ORIGINAL PAPER

A web-based distributed system for integrating mobile music in choral performance

Luca Turchet[1] [ID] · **Martina De Cet**[1]

Abstract

This paper presents an Internet of Musical Things system designed to enhance the singing practices of conventional vocal ensembles with electronic sounds generated by smartphones. The system comprises a small loudspeaker connected to a smartphone running a web-based app that generates sounds locally to the chorister, who uses it while singing. An evaluation of the user experience was conducted through three experiments involving a small choir of 9 choristers and a conductor. In the first experiment, the system was utilized in a co-located setting, where choristers shared the same physical space. In the second experiment, the system was employed in a remote setting, where geographically displaced choristers were connected through a networked music performance system. In the third experiment, a hybrid condition was implemented where part of the choir was co-located and part was remotely connected. Overall, results show that the application can be successfully utilized to augment the practice and experience of choir singing, leading to novel forms of musical expression. We provide a critical reflection where we discuss the lessons learned, strengths, limitations, and possible future developments.

Keywords Mobile music · Internet of musical things · WebAudio · Choir · Smartphones · Singing

1 Introduction

Mobile Music is a field of artistic and scientific research exploring the use of mobile devices, such as smartphones and tablets, for creating novel musical forms [1]. Music-making using mobile platforms, either individually [2] or collaboratively [3], has been a strong emerging community for a number of years, and today represents an established field of research. Indeed, the use of mobile devices as musical instruments has been addressed by numerous works, spanning software tools and architectures [4], as well as their applications in performance settings [1, 5, 6]. Mobile Music may occur in both co-located and remote settings. In the former, the performers share the same physical space

✉ Luca Turchet
  luca.turchet@unitn.it

  Martina De Cet
  martina.decet@unitn.it

1   Department of Information Engineering and Computer
    Science, University of Trento, Via Sommarive 9,
    Trento, 38123, Italy

(see, e.g., [7]). In the latter the communication between geographically displaced performers is enabled by the network (see, e.g., [8]) and facilitated by a networked music performance system [9].

Various Mobile Music instruments have been developed to complement, support, or replace existing traditional musical practices (for a review see [10]). One of these practices is choir singing, which is the object of the present study. For example, a smartphone app supporting choir artificial voice performance is reported in [11]. Specifically, the system allowed a group of smartphone users to create synthetic choir music by interacting with the app GUI with a variety of gestures, which were mapped to a voice synthesis algorithm. In a different vein, the study reported in [12] describes SmartVox, a web-based distributed media player serving as a notation tool for choral practices. Such a system delivers audiovisual scores through the performer's mobile devices while singing. However, to the best of authors' knowledge, the integration in the actual singing practice of electronic sounds interactively generated by smartphones to complement the vocal sounds of the ensemble represents today an unexplored topic.

In this paper, we describe a web-based distributed system for enhancing choral practices with electronic sounds and its evaluation with a vocal ensemble. This study aimed at

exploring novel forms of group expressivity merging vocal and electronic sounds. Considering a traditionalist perspective, the choir is a constrained performance configuration, which is characterized by a strong leadership from the conductor, in both timing and dynamics. Traditionally, choristers mainly work with pre-composed material, and use codified vocal registers (soprano, mezzo-soprano, contralto, tenor, baritone, bass) and standard vocal effects. On the other hand, contemporary choral music and singing practices have challenged traditional choir forms in various ways [13]. The challenging aspect of the research reported in this paper is the exploration of the intersection of choir singing with experimental electronic music, which requires addressing both the expression of individuals and their role in the ensemble. The proposed system was conceived for achieving new forms of expression in the context of electroacoustic music, especially in live performance settings.

In more detail, our investigation was based on the following research questions:

1. Can we successfully augment the practice of choir singing with a web-based distributed system producing concurrent electronic sounds?
2. Does a system like the one we propose provide a satisfactory user experience, adequate to support choristers' expressivity?
3. What are the strengths and weaknesses of the system?

In the reminder of the paper, we present the technical aspects of our system, which comprises a web app for both the conductor and the choristers. We also provide an evaluation of the developed technology conducted with a vocal ensemble. Specifically, the system was evaluated through three experiments each assessing a different condition. In the first experiment, the system was utilized in a co-located setting. In the second experiment, it was employed in a remote setting. In the third experiment, a hybrid condition was implemented, where part of the choir was co-located and part was remotely connected. Finally, we provide a critical reflection where we discuss the lessons learned, strengths, limitations, and possible future developments.

2 Related work

2.1 Web audio

One of the most widespread technologies utilized in the Mobile Music field is Web Audio [14]. To date, a number of promising musical projects have demonstrated how audiobased applications can be bridged into the web browser via the Web Audio API (see, e.g., [15–17]). The Web Audio API is one of the most recent among the technologies for audio applications on the web and its use is becoming increasingly widespread [18]. Specifically, Web Audio is a JavaScript API that enables audio analysis and synthesis inside a standard web browser. Differently from Java or Flash, which are implemented in the form of browser plugins, the Web Audio API is implemented by the browser itself. Moreover, the Web Audio API is a World Wide Web Consortium proposed standard.

Given its features, Web Audio represents a promising basis for the creation of distributed audio applications. One of the advantages of Web Audio is the accessibility of web-based projects, given the fact that applications developed on the web are hosted on a server and can be accessed at any time over the Internet. This allows one to avoid the installation of applications over mobile devices. Another advantage is that web-based applications only depend on common technologies that are implemented by web browsers (like HTML, CSS and JavaScript), rather than relying on technologies specific to a particular underlying hardware or software platform (e.g., an operating system).

Web Audio technologies are envisioned to play a relevant role within the emerging paradigm of the Internet of Musical Things (IoMusT) [19]. The IoMusT is an extension of the Internet of Things to the musical domain, and refers to the network of objects serving a musical purpose (Musical Things). Mobile devices used to generate musical signals are an example of these objects. Musical Things are particularly relevant in the context of networked music performances [9, 20], where musical interactions among performers are mediated by the network. Several applications are emerging in this space [21–23]. For instance, IoMusT research has recently developed interactions between performers that leverage the connection of smart musical instruments [24] with smartphones [25].

Web Audio technologies have also made inroads in a specific area within the Mobile Music field that has focused on participatory live music systems, where audience members contribute to the music creation process [26, 27]. Some systems within this category have prioritized the use of a large amount of mobile devices, gathering a lot of participants (see, e.g., [28, 29]).

2.2 Networked music performances

Networked music performance (NMP) systems aim at enabling musicians to interact and perform together through a telecommunication network. Today, this is an established area of research, whose primary goal is to ensure realistic performance conditions, a significant engineering challenge due to the extremely strict requirements in terms of network latency and audio quality. Several NMP systems have been created for this purpose (see, e.g., [30–33]), along with studies assessing the resulting technical performances

and quality of experience (see, e.g., [34–36]). In [9], the authors provide a comprehensive overview of hardware and software technologies enabling NMPs, including low-latency codecs, frameworks, protocols, as well as a discussion of perceptually relevant aspects.

Musical interactions over the network, such as collaborative music creations, can occur locally and remotely. They can occur either over wired networks such as Local Area Networks (LANs) or Wide Area Networks (WANs), or over wireless networks such as Wireless Local Area Networks (WLANs) [20] or cellular data networks [10]. Substantial research has been conducted to find solutions to minimize latency and jitter of audio or control messages transmission while maximizing audio quality. Regarding NMPs over WLANs, a set of recommended Wi-Fi configurations to reduce latency and increase throughput for live performance scenarios has been proposed in [37]. Noticeable examples of audio streaming services for NMPs over WANs are Jack-Trip [32] and Lola [33]. More recently, the Elk company developed Aloha, a NMP system for low latency and high quality audio communications that uses a dedicated box and leverages the Elk Audio OS operating system for embedded audio [38] to connect up to four performers.

3 Design

This research aimed at integrating the use of smartphones as musical instruments into the compositional and performative practices of choirs. Moreover, we aimed at investigating the distributed nature of IoMusT ecosystems based on smartphones, which could allow choirs members to interact in both co-located and/or remote settings (i.e., where musicians are in the same physical location or are geographically displaced). To achieve this purpose, we worked closely with a non-professional vocal ensemble composed by 10 musicians from the city of Trento (Italy), following a participatory design approach. Specifically, the ensemble was composed by 1 conductor (with 25 years of conducting experience) and 9 choristers, specifically 2 basses, 2 tenor, 2 contralti, and 3 sopranos (aged between 20 and 27, with an average choir experience of 5 years).

An iterative design process was conducted with three cycles of design-evaluation-tests, which led to the following design requirements:

- The ecosystem should be able to preserve the role of conductor and the choristers, by supporting their respective practices in different ways;
- The ecosystem should provide electronic sounds locally to each chorister, so for them to understand their contribution to the ensemble in the same way as it happens for the sounds produced by their voice;
- The ecosystem should be able to support distributed musicians in both co-located and remote settings;
- The smartphone app GUI should have a minimal set of functions in order to minimize the conductor's and choristers' cognitive overload due to simultaneously singing and playing with the smartphone;

Fig. 1 Local ecosystem architecture, with the indication of the involved musicians, devices, and data flow
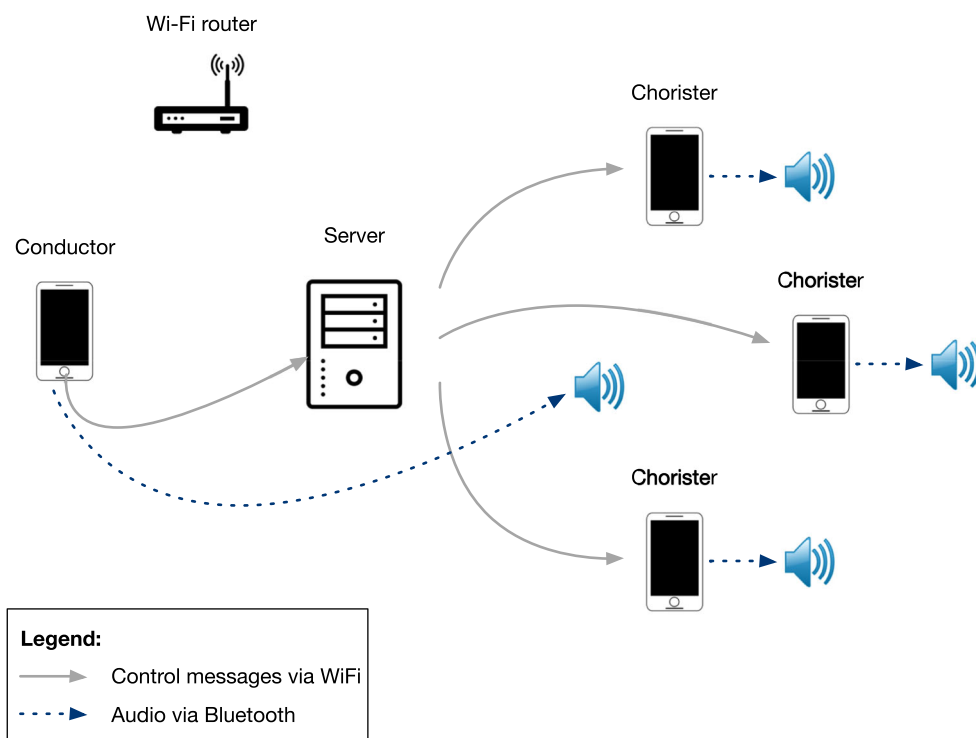
- The ecosystem should not support the tight synchronization of devices, as the role of the smartphone sounds is based on the tenets of the *aleatoric compositional style* [39], with the end goal of creating a distributed sonic texture serving as a background for the vocal sounds.

Such requirements translated into three kinds of ecosystem's architectures:

1. **Local architecture:** it supports interactions between co-located musicians (see Fig. 1);
2. **Remote architecture:** it supports interactions between geographically displaced musicians (see Fig. 2);
3. **Mixed local-remote architecture:** it supports interactions between both co-located and geographically displaced musicians (see Fig. 3).

Moreover, the requirements translated into two smartphone apps having a different GUI, one for the conductor and one for the choristers, each preserving their respective roles present in conventional choir practices. The architectures and involved apps were conceived to allow the conductor and the choristers to interact with the app GUI while performing their usual activities, i.e.,

respectively, directing the choir with gestures, and singing while following the conductor and the score. Specifically, the ecosystem was designed to enable the conductor to select at any time the beat per minute (BPM) of the piece as well as the timbral and harmonic content that the electronic sounds generated by the ensemble will have to adhere to.

Each member of the choir was envisioned to use a small loudspeaker to reproduce next to himself/herself the sounds generated by the smartphones. This was due to the specific compositional choice of creating a distributed sound texture that could seamless merge electronic sounds with the voices. At the same time, such a configuration would have enabled each chorister to understand its own individual contribution to the overall sonic effect.

4 Implementation

We implemented the three kinds of architectures described in Section 3 which were envisioned to support smartphone-based interactions between conductor and choristers in co-located, remote, and mixed settings (see respectively Figs. 1, 2, and 3). The following details the involved hardware and software components.
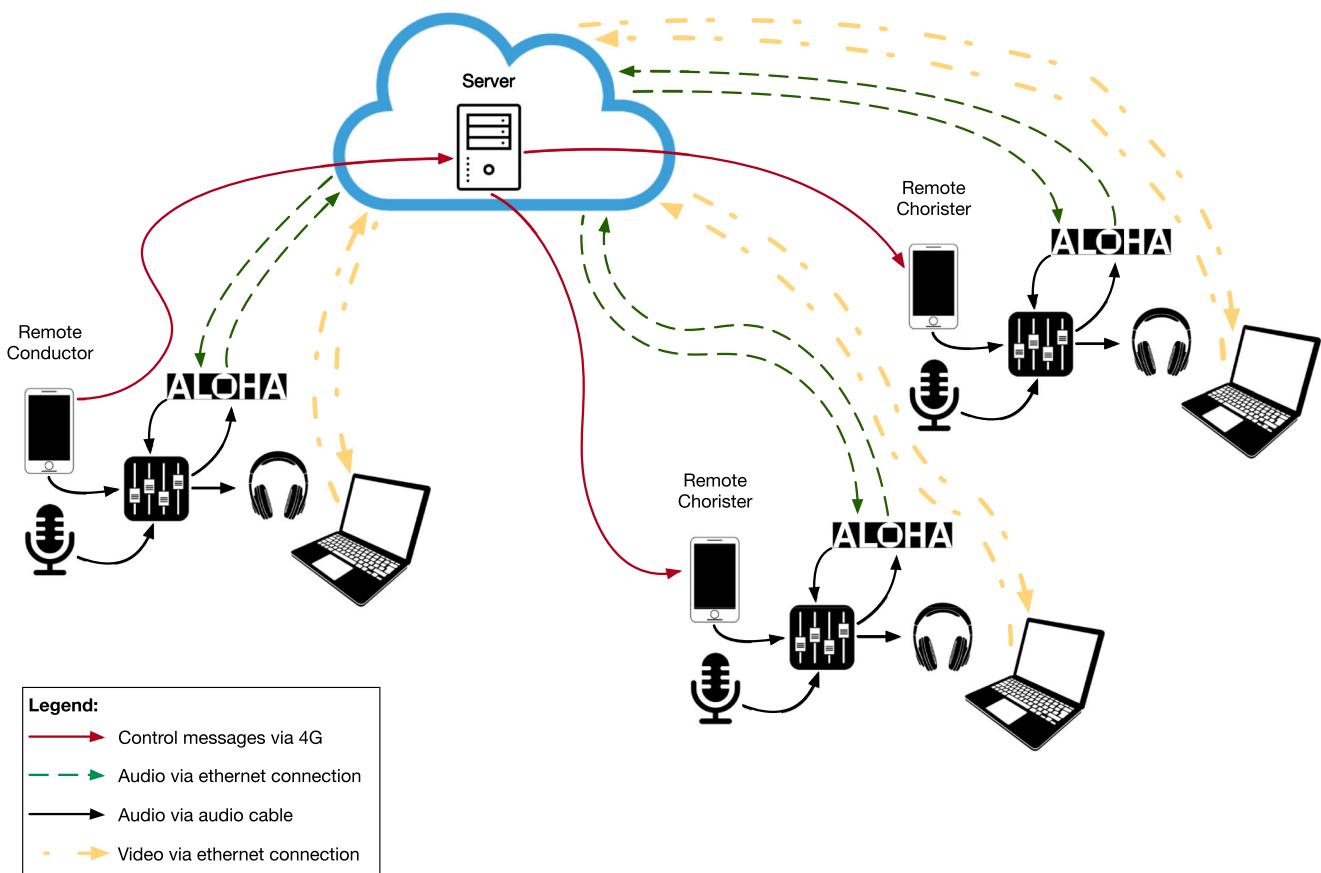
Fig. 2 Remote ecosystem architecture, with the indication of the involved musicians, devices, and data flow
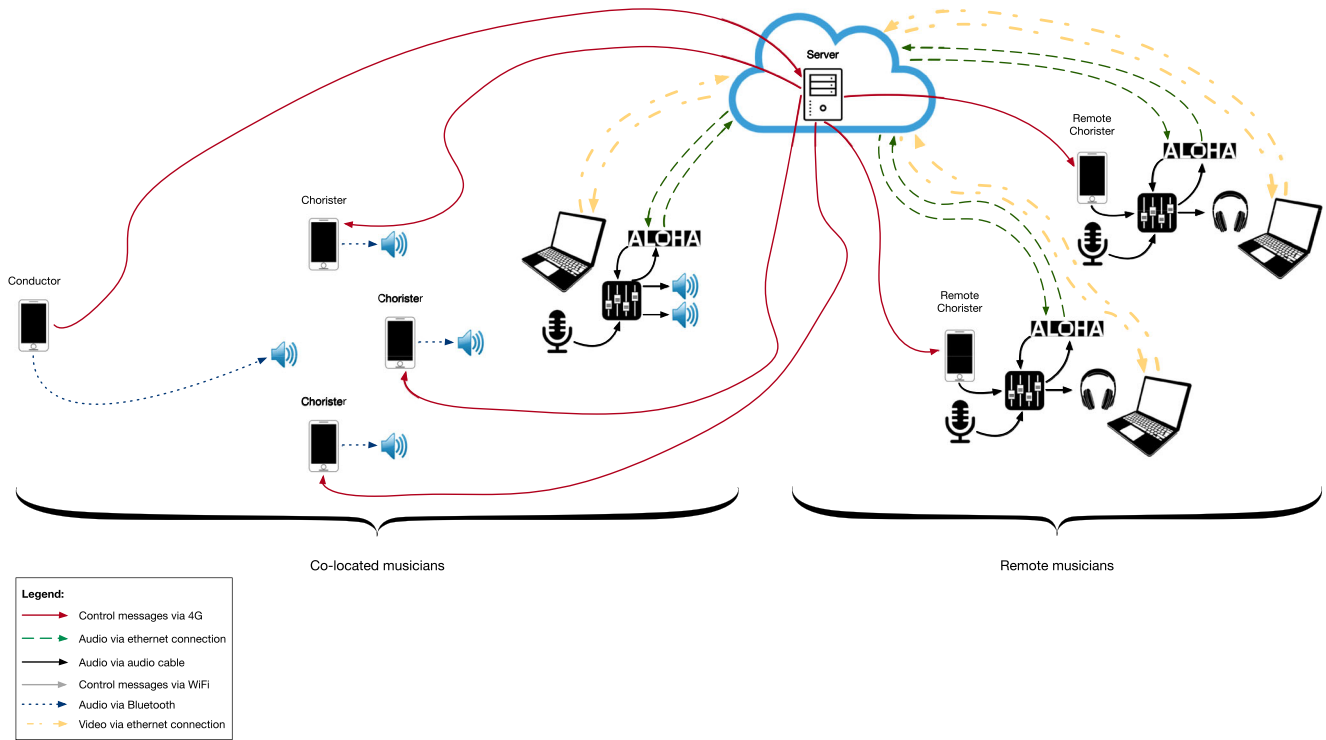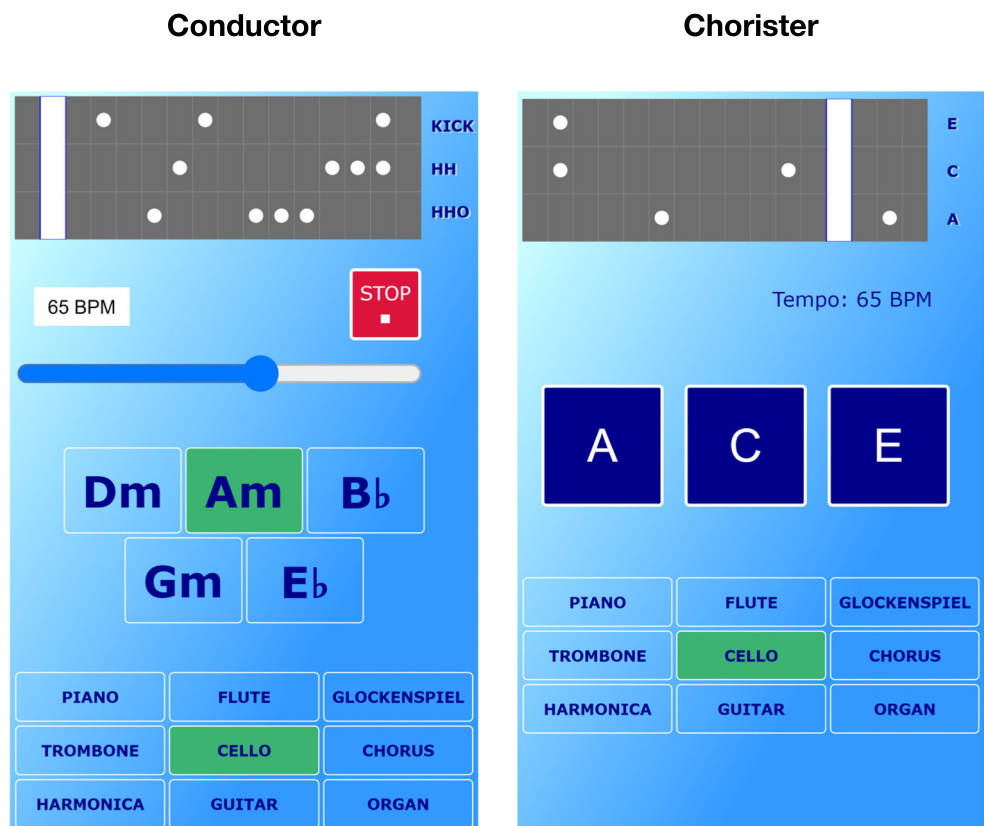
Fig. 3 Mixed local-remote ecosystem architecture, with the indication of the involved musicians, devices, and data flow

Fig. 4 Screenshot of the app GUI of the conductor (left) and the corresponding one of the chorister (right)

4.1 Hardware

Common to all architectures is the server-based nature of the ecosystem, which is utilized to handle the communication between smartphones. Specifically, each ecosystem comprises a server, a smartphone for both the conductor and the choristers, and a small loudspeaker connected to each smartphone. As a web-based application was developed, any Android- and iOS-based smartphones could be utilized. Therefore, the vocal ensemble could use their own smartphones.

Whereas any small loudspeaker can be utilized in conjunction with a smartphone (either connected wireless or via a cable), during the design process we provided the choristers with identical JBL bluetooth speakers. In the local architecture, for the conductor we used a bigger loudspeaker to achieve a louder sound than that produced by the loudspeakers used by the choristers. It was was placed in front of the choir, at the center. This choice was due to the fact that the conductor using such speaker delivered percussive sounds that provide the whole ensemble with a rhythm to follow.

Where the architectures differentiate is the use of specific equipment enabling remote interactions between members of the vocal ensemble. For this purpose, we leveraged the Elk's Aloha networked music performance system [38], which was used in conjunction with a microphone, headphones, and a mixer to merge the vocal end electronically generated sounds before being sent over the Internet. Notably, we explored the use of other networked performance systems such as JackTrip [32] and Lola [33], but from a preliminary comparison we found better performances with Aloha. Moreover, whereas a Wi-Fi network and a local server could be utilized in the local architecture, for the remote and mixed architectures a cloud-based server was involved. To access such server the smartphones used the 4G communication infrastructure. It was observed that the actual latency was fully tolerable by both conductor and choristers (note that there is no synchronicity between the sounds generated by the smartphones, due to the artistic choice of creating textures following the paradigm of the aleatoric compositional style).

In both the remote and mixed architectures, the remote musicians were using a laptop in conjunction with the Aloha box, which jointly provide audio and video streaming (the visual component is particularly relevant for choristers who need to follow the conductor). Choristers were also using a directional microphone to capture their voice. In the case of the mixed architecture, we used a pair of omnidirectional microphones to capture the local ensemble sound, and bigger loudspeakers each placed on a stand to render the voice of the remote singers.

Notably, to enable such kind of low-latency audio and video communications a high speed and large bandwidth network is necessary, while at the same time the distance between remote musicians should be kept relatively low. Moreover, our interest was primarily to test the interaction of choristers with the developed systems rather than performing a technical evaluation of them in real settings (i.e., involving a conventional network connection). For these reasons, we performed the experiments in different rooms of the Department of Information Engineering and Computer Science of University of Trento, which allowed to test the system under minimal network latency conditions.

4.2 Software

The developed web application consisted of two web clients, one for the chorister and one for the conductor, that can be executed in any recent web browser on mobile devices (e.g., smartphones, tablets) and laptops. Such an application was built using WebAudio, while the server side was based on Node.js. The real-time communication between clients was achieved through the WebSocket protocol. The system leveraged the MIDI protocol and the soundfont technology for triggering the notes and associated timbres relating to various musical instruments.

4.2.1 The conductor client

Figure 4 (left) shows a screenshot of the app of the conductor. Through the app, the conductor determines not only the rhythm of a piece, but also the affordances of the app used by the choristers. This is achieved by means of a 16-beat step sequencer that can trigger various pre-defined drums sounds (in the figure, the kick, the hi-hat closed, and the hi-hat open), which the conductor manipulates by easily placing or removing a dot in the corresponding cell. The BPM is set via a slider and displayed, allowing for gradual or abrupt accelerandi, rallentandi, and tempo changes. A play/stop button allows the conductor to start and stop the sequencer.

Moreover, the app allows the conductor to select at any time one of the chords pre-configured according to the piece at hand. While conducting, the conductor is required to change the chord corresponding to the score in order to avoid cacophony due to mismatches between what is sung by the ensemble and the generated electronic sounds (of course, if this is not deliberately the wanted effect). Furthermore, the conductor can select one of 9 instruments (piano, flute, glockenspiel, trombone, cello, chorus, harmonica, guitar, and organ), which relate to the timbres the whole ensemble will generate. Those specific timbres were chosen for their capacity of mixing well with the voices of the ensemble according to the expressive

needs of the conductor and in relation to the played pieces. Notably, the conductor does not generate from its associated speaker any sound other than the percussive ones that provide the rhythm of the piece to the whole ensemble.

4.2.2 The chorister client

The configurations of the conductor's app (i.e., sequencer for the percussive sounds, BPM, harmony, and timber) are reflected in real-time in the app of the choristers. Figure 4 (right) shows a screenshot of the app of the chorister. Such a screenshot corresponds to the settings of the conductor's app displayed in Fig. 4 (left). Each chorister is empowered to modify the app widgets according to his/her expressive intent.

Specifically, the chorister's app displays both the sequencer of percussive sounds (unmodifiable by the chorister) and a sequencer that enables the singer to trigger notes belonging to the triad of the chord set by the conductor. The BPM of this second sequencer is the same of the global BPM set by the conductor. Moreover, the singer can also trigger interactively the same three notes thanks to corresponding buttons. For instance, the figure shows the selection of the A minor chord in the conductor's app, while the chorister's app shows that the notes of the sequencer and of the buttons are those that form the A minor triad (i.e., A, C, E).

5 Evaluation of the local architecture

We conducted a user study to evaluate the local architecture with the aim of assessing the usability of the system and musicians' experience in interacting with it. Notably, the evaluation of the system was conducted throughout the design process in conjunction with the conductor and the choir, which led to various refinements at interface and interaction level. The following reports the experimental evaluation conducted during the final rehearsal, when the ensemble had already a certain degree of familiarity with the system and its usage. The user study took place in their normal rehearsal room.

Two pieces, each of duration of 5-min, were involved during the rehearsal, which were in the repertoire of the ensemble. The pieces were characterized by distinct harmonies, tempo, and character, but both were tonal music (precisely, a canon and a 4-voice polyphonic composition). The rehearsal lasted 3 h, during which the two pieces were performed 6 times. The conductor and the choir explored all the expressive possibilities of the system, including tempo changes, various co-ordinated and non co-ordinated texture densities, as well as dissonances resulting from mismatches between the harmony sung by the ensemble and that of the app configuration.

At the end of the rehearsal the ensemble was asked to fill a questionnaire (inspired to that reported in [25]), which was composed by four sections. The first section included demographic questions about gender, age, and musical experience. The second section consisted of the ten System Usability Scale (SUS) questions measured using 5-point Likert items [40]. The third section presented the eleven Creativity Support Index (CSI) questions measured using 11-point Likert items [41]. The CSI section comprised 15 paired comparisons to determine the relative importance of the six creativity factors in musical practice tasks (Collaboration, Enjoyment, Exploration, Expressiveness, Immersion, Results Worth Effort). The final section of the questionnaire gathered reflective feedback consisting of two parts. The first part comprised two questions to be evaluated on a 7-point Likert scale (1 corresponds to *not at all* and 7 stands for *very much*), which addresses engagement and novelty:

- How engaged you felt when playing with the other musicians using the system?
- How novel was your musical experience with this system?

The second part consisted of the following open questions:

- What was your experience in interacting with the system?
- How would you improve the system?
- Do you have any comments?

Furthermore, after the completion of the questionnaires, a focus group with all members of the choir was conducted. Finally, we conducted an interview with the conductor in order to get further insights about her experience of conducting in presence of the developed technology.

5.1 Results

We report the combined results of quantitative analysis for the conductor and for the choristers (throughout the design process the conductor also extensively tested the app designed for the choristers). The SUS metric assesses the usability of a system on a scale from 0 to 100. As a point of comparison, an average SUS score of about 68 was obtained from over 500 studies. Our system obtained a mean SUS score of 75.75 (95% confidence interval: [69;82.49]) which is above average. Figure 5 shows the breakdown of the result across the topics of the system usability scale. The results reported in the figure indicates that on average participants found the system easy to use, simple, quick to learn and to use without technical support.
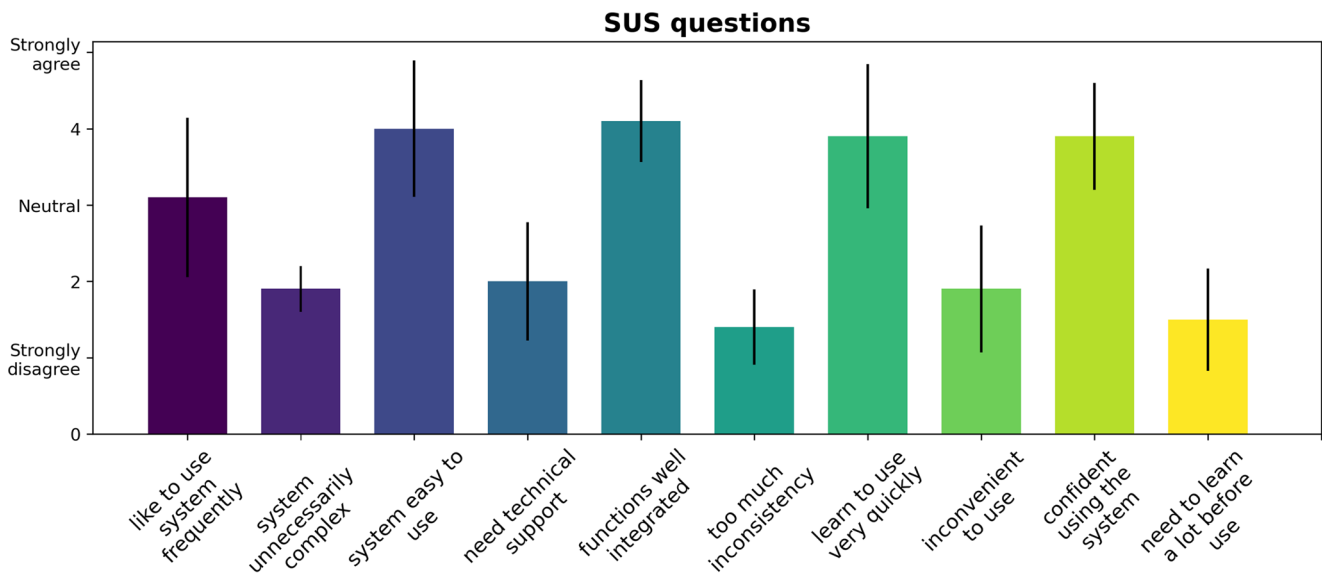
Fig. 5 Scores of the system usability scale topics for the co-located architecture tested

The CSI metric, ranging in [0, 100], enables to assess the ability of a tool to support the open-ended creation of new artifacts. Our system obtained a mean CSI of 73.89 (standard deviation = 10.91), which highlights a good creativity support. Table 1 presents the average CSI results broken down into factor counts (the number of times a creativity factor was judged more important than another for the task, as based on paired comparisons), factor scores (the ratings of the various factors irrespective of their importance for the task), and the weighted factor scores, which combine the factor counts and scores to make it more sensitive to the factors that are the most important to the given task.

The creativity factor which was judged the most important for the task of singing in group while generating electronics sounds with the smartphones is Exploration, closely followed by Expressiveness. This suggests that such factors are important to users engaged in the task. Notably, Exploration and Expressiveness received the highest scores also for both the average and weighted factor scores. The

lowest average weighted factor score was reported for Results Worth Effort, evidencing a certain difficulty in performing the task.

Regarding the ad-hoc questions on novelty and engagement, the experience was deemed as very new (mean = 6.55, standard deviation = 0.72), and the level of involvement was also generally high (mean = 5.2, standard deviation = 1.68), although we note that one chorister reported a very low score for it.

Participants' answers to the open-ended questions and the final focus groups were analyzed using an inductive thematic analysis [42]. Such analysis was conducted by generating codes, which were further organized into the following themes that reflected patterns.

Novelty and fun Five participants stated to have strongly appreciated the idea behind the system, i.e., the possibility to seamless integrate the vocal sounds with the electronic ones in a distributed fashion (e.g., *"I found this experience very interesting, fun, and engaging". "This experience is*

Table 1 Average CSI results for the co-located architecture (SD reported in brackets)

Creativity factor	Avg. factor counts	Avg. factor score	Avg. weighted factor score
Collaboration	2.7 (1.9)	13.3 (4.4)	40.6 (29.6)
Enjoyment	2.4 (1.2)	14 (3)	32.7 (17.1)
Exploration	**3.5** (1.2)	**15.5** (2.1)	**54.4** (21.4)
Expressiveness	3.2 (1.1)	15.5 (2.9)	50.4 (21.8)
Immersion	2.5 (1.2)	13.1 (3.3)	34.7 (23.5)
Results Worth Effort	0.7 (1.4)	14.7 (1.7)	8.9 (19.2)

The highest average value is reported in bold in each column. The mean CSI score is 73.89 (SD = 10.91). Ranges: Avg. factor counts (0 to 5), avg. factor score (0 to 20), avg. weighted factor score (0 to 100)

original and stimulating for the creativity". "This system allows one to deeply immerse himself in the played music while singing with the others"). The concept of the system was considered innovative, especially because it enables to play an electronic instrument while singing a conventional choir piece (e.g., *"It has been interesting to discover this new way of making music, and especially to understand how the two pieces could drastically change on the basis of the used instrument, the dynamics, and tempo variations"*).

Time for getting familiar A recurrent comment from six choristers and the conductor was that despite the interface was easy to use and to learn, a level of familiarity with it is necessary in order to integrate properly its usage in the conventional practice of choir singing (e.g., *"Initially I had some difficulties, but in this last rehearsal I noticed that I could express myself much better via the app, as I was getting more used to it"*). It was reported that the greatest difficulty was that of singing and playing the smartphone while a tempo change occurred. Moreover, difficulties in the interaction with the technology were experienced by the singers when dissonances were introduced by the director, between the electronic sounds and the vocal sounds resulting from the score.

Feature requests While the minimalistic nature of the design was very appreciated, especially by musicians with less expertise with music technology, six participants felt limited in the interaction possibilities afforded by the system, requesting more features. Specifically, five participants suggested to allow the choristers to select themselves the instrument, while one participant suggested to provide a larger set of controls on the sounds parameters and to add effects. Nevertheless, they pointed out that increasing the number of possible controls necessarily would entail a bigger touchscreen, such as that of a tablet. These comments indicate the need for more functionalities at the service of expressivity and creativity when the musicians become expert users of the application.

Suggestions for other usages Participants provided some suggestions of the possible uses of the systems. Two participants suggested to enable the system to be used by a large ensemble in order to explore the mixture of vocal and electronic sounds textures with a higher number of sound sources, which could lead to interesting sonic results especially in settings such as churches or concert halls. Interestingly, three participants envisioned the use of the system by the audience or by both the choristers and the audience, thus leading to a system for audience participation.

Regarding the final interview with the conductor, she deemed her experience as being very new as well as capable of leading to a very high level of involvement and to a novel way of communicating with the choristers. However, she highlighted the initial difficulties in learning the system and the novel way of conducting in presence of it, confirming the need for a period of training. It was found to be challenging to change tempo, harmony and timbres while conducting, with the aim of achieving interesting sonic effects. She also reported on the need of a bigger screen, such as that of a tablet, to facilitate such changes. Nevertheless, she confirmed that the results achieved worth the effort.
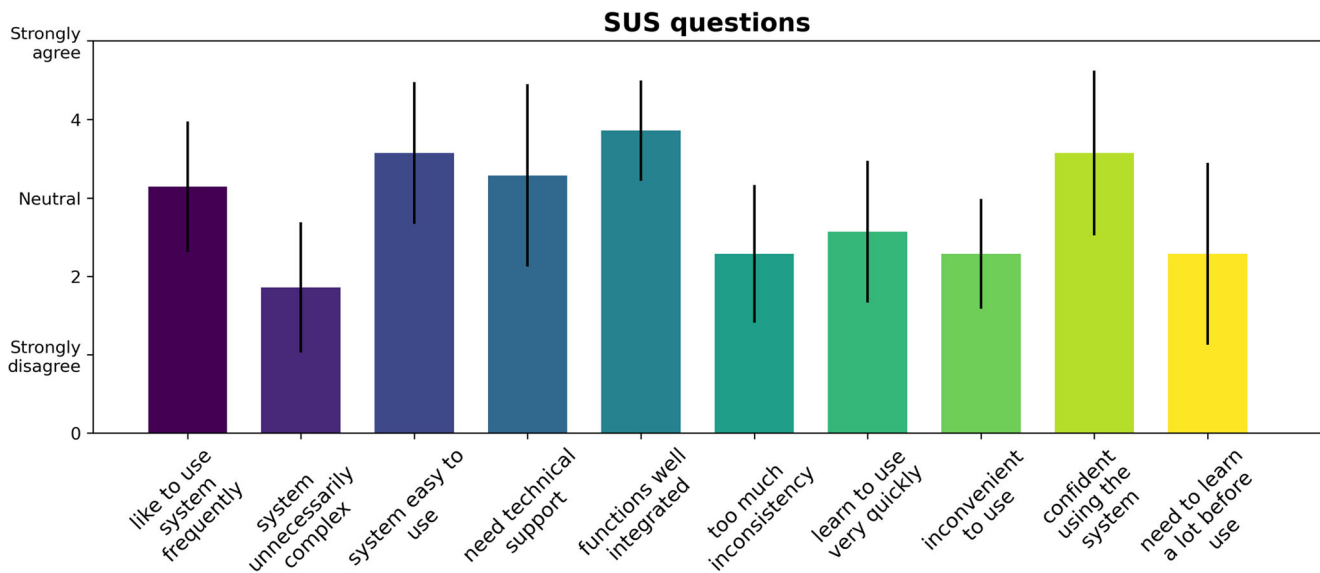
6 Evaluation of the remote architecture

In the second experiment, we tested the distributed usage of the app in conjunction with the Aloha networked music performance system, which allows to stably connect up to 4 performers streaming video and low-latency audio. The experiment were thus conducted in groups of 4 musicians, namely the conductor plus 3 choristers (for a total of 3 groups). The latency of the audio communication was 8 ms, as measured by the internal tools of Aloha. The latency of the video depended on the cameras and computers utilized for capturing the video and representing it on the screen, as well as on the network transmission. A rough estimate of the end to end video latency was about 200 ms. The evaluation of the remote architecture built upon the results obtained for the evaluation of the local architecture. Therefore, we were not interested in re-assessing the interaction of musicians with the app, but rather in understanding the feasibility of the musical interactions afforded by the whole system.

Three sessions were performed involving a total of 9 choristers and the conductor. Each of the three groups of musicians explored the use of the system for 1 h, singing the same pieces involved in the first experiment. After the experiment, they were provided with the same questionnaires used for the first experiment. Finally, a focus group with all members of the choir was conducted, followed by an interview with the sole conductor.

6.1 Results

Our system obtained a mean SUS score of 61.78 (95% confidence interval: [47.05;76.51]) which is below the standard reference of 68%. Figure 6 shows the breakdown of the result across the topics of the system usability scale. The results reported in the figure indicates that on average participants found the system a bit complex to use, as well as not very quick to learn and to use without technical support.

Our system obtained a mean CSI of 68.71 (standard deviation = 10.25), which highlights a good creativity support. Table 2 presents the average CSI results. The

Fig. 6 Scores of the system usability scale topics for the remote architecture

creativity factor which was judged the most important for the task of singing in group over the network while generating electronics sounds with the smartphones is Expressiveness, closely followed by Exploration. This suggests that such factors are important to users engaged in the task. Notably, Exploration and Expressiveness received the highest scores also for both the average and weighted factor scores. The lowest average weighted factor score was reported for Results Worth Effort, evidencing a certain difficulty in performing the task.

Regarding the ad-hoc questions on novelty and engagement, the experience was deemed as new (mean = 5.42, standard deviation = 2.14), and the level of involvement was also generally high (mean = 5.57, standard deviation = 1.27).

The inductive thematic analysis on participants' answers to the open-ended questions and the final focus groups resulted in the following themes.

Setup complexity Three participants reported that the whole system at first glance was difficult to setup. A training period was deemed necessary not only to set properly all components, but also to use them together (e.g., *"I need some time to get familiar with the wires, the interface, the volume adjustments and the usage of the app"*). A shared issue was that of properly adjusting the volumes of the various involved components: overall sound provide to headphones, sound of the JBL speaker, own sound from the microphone, and sound from the connected musicians. This process took some time at the beginning of the session in order to achieve a comfortable sonic experience, and refinements were conducted through the whole duration of the experiment, in particular when a different piece was performed. However, whereas setting up properly the system was deemed laborious by some participants, other three choristers reported that the system was ultimately simple to use once the setup phase was completed.

Table 2 Average CSI results for the remote architecture (SD reported in brackets)

Creativity factor	Avg. factor counts	Avg. factor score	Avg. weighted factor score
Collaboration	3.0 (1.6)	13.5 (3.1)	41.7 (22.9)
Enjoyment	2.4 (1.6)	14.1 (2.5)	33.4 (21)
Exploration	2.7 (1.3)	14 (2.5)	38.1 (18.7)
Expressiveness	**3.4** (1.4)	**14.5** (3)	**52.1** (28.1)
Immersion	2.5 (1)	11.5 (3)	30.4 (15.1)
Results Worth Effort	0.8 (0.8)	11.4 (3.92)	10.2 (10.9)

The highest average value is reported in bold in each column. The mean CSI score is 68.71 (SD = 10.25). Ranges: Avg. factor counts (0 to 5), avg. factor score (0 to 20), avg. weighted factor score (0 to 100)

Audio-video synchronization Three participants reported to not have perceived any latency in the audio streams, but noticed that sometimes the audio stream was arriving before the video counterpart. This issue was reported to sometimes impact their ability to sing on time with the other choristers, and to properly follow the conductor (e.g., *"A few times the delay of the video was disturbing and confused me in following the conductor; however, I was able to tolerate such a delay"*). Interestingly, those participants reported that in order to sing synchronously they tended to rely more on the sound they were listening rather than the conductor's gestures.

Cognitive load Two participants reported that sometimes it was demanding to follow the director on the screen while interacting with the app and the sounds arriving through the headphones.

Features requests A shared issue concerned the size of the screen of the used laptops, which were deemed too small to ensure a proper visual communication not only with the conductor but also with the other choristers (e.g., *"The videos on the screen were too small, I could not read the eyes of the others, the understanding of the facial expressions was missing"*). The conductor also reported on the need of a system that could visually provide the videos of the choristers as placed in real world settings (e.g., all sopranos grouped together and placed in one area distinct from the basses).

Advantages at sonic level Interestingly, especially during the final focus group the choristers reported that the system facilitated their comprehension of what the other connected musicians were singing. This was in contrast with the case of singing in co-located settings, because in such a scenario the own sound and that of every other chorister and app was mixed, instead through listening via headphones and the networked music performance system each singer was more aware of its own contribution and of that of the others (e.g., *"I could better hear the contribution of the others and of the loudspeakers, and use this as a reference for my singing"*; *"With respect to when we sing in presence I could better understand when to sing louder and less loud"*).

Enjoyment Four participants reported that overall the experience to sing over the network together, while using the app, was fun and engaging. Their comments were in line with those reported for the evaluation of the co-located architecture (e.g., *"It was a very positive experience, I had a lot of fun"*).

More participants A shared consideration was that a choir of just three people plus the conductor was deemed not enough for applications of the system in real world scenarios, where typically a higher number of choristers compose a choir. This calls for networked music performance systems actually capable of connecting several musicians, while still ensuring appropriate audio quality and low-latency communications.

7 Considerations on the mixed architecture

We could not implement fully the mixed architecture as conceived in the design phase. Indeed, this required that the remotely connected singers could be rendered, in the physical space where the rest of the choir is present, by means of loudspeakers placed in the positions where such singers would have belonged to (e.g., a tenor would have been rendered as a loudspeaker on a stand, placed near where the tenors sing). However, the networked music performance system utilized was not equipped with four or more independent channels, one for each chorister (like the vast majority of similar systems available in commerce or as research project [9]). Instead, a stereo mix of the remotely connected musicians is provided. This technical issue represents, therefore, a requirement for advancing the hardware and software components of networked music performances systems, which typically provide only a stereo mix instead of a set of independent channels.

On the other hand, it is highly plausible to expect that an evaluation experiment on the mixed architecture would have yielded similar results to those of the previous two experiments for co-located and remote musicians, given the fact that the mixed architecture builds upon the co-located and remote architectures. Nevertheless, the system was implemented and briefly tested, as a proof of concept at technical level, by involving one single remote chorister using the app, who was rendered through a single loudspeaker (Genelec 8010a) placed on a stand. The main outcome of this technical validation was that it is actually possible to involve a connected chorister into a co-located choir, but the volume of the loudspeaker rendering the remote chorister has to be kept at an appropriate volume: on the one hand, this is necessary to avoid feedback loops resulting from the microphones used to capture the whole ensemble, on the other hand to render a sound in line with that of the other voices composing the co-located choir.

8 Discussion and conclusions

This study explored how smartphones could be integrated in the rehearsal and performance practices of a choir via a distributed web based application. Overall, the results of the conducted evaluations show that the application proved

to be successful in augmenting the practice and experience of choir singing. The obtained positive results thus far encourage us to believe that our system has the potential to become a platform capable of leading to innovative forms of musical expression. This is in line with other studies showing that, although smartphones are devices not primarily conceived for musical interactions, mobile-based collaboration is a powerful approach for developing musical applications, including those involving choirs (see, e.g., [12]).

Regarding the experiment with the co-located architecture, the quantitative and qualitative results presented in Section 5.1 show that, on average, participants found the system easy to use and judged their experience during their vocal practice to be fun. Overall, participants felt very engaged during the activity and found their interaction with the other musicians rather novel. For some musicians, the novelty came mostly from the blending of vocally and digitally generated sounds. The developed system demonstrates various promising assets. Firstly, using web standards and Node.js, the application can be executed on a large range of platforms and devices. Secondly, regarding co-located settings, it can be setup effortlessly and quickly during rehearsals and concerts using a laptop running the server and a Wi-Fi router.

The study also enabled to shed light on several frictions that hindered creativity support, both for the conductor and the choristers. While the actual interface was found to be usable, its actual usage while singing required a period of training in order for the ensemble to be able to achieve its expressive intents. Especially the changes in tempo, harmony and timbre, required more practice for the choristers as they entailed a higher level of shared attention between the conductor and the act of singing-playing. Nevertheless, it is also true that practice is needed for any complex musical interface and novel activity. After various rehearsals, indeed, a good level of familiarity was achieved and the resulting sonic effects were found by the ensemble to be very interesting.

Results of the evaluation of the remote architecture were also encouraging, but a certain degree of complexity in setting up the system was highlighted by participants. In particular, the adjustment of volumes of the own vocal sounds, the sounds of the app, and the sounds arriving from the rest of the choir required some practice in order to use the system with confidence and achieve expressiveness. Moreover, an issue was identified in the lack of stable and constant synchronization between the audio and visual content, which sometimes hampered choristers to sing in synch and properly follow the conductor. With respect to this, it is worth noticing that these results were conducted in ideal conditions, i.e., over a low-latency and high bandwidth network. It is therefore plausible to expect that in real-world conditions the perceived delay between audio and visual content could be noticed even more. Therefore, there is a need for progressing both compression and transmission methods for low-latency audio-visual content as well as for advancing network infrastructures capable of ensuring low-latency communications.

Overall, we believe that the results presented in this paper are useful to provide directions to designers of mobile-based interactive systems for electro-acoustic choir singing practices, in co-located, remote, and mixed settings. However, it is worth noticing that our study has some limitations. Firstly, a small sample size was involved, although we performed multiple in-depth sessions with the involved choristers. We plan to continue our evaluation of the system by involving more and larger choirs. Secondly, the found technical difficulties did not allow us to properly test the third system, although it is plausible to derive useful conclusions about it from the results of the co-located and remote architectures. The actual implementation of the envisioned mixed architecture at present remains impractical due to the lack of a proper technological infrastructure: a networked music performance system capable of providing several independent channels is required to render the various remotely connected choristers as physically present with the rest of the co-located choir. This calls for more research in this space. Thirdly, in principle, it is possible that the reported results have been confounded by a novelty effect. However, the app was evaluated only after a series of design iterations following a participatory design approach. Participants were exposed several times to the system before evaluating it in its final version. Our methodology, therefore, reduced the potential for a novelty effect. Moreover, a learning effect could have occurred in the second experiment as the same participants were exposed to the app during the first experiment. However, such a learning effect would have concerned only the app interface, not the overall experience of interacting remotely over the network with a NMP system, which was our aim with the evaluation performed in second experiment.

In future work, we also plan to use the developed application in actual performances, both co-located and remote, and involving choirs with a higher number of choristers. Another interesting research direction, emerged from the comments of participants, concerns the use of the app by the audience, an application falling in the remits of the technology-mediated audience participation field [27, 29]. The app could be used by the general public to provide an accompaniment to the choristers, creating a background sound texture layer over which choristers sing with or without the app. Composers could also create pieces where the audience would have a soloist part alternating with the choir. This could entail both co-located and remote

settings where audience members interact with each other and with the choir. The web-based distributed nature of the developed system supports such scenarios even for hundreds of participants. For these novel art forms, it will be important to conduct further formal evaluations of the success of the interfaces in creating a rich musical experience.

Declarations

Conflict of interest The authors declare no competing interests.

References

1. Gaye L, Holmquist LE, Behrendt F, Tanaka A (2006) Mobile music technology: Report on an emerging community. In: Proceedings of the conference on new interfaces for musical expression, pp 22–25

2. Wang G (2009) Designing smule's iPhone ocarina. In: Proceedings of the international conference on new interfaces for musical expression, vol 291–296

3. Oh J, Herrera J, Bryan NJ, Dahl L, Wang G (2010) Evolving the mobile phone orchestra. In: Proceedings of the international conference on new interfaces for musical expression, pp 82–87

4. Wyse L, Subramanian S (2013) The viability of the web browser as a computer music platform. Comput Music J 37(4):10–23

5. Tanaka A (2004) Mobile music making. In: Proceedings of the conference on new interfaces for musical expression, pp 154–156

6. Poirier-Quinot D, Matuszewski B, Schnell N, Warusfel O (2017) Nü soundworks: Using spectators smartphones as a distributed network of speakers and sensors during live performances. In: Proceedings of the Web Audio Conference

7. Lambert JP, Robaszkiewicz S, Schnell N (2016) Synchronisation for distributed audio rendering over heterogeneous devices, in html5. In: Proceedings of the web audio conference

8. Wang G (2014) Ocarina: designing the iPhone's magic flute. Comput Music J 38(2):8–21

9. Rottondi C, Chafe C, Allocchio C, Sarti A (2016) An overview on networked music performance technologies. IEEE Access 4:8823–8843

10. Essl G, Lee SW (2017) Mobile devices as musical instruments - state of the art and future prospects. In: Proceedings of the international symposium on computer music multidisciplinary research

11. D'Alessandro N, Pon A, Wang J, Eagle D, Sharlin E, Fels SS (2012) A digital mobile choir: joining two interfaces towards composing and performing collaborative mobile music. In: Proceedings of the conference on new interfaces for musical expression

12. Bell J, Matuszewski B (2017) SMARTVOX – a web-based distributed media player as notation tool for choral practices. In: Proceedings of TENOR 2017

13. Alwes CL (2015) A history of Western Choral Music, vol 2. Oxford University Press, Oxford

14. Lee SW, Taylor B, Essl G (2019) Interactive music on the web. In: Filimowicz M (ed) Foundations in sound design for interactive media: a multidisciplinary approach, pp 200–228. Routledge

15. Roberts C, Kuchera-Morin J (2012) Gibber: live coding audio in the browser. In: Proceedings of the international computer music conference

16. Roberts C, Wakefield G, Wright M (2013) The web browser as synthesizer and interface. In: Proceedings of the international conference on new interfaces for musical expression, pp 313–318

17. Clark CBD, Tindale A (2014) Flocking: a framework for declarative music-making on the web. In: Proceedings of the sound and music computing conference, pp 1550–1557

18. Smus B (2013) Web audio API: advanced sound for games and interactive apps. O'Reilly Media Inc.

19. Turchet L, Fischione C, Essl G, Keller D, Barthet M (2018) Internet of musical things: vision and challenges. IEEE Access 6:61994–62017

20. Gabrielli L, Squartini S (2016) Wireless networked music performance. Springer, Berlin

21. Keller D, Lazzarini V (2017) Ecologically grounded creative practices in ubiquitous music. Organised Sound 22(1):61–72

22. Migicovsky A, Scheinerman J, Essl G (2014) Moveosc - smart watches in mobile music performance. In: Joint proceedings of the international computer music conference and the sound and music computing conference, pp 692–696

23. Matuszewski B (2020) A web-based framework for distributed music system research and creation. J Audio Eng Soc 68(10):717–726

24. Turchet L (2019) Smart musical instruments: vision, design principles, and future directions. IEEE Access 7:8944–8963

25. Turchet L, Barthet M (2019) An ubiquitous smart guitar system for collaborative musical practice. J New Music Res 48(4):352–365

26. Hödl O, Kayali F, Fitzpatrick G (2012) Designing interactive audience participation using smart phones in a musical performance. In: Proceedings of the international computer music conference

27. Xambó A, Roma G (2020) Performing audiences: Composition strategies for network music using mobile phones. In: Proceedings of the conference on new interfaces for musical expression. Birmingham City University

28. Weitzner N, Freeman J, Garrett S, Chen YL (2012) massMobile-an Audience Participation Framework. In: Proceedings of the conference on new interfaces for musical expression, pp 21–23

29. Hödl O, Bartmann C, Kayali F, Löw C, Purgathofer P (2020) Large-scale audience participation in live music using smartphones. J New Music Res 49(2):192–207

30. Barbosa Á (2003) Displaced soundscapes: a survey of network systems for music and sonic art creation. Leonardo Music J 13:53–59

31. Weinberg G (2005) Interconnected musical networks: toward a theoretical framework. Comput Music J 29(2):23–39

32. Cáceres JP, Chafe C (2010) Jacktrip: under the hood of an engine for network audio. J New Music Res 39(3):183–187

33. Drioli C, Allocchio C, Buso N (2013) Networked performances and natural interaction via lola: low latency high quality a/v

streaming system. In: International conference on information technologies for performing arts, media access, and entertainment, pp 240–250. Springer

34. Hupke R, Nophut M, Preihs S, Peissig J (2021) Toward professional distributed performances: effects of a global metronome on networked musical ensemble interactions. J Audio Eng Soc 69(10):720–736

35. Battello R, Comanducci L, Antonacci F, Cospito G, Sarti A (2021) Experimenting with adaptive metronomes in networked music performances! J Audio Eng Soc 69(10):737–747

36. Verma P, Mezzay AI, Chafe C, Rottondi C (2020) A deep learning approach for low-latency packet loss concealment of audio signals in networked music performance applications. In: 2020 27th Conference of open innovations association (FRUCT), pp 268–275. IEEE

37. Mitchell T, Madgwick S, Rankine S, Hilton GS, Freed A, Nix AR (2014) Making the most of wi-fi: optimisations for robust wireless live music performance. In: Proceedings of the conference on new interfaces for musical expression, pp 251–256

38. Turchet L, Fischione C (2021) Elk Audio OS: an open source operating system for the Internet of Musical Things. ACM Trans Internet of Things 2(2):1–18

39. Boulez P, Noakes D, Jacobs P (1964) Alea. Perspectives of new music. 42–53

40. Brooke J et al (1996) Sus–a quick and dirty usability scale. Usability Evaluation in Industry 189(194):4–7

41. Cherry E, Latulipe C (2014) Quantifying the creativity support of digital tools through the creativity support index. ACM Trans Computer-Human Interaction 21(4):21

42. Braun V, Clarke V (2006) Using thematic analysis in psychology. Qual Res Psychol 3(2):77–101