Augmented Practice-Room - Augmented Acoustics in Music Education

Matthias Frank, Daniel Rudrich, Manuel Brandner
Institute of Electronic Music and Acoustics, University of Music and Performing Arts Graz, Austria
Emails: frank@iem.at, rudrich@iem.at, brandner@iem.at

Introduction

The acoustic properties of a room affect not only the sound of an instrument in the audience, but also its perception by the musician playing on stage. This can result in a different style of playing [1, 2], especially if one is not used to adapting to different acoustics. In music education, playing an instrument in concert halls often comes too short and happens rather in a late stage of training.

In order to study the effects of different acoustic environments on playing, augmented acoustic reality can be employed. The basic idea is to extend the sound of an instrument by simulating a virtual room with adjustable acoustic properties while the direct sound reaches the ear of the musician unalteredly. In research studies, augmented acoustics is typically implemented with large effort to facilitate unimpaired musical playing [3] under controlled conditions while also taking into account the time-variant directivity of the instrument [4]. Typically, such systems are capable of low-latency processing and are used in anechoic chambers equipped with optical tracking system and surrounding microphone arrays. In contrast, there are simple smartphone apps, such as [1], that allow playing around with real-time reverberation of lower quality and higher latency.

During the last years, augmented reality has entered the field of music education, e.g. in Serafin's study [5], in Keebler's guitar learning system [6], and in Orman's use for the purpose of enhancing music conducting skills [7]. Pätynen employed augmented acoustic reality for the creation of virtual concert hall acoustics within a small practice room [8]. However, the authors are not aware of any study that evaluates its effects when actually applied in teaching of music students.

Our project "Augmented Practice-Room" tries to fill this gap. We developed a tool for augmented acoustic reality, the "Augmented Practice-Room App", cf. Fig. 1, using open headphones or loudspeakers to play back the virtual acoustics while maintaining the direct sound path of the instrument. The app is interactive in terms of location and orientation of the student and the teacher within the virtual room and also incorporates the directivity of the played instrument. The software implementation is optimized for efficient and zero-latency processing on standard computers with usual audio equipment. It employs the JUCE framework[2] and is based on open-source audio plug-ins[3] that have been developed at our institute.

[1] https://play.google.com/store/apps/details?id=com.bigbeard.svreverb
[2] https://juce.com
[3] https://plugins.iem.at

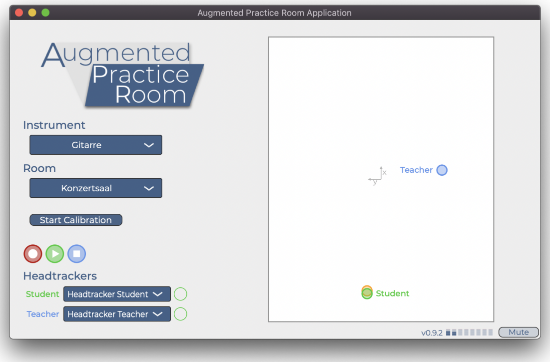

Figure 1: Screenshot of the Augmented Practice-Room App.

The tool is currently installed in six practice rooms at the Johann-Joseph-Fux conservatory of Graz and has been used in teaching since winter term 2019/2020. The teachers are a part of our research team as they help us to choose suitable virtual room settings and pick-up positions for each instrument. During a demonstration in our anechoic chamber, the teachers were playing their own instruments in combination with the application and they were generally convinced by the quality of the virtual rooms. They also assist in developing and carrying out the appropriate evaluation approach. So far, we've been using a semi-structured "research diary" consisting of different types of short questions, rating scales, and free annotations to collect data from the teachers and the students after each utilization. The diaries are accompanied by group discussions every few months. The evaluation is going to continue in the summer term 2020. Results, experiences, and feedback from the first term are going to be incorporated in the evaluation design and updates of the software.

This paper focuses on the technical description of the signal processing blocks in the software that is used to create the augmented acoustic reality in the practice rooms at the conservatory of Graz.

The Augmented Practice-Room App

The entire signal processing within the Augmented Practice-Room App employs Ambisonics technology [9], as it provides maximum flexibility for the choice of playback device (loudspeakers and headphones), scalability of spatial resolution to adjust to available processing power, and easy application of rotations to incorporate the orientation of instrument/student and teacher. The following paragraphs describe the particular parts in the block diagram of the app, as depicted in Fig. 2.

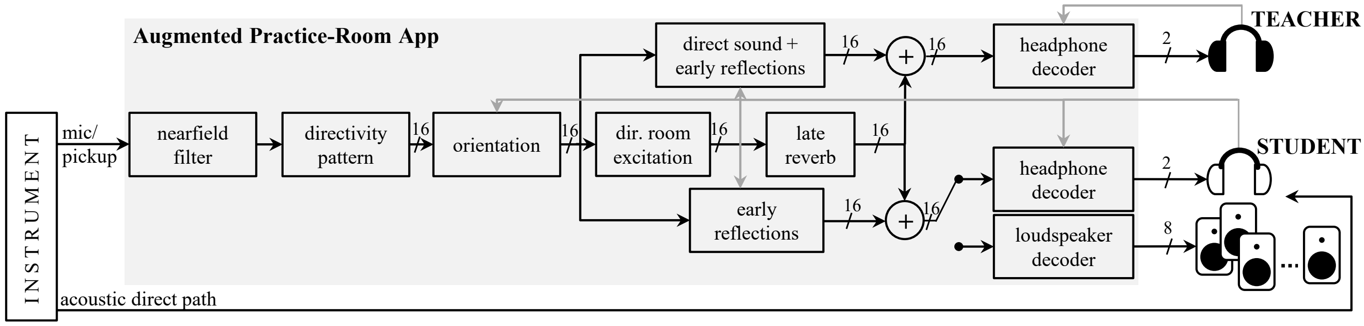

Figure 2: Block diagram with the signal processing from close-microphone/pickup of the instrument to headphones/loudspeakers.

The direct sound of the instrument propagates to the student's ears via the acoustic direct path. The unattenuated acoustic path to the student's ear is facilitated by using either open headphones [10] to play back the virtual room, or no headphones at all with playback over loudspeakers. Simultaneously, the direct sound is captured with a close-up microphone or instrument pickup and fed into the app for generating the virtual room acoustics.

Nearfield Filter

A nearfield filter compensates for timbral differences between the close-up capture of the direct sound and the farfield sound that is defined by the directivity pattern. The filter is based on transfer path measurements between close-up microphone / pickup and three microphones positioned in the farfield near the dominant direction of the directivity pattern.

Directivity Pattern

We use directivity patterns with 3$^{\mathrm{rd}}$-order resolution to reduce computational load, as our previous studies revealed that higher resolution is not necessary for this application [11, 12]. The employed directivity patterns are based on measurements from TU Berlin [13], except for the grand piano [14]. In order to facilitate minimum latency, the directivity patterns are represented as minimum-phase filters. This is done by frequency peak-picking of all single tone recordings of the TU Berlin measurement for each of the 32 microphones, and subsequent interpolation. The 32 magnitude spectra are normalized to the level of the microphone in the dominant direction (same as for the nearfield filter), e.g. normal to the soundboard of a violin, cello, guitar, or the bell of a trumpet. Subsequently, the spectra are transformed into 32 minimum-phase impulse responses which are finally decomposed into 16 spherical harmonic responses. In the current implementation the nearfield filter and the directivity pattern are convolved offline yielding a 512 (samples) $\times$ 16 (channels) impulse response for each instrument.

Orientation

A head-tracking device based on [15] is not only used to incorporate head rotations of student and teacher, but also for the positioning of the instrument and rotation of its directivity pattern.

Early Reflections

Early room reflections are strongly position-dependent up to an image-source order of 2 or 3 [16]. This agrees with the mixing time, where the early reflections transition into the diffuse, late reverberation [17, 18]. Thus, our app employs a dynamic image-source model with 40 reflections (image-source order 2.5) of the directive source in a shoebox room. The overall reflection coefficient of the room can be adjusted frequency-dependently. Moreover, each surface can have an additional broadband attenuation. The time delay of each reflection is reduced by the delay introduced by A/D-, D/A-conversion, processing, and audio buffers to facilitate zero-delay playback. In the case of a reflection arriving earlier, e.g. when a listener is close to a wall, resulting negative delays are clipped to zero. Both student and teacher have their own early reflection processor, whereas only the one for the teacher includes the direct sound.

Late Reverb

As late reverberation is largely independent of the listening position [16], a single instance of that processor can be used for both student and teacher. The **Directional Room Excitation** block takes into account how much energy contributes to the late reverb, depending on the instrument's directivity pattern, its position and orientation, and the wall absorption characteristics. This is done by sampling the directivity signal at 24 points of a t-design [19], frequency-dependent attenuation of the sampled signals according to the room properties, and subsequent summation of the signals. The resulting signal is then fed into a 64-channel feed-back-delay network [20, 21] to create the late reverb. A fade-in enables an increased diffuse envelopment [22] and smoother transition into the early room reflections.

Decoder

Early reflections and late reverberation are finally summed up for student and teacher. For the student, the virtual room acoustics is played back over transparent headphones or loudspeakers so that the acoustic path of the instrument's direct sound reaches the student's ear without attenuations. For the teacher, closed headphones are used to strongly attenuate the acoustic direct path, because position- and orientation-dependent direct sound is already included in his or her virtual room acoustics. Headphone playback employs state-of-the-art binaural Ambisonic decoding [23, 24] with head-tracking, while loudspeaker playback employs the AllRAD approach [25] for maximum flexibility.

Setting

So far, we have created five virtual rooms: a small room, two chamber music halls, a larger concert hall, and a cathedral. The parameters of each room are listed in Table 1. Upon request of the teachers, the second chamber hall is a model based on measurements of one of the conservatory's halls, where the students typically perform concerts. We also created five instrument presets which differ in the used directivity impulse response, and the positioning of the instrument relative to the student, cf. Table 2. While violin and clarinet are rotated with the student's head, all other instruments have a fixed orientation towards the positive x-axis, cf. Fig. 1.

virtual room	RT_{60}	x/m	y/m	z/m	Γ/dB
small room	0.3	5.5	6.5	3.5	-5
chamber music 1	1.0	13	9	6	-0.5
chamber music 2	1.0	15.3	8	5.1	0
concert hall	2.2	30	23.8	20	0
cathedral	5.1	30	23.8	20	-2

Table 1: Available virtual rooms with reverberation time RT_{60}, room dimensions, and overall reflection coefficient Γ.

instrument	rotates with student	r/m	$\varphi/°$	$\vartheta/°$	mics
violin	yes	0.3	70	-45	1
cello	no	0.9	0	-70	1
guitar	no	0.6	0	-60	1
piano	no	1.0	90	-30	2
clarinet	yes	0.6	0	-45	1

Table 2: Available instruments and their position relative to the student, as well as application of student-dependent rotation.

In order to minimize adjustments during the lessons and according to the teaching practice, each instrument has been assigned to a specific real room with its own Augmented Practice-Room system. Moreover, for all instruments that are captured with a single microphone/pickup, an additional reference microphone in the room is used to calibrate the level of the instrument in order to ensure similar conditions on the instruments of different students. As all piano students play on the same instruments, the piano system has been calibrated once and two microphones can be used to capture the sound of the piano.

For each instrument, headphone playback is used, however for the violin, loudspeaker playback is employed as the headphones could touch the instrument, especially when children are playing. In this case, six loudspeakers have been installed more ore less uniformly distributed on the walls of the room. All rooms have been treated to achieve reverberation times below 500 ms within each octave band.

We have also created a modified version of the application for vocalists in order to simulate a sound reinforcement system consisting of two loudspeakers directed to the audience area. Instead of instrument directivity patterns we use measured loudspeaker directivity patterns from [26], with an additional rendering of the direct path from loudspeakers to the student's position on stage.

Installation and First Results

Currently, six systems (one for each instrument in Table 2 plus vocals) are installed at the conservatory of Graz. Seven teachers (one for each instrument, two for violin) with five students each regularly utilize the system for around 10 minutes per lesson. They document their experiences in a semi-structured from of a research diary. Results from the first utilization phase during the winter term 2019/2020 reveal that for some younger students the augmented acoustics were not perceivable in the beginning, however their sensibility increased over time. Moreover, the system increased motivation and love of experimentation. Agreeing with findings from literature using trained musicians [1], the students tend to play slower under conditions with more reverberation.

Conclusion and Outlook

This paper presented the Augmented Practice-Room application, a system that uses augmented acoustic reality to create the acoustic properties of a concert hall inside a small practice room. The system allows for interactive position and orientation changes of student and teacher and incorporates the instrument's directivity. It employs either loudspeaker or headphone playback of the simulated acoustics, while the direct sound of the instrument is not played back additionally, but reaches the student's ears barely attenuated because of acoustically transparent headphones. Thus, zero-latency can be achieved by an appropriate time shift of image sources, as long as the walls of the virtual room are far enough from the student. Nevertheless, the tool runs on standard PCs with small audio interfaces that provide two microphone inputs and two independent headphone outputs.

Based on the feedback of the teachers, a software update is going to add a recording function that enables the students to listen to the recorded instrument from variable perspectives and under different acoustic conditions. For now, the software is designed for a single student with a teacher. In order to enable several students to play in the same room we are currently developing an ensemble version of the Augmented Practice-Room. Several microphones around the ensemble area will pick up their sounds, excite the virtual room, and playback happens over several loudspeakers surrounding the musicians. Additionally, we plan to release a simplified version of the tool in the form of a smartphone app, allowing to sing/speak in switchable virtual rooms with a pair of headphones and an external or the built-in microphone.

Acknowledgments

This work is supported by the project Augmented Practice-Room (1023), which is funded by the local government of Styria via Zukunftsfonds Steiermark (future fond of Styria). The authors thank all students and teachers for their participation in the project.

ZUKUNFTSFONDS STEIERMARK

Das Land Steiermark

References

[1] Z. Schärer Kalkandjiev and S. Weinzierl, "The Influence of Room Acoustics on Solo Music Performance: An Empirical Case Study," *Acta Acustica united with Acustica*, vol. 99, no. 3, pp. 433–441, 2013.

[2] T. Fischinger, K. Frieler, and J. Louhivuori, "Influence of virtual room acoustics on choir singing." *Psychomusicology: Music, Mind, and Brain*, vol. 25, no. 3, pp. 208–218, 2015.

[3] G. Marentakis, C. Kranzler, M. Frank, M. Opitz, and A. Sontacchi, "Latency Tolerance Enhancement in In-Ear Monitoring Systems," in *Fortschritte der Akustik, DAGA*, 2012.

[4] J. M. Arend, T. Lübeck, and C. Pörschmann, "A Reactive Virtual Acoustic Environment for Interactive Immersive Audio," in *Audio Engineering Society Conference: 2019 AES International Conference on Immersive and Interactive Audio*, Mar. 2019.

[5] S. Serafin, A. Adjorlu, N. Nilsson, L. Thomsen, and R. Nordahl, "Considerations on the use of virtual and augmented reality technologies in music education," in *2017 IEEE Virtual Reality Workshop on K-12 Embodied Learning through Virtual & Augmented Reality (KELVAR)*. IEEE, 2017, pp. 1–4.

[6] J. R. Keebler, T. J. Wiltshire, D. C. Smith, S. M. Fiore, and J. S. Bedwell, "Shifting the paradigm of music instruction: implications of embodiment stemming from an augmented reality guitar learning system," *Frontiers in psychology*, vol. 5, p. 471, 2014.

[7] E. K. Orman, H. E. Price, and C. R. Russell, "Feasibility of using an augmented immersive virtual reality learning environment to enhance music conducting skills," *Journal of Music Teacher Education*, vol. 27, no. 1, pp. 24–35, 2017.

[8] J. Pätynen, "Virtual acoustics in practice rooms," Master's thesis, Helsinki University of Technology, 2007.

[9] F. Zotter and M. Frank, *Ambisonics - A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality*, ser. Springer Topics in Signal Processing. Springer, 2019.

[10] N. Meyer-Kahlen, D. Rudrich, M. Brandner, S. Wirler, S. Windtner, and M. Frank, "DIY Modifications for Acoustically Transparent Headphones," in *AES 148th Convention, e-Brief 61*, 2020.

[11] M. Frank and M. Brandner, "Perceptual Evaluation of Spatial Resolution in Directivity Patterns," in *Fortschritte der Akustik, DAGA*, Rostock, Mar. 2019.

[12] ——, "Perceptual Evaluation of Spatial Resolution in Directivity Patterns 2: coincident source/listener positions," in *International Conference on Spatial Audio, ICSA*, 2019.

[13] S. Weinzierl, M. Vorländer, G. Behler, F. Brinkmann, H. von Coler, E. Detzner, J. Krämer, A. Lindau, M. Pollow, F. Schulz *et al.*, "A database of anechoic microphone array measurements of musical instruments," 2017. [Online]. Available: http://dx.doi.org/10.14279/depositonce-5861.2

[14] M. Brandner, N. Meyer-Kahlen, and M. Frank, "Directivity pattern measurement of a grand piano for augmented acoustic reality," in *Fortschritte der Akustik, DAGA*, 2020.

[15] M. Romanov, P. Berghold, M. Frank, D. Rudrich, M. Zaunschirm, and F. Zotter, "Implementation and Evaluation of a Low-Cost Headtracker for Binaural Synthesis," in *Audio Engineering Society Convention 142*, May 2017.

[16] S. Clapp and B. U. Seeber, "Sound Localization in Partially Updated Room Auralizations," in *Fortschritte der Akustik, DAGA*, 2016, pp. 558–560.

[17] A. Lindau, L. Kosanke, and S. Weinzierl, "Perceptual Evaluation of Model- and Signal-Based Predictors of the Mixing Time in Binaural Room Impulse Responses," *J. Audio Eng. Soc*, vol. 60, no. 11, pp. 887–898, 2012.

[18] S. Pelzer, B. Sanches Masiero, and M. Vorländer, "3d reproduction of room acoustics using a hybrid system of combined crosstalk cancellation and ambisonics playback," in *Proceedings of ICSA*, 2011, pp. 297–301.

[19] M. Gräf and D. Potts, "On the computation of spherical designs by a new optimization approach based on fast spherical fourier transforms," *Numer. Math.*, vol. 119, 2011.

[20] J. Stautner and M. Puckette, "Designing Multi-Channel Reverberators," *Computer Music Journal*, vol. 6, no. 1, pp. 52–65, 1982.

[21] J.-M. Jot and A. Chaigne, "Digital delay networks for designing artificial reverberators," in *90th AES Conv., prepr. 3030*, Paris, February 1991.

[22] M. Blochberger, F. Zotter, and M. Frank, "Sweet area size for the envelopment of a recursive and a non-recursive diffuseness rendering approach," in *International Conference on Spatial Audio, ICSA*, 2019.

[23] M. Zaunschirm, C. Schörkhuber, and R. Höldrich, "Binaural rendering of Ambisonic signals by head-related impulse response time alignment and a diffuseness constraint," *J. Acoust. Soc. Am.*, vol. 143, no. 6, pp. 3616–3627, 2018.

[24] C. Schörkhuber, M. Zaunschirm, and R. Höldrich, "Binaural Rendering of Ambisonic Signals via Magnitude Least Squares," in *Fortschritte der Akustik - DAGA*, Munich, March 2018.

[25] F. Zotter and M. Frank, "All-Round Ambisonic Panning and Decoding," *Journal of the Audio Eng. Soc.*, vol. 60, no. 10, pp. 807–820, 2012.

[26] M. Brandner, M. Frank, and D. Rudrich, "DirPat-Database and Viewer of 2D/3D Directivity Patterns of Sound Sources and Receivers," in *Audio Engineering Society Convention 144*, May 2018.