





To answer this important question and to shed light on the artists' creative practice we develop, document and evaluate a series of interfaces and musical performances together with popular music artists. The focus is on providing playful game-like interaction, facilitating collaborative improvisation and giving clear feedback as well as traceable results. The interfaces will be deployed in three popular music live performances at one event. The art-based research approach uses mixed methods, including a focus group and surveys as well as quantitative data logging and video analysis to identify parameters of acceptance, new ways of artistic expression, composition and musical experience. The evaluation will allow to present structured guidelines for designing and applying systems for audience participation.

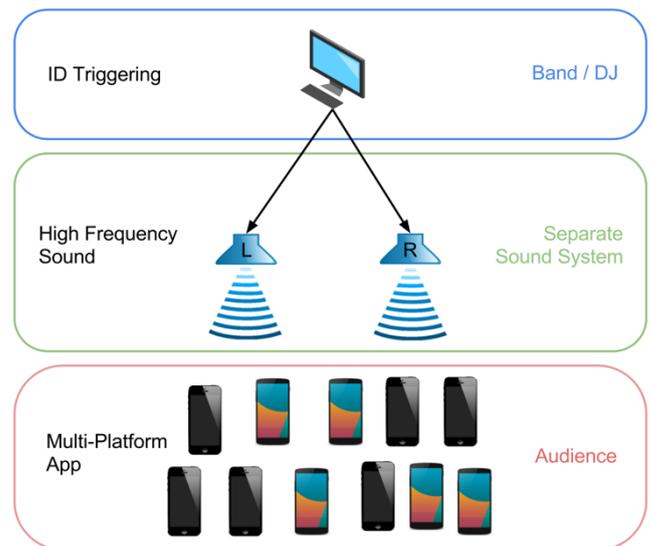
The project team is comprised of popular music artists, and researchers covering diverse areas such as media arts, computer science, human-computer-interaction, game design, musicology, ethnomusicology, technology and interface design.

The results of the project will be situated at the interdisciplinary intersection of art, music and technology. It will present structured and evaluated insights into the unique relation between performers and audience leading to tested and documented new artistic ways of musical expression future performances can build on. It will further deliver a tool-set with new interfaces and collaborative digital instruments.

### 3. Implementation

The technical basis of Poème Numérique is the use of high-frequency sound IDs to trigger events on the audience's smartphones. The use of high-frequency sound or "Ultra Sound Communication" for audience participation has first been documented in [14]. In this approach frequencies above the average human hearing spectrum are transmitted by dedicated speakers and are used to quasi silently trigger events. An app that has to be downloaded before the performance listens for these sound IDs using the smartphone's microphone. Figure 1 shows the full setup with a computer used to send the sound triggers to a sound system and the audience's smartphones, which listen for these triggers using a cross-platform Android / iOS app. The cross-platform app has been implemented using Xamarin Forms [15].

Each Sound ID is composed of two distinct frequencies between 18 kHz and 20,7 kHz. For recognizing the specific frequencies on the smartphones fast fourier transformation (FFT) was used. In our case we use the frequency spectrum of 22050 Hz and divide it into 256 pieces. Each piece consists of a frequency spectrum of 86.13 Hz. To do this 512 time-based values have to be used. The information provided by the array of frequency-base values is crucial for recognizing the intensity of a specific frequency interval and therefore crucial for identifying high frequency sound IDs.



**Figure 1.** The technical setup of using high-frequency sound IDs.

Two frequencies are recognized as an ID if both of them are recognized in at least eight out of the ten last FFTs. That means that for both frequencies a peak has to be recognized eight out of ten times. For a peak to be recognized the following two conditions have to be fulfilled:

- The value in the respective array entry has to be higher than a threshold which is platform and frequency specific.
- The value in the respective array entry has to be higher than the surrounding array entries.

Two speakers are used to transmit the two frequencies simultaneously. The IDs always are played back for three seconds. Much smaller playback timeframes are theoretically possible, but our application does not need to allow for fast sequences of triggers. Within the above frequency range we managed to implement 15 unique IDs. To reduce false positives and faulty recognition we used one of these for a Sync ID sent before an actual Sound ID. This Sync ID prompts the phone to listen for a Sound ID for nine seconds. After the Sync ID we introduced the option of sending what we called a Change ID used to allow a second bank of triggers. After that the actual Sound ID is transmitted. By this means the system at present supports 26 unique Sound IDs. A PD (Pure Data) [16] patch is used to play back the high-frequency Sound IDs and thus is the central hub for controlling the distributed performance. The PD patch can itself be controlled through any network protocol including MIDI or OSC.

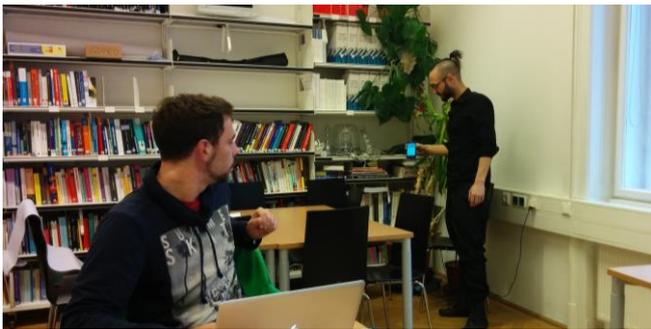
### 3. Formative Evaluation

In the following, we discuss five steps spanning ten months of the iterative and explorative design process of Poème Numérique. For each of the presented prototypes we describe the setting it was tested in, the improvements made based on previous iterations, the musician’s feedback and reflective insights.

#### 4.1 Spatial Sound Distribution using Smartphones (first iteration)

The first prototype successfully realised the use of high-frequency sound IDs guided by Hirabayashi and Eshima [14]. A basic version of the app allowed sounds to be triggered on multiple smartphones in a university meeting room with 42,57m<sup>2</sup> tested on 28. October 2015. The sound recognition worked, but had troubles when the phones were moved. The musician emphasised the need that the system has to be 100% reliable in recognising trigger signals. The recognition problems were caused by a Doppler effect. This technical problem was solved with the next iteration, as explained later.

From a creative perspective the musician appreciated the spatial distribution of sounds in the room [2]. This test also allowed the musician to get a feeling for the technical setup and its creative capabilities. This firstly prompted her decision to make an original composition instead of using an existing piece and secondly allowed her to brainstorm different ways to compose for a distributed smartphone setup. An open question after this first test was how to give the audience more agency beyond moving with their smartphones.

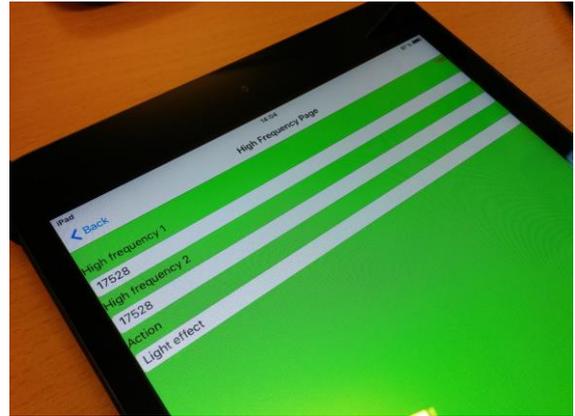


**Figure 2.** Triggering sounds of spatially distributed smartphones

#### 4.2 Second Iteration

The second prototype (see figure 3) mostly included technical improvements in frequency recognition and was again tested in the university meeting room setting. From a design perspective we discussed the need for moderation and feedback of audience interaction. Later versions of the

app use distinct colors for each sample to make clear what is played back on each phone. We also added capabilities for an introductory message and to display text with each sample that is triggered. From a creative perspective we discussed that some sounds could be played on a normal sound system with only a part of the music coming from smartphones. To increase the diversity of sounds and to make the resulting soundscape more interesting we also introduced the use of weighted random to determine which sounds are played on a particular phone.



**Figure 3.** Second prototype version of the app displaying recognised control signals

#### 4.3 Third Iteration and Large-Scale Test with Students

Figure 4 shows a test performance using the system at a lecture at the Vienna University of Technology with informatics students. The lecture hall, one of the biggest in this university, had 404,51m<sup>2</sup>. The app was distributed through the Android platform store. The test performance showed that the transmission of high-frequency sound IDs is mostly robust, but that recognition problems might occur with untested smartphones and with increasing distance from the sound source. Also some Android mods (e.g. Cyanogen) block microphone access due to privacy settings.



**Figure 4.** A test of the system with students during a lecture.

While the large-scale use of the app was impressive visually, the musician was disappointed regarding the loudness of the joint output of the phones. We expected hundreds of phones would increase loudness drastically compared to the meeting room setting, but this was not the case. This may also be due to the special architecture of this auditorium. We consequentially made the decision to amplify the output (of some) of the audience's phones. The test further underlined the necessity of feedback through colours - each sound is associated with a specific (steady or flickering) color displayed on the phone when the sound is played back - which not only made the interaction more transparent for the audience but also added feedback for the musician.

#### 4.4 Fourth Iteration

The fourth prototype implemented the visual feedback changes discussed in the previous iteration. The musician also provided a new set of audio samples and it was the first prototype that was tested in the location planned for the performance (see figure 6). The test first used electromagnetic pick-up coil microphones (see figure 5) to amplify some of the smartphones in addition to the speaker-sounds emitted by all phones. Only electromagnetic waves (in this case from the speaker inside the phone) are amplified to prevent the amplification of any ambient noise. The overall results were appreciated. To ensure a clearer distribution of sounds in the performance space we decided to use a center stage for the performer and four stations in each corner where smartphones can be amplified by the audience.



**Figure 5.** Using a pick-up coil to amplify the audio output of a smartphone

#### 4.5 Fifth Iteration

This iteration marks the first use of the four stations with one speaker each. A new set of sound samples considering

this setup was used. This was also the first time that the musician's composition adapted to the sound of mobile phones by using hi-pass filters and generally making use of the frequencies that work well with phones (e.g. bell sounds).



**Figure 6.** Center stage setting in the final location with four stations, one in each corner.

We used 17 phones overall. Combined with the four stations this lead to an encompassing spatial soundscape where individual sounds are not localizable anymore. Requirements by the musician were less delay when triggering sounds and to be able to trigger more sounds. Both changes are trade-offs means a trade-off with reliability and response time of the system. We were able to drastically reduce the delay in triggering samples from four seconds to 250 milliseconds with the only requirement of playing back the high-frequency sound IDs louder. We could not add a larger number of sounds because that would either make the system less reliable, introduce more latency again, or make control frequencies hearable. Overall this iteration was considered to be the first realistic setting for the performance. Future iterations will not change this setup but only refine details.

### 5. Discussion

In the following, we first draw together artist reflections from the formative evaluation steps (see chapters 4.1 through 4.5) resulting in a list of *learnings for design and conceptualisation of TMAP*. These issues are mainly derived from the presented case study using the audience's distributed smartphones. However, we discuss their implications in context of other studies around TMAP.

Second, we present a series of *learnings concerning the underlying technology* from applying TMAP with smartphones using high-frequency sound IDs thereby reflecting and extending the insights presented in [14].

## 5.1 Learnings for the design and conceptualisation of TMAP

Having a *low latency* and an immediate reaction of the devices, when they are remotely triggered by the DJ, were an issue that concerned the general concept throughout all development steps. This low latency is not only a requirement by our particular and her musical intentions, but was reported by other work in the field of TMAP [17,18] as reported by the audience.

During development we had to spend a certain amount of effort in testing and improving the *reliability* of the whole system. Especially at an early phase, there were several dropouts of smartphones which did not react when being remotely triggered. Given, that every smartphone is held by an individual during a performance, every dropout causes one disappointed participant who is prevented from interaction. As this spectator interaction is central to TMAP, we consider a reliable system (without dropouts in our case) as crucial for the whole concept.

Iterative *testing* is necessary for the involved musicians to get an idea and feeling for the anticipated system. As in our case, a system for TMAP starts with a basic concept. Although, this might include musical and technical details even in an early development phase, it needs much more imagination at the beginning than later, when the system can actually be tested to some extent.

Considering the actual music or musical piece, we identified the *difficulty to use an existing piece of music*. In fact, it was the decision of the artist we collaborated with, that she will write a new piece instead of using an existing one as originally intended. This issue of using or changing existing pieces or create a new composition on purpose for TMAP, has been discussed in other work as well [19,20].

Another issue concerned the *musical quality and the artist's flexibility*. In particular, this concerned the low audio quality of smartphone speakers (e.g. little loudness and bass). The sound quality was not satisfying first hand, but the artist was willing to adapt to the circumstances. Moreover, the artist suggested to adapt the composition with regard of the given smartphone speaker constraints. In practice, she used sounds within a higher frequency range more appropriate for this kind of speakers.

Important for the artist we worked with, was a *maximum of control over the musical result*. This concerned the random distribution of a set of different sounds and colours among all participating smartphones. With the early prototypes, the distribution happened completely random, which caused some silent sounds being covered by louder sounds although equally distributed. Thus, the artist wished to have a certain control over the random distribution to give silent sounds more weight.

During the development, we discussed the *amount of interaction* the spectators have using the application. Our most recent prototype does not allow much interaction except of start and stop the application or moving the phone while sound is triggered remotely. In literature, this requirement was discussed already as the audience wishes a particular amount of interaction [20,21].

## 5.2 Learnings concerning the underlying technology

We developed Poème Numérique based on a particular technology for audience participation using high frequency sound IDs in *Sense of Space* [14]. In relation to the original technology and findings presented by the authors of *Sense of Space*, we discuss identified issues around this particular technology based on our our experience.

In *Sense of Space*, the authors used iPhones only for their application and performance. We tried to increase the number of spectators being able to participate by developing an application for iOS and Android. To reduce the effort in building the same application for two platforms, we used Xamarin as described earlier. This approach turned out to be working from a technical perspective as we could realise the application for both platforms. Although, we observed a problem with privacy settings with the Android modification Cyanogenmod.

We can confirm the spatialization effect reported by *Sense of Space* [14 p.60]. At the same time we can confirm the issues with the smartphone speakers being too weak to hear. An issue we observed concerns the distance of smartphones and the power of the PA speakers to distribute the high frequency IDs. In our case, some speakers we used to trigger the sound IDs were too weak if the smartphone to be triggered was further away. Nothing is reported about this issues in the original paper. Further improvements addressed an occurrence of the Doppler effect. The doppler effect occurs when a smartphone is moved towards or away from the speakers. This effect leads to the fact that the frequency recognized by the respective smartphone is higher or lower than the frequencies played by the speakers. The ID is therefore not recognized. To minimize the chance of a doppler effect disturbing the recognition process, the frequency interval used for recognition is chosen very generously.

## 6 Conclusions and Outlook

This contribution presented the technical basis and iterative explorative design process of a system for TMAP using smartphones and high-frequency sound IDs. The learnings from the perspective of musicians were technical requirements such as low latency, reliability, as well as increasing the number of possible sound samples and sound quality. Musicians accept constraints and a broad range of functions is not necessary. Constraints even seem to further creativity. Further we learned that such a system for TMAP necessitates original compositions. We also identified balancing the musician's control over the musical result with freedom for audience interaction as one of the core challenges in designing such a system. The iterative design and development process not only helped the results presented in this paper, iterative testing also proved invaluable in helping the musician familiarize herself with the system, quite similarly to learning a new musical instrument.

Building on [14] from a technical perspective we could confirm spatialization effects and issues with low volume of smartphone speakers. We also observed issues with decreasing reliability when increasing the distance between smartphone and the sound source of the high-frequency sound IDs or when using speakers that are too weak. In contrast to [14] we did not observe disadvantages from a creative perspective with regard to sounds being not localizable.

Poème Numérique will be refined and finalised in this iterative manner. Further tests in a live music performance setting will determine the acceptance and creative possibilities of such a system from an artist and an audience perspective.

## Acknowledgements

“Breaking The Wall” is a project funded by the Funded by the Austrian Science Fund (FWF), AR 322-G21.

Project team: Chris Bruckmayr, Dietmar Bruckmayr, Geraldine Fitzpatrick, Simon Holland, Susanne Kirchmayr, Johannes Kretz, Peter Purgathofer, Hande Sağlam

## References

- [1] <<http://indigo-inc.at>> [Accessed: 14 December 2017]
- [2] <<http://www.piglab.org/breakingthewall>> [Accessed: 14 December 2017]
- [3] Bartmann, C: Exploring audience participation in live music with a mobile application. Master thesis, Vienna University of Technology (2016).
- [4] Mozart, W.A.: Anleitung so viel Walzer oder Schleifer mit zwei Würfeln zu componiren so viel man will ohne musikalisch zu seyn noch etwas von der Composition zu verstehen, J.J. Hummel, Berlin-Amsterdam (1793)
- [5] Zbikowski, L.M.: Conceptualizing Music. Cognitive Structures Theory, and Analysis, Oxford University Press (2002)
- [6] Pichlmair, M., & Kayali, F.: Levels of sound: On the principles of interactivity in music video games. In Proceedings of the Digital Games Research Association 2007 Conference" Situated play (2007)
- [7] Mazzanti, D., Zappi, V., Caldwell, D., & Brogni, A.: Augmented Stage for Participatory Performances. In Proceedings of the International Conference on New Interfaces for Musical Expression, pp. 29–34 (2014).
- [8] Kayali, F., Hödl, O., Fitzpatrick, G., Purgathofer, P., Philipp, A., Mateus-Berr, R., Kühn, U., Wagensommerer, T., Kretz, J., Kirchmayr, S.: “Playful Technology-Mediated Audience Participation in a Live Music Event”. Proceedings of the ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play (CHI PLAY) 2017. (2017)
- [9] Reichl, P., Löw, C., Schröder, S., Schmidt, T., Schatzl, B., Lushaj, V., Hödl, O., Güldenpfennig, F. and Widauer, C.: The Salome Experience: Opera Live Streaming and Beyond. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*(pp. 728-737). ACM. (2016)
- [10] Wu, Y., Zhang, L., Bryan-Kinns, N. and Barthelet, M.: Open Symphony: Creative Participation for Audiences of Live Music Performances. *IEEE MultiMedia*, 24(1), pp.48-62. (2017)
- [11] Dewey, J.: *Logic: The Theory of Inquiry*. Carbondale: Southern Illinois University Press (1938).
- [12] Schön, D. A.: *The Reflective Practitioner: How Professionals Think in Action*. Smith, T. (ed.). Basic Books, p.374 (1983).
- [13] Zimmerman, J., Forlizzi, J. and Evenson, S. . Research through design as a method for interaction design research in HCI. Proceedings of the SIGCHI conference on Human factors in computing systems CHI 07, 07, ACM Press, p.493 (2007). Online: <http://portal.acm.org/citation.cfm?doid=1240624.1240704>. [Accessed: 14 December 2017].
- [14] Hirabayashi, M., & Eshima, K.: Sense of Space: The Audience Participation Music Performance with High-Frequency Sound ID. In Proceedings of the NIME 2015 International Conference on New Interfaces for Musical Expression. LSU, Baton Rouge, LA (2015)
- [15] <<http://xamarin.com/forms>> [Accessed: 14 December 2017]
- [16] Puckette, M.: Pure Data: another integrated computer music environment. In Proceedings of the Second Intercollege Computer Music Concerts, Tachikawa, Japan (1996).
- [17] McAllister, G. and Alcorn, M. Interactive performance with wireless PDAs. In Proceedings of the International Computer Music Conference. pp. 1–4. (2004)
- [18] Feldmeier, M. and Paradiso, J. An Interactive Music Environment for Large Groups with Giveaway Wireless Motion Sensors. *Computer Music Journal*, 31(1), pp.50–67. (2007)
- [19] Hödl, O., Fitzpatrick, G. & Holland, S. Experimentence: Considerations for Composing a Rock Song for Interactive Audience Participation. In Proceedings of ICMC. pp. 169–176. (2014)
- [20] Hödl, O., Kayali, F. & Fitzpatrick, G. Designing Interactive Audience Participation Using Smart Phones in a Musical Performance. In Proceedings of the International Computer Music Conference. pp. 236–241. (2012)
- [21] Freeman, J. Large Audience Participation, Technology, and Orchestral Performance. In Proceedings of the International Computer Music Conference. pp. 757–760. (2005)