Designing a system for Online Orchestra:

Peripheral equipment

David Prior | Philip Reeder | Michael Rofe | Ian Biscoe | Samuel Murray

Abstract

Online Orchestra is a telematic performance project, aimed at enabling young and amateur musicians in geographically remote locations to make music together over the Internet. This article reports the processes by which the audio and video peripheral equipment used for Online Orchestra was chosen and how the system was designed and used. Starting with an overview of guiding design principles, a description of methods for choosing, integrating and configuring audio and video hardware is presented. Following the development of the project from initial workgroups to the pilot performance of Online Orchestra, this article compares the 'ideal' test scenarios of workgroups with the reality of deploying the technology in a performance context and concludes with an account of using the system on site.

Keywords

Online Orchestra; microphones; speakers; cameras; screens; echo

Introduction

Online Orchestra involved a telematic music system that integrates a range of existing technologies drawn mainly from live sound engineering and network communication. With the exception of Online Orchestra's latency-control programme (see Rofe and Reuben 2017), most of the technologies used were not new in themselves. However, as equipment is brought into a new formulation, each element interacts with others often in new, unpredictable ways, each playing a role in the overall experience of telematic performance. The aim was to design a system for a pilot performance of Online Orchestra, involving musicians in four locations around Cornwall, United Kingdom: strings and female voices in Truro Cathedral, brass in Mullion on the Lizard Peninsula, flutes in Five Islands' School, Isles of Scilly, and a conductor at Falmouth University.

The majority of texts on telematic system design focus on computing and networking, with less detailed reference to peripheral equipment such as microphones, speakers, cameras and screens (see, for instance Bouillot and Cooperstock 2009; Braasch 2009; Meier 2013). However, peripheral equipment clearly plays a significant role in enabling a high-quality experience for users (see Braasch et al. 2009; Chabot 2016; Naugle 2002). This article reports on peripheral equipment used for Online Orchestra, including details of a range of trials undertaken during the design phase of the project, and the decision-making process that led to the final design solution deployed in the Online Orchestra pilot performance (see Rofe et al. 2017b, for an overview; see Rofe et al. 2017a, for an evaluation by participant performers). In all cases, decision-making was informed either by the specific requirements of the project, as detailed in the design principles section below and in Rofe et al. 2017b, or through a process of action research (see Kolb 1984), often involving participant musicians, in which iterations were made, reflected upon and then reiterated, until a suitable solution was reached. The article concludes with a detailed report on the challenges faced and solutions developed, by the local technicians during rehearsals and the final performance.

Design principles

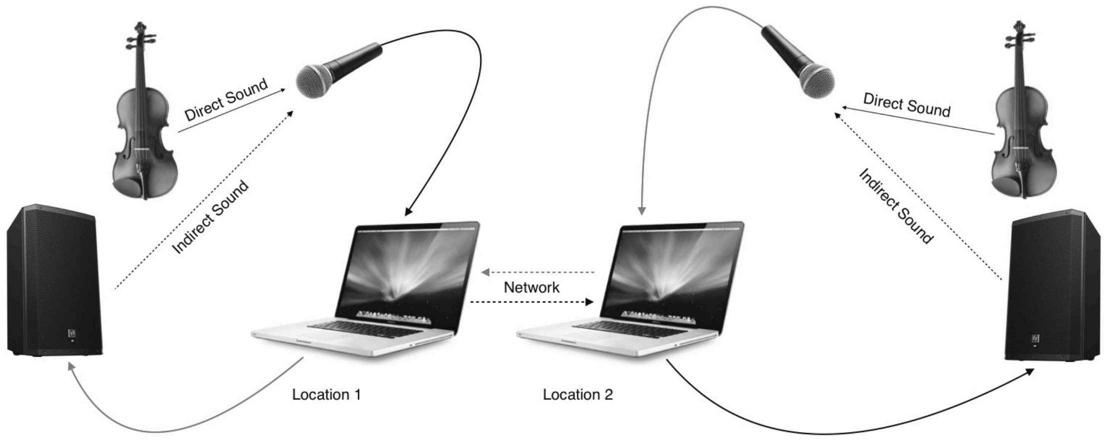
The design for Online Orchestra's audio-visual peripheral system was continually benchmarked against two primary criteria: (1) to create a sense of connection between musicians across nodes and (2) to create a sense of immersion in the musical experience. These benchmarks emerge from a range of starting premises within the Online Orchestra project overall, as detailed in Rofe et al. 2017b. As such, a starting assumption was that the higher the quality of audio and visual streams, the greater the sense of connection and immersion would be. However, as described in Prior et al. 2017, an important feature of Online Orchestra was that it would operate across standard, domestic broadband connections. With this in mind, data bandwidth was of primary concern from the outset of the project: higher quality streams consist of larger amounts of data. As such, various trade-offs emerged, such as that between the benefits afforded by using a greater number of individual audio channels and the requirement that this would bring to operate the system at a lower audio resolution. Similarly, the more bandwidth that was used for audio, the less would be available for video. The need to balance these and similar equations became the point of departure for much of the decision-making.

One particular consideration was the need for clarity in the video feed of the conductor: Online Orchestra's pilot performance had a fairly conventional format, in that it required musicians (albeit distributed musicians) to follow a single conductor. This conductor was located remotely, meaning musicians would need to follow a televisual feed. Understanding what musicians needed to see and identifying and obviating impediments to them doing so – particularly in regard to their following the gestures of the conductor – became an important point of focus in the research. A final consideration was equipment cost: as described in Rofe et al. 2017b, Online Orchestra aimed to design a solution that was repeatable and scalable, so preference was given to low-cost equipment or equipment that potential users might already own.

Audio peripherals

Online Orchestra is based on the premise of instrumental sounds being amplified across a data network in multiple locations. As each location receives instrumental sound through a microphone but also plays the sound from the other locations through a loudspeaker, the possibility for feedback within the system is high because microphones will receive not only the direct sound from the instrument but also indirectly pick up sound from the loudspeaker: see Figure 1.

Figure 1: Direct and indirect sound and the possibility of feedback.



However, in a network environment in which there is latency (time delay; see Rofe and Reuben 2017) between locations, this will manifest as echo, due to the time delay involved in sending data over the network. This enables the listener to hear the original sound and the return from the feedback cycle individually, as opposed to the characteristic 'howl' that is heard when feedback occurs within the instantaneous closed loop of a conventional public address system. Whilst some precedent telematic performance projects approach this challenge through digital signal processing,[1] Online Orchestra adopted basic audio engineering solutions to the challenge of echo, involving microphone and speaker choice and placement and appropriate gain structure. Whilst this approach did not eliminate echo altogether, it did reduce its level to the point where it is masked by other audio information. Managing echo was therefore a core objective in designing the audio peripheral system.

Microphone testing: Audio quality

As shown in Figure 2, initial testing of microphones began by recording a student ensemble consisting of three violins, clarinet and saxophone at Falmouth University using a range of microphone arrays. The main purpose of this exercise was to establish the relative benefits of different microphone types (dynamic, condenser), brands and polar patterns, arranged in different arrays and proximities to the performers. In addition to observing the technical performance of different arrays, a group of 39 students were asked to evaluate the quality of the different recordings, according to a number of characteristics: this experiment and its findings are reported in detail in Geelhoed et al. 2017, in this special issue.

The decision to start testing using 'offline' recordings, with a broad range of microphone types and arrays, was based on the need to establish first a microphone array based on optimum audio quality (as defined by the listening tests) as this would provide a benchmark against which to compare any compromises that had to be made later. Although the distant microphone arrays gave rise to some of the most natural results and were the most successful in conveying the quality of the performance space, it was felt that for Online Orchestra's immediate purposes, priority should be given to the microphones that mediated the instrumental sound with the least influence from the performance environment. In this initial stage of Online Orchestra's development, this was the most direct route to ensuring that each node of the ensemble could integrate with every other node. Giving more emphasis to the disparity between the acoustics of different nodes has been an important theme in several recent telematic performances, with pieces such as Pauline Oliveros's *Dynamic Spaces* (Oliveros et al. 2007) that seek to emphasize and manipulate the acoustic disparities between the different nodes of a network. Andrew Hugill also identifies network space as a ‘new frontier’, which affords the opportunity to create virtual- and mixed-reality spaces (Hugill 2012: 81). Chris Chafe goes further, however, by exploring the idea of the network as a new kind of sound propagation medium, with its own distinct ‘acoustic’ properties (Chafe 2009).[2]

Figure 2: Initial microphone test.

However, the focus for Online Orchestra was different, and during this first phase of the project, priority was given to capturing instrumental sources with as little colouration from their room acoustics as possible. Distant microphone arrays would present significant problems with regard to feedback and, for these reasons, any use of ambient microphones was ruled out in favour of spot microphones. For the smaller ensembles that made up certain nodes in the Online Orchestra pilot performance, one microphone per player was used, and for the larger ensembles, one microphone per two to four players was used.

Based on the results of the listening tests (see Geelhoed et al. 2017) and the principles outlined above, it was decided that two types of microphone warranted further field testing to ascertain their behaviour in a live environment. The first of these was a standard dynamic stage microphone with a cardioid polar pattern. The model used in the listening test was a Sure SM57, and, as detailed in Geelhoed et al. 2017, this significantly outperformed expectations. Given that the SM57 is cheap, commonly available and extremely durable, it was felt that a microphone of this type should be taken forward to the next round of tests. However, despite its positive performance in the first test, and its ubiquity in live sound environments, the SM57's frequency response is far from flat and its output is relatively low. Alternative dynamic microphones were therefore explored that might exhibit the SM57's relatively narrow pickup pattern (minimizing both feedback and colouration from room acoustics), its durability and its relative low cost. For these reasons, the Electro-Voice N/D967 was chosen: a dynamic microphone that exhibits remarkable feedback rejection, has both a higher output and flatter frequency response than the SM57 and is only marginally more expensive.

Another microphone that performed well in the initial test was the DPA VO4099 super-cardioid, clip-on microphone. In fact, this microphone outperformed the SM57 in most categories, though only marginally. However, the DPA VO4099s were too expensive to consider, given the ambition to create a scalable design solution. For this reason, the Sontronics STC-1 was chosen as a small diaphragm condenser microphone to take forward for further testing: it is one of the only microphones of its type, at a relatively low budget, and available with a hyper-cardioid capsule (necessary to limit spillage between instruments and to avoid capturing too much of the loudspeaker signal).

It should be noted that in broadcast, recording and live amplification contexts, a wide variety of microphones are usually deployed specific to the instruments being used. For Online Orchestra, however, part of the aim was to create a scalable infrastructure that could ultimately be used by non-professionals without the need for complex additional peripheral equipment. Having established a benchmark for audio quality through the first tests, the aim was to arrive at a solution for microphone capture that could be used across a broad range of instruments.

Microphone testing: Source isolation and feedback rejection

In a second round of tests, limited to the Electro-Voice N/D967s (henceforth referred to as the 'EVs') and the Sontronics STC-1s with hyper-cardioid capsules (henceforth 'Sontronics'), relative gain-before-feedback performance was investigated by setting up both microphones above a 'wedge' stage monitor in free field with identical gain structures. With no acoustic input to the microphone, the EVs performed noticeably better, but with spoken word tests, their performance was more similar. Although the Sontronics tended to feed back slightly sooner than the EVs, their feedback rejection performance was better than expected overall. At the threshold where feedback started to occur, it was also noted that the Sontronics exhibited a far narrower resonant peak than the EVs, and this was consistent across a number of acoustic spaces. With the judicious use of narrow-Q notch filtering performed on the mixing desk, the Sontronics were able to operate at levels quite close to the EVs. As is characteristic of condenser microphones, the Sontronics had an ‘airy’, open sound with more high-frequency presence, and their higher output meant that they could operate at lower gain levels, thus significantly reducing system noise on quieter sources. With this said, the behaviour of the EVs was more consistent when used with different sound sources, less dependent on positioning and proximity to the sound source, and less influenced by the acoustic of the room in which it was being used. As well as being physically more robust, the EVs were also capable of withstanding higher sound pressure levels before distorting, making them more suitable for louder sources.

Having established something of the behaviour of the microphones through the tests outlined above, a third test was initiated that investigated the two shortlisted microphones in the context of a networked performance. In the first instance, the network in question was a local area network (LAN) created within the *Academy of Music and Theatre Arts* at Falmouth University. Brass, woodwind and string players in three rooms were linked by a prototype of the Online Orchestra software, with a conductor in a fourth room. Each player was mic’d with both a Sontronics and an EV microphone, and where there was more than one player per room, signals from both players were mixed to create one, mono Sontronics mix and one, mono EV mix from each room. Each room also had one video camera to capture all of the players, three video screens and three loudspeakers (one for each of the other spaces), with the option to choose between the Sontronics and the EV mix from each of the sending spaces. In addition to the observations made on the day by the performers themselves and members of the Online Orchestra team, multichannel recordings were made for later reference.

On the basis of subjective evaluations of this test, and due to the respective technical benefits of each microphone outlined above, it was decided to proceed with using both microphone types for the remainder of the project. In the pilot performance, Sontronics microphones were used for the flute ensemble on the Isles of Scilly, where their higher output and more transparent 'top end' were advantageous; EV microphones were used for the brass ensemble in Mullion, where their ability to withstand higher sound pressure levels and their greater independence from room characteristics was a significant advantage. In Truro, the larger ensemble was captured by a mixture of microphone types according to the characteristics of the instrument and its location in relation to loudspeakers and other instruments.

Loudspeaker testing

The recording from test one above featured both distant 'ambient' microphone techniques and discreet, per-instrument spot microphones. These discreet multichannel recordings were later used as the basis for testing various loudspeaker arrays. In the same room as the recordings were made, five Genelec 8040 loudspeakers were arranged in the same positions as the performers had played. Either side of this speaker array was placed a pair of Electro-Voice ZLX-12P PA speakers on higher PA tripod stands. At the centre of the Genelec array was placed a Bose L1 articulated line array loudspeaker. The speakers were calibrated to ensure that playback level was the same for each configuration. The recordings made with the DPA VO4099 hyper-cardioid spot microphones were used throughout the listening tests as these were felt to be the higher quality of the two discreet recordings that were available and had the least amount of crosstalk between channels. The Bose L1 speaker received a mono mix, the Electro-Voice PA received a stereo mix and the Genelec five-speaker array received the discreet recordings of each individual instrument. Other than dynamically balancing and panning in the mono and stereo mixes, no other processing was applied.

The differences in speaker design meant that not only were mono vs. stereo vs. multichannel arrays being compared but also the loudspeakers themselves. However, the exercise was useful nevertheless.[3] It should be noted though that the Genelec 8040 speakers are studio monitors designed for critical listening at relatively close distances. At distances greater than 2–3m, they become relatively diffuse and this made them particularly suitable for performing the role of 'absent performers', as they blended with both the room acoustic and each other more than the Electro-Voice and Bose speakers. The Electro-Voice speakers have a conventional 120-degree radiation pattern and therefore retain a stable stereo image further back into the listening space than the Genelecs. The Bose L1 has an unusually wide 180-degree horizontal dispersion pattern and is designed specifically to sound consistent even when listening at the extreme sides of the loudspeaker, making it an ideal choice for the mono mix.

The core members of the Online Orchestra team performed listening tests in which the recording made with the DPA VO4099s was listened to through each speaker array in turn. It was agreed by all members of the research team that the five-speaker Genelec array provided the most life-like presentation of the performance. Interestingly, however, there was a marked difference of opinion as to whether the next best option was the stereo mix or the mono mix. One listener expressed the sentiment that if five-channel playback was not an option, he would rather hear an 'honest' mono reduction than an artificial sounding stereo approximation of the performance. It may well be that the juxtaposition with the five-channel playback led to a different ranking between the mono and the stereo playback configurations, but the overall consensus erred towards the mono system in second place.

Budget and data bandwidth did not allow for multichannel transmission from each node, so, for the Online Orchestra pilot performance, spot microphones were mixed live and summed to a mono feed sent from each node. Nodes then featured a single Electro-Voice ZLX-12P loudspeaker to represent the audio being sent from the other locations, with speakers situated underneath corresponding video screens. The exception to this was in Truro Cathedral: the venue in the pilot performance with both the largest ensemble and the largest audience. Here, a rather larger and more powerful speaker system was needed, although it too was designed according to the same underlying principles outlined above.

Mixing desk and audio interface

In the context of the rest of the audio system, requirements for a suitable mixing desk and audio interface were relatively generic and straightforward, with the primary requirement of enabling multiple live inputs and sufficient returns and outputs to route audio from other locations. Allen & Heath's MixWizard 14:4:2 provided this flexibility in terms of input channels, auxiliary and bus sends. A MOTU Ultralite MKIII was chosen as one of the cheapest products that offer the necessary number of I/O channels.

Taking the example of Five Islands' School, which contained a group of eight flautists in the pilot performance, the first eight inputs of the mixing desk were used for the live microphones, one per instrument. Gain levels were set so that peaks would reach around −6dB, allowing for plenty of headroom when signals were mixed. Each microphone channel was fed to Aux 1, and it was this mono signal that was sent to the other locations via the MOTU. As Aux 1 is a post-fade send on the Allen and Heath desk, the amount of signal sent to the Aux bus was relative to the position of the fader. The Aux send pot on each channel was set to ±0dB, and the balance between these inputs was set by means of the faders. The master Aux send mix was adjusted so that, once digitized at the audio interface, it peaked at −18dBFS.[4] Output busses 1–3 were used to feed three local speakers corresponding to the return audio from the other three locations. As channels 1–8 were not routed to the speakers, headphones were used to monitor the mix of the live instruments.

Inputs 9–12 of the mixing desk were used to receive the audio returns from the server, via the MOTU. Channel 9 was dedicated to talkback from the engineers at the other venues and was again routed to headphones. Channel 10 contained a metronome signal that corresponded to the latency, for the benefit of system calibration and rehearsals. Channels 11–13 carried the mono audio feeds from the other three nodes and were sent to busses 1–3 in order to feed the speakers. In order to keep the gain structure optimized, fader levels for channels 9–13 were kept around 0dB with listening levels controlled by means of the bus output faders.

Video peripherals

In broadcast standards such as those set by the Advanced Television Systems Committee (ATSC),[5] notions of 'high definition' assume notional parity between video and audio. Given the primacy of audio in musical performance, greater focus was placed throughout the Online Orchestra project to optimizing the audio quality over the video quality: notionally, it is less significant that musicians can see one another at high resolution than it is that they can hear each other at high resolution. Moreover, the data size required to stream high-definition video would be a significant challenge in the context of available bandwidths at the remote locations used for the pilot performance (see Prior et al. 2017). The exception to this was in the case of the conductor, whose image needed to be captured and reproduced clearly in order to enable musicians to see gestures clearly. This resulted in an asymmetrical architecture in which the conductor was realized at higher resolution than the musicians, in order to keep overall bandwidth usage down. Time and budgetary constraints meant that less detailed testing and evaluation of video peripherals took place than was the case for their audio counterparts. More research is needed in this area, as video clearly has the capacity to contribution to achieving a sense of connectedness and immersion in the musical experience.

Video acquisition: Cameras and interfaces

An early decision was taken to use SDI (serial digital interface) as the means of connectivity between cameras and capture devices. SDI uses robust connectors and cables and allows cable runs of up to 50m, making it easy to position cameras optimally, rather than having to locate them near to the main computers. SDI is also well supported by a wide range of industry standard solutions for video acquisition, routing and computer interfaces. For Online Orchestra's development workshops, and its pilot performance, two types of camera were used across the four locations: a Sony EVI-HD1 was used in Truro and Mullion, and a Black Magic Cinema Camera (BMC) was used in Falmouth and the Isles of Scilly. The Sony EVI-HD1 is a professional videoconferencing pan/tilt/zoom camera designed for large meeting rooms. It is capable of resolutions from SD PAL/NTSC up to HD 1080i/59.94 and supports live SDI output at all of these resolutions. The BMC is a lower cost professional quality camera primarily orientated towards independent cinema makers and can capture at resolutions of up to 2.5K in frame rates from 23.98 to 30. The live SDI output is fixed at either HD 1080i/25 or 1080i/50. In the pilot performance, bandwidth requirements meant the need to compress the video, with the conductor being broadcast at 720p and the three groups of musicians at 480p. As such, the cameras significantly outperformed requirements in the final system.

Several different SDI solutions were used to connect the cameras with the computing platforms. A Black Magic Decklink Quad was used in Truro Cathedral. This interface provides acquisition of up to four SDI channels on a single slot PCIe card, allowing switching between camera feeds if required. In other locations, a Black Magic SDI Intensity Shuttle was used. This is a relatively low-cost external unit that captures a single SDI stream and makes it available to the host computer via a USB3 interface. The device also has SDI and HDMI outputs for monitoring of the SDI input. It is an ideal solution for laptops or host PCs that do not have a spare PCIe slot available.

As detailed in Rofe et al. 2017a, in this special issue, participants in the pilot performance noted that they would have preferred greater clarity in the image of other musicians. This could be achieved by increasing resolution given suitable bandwidth, or through the use of an interface such as the Black Magic Decklink Quad, which would enable multiple, switchable camera feeds. However, given limited bandwidths in the locations used for the pilot performance, particularly upload limits in several remote locations, lower video resolution was a necessary compromise.

Video display configuration

Again, following the design principles described above, trials were undertaken to decide upon display screen configuration. In particular, the aim was to establish musicians' preferences of (1) size of screen and (2) display screen configuration. Four rooms at Falmouth University were selected for the trial, all benefitting from access to a CAT6 LAN independent of the University's Internet connection. Similar to the ideal conditions in which microphone tests began, the University's LAN enabled initial experimentation with different display configurations in a 'control environment' where bandwidth was not an issue. Five participants took part in the trial: a conductor and four undergraduate musicians from Falmouth University, a vocalist, a violinist, a flautist and a saxophonist, thus modelling the final pilot performance. Each room was prepared with a different display screen configuration:

1. a single, large projection screen, containing a large image of the conductor and small tiles of the three musicians to the side

2. a single, large projection screen, split into equal quadrants, containing equal-sized images of the conductor and three musicians
3. three large projection screens, with each screen dedicated to the video feed of one of the three remaining rooms
4. three 40-inch flat-screen televisions, with each screen dedicated to the video feed of one of the three remaining rooms

A short piece of music was written by one of the project composers, and this was performed by the four musicians and the conductor. The conductor, flautist and saxophonist were assigned to one room each, with the vocalist and violinist sharing a final room, modelling the scenario in Truro Cathedral in the pilot performance. Participants then rotated through the rooms, performing the same piece on each rotation, until all participants had experienced performing using all four display screen configurations. Audio was streamed through JackTrip and video through VSee throughout the trial, with no changes to software settings. Following all rotations, participants met together with members of the project team for an informal discussion on the relative benefits and drawbacks of each configuration, particularly in relation to benchmarks of quality, perceived sense of connectedness and immersion, and invasiveness of the technology. These discussions were recorded and then transcribed.

The issue of image size emerged quickly in the discussion, with mixed opinions expressed on the issue. The size of the conductor was particularly crucial: it was felt especially in configuration 1 (large conductor, tiled musicians) that the disproportionality of size placed too much focus on the conductor, reducing the sense of connection between musicians; one musician noted that this felt ‘a bit alienating to me; I felt a little bit uncomfortable' (musician 2). It was also felt that, when the conductor was rendered larger than real life, it made movements somewhat hard to follow, as peripheral vision was not sufficient to capture the full horizontal strokes of the conductor's arm. Conversely, regarding the smaller screens of configuration 4, musician 4 stated that she 'could barely see the conductor […] you would have to be more exaggerated in the motions' (musician 4). Of the configurations tested, none were considered ideal with respect to size; rather, all participants agreed that something between the large and small screen size for the conductor might work well, ideally of a size similar to real life. With respect to size of musicians, concern was expressed by all participants over the size of musicians in the small tiles of configuration 1, with a preference emerging for larger image sizes. Musician 1 noted that larger screens would be particularly effective for larger ensembles: 'if you had a whole section of an orchestra large screens would be really good' (musician 1).

With respect to image arrangement, configuration 2 (quadrants) was preferred over configuration 1 (large conductor with tiles of musicians) in terms of generating a sense of connection between musicians: 'We were all the same size. It's more like a team working thing: I could see everybody as much as I could see the conductor' (musician 2). Likewise, musician 4 'really liked the four tiles; it's made a huge difference to me, a real sense of connection' (musician 4). However, as musician 1 noted, 'you still focus on the corner with the conductor in more' (musician 1), which, given the smaller size of the conductor that resulted from equal partitioning, in turn created some of the challenges described above.

Opinion was split with respect to the number of screens: configurations 1 and 2 using a single screen, split into component images, and configurations 3 and 4 having a screen per image. The key differentiating factor seemed to be related to having to turn to see the screens, and this was particularly an issue when the screens were large or positioned close to the performers. As such, musician 2 preferred a single screen 'because everything is together in one place' (musician 2). By contrast, musicians 3 and 4 had a strong preference for three screens. Describing the experience as 'immersive', musician 4 found three screens particularly engaging: 'it was fantastic; that was an absolutely amazing experience. It made it exciting' (musician 4). For musician 4, 'Four tiles was more "workhorse": it worked and you were able to see everybody equally, but it just didn't give you that "wow factor"' (musician 4).

Although opinion was split with respect to image arrangement, issues of image size were perceived to be more problematic in the context of configurations 1 and 2: with three screens, there is more flexibility to determine the size of conductor and musicians independently. Additionally, a three-screen arrangement enabled spatialized sound more effectively: as described above, it was possible to position speakers below their respective screens, enabling a sense of width in the sound; having images of nodes tiled on a single screen did not allow for this effect. As such, it was decided to take the three-screen arrangement forwards into the final pilot performance, not least because of the sense of immersion participants reported that this arrangement created.

In Falmouth, Mullion and the Isles of Scilly, 50-inch flat panel displays were used, enabling a slightly larger image than that used in the aforementioned trial, but using equipment that was not too costly (one of the aims of Online Orchestra being the design of a scalable system). In Truro Cathedral, the size of the venue required a slightly different approach. Three screens were still used, but these were large, projected screens, in order to enable audience, as well as musicians, to see streamed video content. However, following the feedback from participants in the trial, a smaller screen was used for the conductor, such that the size of the image remained roughly equivalent to real life.

Using Online Orchestra: The node technician

Having established the peripheral equipment to be deployed in Online Orchestra, it fell to a technician in each node to optimize and monitor that equipment during rehearsals and the final pilot performance. The technician was responsible for ensuring the quality of the audio-visual signal from their local node; maintaining the audio-visual signal flow between the local and remote nodes; acting as an intermediary with participants at other nodes when required; and liaising with other remote technicians.

Audio engineering

In terms of audio, remote technicians had to balance their microphones, creating a single mono mix, which was then streamed to the other nodes. As technicians were situated in the performance space – and therefore party to the live sound of the ensemble – they were not ideally located to perform this task. However, with the benefit of high-quality, closed-back headphones, reference to the scores, and communication with receiving nodes, it was possible to create adequate mixes on site. Remote technicians initially created a mix that was neutral in their headphones and responded to requests from other remote technicians (e.g. 'please send less bass') to enable a more desirable signal at the receiving end. Remote technicians in each node were also responsible for adjusting the timbral equalization of the audio streams they received according to the acoustic characteristics of the venue in which they were working.

Typically, a live sound engineer is concerned with the amplification of a local sound source, to a local audience. In a telematic system like Online Orchestra, the familiar challenges of balancing multiple sources and correcting suboptimal room acoustics remain, but new problems also arise. First, due to the distributed nature of Online Orchestra, with sub-mixes being sent from each venue, the venues to which that sub-mix is sent are different, and indeed unknown to the technician who mixed it.[6] Considering the fact that several of these 'foreign' mixes are layered on top of one another in each venue, the problem starts to multiply. A second issue, related to the first, is the challenge of distributed echo. Earlier in this article, it was seen that the phenomenon of feedback, which occurs almost instantaneously in the 'closed loop' of a microphone or pickup and loudspeaker, manifests as an echo once the parameter of latency is introduced into the equation. As each iteration of the echo multiplies any resonant frequencies in either the sending or receiving nodes, the dangers identified in the previous first challenge are multiplied further.

Distributed echo had a number of consequences and was particularly difficult to control in Mullion, where the dynamics of the acoustic sound of the brass section and the amplified sound of the remote nodes were difficult to balance against one another. Indeed, minimizing distributed echo while providing enough level for the local brass ensemble to hear the remote nodes over their own sound was perhaps one of the greatest challenges for the remote technicians. In a typical live sound engineering scenario, the live sound level is set at least 6dB below that at which feedback occurs (Davis and Jones 1990: 47). In Online Orchestra, however, due to the absence of a live source – which would usually be the cause of the feedback in a live amplification context – echo became perceptible at more like 20–30dB below the overall listening level.

Typical live sound engineering techniques for feedback prevention continued to play a role here, and the maximization of what Davis and Jones refer to as 'acoustic gain' (Davis and Jones 1990: 49): the degree to which an acoustic source can be maximized before it is amplified. In the case of recording brass instruments, microphones are often placed some distance from the bell, but in order to maximize acoustic gain, the microphone needed to be placed considerably closer. Similarly, great care was taken to maximize the distance between loudspeakers and microphones and to ensure that microphones were as close to pointing directly away from speakers as possible.

In a small ensemble context, any loudspeaker would ideally act both as a public address for the audience and as a monitor for the performers. At Five Islands' School on the Isles of Scilly, this was possible and a single speaker was placed beneath its corresponding video screen, angled up 45 degrees. This provided enough coverage for both audience and performers and made for both a simpler technical set up and a radiation characteristic more similar to that of the acoustic ensemble it was designed to represent. However, in both Truro (due to the size of the venue and the number of performers and audience members) and Mullion (due to the acoustics of the building and the difficulties of balancing the brass ensemble against the amplified sound of the remote nodes), this was not possible. In these venues, additional sets of speakers were set up for the audience as a necessary compromise. However, this did present new problems of its own in that whereas the priority for the audience was to achieve a good balance overall, the performers often needed to hear the sound of the other nodes at a disproportionately high level over their own unamplified sound in order to hear musical cues.

Distributed dynamics

In this distributed performance context, it was all too easy for musicians to adjust their dynamics to the perceived level of the other ensembles, rather than adhering to the score and the cues given by the conductor. If the level of the other ensembles was too loud or too quiet, the local musicians would tend to compensate, sometimes without knowing they were doing so. It was therefore essential that the remote technician created a balance for the musicians that felt natural for them, as failure to do so could subject the whole system to the familiar multiplying effects described earlier in relation to distributed echo: musicians in one node, perceiving themselves to be too loud in relation to the others, play more quietly; musicians in the other nodes subsequently do the same; and so on, creating an overall decrease in volume throughout the orchestra.[7] The solution was relatively straightforward, requiring the technician to mediate between the musical leader of their local ensemble and the conductor, as well as with the other node technicians.

The music commissioned for the pilot performance involved numerous rapid changes in dynamic. Score reading was therefore a useful asset for the node technicians, as this enabled them to ascertain whether the signals being received were appropriate to the expression markings. The remote technician could therefore understand whether dynamic shifts in the material needed compensating for on the receiving mixer or whether changes in balance needed to be made at the sending mixer.

Audio-visual coherence and mediation

Node technicians were also responsible for ensuring that an appropriate visual signal of the ensemble was transmitted and that video signals from sending nodes were received and displayed appropriately. The field of view needed to capture the breadth and depth of the ensemble and ideally be free of any superfluous visual distractions. This was often difficult, particularly in educational or community venues, as multifunctional rooms often have a variety of paraphernalia installed. Mounting the camera on a tripod and pointing downwards towards the players tended to exclude most undesirable elements, while facilitating the inclusion of the ensemble. However, this perspective had to be balanced against the communicative advantage of the conductor clearly viewing the players' faces. It was important for the conductor to be able to view the ensemble leader in particular, as this leader tended to be the primary diplomat for verbal communication between nodes.

Conclusion

Attempts were made throughout the design phase of the project to limit costs, in order to derive a solution that could potentially be scaled, enabling schools and community groups in the future to make use of already available, or relatively cheap, equipment. With the exception of cameras used on the project – which were overly costly and outperformed requirements – the remainder of equipment could be purchased at relatively modest cost and indeed could be scaled down or up depending on the number of musicians in the ensemble.

Trialling of peripheral equipment led to the decision to deploy in each node a single camera, three screens, three speakers and a microphone per instrument. This proved most effective in enabling a high-quality audio-visual experience and also sought to maximize the potential of connection and immersion between and by musicians. Online Orchestra required a wide array of familiar equipment, but the deployment of this equipment in a telematic environment brought about unfamiliar challenges, particularly in the audio domain. For this reason, having technicians in each node who understood not only the basics of sound engineering but also the behaviour of the telematic system overall was vital.

More generally, it can be seen that peripheral equipment can impact significantly upon the experience of telematic performance. With respect to audio quality, microphone choice, placement and usage need careful consideration in order to isolate instruments and avoid echo. Video resolution needs to be sufficiently high to enable clear and smooth movements by the conductor; lower resolution is acceptable, though not ideal, for video streams of musician. Speaker and screen choice/placement can also significantly impact upon the experience of immersion and connection between distributed performers. Online Orchestra established a design solution that enabled its pilot performance, but more research is now possible to optimize this system and, in particular, to define minimum requirements that give rise to meaningful musical experiences on the parts of its users.

References

Bouillot, N. and Cooperstock, J. R. (2009), ‘Challenges and performance of high-fidelity audio streaming for interactive performances’, *Proceedings of the New Interfaces for Musical Expression Conference*, Pittsburgh, PA.

Braasch, J. (2009), ‘The telematic music system: Affordances for a new instrument to shape the music of tomorrow’, *Contemporary Music Review*, 28:4&5, pp. 421–32.

Braasch, J., Chafe, C., Oliveros, P. and Van Nort, D. (2009), ‘Mixing console design considerations for telematic music applications’, *Audio Engineering Society*, http://www.aes.org/e-lib/browse.cfm?elib=15136. Accessed 12 December 2016.

Chabot, S. (2016), ‘Spatialized sound reproduction for telematic music performances in an immersive virtual environment’, *Journal of the Acoustical Society of America*, 140:4, http://dx.doi.org/10.1121/1.4970468. Accessed 12 December 2016.

Chafe, C. (2009), ‘Tapping into the internet as an acoustical/musical medium’, *Contemporary Music Review*, 28:4&5, pp. 413–20.

Chafe, C. and Niemeyer, G. (2001), *Ping*, http://chrischafe.net/portfolio/ping/. Accessed 14 October 2016.

Chafe, C., Gurevich, M., Leslie, G. and Tyan, S. (2004), ‘Effect of time delay on ensemble accuracy’, *Proceedings of the International Symposium on Musical Acoustics (ISMA2004)*, Nara, Japan, 31 March-3 April.

Davis, G. and Jones, R. (1990), *The Sound Reinforcement Handbook*, Milwaukee, WI: Hal Leonard Publishing Corporation.

Geelhoed, E., Prior, D. and Rofe, M. (2017), 'Designing a system for Online Orchestra: Microphone evaluation and cost-benefit analysis', *Journal of Music, Technology and Education*, 10: 2-3, pp. 213-30.

Hargreaves, J. J. (2017), 'Notes from the podium of an Online Orchestra', *Journal of Music, Technology and Education*, 10: 2-3, pp. 277-87.

Hugill, A. (2012), *The Digital Musician*, 2nd ed., Oxford and New York: Routledge.

Kolb, D. (1984), *Experiential Learning: Experience as the Source of Learning and Development*, Eaglewooods New Jersey: Prentice-Hall.

Meier, F. (2013), *Low-Latency Audio over IP on Embedded Systems*, master's thesis, Hamburg: Technische Universitat.

Naugle, L. M. (2002), 'Distributed choreography: A video-conferencing environment', *Journal of Performance and Art*, 24:2, pp. 56–62.

Oliveros, P., Bahn, C., Braash, J., Chafe, C., Hahn, T., Soundwire Ensemble, Tintinnabulate Ensemble, Valente, D. and Woodstrup, B. (2007), 'Dynamic spaces', *Proceedings of SIGGRAPH 2007*, San Diego, CA, 5-9 August, http://www.siggraph.org/s2007/attendees/art/performance.html. Accessed 5 January 2017.

Prior, D., Reuben, F., Biscoe, I. and Rofe, M. (2017), 'Designing a system for Online Orchestra: Computer hardware and software', *Journal of Music, Technology and Education*, 10: 2-3, pp. 185-96.

Rofe, M. and Reuben, F. (2017), 'Telematic performance and the challenge of latency', *Journal of Music, Technology and Education*, 10: 2-3, pp. 167-84.

Rofe, M., Geelhoed, E. and Hodsdon, L. (2017a), 'Experiencing Online Orchestra: Communities, connections and music-making through telematic performance', *Journal of Music, Technology and Education*, 10: 2-3, pp. 257-76.

Rofe, M., Murray, S. and Parker, W. (2017b), 'Online Orchestra: Connecting remote communities through music', *Journal of Music, Technology and Education*, 10: 2-3, pp. 147-66.

Rohrhuber, J. (2007), 'Network music', in N. Collins and J. D'Escrivian (eds), *Cambridge Companion to Electronic Music*, Cambridge: Cambridge University Press, pp. 140–261.

Tanaka, A. and Bongers, B. (2001), *Global String: A Musical Instrument for Hybrid Space*, https://csl.sony.fr/downloads/papers/2001/Atauglobalstring.pdf. Accessed 1 May 2015.

Tremblay, P. A. and McLaughlin, S. (2009), 'Thinking inside the box: A new integrated approach to mixed music composition and performance', *Proceedings of the International Computer Music Conference (ICMC 2009)*, Montreal, Quebec, Canada, 16-21 August, Michigan: University of Michigan, pp. 379–86.

Notes

1. For instance, Brian Shepard's *EchoDamp* has been used in a range telematic performances using LOLA and UltraGrid; see http://www.echodamp.com.

2. A number of telematic works have sought to sonify the acoustic properties of the network as a medium in its own right. Chris Chafe's own piece, *Ping* (Chafe and

Niemeyer 2001) and Atau Tanaka's piece, *Global String* (Tanaka and Bongers 2001) – both installation pieces – are good examples of this.

3. For an in-depth enquiry into optimizing the relationship between live and loudspeaker sources in electro-acoustic performance, see Tremblay and McLaughlin (2009).

4. Although there are numerous technical standards in operation, −18dBFS is used throughout Europe and the United Kingdom as a calibration level, due to its equivalence to 0VU in the analogue domain.

5. See http://atsc.org/standards/atsc-standards/.

6. Julian Rohrhuber offers valuable insight into the tendency within networked music to delocalize causation (see Rohrhuber 2007).

7. This global decrease in dynamics mirrors a similar effect observed in tempo decrease in the presence of small amounts of latency (see Chafe et al. 2004: 1).