Online-Tool for interactive sound analysis of orchestra instruments

Malte Kob, Sebastià V. Amengual, Banu Sahin, Aristotelis Hadjakos, Maria Saulich
[2] *Hochschule für Musik Detmold, 32756 Detmold, Deutschland, Email: {kob, amengual, hadjakos, saulich}@hfm-detmold.de,*

Motivation

How does the sound of a music instrument change at different locations of the player and the listener? How can I optimise that sound as a musician or as an acoustician? This contribution presents an interactive web concept (see Fig. 1) that allows to assess the effect of several performance parameters on the sound of a music instrument. In a cooperation between the "Netzwerk Musikhochschulen", funded by the BMBF, students and docents of the Hochschule für Musik Detmold a long-term project was initiated that addresses educational demands of musicians, sound engineers and acousticians.

Education

Musicians, when playing during lessons or practicing, spend most of their time in small rooms that have nothing in common with the concert halls in which they will perform later on. They act on what they perceive in their practicing surroundings. However, this can be deceptive. Wouldn't the chance to experience one's own playing from different seats in the audience lead to other corrections in playing style and dynamics, for example? To what extent will somebody sitting in the last row be able to understand the soft passages? What would the perfect balance be between different chamber music partners?

By offering insight into the audience's perspective, the interactive web tool aims at addressing these questions. It intends to extend listing skills and to train the awareness to what extent one's own perception differs to the perception of others. In preparation for exams or concerts the platform supports the development of a tonal sense of different surroundings. It depicts the specific characteristics of different instruments, such as tonal range and ways of articulation, and enables the user to listen to these characteristics from various positions in different concert halls in Detmold, namely the Detmold Konzerthaus, the Sommertheater and the Brahmssaal of the University of Music.

The platform offers the perspective of feeding own recordings into the online tool. Whether integrated into lessons or used individually while practicing, the web tool thus can be of interest to a wide range of subjects, such as artistic one-to-one teaching, instrumental pedagogy, composition, musical acoustics and sound engineering.

Implementation

This projects spans several sub-projects that have been initiated by the need to provide reliable recordings of music instruments that can be used for teaching in the programs at the University of Music Detmold. The first project aims at video and audio recordings of sounds in various conditions. The second uses measured room impulse responses (RIR) and performs convolutions of these recordings with the RIRs. The third project implements both projects into a web-based user interface that provides easy access to the recordings and performs visualisations of the sound data.

Audio and video recordings

Audio and video recordings were held in the Wave Field Synthesis room of the Erich-Thienhaus-Institut (see Fig. 2). Musicians were asked to sit in the center [1] of a spherical construction covered with a molton textile and play an arpeggio, a chromatic scale and a short melody varying in dynamics (piano, mezzo-forte, forte, sforzato) and articulation (legato, non-legato, pizzicato, staccato, vibrato).

The microphone setup for the directivity measurements was chosen by taking into account the most effective configuration with available equipment [2], [3], [4]. In total, eight omnidirectional and two cardioid pattern condenser microphones were used for the audio recordings using a 24-bit quantization and 48.1 kHz sampling rate. Six omnidirectional microphones (NTi M 2010) were placed at an angle of respectively 0°, 45°, 90°, 180°, 270°and 315°, 90 cm from the center and 105 cm from the ground. Additionally, two omnidirectional Brüel-Kjaer microphones were placed with an elevation angle of -45°and 45°and 0°azimuth. They were both positioned 90 cm from the center, the bottom one 20 cm from the ground and the upper one 175 cm from the ground.

To provide material for the sound auralization two Neumann cardioid microphones were specifically positioned depending on each instrument's features. Furthermore, a video recording was held simultaneously using a Sony HD Camera.

Six instruments from five different instrument families (woodwind instruments, brass instruments, string instruments, keyboard instruments, percussion) were recorded as well as female and male voice. The glockenspiel (shown in Fig. 3) was recorded both, with hard and soft mallets, to provide further sound distinctions. This is a list of the recorded instruments and musicians:

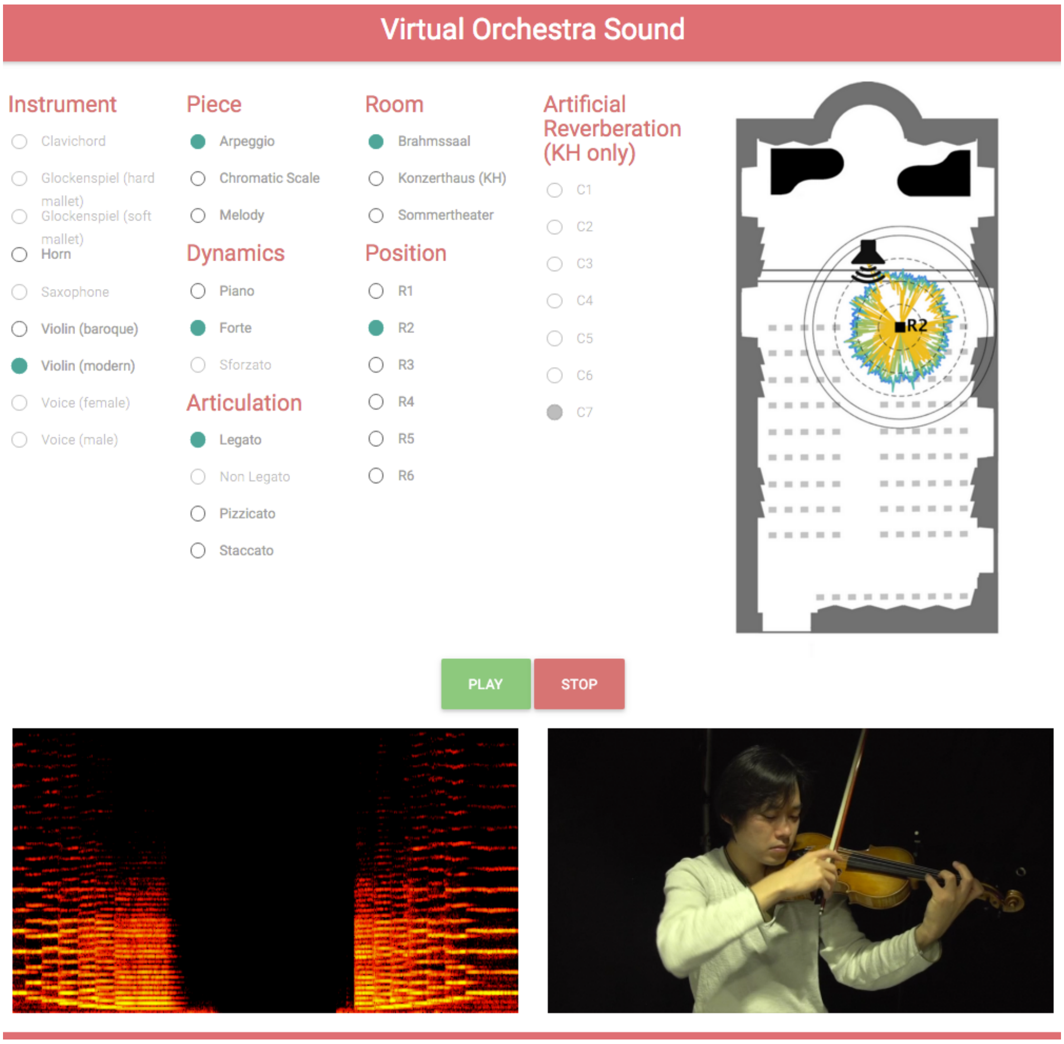

Figure 1: The Virtual Orchestra Sound website

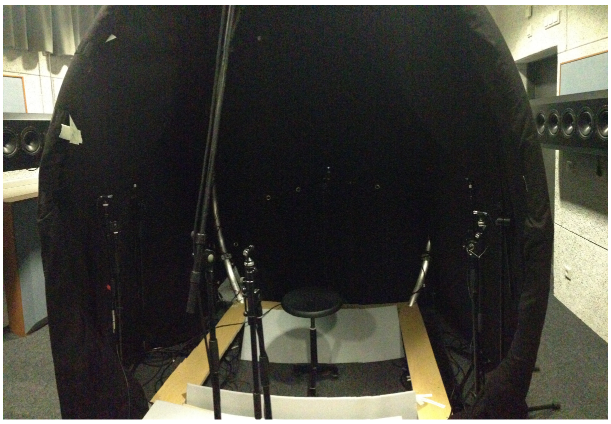

Figure 2: Recording set-up.

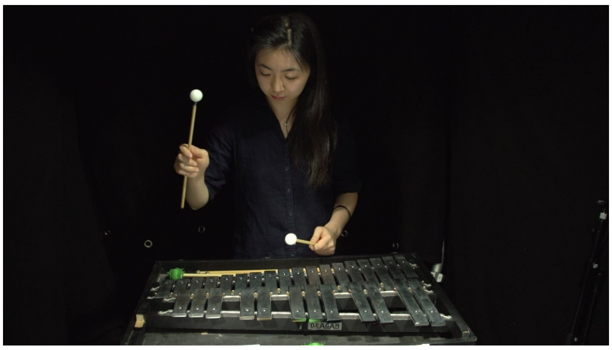

Figure 3: Recording session with Kazuyo Tsunehiro, Glockenspiel

1 – Clavichord - Maria Saulich

2 – Glockenspiel - Kazuyo Tsunehiro

3 – Horn - Anton Langer

4 – Saxophone - Dustin Eddy

5 – Violin-Barock - Tetsuro Kanai

6 – Violin-Modern - Tetsuro Kanai

7 – Voice-Female - Amanda Kyrie Ellison

8 – Voice-Male - Manuel Grunden

More details about the audio and video recordings are given in [5].

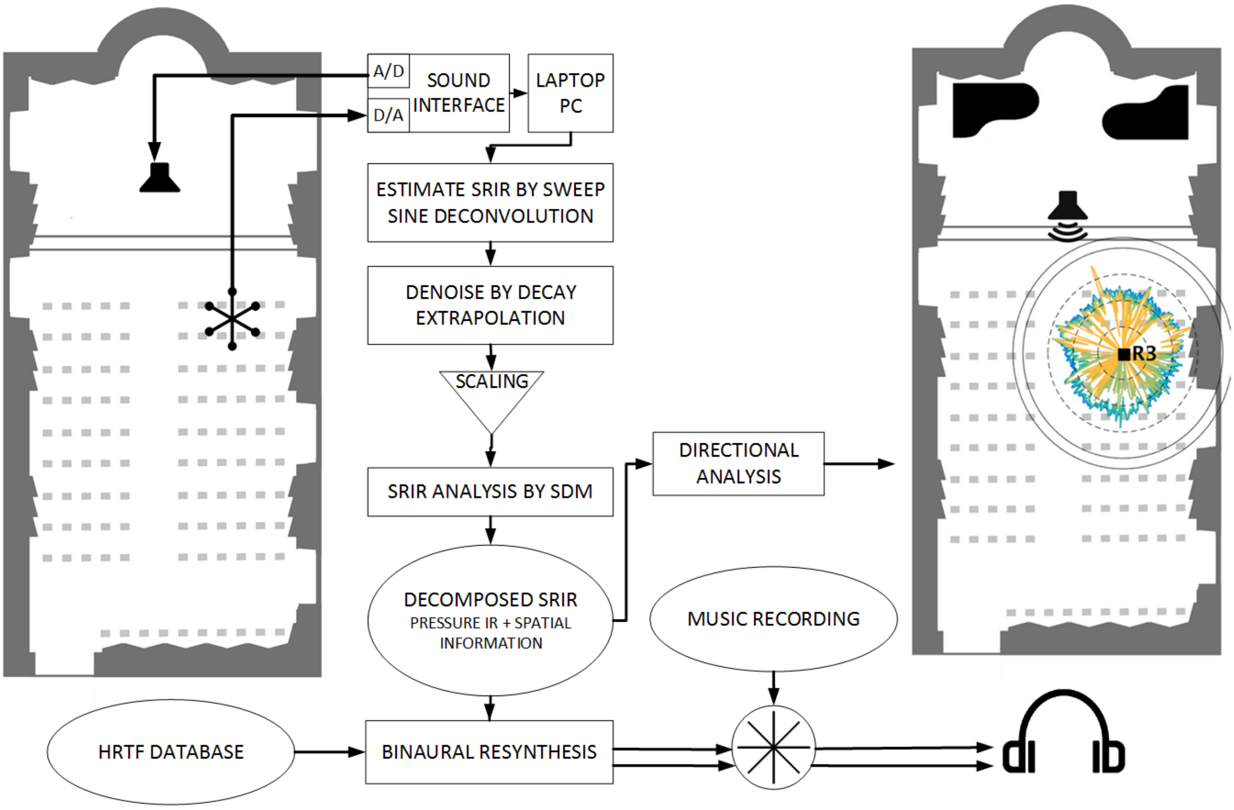

Figure 4: Operations performed in the room auralization and spatiotemporal analysis of spatial room impulse responses

Room auralization

To include the room effects on the dry recordings it is necessary to convolve the music signals with impulse responses of different rooms. Spatial Room Impulse Responses (SRIR) are obtained by sweep sine deconvolution [6] using a studio monitor (Neumann KH120A) placed on stage and an open compact array composed of 6 omnidirectional matched microphones (NTi M2010) placed in several positions of different halls.

The impulse responses are denoised by energy decay extrapolation in octave frequency bands [8] in order to remove convolution artifacts and they are scaled appropriately to ensure the correct level at every measurement position. They are then analyzed using the Spatial Decomposition Method for Room Impulse Responses [7]. After the analysis they are resynthesyzed into a binaural format using a Head-Related Transfer Function (HRTF) database [10]. The binaural impulse responses are finally saved into a database for further usage within the web interface.

The analyzed SRIR are used also to represent the spatio-temporal information of the incident sound at every measurement position. The spatio-temporal shows an energy histogram as a function of the angle of incidence of the sound and it is calculated using time windows of different length, providing an intuitive view of the sound spatial behavior. The procedure presented in [9] is followed in this work to generate the spatio-temporal representation.

A block diagram depicting the main signal processing operations as well as the spatio-temporal representation is shown in Fig. 4

User interface implementation

Figure 1 shows the user interface. The user selects an instrument, one of three different pieces, dynamics, articulation, and a room. The floor plan on the right side changes according to the room selection. Each room has up to six listening positions which are shown in the floor plan as R1 to R6. Furthermore the interface lets the user control the artificial reverberation system of the "Konzerthaus" on a seven point scale from off (C1) to maximum artificial reverberation (C7). After hitting the "Play" button, the user hears the selected instrument and piece at the selected room and listener position. A spectrogram and the video of the player are shown in the bottom part of the interface. Also, the floor plan changes to show only the selected listening position together with the receiver characteristics at that position.

The interactivity has been realized with JavaScript using the web audio API for audio processing. Each combination of instrument, piece, dynamics, and articulation corresponds to one sound file; and each combination of room, position, and artificial reverberation setting corresponds to an impulse response stored in another sound file. When the user hits the "Play" button, the two sound files (music and impulse response) are fetched from the server. Once both have been downloaded, the music is convolved with the impulse response in real-time. At the same time, the spectrogram is drawn based on a FFT of the convolved audio signal.

Discussion and Outlook

The convolution of the room with the dry recording does not represent an actual recording under live conditions. Therefore, the room influence on the player is not taken into account, and all music and video files represent the same musical performance. Future instrument recordings will be performed under acoustic conditions that represent the actual situation when performing in the appropriate room. A concept is currently developed in the frame of the BATWOMAN project [12].

The sender directivity will be made available in the GUI as well. The current idea is a small image that represents an average sender characteristics as in [1]. A more complex visualization will show the note-by-note directivity during performance. The need for this representation has been discussed e.g. in [11].

Future implementations will provide upload facilities and more comprehensive sound analysis tools. Composers and musicians will be able to provide own recordings of instrument sounds that can be rendered in any room. This approach will give insight into the effect of room acoustics on the perceived instrument sound and help to adopt instrumentation, playing style and position of musicians to rooms. The simultaneous reproduction of ensemble sounds together with dedicated instruments can provide further understanding of sound blending of an orchestra or choir.

Currently, the convolution is performed on the user's computer in the browser which requires the transfer of both audio and room impulse files through the Internet. Alternatively, the audio files could be convoluted with the room impulse responses on the server. Such a server-based solution would allow to save bandwidth by transmitting compressed audio, adopted to the wishes to the user and the network conditions. On the other hand, a browser-based convolution simplifies the implementation greatly by making server-side code and client-server communication unnecessary. The pros and cons of the two approaches are currently evaluated.

For research purposes, further analysis tools will be implemented that quantify the sound quality of the instruments. A comparative presentation of varied sounds provides a visual and auditive assessment of differences between two or more interpretations or playing styles. This use case would be desirable for education in various programs as well as for research [12].

Acknowledgements

Thanks to all the musicians and colleagues who contributed with their music and good ideas to the project. The Netzwerk Musikhochschulen and the Deutsche Zentrum für Luft- und Raumfahrt e. V. (DLR) are acknowledged for support of the project. A part of this project is funded by the European Commission within the ITN Marie Curie Action BATWOMAN under the 7th Framework Programme (EC grant agreement no. 605867).

References

[1] Meyer, J.: "Acoustics and the Performance of Music", Das Musikinstrument, Frankfurt/Main, 1978.

[2] Hagai, I. B.; Pollow, M.; Vorländer, M.; Rafaely, B.: "Acoustic centering of sources measured by surrounding spherical microphone arrays", *J. Acoust. Soc. Am. 130(4)*, 2003-2015, 2011.

[3] Pollow, M.; Behler, G. K.; Schultz, F.: "Musical Instrument Recording for Building a Directivity Database", *Fortschritte der Akustik – DAGA 2010: 36. Deutsche Jahrestagung für Akustik*, Berlin, 2010.

[4] Fletcher, N. H.; Rossing, T. D.: "The Physics of Musical Instruments", 2nd ed., Springer Science and Business Media, Inc., New York, 2008.

[5] Sahin, B.; Saulich M.; Kob, M.: "Video and audio recordings for a new interactive method to teach instrument and room acoustics", *Fortschritte der Akustik – DAGA 2016: 42. Deutsche Jahrestagung für Akustik*, Aachen, 2016.

[6] Farina, A.: "Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique", *108th AES Convention*, Paris, France, 2000.

[7] Tervo, S.; Pätynen, J.; Kuusinen, A.; Lokki, T.: "Spatial Decomposition Method for Room Impulse Responses", *J. Audio Eng. Soc.*, vol. 61, pp. 17-28, 2013.

[8] Cabrera, D.; Lee, D.; Yadav, M.; Martens W. L.: "Decay Envelope Manipulation of Room Impulse Responses: Techniques for Auralization and Sonification", *Proceedings of ACOUSTICS 2011*, Gold Coast, Australia, 2011.

[9] Pätynen, J.; Tervo, S.; Lokki, T.: "Analysis of concert hall acoustics via visualizations of time-frequency and spatiotemporal responses", *J. American Soc. Ac.*, vol. 133(2), pp. 842-857, 2012.

[10] Bernschütz, B.: "A Spherical Far Field HRIR/HRTF Compilation of the Neumann KU 100", *AIA-DAGA 2013*, Merano, Italy, 2013.

[11] Grothe, T.; Kob, M.: "Investigation of bassoon directivity" , *Proceedings of the Stockholm Music Acoustics Conference 2013, SMAC 13, edited by Roberto Bresin and Anders Askenfeldt*, Stockholm, Sweden, pp. 391-397, 2013.

[12] Amengual, S. V.; Eddy, D.; Kob, M.; Lokki, T.: "Real-time auralization of room acoustics for the study of live music performance", *Fortschritte der Akustik – DAGA 2016: 42. Deutsche Jahrestagung für Akustik*, Aachen, 2016.