading to better reception than passive viewing.

Wearable sensors and actuators have also been explored for both capturing data from performers and delivering them to the audience. RaveNet [28] presents an approach to interactive musical performances through the use of wearable technology which enables musicians to modulate sound through their breath, pulse, movement, and facial expressions. The authors present three prototypes of wearable devices that can capture the various stimuli of the musicians.

A study by [18] showed the potential of integrating haptic wearable devices into live music experiences. Through a series of user studies with electronic music performers and audiences equipped with haptic wearable devices, authors revealed that the haptic sensations increased arousal, valence, enjoyment, and engagement for some participants.

Another strand of research have explored the use of SMIs in the context of multi-sensory music performances. One example is a prototype of a system that uses a smart guitar for controlling virtual musical experiences with Head Mounted Displays (HMDs) worn audience members [29]. Through some dedicated sensors placed on the guitar, a player can control the temporal evolution of virtual objects provided concurrently with the music heard.

However, the integration between SMIs, stage technologies, and real-time audience participation systems remains an under-explored area, particularly in terms of how these technologies can work together to enhance both performer and audience experiences.

III. Methodology

A. Experiment Design

To explore the relationship between music and different sensory elements (vision and haptic), we devised an experiment in medium sized room equipped with speakers, stage lights, and musical instruments that can resemble a music club. We conducted the experiment in a ecologically valid setting to enhance the generalizability of our findings to real-world situations and to allow for authentic and realistic user interactions that are often absent in controlled laboratory environments [30], [31]. The experiment compared two music performances held over two days. Each performance maintained identical structure but featured different performers and audiences to enable comparative analysis. The two performances featured a guitarist, keyboardist, and drummer, who were playing SMIs with RTPD capabilities.

The perception of the performance and the RTPD, both from the performers and audience members were captured through questionnaires and semi-structured interviews. The questionnaires were created to evaluate the enjoyment, distraction, and coherence felt by the audience members. The performers were also given questionnaires to assess their opinion of the RTPD system, and semi structured interviews were conducted to identify potential ways they would use the system.

1) Musical Repertoire and Performance Structure: The performance featured two pieces created to complement the technical setup. Several requirements had to be met when composing the pieces to ensure that they were suited to evaluate the performances and RTPDs:

- The compositions must be simple and easy to understand. As the performance includes multiple types of sensory stimuli, careful attention must be paid to avoid subjecting the audience to a *cognitive overload* [32];

[1] https://youtu.be/axEHkdnxFB8

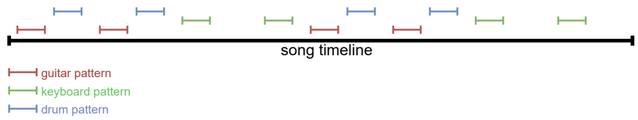

Fig. 1. How each instrument's patterns can be spread within the timeline of the piece.

- The compositions must have easily recognizable patterns to evaluate the pattern detection system;
- The composition must allow all performers to contribute equally with their instrument. All solos, improvisations, and melodic segments must be distributed equally among all instruments;
- Careful attention must be paid to the locations of musical patterns within the composition, so as they are spread equally.

Keeping these requirements in mind, we created the two pieces in collaboration with a composer (male, age: 33 years, professional music experience) holding a diploma from Trinity College London in Rock and Pop guitar, and having extensive experience with live performances of various musical styles. The two pieces can be described as follows:

1) Piece 1 is inspired by grunge and electronic music with several melodic elements are played over a 4-chord loop. The piece employed 1 pattern each for the guitar, piano, and drums.
2) Piece 2 is an upbeat track, with inspiration from 80's pop music. This piece was a high energy piece with an upbeat tempo and vibrant instrumentation. Similar to piece 1, this too had 1 pattern each for the guitar, piano, and drums.

The performers were set up as two musical trios, each consisting of a guitarist, keyboardist, and a drummer. The instruments were selected as they are arguably the most widespread and common musical instruments in many musical styles.

The patterns were strategically distributed within each piece to minimize confusion (i.e., if the guitar pattern and the keyboard pattern were immediately adjacent to each other, it would cause some confusion to the audience member). Each pattern repeated multiple times within the piece to allow the system to trigger sensory effects.

Figure Figure 1 shows an example where each instrument's patterns are spread within the timeline of the piece.

B. Components of the IoMusT ecosystem

1) Smart Musical Instruments and Pattern Detection: Our previous work in real-time musical pattern detection has led to the development of different SMIs with embedded pattern detection capabilities for both monophonic [33] and polyphonic streams [34]. These instruments are equipped with neural networks trained to detect specific musical patterns as soon as they are played on the instrument. The neural networks are trained using a synthetically created dataset which reflects

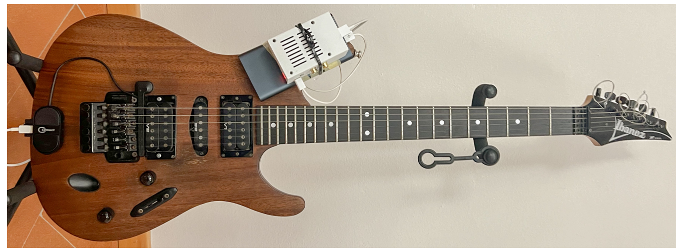

Fig. 2. The developed smart electric guitar.

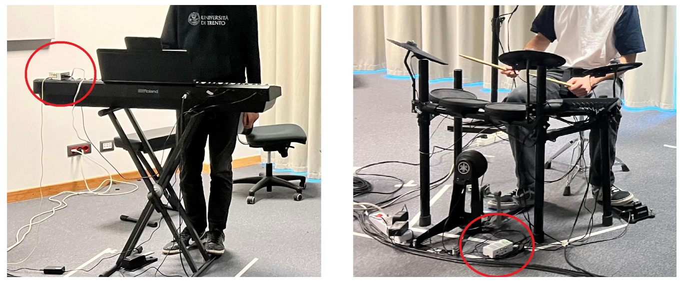

Fig. 3. The smart keyboard and the smart drum kit.

expressive and artistic variations that a musician might add to each pattern.

The core of the experiment revolved around the SMIs outfitted with embedded computing devices with the integrated pattern detection algorithm. Each SMI had an Recurrent Neural Network (RNN) model trained to detect the pattern corresponding to the instrument. Figure 2 shows the smart electric guitar with the embedded system attached, and Figure 3 shows the smart keyboard and the smart drum kit, where the embedded computing devices are highlighted in red.

Each instrument was wirelessly connected to a local area network created by Wi-Fi router as shown by Figure 4. Upon detection of a pattern, the corresponding message was then transmitted to a central server using the Open Sound Control (OSC) protocol. the server decoded the OSC message and triggered the relevant musical thing. The audio output of each instrument was sent to a cental audio mixer and then to stereo loudspeakers.

C. Musical Things for the audience

Musical things refer to the peripheral devices synchronized to the performance. We use a heterogeneous set of musical things controlled by the SMIs to create multi-sensory stimuli. Studies have shown that the combination of different sensory modalities improves the experience related to music, as individuals process information more holistically when multiple senses are engaged [35], [36]. Beyond the sound from the SMIs that create an auditory stimulus, the visual stimuli from the lights can amplify the emotional messages conveyed through the music [35], [37]; the smoke creates a visual and a somewhat olfactory stimulus; the MR headsets create a visual stimulus that allows users to feel as they are a part of the performance as opposed to mere spectators [36]; and the vibrations create a haptic stimuli which allow the audience to *feel* the music [38].

The Musical Things for the audience consisted of 10 *Meta Quest 3* headsets, and wearable haptic devices. The MR headsets displayed interactive 3D contents that was triggered by the patterns. A series of randomly colored spheres appear in space. The virtual objects were rendered using the color passthrough of the headset (see Figure 8). The application was developed with Unity (v2022.3.50f1), using the Meta XR Core SDK (v 69.00). The headsets were connected through the local network using Wi-Fi. The SMIs, upon detecting a pattern broadcasts an OSC message corresponding to the pattern to the subnet assigned to the MR headsets.

Moreover, each audience member held an Android smartphone that provided tactile feedback through vibrations triggered by the RTPD system. We developed an application running on web browsers (Google Chrome) using the Vibration API together with a Node.js server (v22.12.0). Smartphones were connected through the Wi-Fi. Similar to the MR headsets, the SMIs broadcast an OSC message to the subnet assigned to the smartphones.

1) Musical Things for the stage: The Musical Things for the stage are:

- **Moving Head Lights**: Two *Varytec Hero Wash 715 HEX LED* lights capable of dynamic color changes, rotations, and beam adjustments;
- **LED Bar**: A *Stairville Stage TRI LED*, with 4 individually controllable RGB lights;
- **Smoke Machine**: *AF-180 LED Fogger* to enhance dramatic moments;

The stage lights and smoke machine were connected in a daisy chain configuration via their interfaces. The lights and smoke machine were controlled through the open source software Q Light Controller+[2] (QLC). Upon detecting a pattern, the SMIs transmit the OSC message to the IP address assigned to QLC, and the OSC message handling was programmed within QLC.

D. Router, central server and tablet

We created a local area network via a Wi-Fi router (TL-WR902AC by TP Link), which interconnected all the musical devices in the ecosystem. The Wi-Fi was configured according to the guidelines reported in [39] to optimize the latency.

A central server (an Alienware M15 laptop) was used to run the mapping system described in Section III-E. This was dynamically controlled by a web app which could be accessed remotely via a tablet connected to the same network. The web app enables users to associate a pattern of a given smart musical instrument to a stimulus provided by the musical things for the audience. Figure 4 illustrates the main components of the ecosystem and the data flow.

E. Hot Licks Mapper

The Hot Licks Mapper is an application designed to facilitate the dynamic mapping between SMIs with RTPD capability

[2]https://www.qlcplus.org/

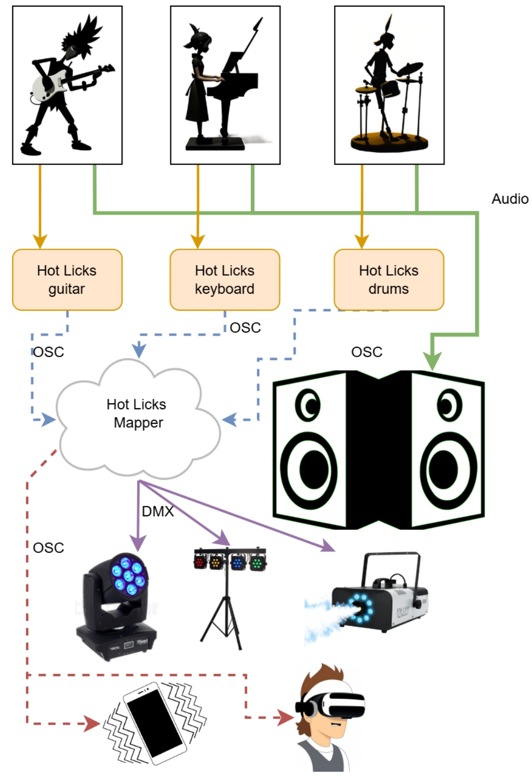

Fig. 4. The connections between SMIs and musical things.

and external peripherals, to allow an interactive experience for the audience. Hot Licks Mapper serves as a visual and interactive tool that allows a composer to connect multiple patterns detected through different SMIs to multiple stimuli created by different output devices. Specifically, the application enables a composer to define 4 types of mappings:

1) **one-to-one:** one pattern of an SMI to one stimulus on a single type musical thing for the audience;
2) **one-to-many:** one pattern of an SMI to many stimuli on a single or multiple types of musical things for the audience;
3) **many-to-one:** a single pattern for each of multiple SMIs to one stimulus on a single type of musical things for the audience);
4) **many-to-many:** multiple patterns played on multiple SMIs to many stimuli on a single or multiple types of musical things for the audience.

As elaborated in Section III-B1, the developed smart musical instruments are able to output OSC commands upon detection of various patterns. Consider N smart musical instruments, each equipped with its own model trained to identify n patterns. Thus, the inputs to the Hot Licks Mapper will be $pattern_1{}^1, pattern_1{}^2, \ldots, pattern_N{}^{n-1}, pattern_N{}^n$, where $pattern_N{}^n$ is the n^{th} pattern produced by the N^{th} smart musical instrument.

Similarly, the outputs will be $stimulus_1{}^1, stimulus_1{}^2, \ldots, stimulus_M{}^{m-1}, stimulus_M{}^m$, where $stimulus_M{}^m$ is the m^{th} stimulus produced by the M^{th} output device. Output devices may be stage lights, smoke machines, MR headsets, projectors, mobile phones, wearable haptic devices and so forth. Different stimuli may be colors and strobe patterns of lights, smoke patterns, and colors, projections on screens, different vibration patterns on wearable haptic devices, etc.

To enable a truly immersive experience, all communications are wireless as illustrated by Figure 4, where each SMI emits OSC commands via Wi-Fi upon detecting a pattern. The User can create a one-to-one, or one-to-many routing between patterns and stimuli. Moreover, the user can dynamically create or remove connections in real-time with a handheld device, which communicates wireless via Wi-Fi. The OSC commands are routed to each output device wirelessly via Wi-Fi. Below are the routing between SMIs and musical things used in the performance.

- **Smart Guitar**
  - *Moving Heads and LED Bar*: The light rotates along the x and y axes to illuminate the guitarist and turns blue for a duration of 3000 ms, then switches off and rotates to the original position.
  - *MR headsets*: Virtual objects spawning on the headset. The users can use their hands to move the objects around the space.
- **Smart Keyboard**
  - *Moving Heads and LED Bar*: The light rotates along the x and y axes to illuminate the guitarist and turns green for a duration of 3000 ms, then switches off and rotates to the original position.
  - *MR headsets*: Virtual objects spawning on the headset. The users can use their hands to move the objects around the space.
- **Smart Drumkit**
  - *Moving Heads and LED Bar*: The light rotates along the x and y axes to illuminate the guitarist and turns red for a duration of 3000 ms, then switches off and rotates to the original position.
  - *Vibrations*: The android phones emit a series of pulses of 200ms, with an interval of 100ms.
  - *Smoke*: The smoke machine produces smoke for 1500ms.

The composer interacts with the hot licks mapper through a web interface where inputs and outputs are represented as blocks. These blocks can be visually connected with lines (arrows), symbolizing the data flow from inputs to outputs. Each input block can be associated with an SMI, and each output block can associated with an output device, allowing for complex routing and manipulation scenarios. Figure 5 shows the routing we used during the experiments, where the composer has routed several inputs with one or multiple output devices. The composer is also able to create new mappings or adjust existing ones by drawing arrows from patterns to stimuli, facilitating a customizable routing environment in real-time.

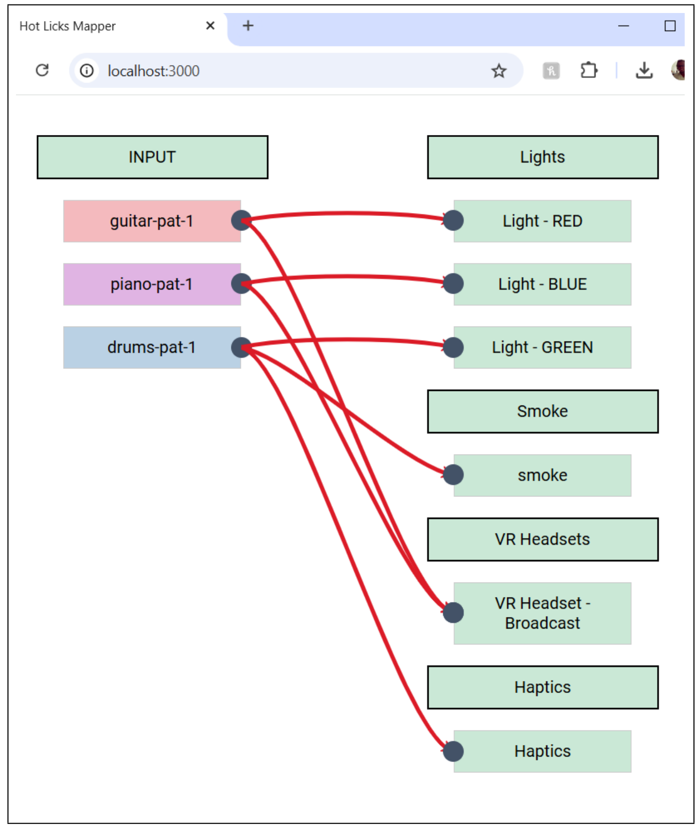

Fig. 5. A screenshot of the initial mapping between the SMIs' patterns and musical things' parameters.

IV. Evaluation

To evaluate the SMIs-based performance ecosystem we relied on the framework for the evaluation of Digital Musical Instruments proposed by [40]. This framework identifies multiple stakeholders that need to be involved in the evaluation process of a novel Digital Musical Instrument: performers, audience members, composers, and designers. Our evaluation considers 3 of these stakeholders. Qualitative and quantitative feedback was obtained from a composer, as well as the audience and the performers involved in the performances created by the composer. Instrument designers were excluded as they are not applicable in our case study. Qualitative responses were analyzed using a reflexive thematic analysis [41].

A. Participants

Audience members. Each performance involved 10 audience members, for a total of 20 participants (3 female, 17 male). They were aged between 22 and 31 years (mean = 27.65, standard deviation = 3.23). The position of the audience, performers, and all Musical Things are shown by Figure 6. They were recruited from the university of Trento, and their informed consent was received prior to the start.

Performers. A total of 6 performers were involved, 2 guitar players, 2 drummers, and 2 keyboard players (all male, aged between 26-32). They were recruited from the authors'

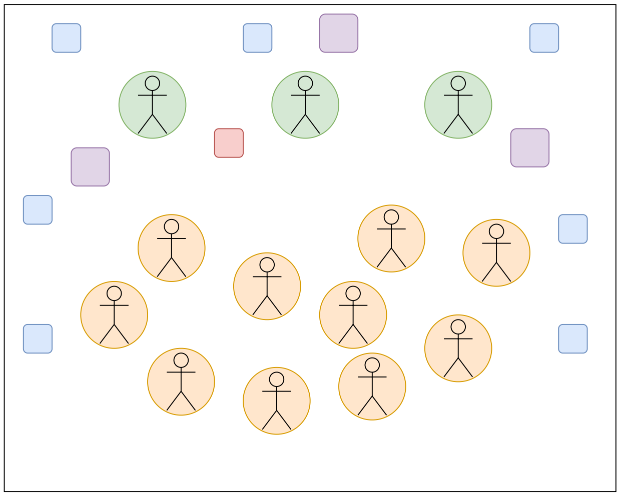

Fig. 6. A diagram of the experiment setting. The performers are shown by the green circles, the audience members are shown by the orange circles. The purple boxes indicate stage lights, the red box indicates the smoke machine, and the blue boxes indicate the position of the loudspeakers.

personal network, and their informed consent was received prior to the start.

Composers. A professional composer (male, 33 years old). They were recruited from the authors' personal network, and were involved in creating the pieces and establishing the mappings between the patterns and Musical Things via the Hot Licks Mapper. Their extensive knowledge and experience in live performances, stage lights, effects, and virtual/mixed reality environments was essential in creating a performance that was both enjoyable and suitable for a technical analysis.

B. Procedure

1) Audience: Each audience member was equipped with a Meta Quest 3 MR headset, and an android phone connected to the local network. Figure 7 shows an audience member wearing the MR headset and holding an android phone. The MR headset displayed a pass-through view with multi colored spheres spawning on the view upon receiving a trigger message. Figure 8 shows a screen capture from the MR headset where the spheres are present. The audience members were standing, facing the band. They could see the performers, lights, as well as the other audience members through the pass-through view of the MR headset.

Each of the two experimental sessions began with a brief introduction to the nature of the performance. We instructed the audience members on how to connect to the local network and open the web-based application to receive OSC commands that triggered the vibrations. They were then instructed to wear the MR headset.

Each experiment involved the two music pieces described in section III-A1. Two versions of each piece were performed, for a total of 4 trials:

Fig. 7. The audience of the performance, wearing the MR headsets and holding the mobile phones.

Fig. 8. Screenshot of the MR headset view.

- **Coherent condition**: The musical things were triggered by the RTPDs system. This created a coherent and synchronized relationship between the music and sensory effects.
- **Random condition**: The musical things were manually triggered at random times during the performance. This provided a contrasting experience, which allowed the audience to have a clear distinction between the two versions.

We hypothesized that the coherent triggering of musical things from SMIs's patterns (trials in Coherent condition) would have been appreciated more than the randomized counterparts (trials in Random condition). To prove this hypothesis, as well as get insights on the audiences' experience, after each trial, the audience members were asked to complete a questionnaire. The questionnaire consisted of the statements presented in Table I to be rated on a 11-point Likert scale [0 = not at all, 10 = very much].

Moreover, they were asked to express their preference via the following question:

- Did you prefer the first performance or the second performance?

The questionnaire was designed to obtain an understanding of the audience's experience of each performance. Upon completing all performance sessions, the audience members were asked to fill a post-experiment questionnaire involving the following open-ended questions:

- What was the most memorable experience for you during the performance? And why?
- How clear was the relationship between what you heard and what you saw or felt?

2) Performers: The performers involved in the performance were given a questionnaire, in which they were asked to rate the statements presented in Table II on a 11-point Likert scale [0 = not at all, 10 = very much].

The performers were also asked to answer the following open-ended questions to understand their perception about the system, acquiring feedback on how they would use the system, and obtaining suggestions for improvements:

- In what musical activities would you use this system?
- Please explain any other way you would use this system?
- If you had access to this system (where another external output can supplement the music) when composing a piece, will there be any changes in the way you would approach the composition?
- Would you associate specific musical phrases with specific output devices?
- How would you improve the system?
- In your opinion, what is the added value of the system?
- Do you have any other comments?

V. Results

A. Audience: Quantitative Analysis

Figure 9 shows the participants' rankings to the questionnaire items for the conditions of musical things triggered with pattern detection vs. musical things triggered randomly. In general, we can make the following observations: 1) lighting effects were the most engaging and least distracting in both conditions; 2) vibrations showed a promising engagement score; 3) the 3D graphics in the MR headset showed mixed reactions, with higher distraction scores compared to lights and vibrations.

We analyzed the results with a Wilcoxon Signed-Rank Test, which revealed that only two questionnaire items were rated significantly higher for the synchronous triggering condition than the randomized condition: Lights Coherence and Vibrations Coherence (both $p < .05$). The mean, standard deviation, p-values and the effect size of each audience question are presented in Table I.

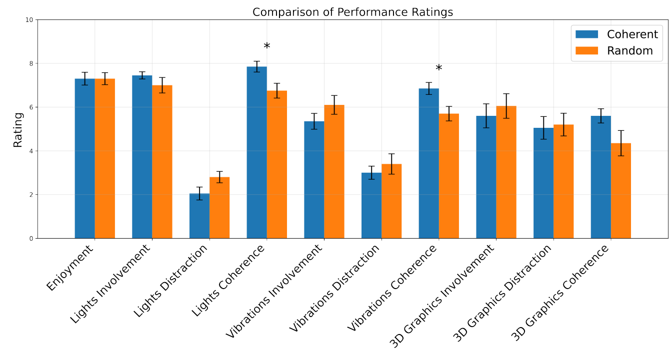

Fig. 9. Mean and standard error of the Audience questionnaire items for the Coherent and Random conditions. Refer Table I for p-values and effect size.

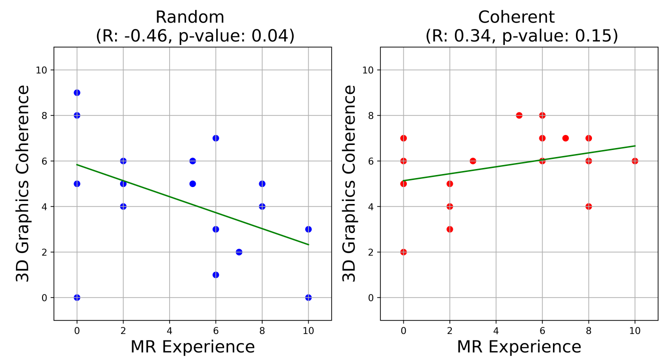

Fig. 10. 3D Graphics coherence measured against participants experience with MR headsets. The correlation line in shown in green.

Concerning the question about the preference, 11 participants preferred the Coherent condition while 9 preferred the Random one. These results show that that the audience had no significant preference for the RTPDs over the random version. However, the statistical analysis reported above revealed that when the RTPD system was used, the musical things were perceived to be more coherent and less distracting. We conducted a Pearson's correlation test to assess whether the experience with VR/MR technology could affect these results. A significant negative correlation (r = -0.462 , $p < 0.040$) was found for the Random condition, while no significant correlation was found for the Coherent condition (see Figure 10). This indicates that participants with a high experience in MR headsets reported a higher coherence for the 3D Graphics during performance with RTPD-based triggering as opposed to that with random triggering.

B. Audience - Qualitative Analysis

The reflexive thematic analysis yielded two themes.

Multi-sensory coherence. Multi-sensory coherence emerged as a critical aspect that highlighted how the audience members noticed in the different experimental sessions varying levels of coherence between the music and various sensory modalities involved. Some participants expressed perfect synchronization (e.g., *"In the last performance there*

	Mean	SD	p-value	effect size
Enjoyment: To what extent did you enjoy the performance experience?	7.3	1.26	0.89	3.67
Lights Involvement: How much did the Lights make you feel involved in the performance?	7.22	1.25	0.19	2.48
Lights Distraction: How much did the Lights make you feel distracted from the performance?	2.42	1.30	0.13	1.70
Lights Coherence: How coherent were the Lights with the music during the performance?	7.3	1.43	**0.02**	1.45
Vibrations Involvement: How much did the Vibrations make you feel involved in the performance?	5.72	1.81	0.28	3.05
Vibrations Distraction: How much did the Vibrations make you feel distracted from the performance?	3.2	1.76	0.51	2.77
Vibrations Coherence: How coherent were the Vibrations with the music during the performance?	6.27	1.48	**0.02**	1.57
Graphics Involvement: How much did the 3D Graphics in the headset make you feel involved in the performance?	5.82	2.49	0.38	2.55
Graphics Distraction: How much did the 3D Graphics in the headset make you feel distracted from the performance?	5.12	2.31	0.91	3.70
Graphics Coherence: How consistent was the 3D Graphics in the headset from the performance?	4.97	2.19	0.12	2.85

TABLE I

THE QUESTIONNAIRE GIVEN TO THE AUDIENCE MEMBERS, MEAN, STANDARD DEVIATION (SD), P-VALUE, AND EFFECT SIZE. THE SIGNIFICANT COMPARISONS ARE PRESENTED IN BOLD TEXT.

were several moment with good lights, smoke and the virtual ball actually came in sync with the music"; "In some music, there was a clear relationship between what I heard and what I saw and felt"; "The haptic feedback and the lights were clear between the performance and what I saw and felt"; "Feeling the vibrations with the kick drum was a very enjoyable experience") others felt disconnected (e.g., *"There was a very clear relationship between the music and the vibrations only in the first performance"; "The balls also disappeared fairly quickly"; "Ultimately I was not looking at the lights so much any more as the headset made my eyes hurt, and it gave me a slight headache"; "having to lower myself to touch the balls in the headset was uncomfortable"*).

Social Interaction. The social interaction was identified as an unanticipated theme that emerged from the shared experience of using this novel system in group setting. The presence of other participants and their interactions with the virtual objects proved to be an additional level of enjoyment.

Moreover, several statements reflected shared experiences, suggesting that the social aspect of the performance enhanced the experience (e.g., *"Watching others interacting with things that are not there is very funny"; "Most of the audience were trying to touch the balls. It was very engaging to see the others play with the balls in their headsets"; "It was interesting to see the combination of the real atmosphere and virtual reality. Seeing everyone with a headset and playing with nonexistent things, while I am playing with the balls in my headset was very cool"*).

C. Performers: Quantitative Analysis

Upon completing the performance, each musician was asked to complete the questionnaire devised to assess their perception of the system, and identify their interest in involving AI-based solutions into their performances.

The average results illustrated in Figure 11, are all above the neutral point. The means and standard deviation for the questions are presented in Table II. Performers expressed a high satisfaction in the behavior of the system. Moreover, they ranked a high enjoyment in using the system and a strong interest in using AI-based tools in their performances. The performers also reported a high interest in incorporating the RTPD system into their existing setups, and they expressed willingness to recommend this system to fellow musicians.

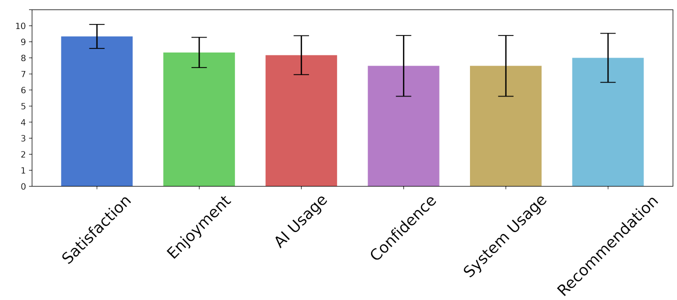

Fig. 11. Results of the questionnaire given to the performers. The questionnaire, the mean and the standard deviation of the answers are presented in Table II.

D. Performers: Qualitative Analysis

The following themes were identified.

Performance Applications. All the performers expressed a strong interest in using the system in live performances. Solo performances and improvisations were some of the potential use cases mentioned (e.g., *"I would use this system during a live performance, specifically during a section where I would improvise. I can use some licks to involve lights as a part of my playing"; "I would use this during live performances with my band. It would be great to see how every instrument can be used to trigger different light and smoke effects"*).

Educational Uses. Some of the performers were involved in teaching activities. They expressed interest in adapting the RTPD system as a teaching tool. Specifically, the envisioned potential of using the classification result as a measure to detect the accuracy in playing (e.g., *"If I can use the output of this system to trigger a red light and a green light, I could use this as a teaching tool. When the student plays a phrase with a specified accuracy, I can use the green light to signal it. If the playing is not accurate enough, the red light will be lit. Students can use this tool by themselves to practice correct technique and phrasing"*).

	Mean	SD
Satisfaction: I am satisfied with the behavior of the system	9.33	0.74
Enjoyment: I enjoyed using the system	8.33	0.94
AI usage: I would use AI tools in my performances	8.16	1.21
Confidence: I was confident when using the system	7.50	1.89
System usage: To what extent would you consider using this system into your performance?	8.00	1.52
Recommendation: Would you recommend this to a fellow musician?	7.83	1.57

TABLE II
THE QUESTIONNAIRE GIVEN TO THE MUSICIANS, MEAN AND STANDARD DEVIATION (SD).

Creative Impact. Performers mentioned the possibility of using this system in their composition and creative process. They expressed an interest in integrating the RTPD system with commercial software for music production (e.g., *"Do you think I can use this with a DAW software? It would be a great tool when producing music. I can use this to spur my creativity in the studio, and also while improvising"*).

VI. DISCUSSION

This study presented the design, implementation and evaluation of IoMusT-based performances, which involved SMIs to trigger connected musical things for the audience. We explored the interplay between auditory, haptic, and visual stimuli and we discussed how the coherence of these elements impacted audiences' and performers' engagement.

The conducted evaluation aimed at assessing the ability of the ecosystem to support the multi-sensory performance from the technological standpoint. By performing a real live concert, we showed that the developed technological ecosystem was perceived effective in detecting in real-time the patterns on each SMI and routing them to the different stimuli of the various musical things involved, based on the defined mappings.

A. Audience Members

We focused on the assessment of the experience of two stakeholders involved in a performance: audience members and performers. Concerning the audience experience assessment, we provided a controlled framework by a comparison between a pattern detection enabled operation and a random operation to evaluate the effects of coherence.

The results showed that lighting effects were deemed as the most coherent with a high engagement and minimal distraction. Haptic feedback showed a moderate level of engagement. The virtual elements had mixed results, where some participants experiencing a high involvement and some experiencing a high distraction. The mixed results may be attributed to the audience members' diverse experience with MR headsets (they reported a mean familiarity with Virtual Reality (VR) of 4.72 with a standard deviation of 3.46, on a 0-to-10 point Likert scale). These findings show the importance of balancing sensory stimuli to avoid a cognitive overload while maximizing the audience engagement.

The qualitative feedback from the audience highlighted the critical role of coherence in creating engaging multi-sensory experiences, as well as showcased the influence of shared social dynamics in enhancing enjoyment. These insights not only validate the potential of SMIs coupled with multi-sensory technologies for creating novel art formats, but also provide a starting point for future designs that consider both technological integration and audience interaction in performance settings.

B. Performers

The six performers involved expressed a high satisfaction for the system. They also expressed their enjoyment in using the system and their interest in incorporating AI-based tools into their performances. Moreover, the performers expressed their interest in using this system during live performances, and during studio activities. According to the performers interviewed, usage during solo performances and usage during live improvisational settings were some of the key potential activities of the RTPD system.

Several performers involved in teaching activities expressed their interest in adapting such technologies into an educational setting. The classification result as a metric of accuracy when teaching scales, phrases, techniques etc, are some key potential applications in a teaching context.

Notably, their responses were in line with the themes identified in our earlier user study reported in [33], which was conducted at the University of Visual and Performing Arts, Sri Lanka with a different population of musicians.

C. Composer

As far as the involved composer is concerned, the evaluation focused primarily on the assessment of the Hot Licks Mapper in supporting the composition process. Overall the feedback of the composer was very positive about the novelty and the concept. Based to the composer's comments, the Hot Licks Mapper was proven to be intuitive to use and, ultimately, successful in allowing the user to create a new artistic format based on the created performance ecosystem. The test showed that the composed ecosystem have the potentials of supporting novel outputs for the audience, driven by performers' expressions over the course of live performances. In particular, the real-time control by SMIs of MR headsets, haptic vests and smoke machines represents a new avenue for interactive live performances, where virtual content can blend with the musical and embodied expressions performed on stage.

The composition process revealed the need to optimally associate musical information with sensory content for a successful performance where Musical Things can act as a complement to the music. There is a need to establish a set of guidelines to support musicians in their creative processes. With respect to this, as proposed by [42], a promising avenue is represented by the identification of common associations between musical features and stimuli. This falls in the remits of research on crossmodal correspondences [43], i.e., people's tendency of matching information presented in one sensory modality to that presented in another.

D. Limitations and Future Work

Our study focused only on the unidirectional communication from performers' to audience members' devices. Nevertheless, the proposed IoMusT communication infrastructure can easily support a bidirectional communication were also the audience can send messages to the performers. The latter could for instance be triggered by some gestural patterns, tracked by some wearable motion sensors. This avenue falls in the remits of the field of technology-mediated audience participation [44].

Another notable limitation was the usage of only two pieces, both composed by a single composer. Further studies involving multiple composers and compositions, multiple musical styles involving both beat-based and non beat-based, as well as the adoption of a longitudinal study as a methodological approach, are needed to reveal more details about the strengths and weaknesses of our concept, and identify actual needs and desires of composers, performers, and audience members.

Our experiments involved a composer and an ensemble of performers, who were all males. The audience members were mostly male too, This was primarily due to the availability of participants. A balanced gender split among participants would yield more generalizable results. Moreover, an audience with a high familiarity with MR headsets could provide more accurate findings on coherence. Nevertheless, our results concretely showed the potential of the IoMusT paradigm in enabling unprecedented forms of musical expression. This calls for further interdisciplinary research at the confluence of the technological and artistic domains.

Several avenues exist for future work. First, we plan to conduct more investigations with composers, and to extend the range of SMIs supporting our real-time pattern detection algorithm. Moreover, we plan to focus on creating new ecosystems where also the audience can participate to the music creation process via a bidirectional communication infrastructure. Furthermore, we plan to test the system in the context of networked music performances were performers and audience members are geographically displaced. Notably, the experiment revealed the need for music composers, lighting directors, and 3D graphic artists to work in tandem in order to deliver a successful experience for the audience. Therefore, in future works we plan to explore the establishment of such interdisciplinary teams.

VII. Conclusions

This paper presented an experiment consisting of two live concerts performed by two group of performers, each group playing a smart electric guitar, a smart keyboard, and a smart drumkit. Each instrument was outfitted with an embedded computing device running the RTPD algorithm, and each instrument had a machine learning model trained to detect a pattern played on the respective instrument. Several stage lights and a smoke machine were triggered by the SMIs. Each performance was attended by 10 audience members who wore MR HMDs, and held Android Smartphones in their hands providing vibrotactile stimuli. The RTPD system was used to spawn 3D objects on the MR headsets and to trigger vibrations on the smartphones. These musical things were triggered by the SMI when each performer played the corresponding pattern.

The performers played two pieces, arranged by a composer recruited for the experiment. Each piece was played twice, the first version with the RTPD system being used to trigger Musical Things, and the second version having them triggered randomly. The results revealed that a higher coherence between the music and Musical Things was observed by the audience members when the RTPD was in place.

Furthermore, the qualitative and quantitative feedback collected from both performers and audience members indicated a high appreciation for the developed IoMusT ecosystem. Performers expressed enthusiasm about the potential of RTPD systems to enhance their creativity, and the audience members reported a higher engagement when multi-sensory elements were incorporated into the performances. The findings show the importance of exploring new artistic formats that leverage technology to enhance musical performances.

In conclusion, this study highlighted the potential of the Internet of Musical Things (IoMusT) paradigm to enhance musical performances through the use of SMIs and IoMusT. Through the integration of SMIs equipped with RTPD capabilities into a cohesive performance ecosystem, interactions that engage multiple senses could be incorporated into novel types of musical performances. We hope that the present work could spur the interest of researchers towards novel music performances and inspire other practitioners to leverage artistic possibilities offered by the IoMusT paradigm.

VIII. Acknowledgments

We acknowledge the support of the MUSMET project funded by the EIC Pathfinder Open scheme of the European Union (grant agreement n. 101184379). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Innovation Council. Neither the European Union nor the European Innovation Council can be held responsible for them.

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] 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.

[3] L. Turchet, "Smart Musical Instruments: vision, design principles, and future directions," *IEEE Access*, vol. 7, pp. 8944–8963, 2019.

[4] S. Waters, "Performance ecosystems: Ecological approaches to musical interaction," *EMS: Electroacoustic Music Studies Network*, pp. 1–20, 2007.

[5] A. Di Scipio, "'Sound is the interface': from interactive to ecosystemic signal processing," *Organised Sound*, vol. 8, no. 3, pp. 269–277, 2003, publisher: Cambridge University Press.

[6] R. Masu, M. Bettega, N. N. Correia, T. Romão, and F. Morreale, "ARCAA: a framework to analyse the artefact ecology in computer music performance," in *Proceedings of the 9th International Conference on Digital and Interactive Arts*, 2019, pp. 1–9.

[7] J. Malloch, S. Sinclair, and M. Wanderley, "Libmapper: (a library for connecting things)," in *Extended Abstracts on Human Factors in Computing Systems*. ACM, 2013, pp. 3087–3090.

[8] A. Fraietta, O. Bown, S. Ferguson, S. Gillespie, and L. Bray, "Rapid composition for networked devices: HappyBrackets," *Computer Music Journal*, vol. 43, no. 2, pp. 89–108, 2019.

[9] A. Fraietta, O. Bown, and S. Ferguson, "Transparent communication within multiplicities," in *2020 27th Conference of Open Innovations Association (FRUCT)*. IEEE, 2020, pp. 61–72.

[10] R. Vieira, D. C. Muchaluat-Saade, and F. L. Schiavoni, "Sunflower: An interactive artistic environment based on iomust concepts," in *ACM International Conference on Interactive Media Experiences*, 2022, pp. 245–248.

[11] R. Dannenberg, "O2: A network protocol for music systems," *Wireless Communications and Mobile Computing*, vol. 2019, 2019.

[12] L. Turchet and F. Antoniazzi, "Semantic web of musical things: Achieving interoperability in the internet of musical things," *Journal of Web Semantics*, vol. 75, p. 100758, 09 2022.

[13] B. Matuszewski, "A web-based framework for distributed music system research and creation," *Journal of the Audio Engineering Society*, vol. 68, no. 10, pp. 717–726, 2020.

[14] I. Clester and J. Freeman, "Composing the network with streams," in *Proceedings of the Audio Mostly Conference*, 2021, pp. 196–199.

[15] A. Yaseen and J. Timoney, "Possibilities emerging on the trajectory from iot to iomust: Enabling ubiquitous musical interactions for wellbeing," in *EMPATHY: 3rd International Workshop on Empowering People in Dealing with Internet of Things Ecosystems. Workshop co-located with AVI 2022*, 2022.

[16] K. Mikolajczyk, S. Ferguson, L. Candy, A. Dias Periera dos Santos, and O. Bown, "Space shaping in the design process for creative coding: a case study in media multiplicities," *Digital Creativity*, pp. 1–21, 2024.

[17] F. Visi, T. Basso, B. Greinke, E. Wood, P. Gschwendtner, C. Hope, and S. Östersjö, "Networking concert halls, musicians, and interactive textiles: Interwoven sound spaces," *Digital Creativity*, pp. 1–22, 2024.

[18] L. Turchet, T. West, and M. M. Wanderley, "Touching the audience: Musical Haptic Wearables for augmented and participatory live music performances," *Journal of Personal and Ubiquitous Computing*, pp. 1–21, 2020.

[19] R. Vieira, D. C. Muchaluat-Saade, and P. César, "Towards an internet of multisensory, multimedia and musical things (io3mt) environment," in *2023 4th International Symposium on the Internet of Sounds*. IEEE, 2023, pp. 1–10.

[20] C. Spence, "Musical scents: On the surprising absence of scented musical/auditory events, entertainments, and experiences," *i-Perception*, vol. 12, no. 5, 2021.

[21] J. Park, Y. Choi, and K. M. Lee, "Research trends in virtual reality music concert technology: A systematic literature review," *IEEE Transactions on Visualization and Computer Graphics*, 2024.

[22] H. Yakura and M. Goto, "Enhancing participation experience in vr live concerts by improving motions of virtual audience avatars," in *2020 IEEE international symposium on mixed and augmented reality (ISMAR)*. IEEE, 2020, pp. 555–565.

[23] T. Venkatesan and Q. J. Wang, "Feeling connected: the role of haptic feedback in vr concerts and the impact of haptic music players on the music listening experience," in *Arts*, vol. 12, no. 4. MDPI, 2023, p. 148.

[24] G. W. Young, N. O'Dwyer, M. F. Vargas, R. M. Donnell, and A. Smolic, "Feel the music!—audience experiences of audio–tactile feedback in a novel virtual reality volumetric music video," in *Arts*, vol. 12, no. 4. MDPI, 2023, p. 156.

[25] K. Jung, S. Kim, S. Oh, and S. H. Yoon, "Hapmotion: motion-to-tactile framework with wearable haptic devices for immersive vr performance experience," *Virtual Reality*, vol. 28, no. 1, p. 13, 2024.

[26] Z. Zhou, A. D. Cheok, W. Liu, X. Chen, F. Farbiz, X. Yang, and M. Haller, "Multisensory musical entertainment systems," *IEEE MultiMedia*, vol. 11, no. 3, p. 88–101, Jul. 2004.

[27] E. Lindström, A. Camurri, A. Friberg, G. Volpe, and M.-L. Rinman, "Affect, attitude and evaluation of multisensory performances," *Journal of New Music Research*, vol. 34, no. 1, pp. 69–86, 2005.

[28] R. Freire, V. Martinez-Missir, C. N. Reed, and P. Strohmeier, "Ravenet: Connecting people and exploring liminal space through wearable networks in music performance," in *Proceedings of the Eighteenth International Conference on Tangible, Embedded, and Embodied Interaction*, ser. TEI '24. New York, NY, USA: Association for Computing Machinery, 2024.

[29] L. Turchet, M. Benincaso, and C. Fischione, "Examples of use cases with smart instruments," in *Proceedings of Audio Mostly Conference*, 2017, pp. 47:1–47:5.

[30] P. Saariluoma and H. Isomäki, "Field experiments in hci: Promises and challenges," in *Future Interaction Design II*. Springer, 2009, pp. 257–272.

[31] A. Lund, J. Lazar, and V. Wulf, "Of course i wouldn't do that in real life: advancing the arguments for increasing realism in hci experiments," in *Proceedings of the 2011 CHI Conference on Human Factors in Computing Systems*. ACM, 2011, pp. 419–428.

[32] L. L. Cheong, M. J. Y. Kee, and P. L. Chen, "How attentional resources of the same or across sensory modalities and task load affect cognitive performance? a multi-sensory integration study," *Journal of Cognitive Psychology*, vol. 35, no. 8, pp. 812–820, 2023.

[33] N. Silva and L. Turchet, "Real-time pattern recognition of symbolic monophonic music," in *Proceedings of the 19th International Audio Mostly Conference*, ser. AM '24. New York, NY, USA: Association for Computing Machinery, 2024.

[34] ——, "Proceedings of the immersive and 3d audio conference," *User-Centered Evaluation of Smart Musical Instruments with Embedded Real-Time Pattern Detection*, 2025, submitted.

[35] E. B. Lange, J. Fünderich, and H. Grimm, "Multisensory integration of musical emotion perception in singing," *Psychological Research*, vol. 86, no. 7, pp. 2099–2114, October 2022, epub 2022 Jan 10.

[36] H. Mitchell, D. Blom, and P. Long, "We hear with our eyes! unlocking tacit knowledge about multisensory music performing," *International Journal of Music Education*, vol. 42, no. 1, pp. 3–14, 2024.

[37] L. Ritchie, "Multisensory music performance with cymatic images," *Music & Science*, vol. 6, p. 205920432311590, March 2023.

[38] M. Hirano and S. Furuya, "Multisensory interactions on auditory and somatosensory information in expert pianists," *Scientific Reports*, vol. 12, no. 1, p. 12503, July 2022.

[39] 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.

[40] S. O'Modhrain, "A framework for the evaluation of digital musical instruments," *Computer Music Journal*, vol. 35, no. 1, pp. 28–42, 2011.

[41] V. Braun and V. Clarke, "Reflecting on reflexive thematic analysis," *Qualitative research in sport, exercise and health*, vol. 11, no. 4, pp. 589–597, 2019, publisher: Taylor & Francis.

[42] P. Walker, "Cross-sensory correspondences: A theoretical framework and their relevance to music." *Psychomusicology: Music, Mind, and Brain*, vol. 26, no. 2, p. 103, 2016.

[43] O. Deroy and C. Spence, "Crossmodal correspondences: four challenges," *Multisensory research*, vol. 29, no. 1-3, pp. 29–48, 2016.

[44] O. Hödl, G. Fitzpatrick, and F. Kayali, "Design implications for technology-mediated audience participation in live music," in *Proceedings of the Sound and Music Computing Conference*, 2017, pp. 28–34.