PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Experiencing room acoustics through a library of multichannel high-resolution room impulse responses

Wieslaw WOSZCZYK and David H. BENSON

Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT)

Schulich School of Music, McGill University, Montreal, Quebec, Canada

ABSTRACT

A large room impulse response library being developed at McGill University consists of more than 10,000 individual IRs captured using multichannel microphone techniques in a large variety of acoustic spaces. The challenge is to maximize the learning value of this resource by maximizing the way the library can be experienced. To accomplish this a search interface is developed for the library that uses audio signal features, and statistical and analytical treatment, to add rich metadata describing perceptual qualities, acoustical parameters, and physical characteristics. The user is able to explore 3D graphic models of venues and audition different locations of sources and receivers, using a variety of anechoic sources. Moreover, an integrated convolution reverberation plugin enables the user to employ searched spaces for live music productions and post-production. Strict criteria are followed to maximize the sonic quality of rendering and of the exploration experience.

Keywords: Sound, Room Impulse Response (RIR), Library.

1. INTRODUCTION

There is a practical need to assess and compare acoustical spaces, such as concert halls. Over a number of decades, the acoustical community has worked to develop parametric characterization of room acoustics that would correlate with perceived room quality, and represent a listening experience at least to some extent (1, 2). While many parametric measures were adopted (3), their usefulness meets with skepticism, as halls with similar measures do sound different (4, 5, 6).

Listening is the best manner for anyone to assess the perceived quality of an acoustic space. Analytical listening during a musical performance or testing may reveal important details, which are absent in parametric characterization of the acoustics. Listening augmented with parametric assessment of the acoustics allows for subjective validation of parameters and their correlation with perceived quality characteristics. Listening also allows verification of the acoustics using suitable musical material, instruments, voices, repertoire, which are also absent in parametric characterization of rooms. Listening may include any location within the enclosure and reflect personal preference in balancing attributes to form a judgment.

Listening while being present in a hall during a performance may deliver a strong multisensory impression of the event and the acoustics, but it is usually based on hearing in a single location, at the most on two listening perspectives. Comparing listening impressions in multiple locations or in different halls is not possible live but can be delivered virtually, using auralization. For auralization to be truly a comprehensive substitute of listening in situ, it must provide a comparable quality of experience but excel in flexibility of use. This requires a carefully constructed database of halls.

2. MCGILL'S SPACEBUILDER

A large room impulse response library is being developed at McGill University as a component of Spacebuilder. **Spacebuilder** is a convolution reverberation system, which has been in development for many years at CIRMMT (7, 8, 9). It consists of three components: a large impulse response (IR) library (the *Spacebuilder Library*), a search interface for the library, which is implemented as a web application (the *Spacebuilder Webapp*), and a convolution reverb plugin (the *Spacebuilder Plugin*). The *Plugin* includes an embedded web browser, which can display a version of the *Webapp*'s search interface. When IRs are selected in this interface they are seamlessly downloaded into *Plugin*'s convolution engine, which affords *Plugin* users access to the entire IR library without requiring that it be preinstalled on their machine.

The Library and the Plugin allow for a meaningful and systematic aural comparison of perceived characteristics in different spaces. They offer comparison in different locations of each venue, enable study of the interactions between different sources and enclosures, and permit subjective evaluation of audible alterations made to each virtual space.

Two modes of listening. The Library and the Plugin allow for two modes of listening: stereophonic - directly in the *Spacebuilder Webapp*; and 3D immersive – via the Plugin, installed within a digital workstation (DAW), and using a suitable 3D monitoring environment.

2.1 The Library

At present, two interfaces are provided for searching and browsing the library. These are the Venue View and Feature View. The former displays IRs from one venue at a time and emphasizes information related to how these IRs were measured. This includes the size and shape of the venue, and the positions of the IRs' sources and receivers within it. The latter interface displays various signal features calculated from the IRs. Signal features in this context refer to standard room acoustic parameters such as the Early Decay Time (EDT) and Clarity (C80), as well as lesser-known measures. The two interfaces are each described in more detail below.

2.2 Venue View

The Venue View is designed to allow a spatial exploration of the acoustics of each venue. In this view, venues are visualized as 3D models, and each IR is visualized by a pair of icons representing its source and receiver position. Box-shaped yellow icons indicate sources while round blue icons indicate receivers. IRs can be selected for listening by clicking on a source and receiver icon, or by selecting the IR in a textual list (Figure 1). Either of these actions loads the IR into a convolution engine, where it can be convolved with a sound source and auditioned. The currently loaded source-receiver pair is always highlighted in red.

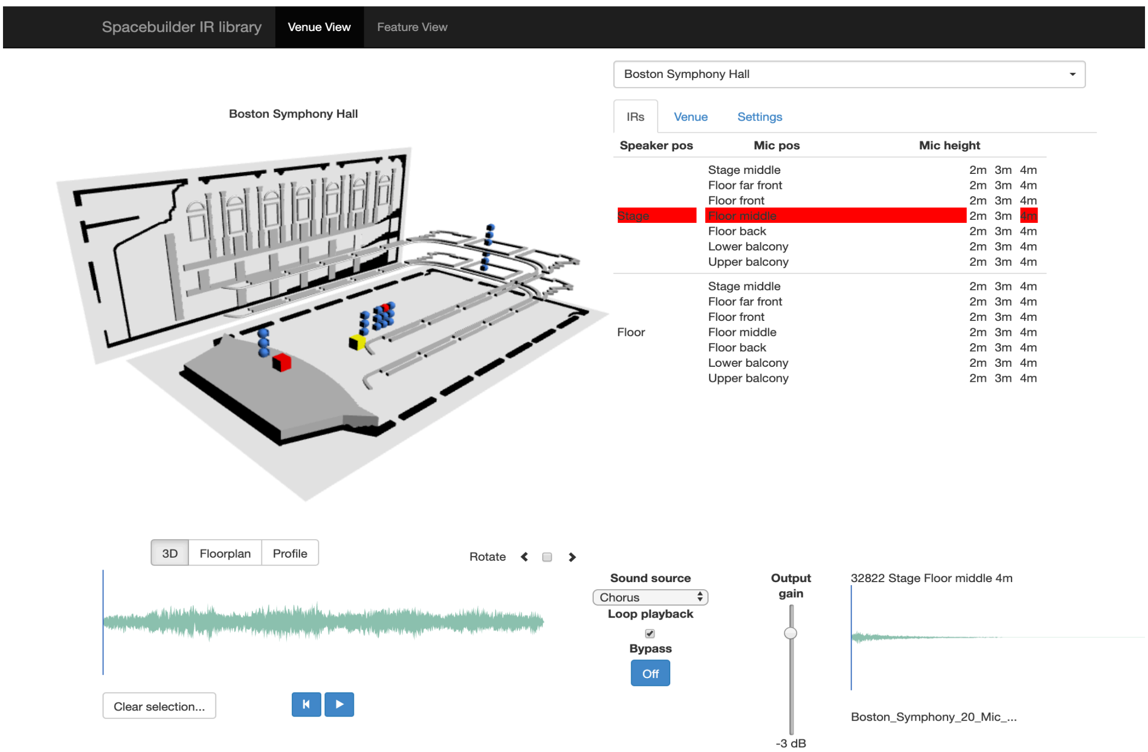

Figure 1 – The Venue View showing the IR list, with one IR selected

The Venue View also displays architectural and cultural information related to each venue. The 3D model is rendered in WebGL and may be freely rotated, zoomed and repositioned in real-time. Additionally, buttons allow the visual perspective to be quickly changed to a “floorplan” view (from directly above) or a “profile” view (from the side). Information about fine details of the venue, such as the nature of wall or floor coverings, can be gathered from a set of included photographs. Finally, a short text explains the venue’s noteworthy acoustic qualities and its cultural significance (Figure 2).

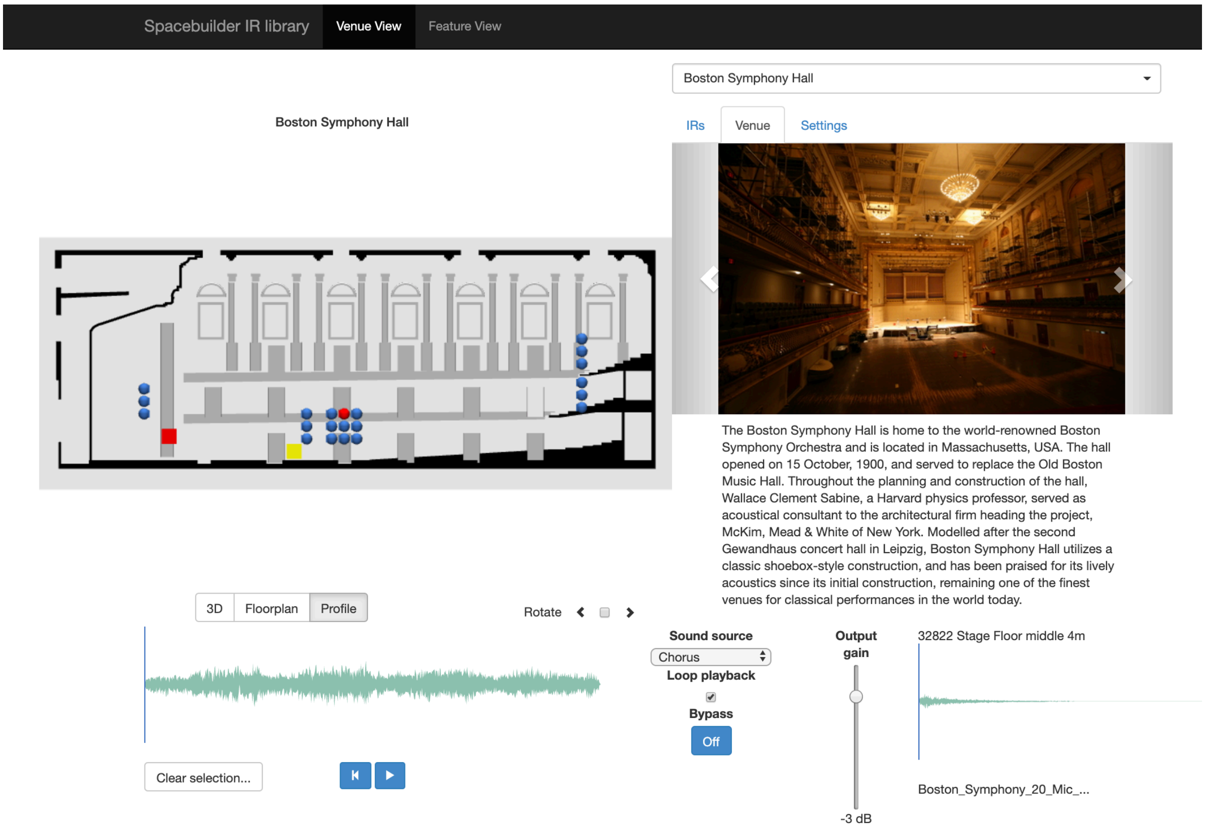

Figure 2 – The Venue View showing a venue photo and descriptive text

2.3 Feature View

Unlike the Venue View, which focuses on communicating information related to where an IR was measured, the Feature View is instead focused on conveying physical and perceptual qualities inherent to the IR’s signal. In this view, each IR is displayed as a point on a two-dimensional scatterplot. The axes of the plot represent features calculated from the IR signals. Any of a large collection of features, described below, can be assigned to the X or Y axis.

While the display shows the entire library by default, functionality also exists to “filter” the data, or to hide IRs according to certain criteria. IRs can be filtered according to feature values, using a set of range sliders, for example to hide IRs outside of a narrow range of EDT values. Finally, the interface also allows for filtering by venue, permitting investigations of inter-venue feature differences (Figure 3).

2.4 Signal features

The features available in the Feature View belong to three broad categories. First are features calculated directly from the impulse response. These features generally originate from the room acoustics community. Second are features from the room acoustics community that are calculated not on impulse responses, but rather on reverberant audio signals. (A reverberant audio signal is taken here to mean the convolution output of an IR and an anechoic musical signal.) Third are features that originate from outside of the room acoustics community.

2.4.1 Impulse-response based features

The first group includes many of the standard room acoustic measures described in the ISO 3382-1 document (10) such as the T30, the EDT and the C80. These measures are available in 8 octave bands, as well as in A, C and linearly weighted versions, as calculated by the AcMus Matlab library (11). This first group also includes the Interaural Cross Correlation measure (IACC) as described by Hidaka, Beranek and Okano (12), as well as measures of IR spectral balance, such as the Early Bass Level and Treble Ratio described by Bradley and Soulodre (13).

Also included are several impulse response-based features that are either more recently developed or whose relationship with perceptual attributes of reverberation is less well understood. These include the Deviation of Level (14), the Normalized Echo Density mixing time of Abel and Huang (15), and the Number of Peaks measure of Jeon and Vörlander (16).

Figure 3 – The Feature View showing the 1 kHz EDT and P_{ASW} for three selected venues

2.4.2 Room acoustic (spatial) reverberant audio features

The second group of features, like the first, has also been designed to predict perceptual attributes of reverberation. Unlike the first group, however, the second group are features calculated on reverberation signals rather than on impulse responses. To compute the features in this group, the IRs in the library were first convolved with a set of four representative monophonic musical signals (solo singing, choir singing, drum kit, and orchestra) to generate four reverberant audio signals. Features were calculated on these four convolution outputs.

The four features in this group developed from recent work towards predicting room acoustic attributes using binaural modeling of van Dorp Schuitman (17). The features P_{REV}, P_{CLAR}, P_{ASW}, and P_{LEV} are designed to correlate, respectively, with perceived reverberance, clarity, apparent source width and listener envelopment. The features are computed over short frames of the audio signals; the average values over the duration of each signal are recorded in the system.

2.4.3 Magnitude spectrum (timbre) reverberant audio features

The third and final group differs from group one in being calculated on reverberant audio signals, and differs from group two in not having been designed specifically to analyze reverberation. These features all describe timbral aspects of the signal's magnitude spectrum, namely the Spectral Slope, Spectral Skew, Spectral Decrease and Spectral Flatness. All are calculated via the open source Timbre Toolbox (18). As with the second group, only the average value over the all signal windows is recorded in the system.

2.5 The Spacebuilder Plugin

The Spacebuilder audio-plugin is software that transforms the input signal via several audio-processing algorithms. It contains multiple "multi-channel processing units" (MCPUs), which are able to perform certain digital signal processing tasks: convolution, linear decay-slope transformation (RT adjustment), frequency-dependent (linearly distorted) decay slope transformation, time delay, equalization, dynamic processing, switchable between Compressor/Expander, or available simultaneously, tremolo (amplitude) modulation at sub-audio modulation rates (slow and very slow oscillations), and diffusion (applied at the front end – to input signals).

Multiple MCPUs can be engaged simultaneously to allow comparing and combining several reverberation streams, including adjustments of their positions in the rendering space. The adjustability of each stream offers an ability to fine-tune reverb quality, including gentle time variance and motion. The editing, trimming the IR in time, is used to aurally evaluate temporal sections of the IR. Segmentation of IRs has proven to be useful for building parts of hybrid IRs (19).

As stated earlier, the *Plugin* is integrated with the *Spacebuilder Library* and the *Spacebuilder Webapp* via an embedded web browser. Both components of the *Webapp*'s search interface (i.e., the Venue View and the Feature View) are available to *Plugin* users for selecting IRs.

2.5.1 The Intended Plugin Applications

When installed in a DAW (digital audio workstation), the plugin may output convolution streams to multiple loudspeakers or binaural processors. Its primary fields of application are post-production for film and music recording, and live sound applications, including concerts in virtual acoustics and auralization of architectural spaces. A unique feature of the library is to allow 3D rendering of virtual acoustics with immersive characteristics including height. Different reverberant streams can be combined into a common spherical space image, and hybrid designs of new spaces having customized perceptual characteristics can be built. Architectural designers and acousticians may find the coordinated aural and parametric approach to be useful for research and development.

2.6 The Quality of Room Impulse Responses

One of the challenges in measuring a representative room impulse response involves the choice and arrangement of electroacoustic radiators emanating analytical signal within the room. Can a loudspeaker properly substitute for voices and musical instruments having different frequency and dynamic range, varied dimensions, and complex variations in directivity as solo, group, or orchestra? Musical instruments are arranged on stage in depth, height, and width, and listeners too are broadly distributed in the hall. Changes in directivity, pitch, dynamics, articulation, gestures, and movement during performance are difficult to represent in a time-invariant capture of a room response triggered by a fixed loudspeaker's signal. Occlusion, diffraction and absorption of sound radiated among musicians on stage are another complicating factor, as is the acoustic impact of audience members in proximity to individual listeners. All of these factors affect audibility of music performance (20).

This abundance of factors cannot be precisely accommodated and accounted for in a measured IR, even less so in an IR of a modeled space. So, the complexity of musical radiation must be somehow approximated in the measurement, and expanded upon later during the rendering process for auralization (21, 22). When choosing and adjusting microphone arrays, the focus must be on capturing the clarity and aural intelligibility of the room, conveying its spatial, temporal, and dynamic uniqueness. This, in principle, is similar to a photographic rendering of a room.

2.6.1 Source representation

Every effort has been made to articulate key characteristics of real acoustic sources (voices, instruments, groups) radiating energy into multiple directions. During IR measurements, multiple radiators were set up to emulate musical sources with two goals in mind: to engage room reflections cultivating directions that enhance room perception (lateral and vertical sound energy) (23), and to deliver a convincing source image to listeners conducting the measurements. Ten sources (bidirectional, unidirectional, omnidirectional) were adjusted in situ to produce a compelling image of voice, trumpet, chorus, French horn, clarinet, acoustic guitar, and percussion, using playback of anechoic monophonic recordings. Radiation synthesis of other virtual sources was previously tested in the lab, including grand piano, strings, and woodwinds. The grouped radiators covered a frequency range from 20Hz to 40kHz to assure adequate representation of physical dimensions and frequency characteristics of all instruments. To minimize interference, the sources were spaced apart and overlaps in frequency and direction were avoided. The 10 sources managed to significantly increase the number of reflections as they represented a volumetric source, not a point source.

2.6.2 Receiver representation

The capture of room response requires a 3D microphone array to record the arrivals of reflections and reverberation from all directions. A channel-based method of capture was employed with omnidirectional and bidirectional high quality microphones receiving room response at three heights, 2m, 3m, and 4m, from near the audience level to higher up in the free field. Channel-based methods were developed over decades to record music, and are accepted as being able to deliver to a human listener a believable image of musical performance in a room. The richness of time- and intensity-based stereophonic cues captured in three dimensions gives compelling auralization results in both loudspeaker and headphone listening (24, 25).

2.6.3 Capturing musical sources for auralization

To take advantage of the 3D room impulse responses, the input signal sent to the convolution engine also needs to be captured in three dimensions, including width, depth, and height. The simplest receiver of spherical sound field at the source is a tetrahedral microphone containing four nearly-coincident capsules arranged in 3 dimensions (26). The Spacebuilder plugin allows the capsule signals to be routed into appropriate convolution channels, so that e.g. the upward radiation of the source is convolved with the elevation channels of the IRs.

3. CONCLUSIONS

Experiencing halls aurally via the Library containing a large number of measured IRs provides educational and research value, as it allows for comparisons of acoustical properties, finding examples of halls with specific perceptual and parametric characteristics, while auditioning a variety of music sources. For example, one may study the aural effect of the shape of the enclosure on its acoustic response, the influence of source and listener location, assess the sound quality on stage and in the auditorium, and simulate reverberation time changes due to air volume, absorption from seat coverings, presence of audience, scattering from acoustic finishes, and even the effect of temperature and humidity.

Independent multichannel convolution engines contained in the plugin may be used to study audibility of aural fusion and segregation (27) of sources when changing parameters of auralization. What room characteristics improve perceptual separation and blending of sources? How do different sources uniquely engage the room and determine an overall spatial image perceived by a listener? To what extent do sources and rooms depend on each other, functionally and aesthetically? We hope to study such questions using these Library resources.

Can auralization replace the experience of hearing live musicians who interact with the acoustics and the audience? Perhaps. Virtual acoustics immerses both the audience and the musicians in a common simulated space, and allows the spaces to be changed. It can sustain an emotional connection between musicians and the audience, and between musicians and the room (28). The comprehensive toolset of the library, the search engine, and the convolution plugin, may be employed to create an acoustic space that articulates the music and enhances meaningful interactions. Virtual acoustics is becoming a formidable tool for designing rooms by experience.

ACKNOWLEDGEMENTS

The authors would like to thank Vanille Patier-Debray for her help curating the Library's metadata. Numerous collaborators have helped to build the McGill's Spacebuilder system and the Library, and we graciously thank them for their commitment. The financial assistance of the Social Sciences and Humanities Research Council and the Natural Sciences and Engineering Research Council of Canada are acknowledged. We thank all musicians and artists who engage to express their music in virtual acoustics.

REFERENCES

1. Schroeder M, Gottlob G, Siebrasse K. Comparative study of European concert halls: Correlation of subjective preference with geometric and acoustic parameters. The Journal of the Acoustical Society of America 56, (1974) 1195-1201.
2. Beranek LL. Concert hall acoustics: Recent findings. The Journal of the Acoustical Society of America **139**, Issue 4, 1548 – 1556 (2016).
3. ISO 3382-1:2009(en) Acoustics — Measurement of room acoustic parameters — Part 1: Performance spaces. International Organization for Standardization, Geneva, Switzerland. 2009.
4. Kirkegaard L, Gulsrund T. In search of a new paradigm: How do our parameter and measurement techniques constrain approaches to concert hall design? Acoustics Today 7(1), (2011), 7-14.
5. Lokki T. Throw away that standard and listen: your two ears work better. Journal of Building Acoustics 20(4), (2013), 283-294.
6. Hochgraf KA. The Art of Concert Hall Acoustics: Current Trends and Questions in Research and Design. Acoustics Today, Spring 2019, volume 15, issue 1, p. 28-36.
7. Woszczyk W, Leonard B, Benson D. Spatial sound design tool for 22.2 channel 3-D audio productions, with height. Audio Engineering Society Convention Paper Number: 8753, in Proceedings of the 133rd Convention, San Francisco, October 26-29, (2012).
8. Benson D, Woszczyk W. Comparing search interfaces for room impulse response libraries. Audio Engineering Society Paper Number: 8758, in Proceedings of the 133rd Convention, San Francisco, October 26-29, (2012).
9. Ko D, Woszczyk W, Chon S.H. Evaluation of a new active acoustics system in performances of five string quartets". Audio Engineering Society Convention Paper Number: 8603, in Proceedings of the 132nd Convention, Budapest, Hungary, April 26–29, 2012.
10. ISO 3382-1:2009(en) Acoustics — Measurement of room acoustic parameters — Part 1: Performance spaces. International Organization for Standardization, Geneva, Switzerland.
11. Masiero BS. AcMus - Room Acoustic Parameters, São Paulo; 2007. [cited May 31, 2019]. Matlab Central, The MathWorks Inc. https://www.mathworks.com/matlabcentral/fileexchange/11392-acmus-room-acoustic-parameters
12. Hidaka T, Beranek LL, Okano T. Interaural cross-correlation, lateral fraction, and low- and high-frequency sound levels as measures of acoustical quality in concert halls. J Acoust Soc of Am. 1995; 98(2): 988-1007.
13. Soulodre G, Bradley J. Subjective evaluation of new room acoustic measures. J Acoust Soc of Am. 1995; 98(1): 294-301.
14. Takahashi D, Togawa K, Hotta Tetsuo. Objective measures for evaluating tonal balance of sound fields. Acoust. Sci. & Tech. 2008; 29(1):1-8.
15. Abel J, Huang P. A Simple, Robust Measure of Reverberation Echo Density. In AES Convention 121; 2006 Oct 5-8; San Francisco, California, United States. Available from: AES E-Library.
16. Jeon JY, Jang HS, Kim YH, Vorländer M. Influence of wall scattering on the early fine structures of measured room impulse responses. J Acoust Soc Am. 2015;137(3):1108-1116.
17. van Dorp Schuitman J, de Vries D, Lindau A. Deriving content-specific measures of room acoustic perception using a binaural, nonlinear auditory model. J Acoust Soc Am. 2013; 133(3):1572-1585.
18. Peeters G, Giordano BL, Susini P, Misdariis N, McAdams S. The Timbre Toolbox: Extracting audio descriptors from musical signals. 2011;130(5):2902-2916.
19. Woszczyk W, Ko D, Leonard B. Convolution-based virtual concert hall acoustics using aural segmentation and selection of multichannel impulse responses, The Proceedings of INTER-NOISE 2009, The 38 International Congress and Exposition on Noise Control Engineering, Ottawa, Canada, August 23-26, 2009.
20. Meyer J. Acoustics and the Performance of Music: Manual for Acousticians, Audio Engineers, Musicians, Architects and Musical Instrument Makers, 5th ed. Springer-Verlag, New York. 2009.

21. Sacks J, Kleiner M. Front-to-rear movement of reverberation in a concert hall. The 148th Meeting of the Acoust. Soc. Am. San Diego, California, 15-19 November 2004.
22. Hong J, Woszczyk W, Begault D, Benson D. Synthesis of moving reverberation using active acoustics—preliminary report", 138th Convention of Audio Engineering Society Paper Number: 9326, Warsaw, Poland. (May 2015).
23. Pätynen J, Lokki T. Concert halls with strong and lateral sound increase the emotional impact of orchestra music. J. Acoust. Soc. Am. 139 (3), March 2016, p. 1214-1224.
24. Ko D, Woszczyk W. Virtual acoustics for musicians: subjective evaluation of a virtual acoustic system in performance of string quartets. J. Audio Eng. Soc., vol. 66, no. 09, pp. 1–12, (2018 September).
25. Kaushik S, Woszczyk W. Investigation of impulse response recording techniques in binaural rendering of virtual acoustics. 141st Convention of Audio Engineering Society Convention Paper 9609, Los Angeles, USA. (October 2016)
26. Woszczyk W, Nulsen J, Hahn E, Degtyareva K, Castillo J, Nagata H. Tetrahedral microphone: a versatile "spot" and ambience receiver for 3D mixing and sound design". Audio Engineering Society Conference e-Brief 57, published and presented at the International Conference on Spatial Reproduction – Aesthetics and Science, 2018 August 7–9, Tokyo, Japan (2018)
27. Bregman, A. S. Auditory Scene Analysis. The MIT Press, Cambridge, MA. 1990.
28. Abel JS, Woszczyk W, Ko D, Levine S, Hong J, Skare T, Wilson MJ, Coffin S, Lopez-Lezcano F. Recreation of the acoustics of Hagia Sophia in Stanford's Bing Concert Hall for the concert performance and recording of Cappella Romana, ISRA 2013, Proceedings of the International Symposium on Room Acoustics, Toronto, Canada, 2013 June 9-11.