# Evaluation of Auralization Results

Tapio Lokki and Lauri Savioja
Helsinki University of Technology, Telecommunications Software and Multimedia Laboratory,
P.O.Box 5400, FI-02015 HUT, Finland e-mail: **{Tapio.Lokki,Lauri.Savioja}@hut.fi**

There are several factors that affect the quality of auralization. Most important of them are the accuracy of room acoustic modeling and the properties of the spatial sound reproduction. These topics have gained lots of research, but still evaluation of the auralization quality is an unsolved problem. In this paper we make a short review on the performed and published evaluations. Most remarkable issues affecting the quality are discussed in more detail. As a conclusion we present that a round-robin comparison for auralization systems should be performed, similarly as earlier done with room acoustic modeling programs.

## 1 Introduction

Auralization is a tool applicable in listening of both measured responses of real spaces and room acoustical simulation results of virtual spaces. The technique is based on capturing or modeling a spatial impulse response and convolving this response with anechoic music, speech, or other suitable signal [1]. Auralization of measured responses is beyond scope of this paper and the term auralization refers always to the process of making results of room acoustic simulations audible.

Auralization provides excellent means for finding acoustical defects of a space, such as, too much reverberation, flutter echoes, focus points, or delayed sound from coupled volumes [2]. It should also help in judging the quality of acoustics of a space, e.g., a concert hall. Since auralization is often used as a demonstration tool for clients [3], it has to be authentic. Such professionals as conductors and musicians are very well trained to listen to small variations in acoustics. They certainly find out even the smallest defect which might be caused by any component of the auralization process. For this reason, it is crucial to keep up the high quality throughout the whole chain starting from selection of the stimulus and ending in the final listening experience.

The evaluation of the auralization quality is needed so that we can be convinced of the authentic auralization results. However, the major problem is to prove authenticity. An obvious method for this auralization verification is to perform listening tests between recordings of a real space and simulations of a virtual space. The recording should be flawless meaning that spatial sound capturing and reproduction should not modify the signals reaching the listeners' ears. As well known such perfect recording and reproduction techniques do not exist and a flawless reference for auralization verification is very difficult to obtain. Despite of this, auralization results have to be evaluated somehow. In this paper we discuss the issues related to both subjective and objective evaluations. We also review a few performed evaluation studies found in the literature. Based on these findings and our own experience, we conclude the paper by highlighting the most important factors affecting the quality of auralizations.

## 2 Evaluation of Auralization Quality

In general, the quality of auralization depends mainly on the properties of the employed room acoustic modeling and spatial sound reproduction techniques. In addition, the applied stimulus sound signal and listening test scheme play important roles in an evaluation. To fully analyze the quality of auralization both objective and subjective evaluation should be performed.

For evaluating room acoustics objectively, standardized parameters, such as rerberation time and clarity, have been defined [4]. Similar objective parameters for auralization quality do not exist yet. To define such parameters a binaural auditory model [5] would be a great tool, but a comprehensive model for human auditory perception is almost impossible to construct. A binaural auditory model could be applied in defining "the most pleasant acoustics", even though it depends on the purpose of use of a space, e.g., a concert hall requires different acoustics than a class room. A binaural auditory model would also serve as a great tool in evaluating the acoustics of existing concert halls.

### 2.1 Listening Test Methodology

Auralization quality can be evaluated to some extent subjectively with listening tests by comparing a reference auralization with an auralization under study. However, listening tests are far form a trivial task to complete. One major problem is the lack of quality metrics for auralization, i.e., what to ask from subjects. Typically, no information about the amount and type of differences between samples is obtained. In addition, there are several factors that should be taken into account to gain reliable results. For example, short-term auditory memory of humans is only a few seconds that makes sequential comparison problematic. In addition, subjects should be trained, hopefully with feedback in groups [6, 7], so that they really can give reliable judgments.

Unfortunately, no recommended listening test methodology for testing the auralization quality exist. In auralization verification, auralization should be indistinguishable from a reference auralization based on measurements that is exactly the same situation than in audio codecs verification. Such congruence lets us borrow the evaluation methodology from audio codec quality testing. One good recommendation is ITU-R BS.1116-1 [8] which presents methods for the subjective assessment of small impairments in audio systems including multichannel sound systems. Other good listening test methods should be found, and studies for comparing listening test methods, such as presented by Parizet et al. [9], are needed also for subjective evaluation of auralization. The discussion on the most suitable listening test method for evaluation of auralizations should continue in the research community. However, this subject is not investigated deeper in this paper.

# 3 Performed evaluations of auralization quality

Even if the optimal auralization verification methodology is not yet defined, some evaluation studies have been completed. We present here studies to exemplify performed evaluations.

## 3.1 Verifications at 1980s and 1990s

The earliest reported auralization verification is presented by Kleiner [10] who studied speech intelligibility in real and simulated sound fields already 25 years ago. He compared the speech intelligibility values, obtained with listening tests in a real theater and with a simulation of the same theater. He reported that simulations did not manage to produce same results than direct listening in the theater, but this work is exemplary and presents methodology for auralization verification.

Kuttruff [11] wrote in 1991 that his subjective comparisons of the music samples showed almost flawless agreement between the binaurally recorded samples and auralizations based on ray-tracing [12, 13]. Unfortunately, we did not find any published details of this verification.

The first work which studied the perceptual authenticity of auralization was reported by Pompetzki and Blauert [14] in 1994. They measured and modeled with the image-source method [15, 16] a lecture hall with an omnidirectional dodecahedron loudspeaker and a dummy head. Subjective comparison was performed with speech stimuli by 26 subjects. Pompetzki and Blauert reported that a reasonable authentic perception was obtained with their auralization system, even though the room acoustic modeling was not flawless.

At the same time, Farina [17] made a pair-wise comparison between measured and simulated responses. He measured and modeled with pyramid tracing one church and a sport arena. The sound source was an omnidirectional loudspeaker and sound was captured with an approximation of a human listener (a sphere with two microphones). The result of the listening test, completed by 14 subjects, was that auralizations differ from the experimental sounds, but auralizations reproduced correctly the most important acoustical effects.

## 3.2 Recent verifications

Recently, Rindel [18] has presented another auralization verification study in which comparison of binaural dummy head recordings in a Byzantine cathedral and a multipurpose hall with auralizations were performed. No accurate results are given, but he points out that recordings and auralizations differed. Already in 1994 Rindel et al. [19] reported objective validation of auralization by comparing measured and simulated room acoustical parameters. Although, the presented evaluation results were good, nothing about the aural differences were discussed at that time.

One of the recent validations of an auralization software has been published by us [20]. We measured and modeled carefully a lecture room. The sound source was one high quality loudspeaker and a real head recording technique was applied for sound capturing. Both objective and subjective evaluation proofed that only with transient sounds the subjects could distinguish the recordings from the auralizations.

We have also studied concert hall acoustics based on binaural measurements [21]. This study relates closely to evaluation of auralization, since similar listening test could be done to compare auralization results of different auralization systems. However, our study on measured concert halls showed that comparison of different concert halls based on binaural listening is very hard. Another relevant issue concerning auralization verification is the auralization in slow motion [22]. This novel way of listening spatial impulse responses could be applied in auralization verification, since with auralization in slow motion a detailed information about the early reflection can be heard.

# 4 Critical issues in high quality auralization

As presented above, few evaluations of auralization quality report that plausible, almost authentic auralizations have been realized. Nonetheless, current auralization softwares do not provide authentic results in all geometries and in all possible situations. In this section some important issues in auralization and its verification are highlighted. The list is not at all comprehensive, but it points out some relevant problems. The issues relate mostly to measurements and modeling of concert halls since they are often used in auralization studies. In addition, our viewpoint assumes that verification would be performed objectively or subjectively by comparing recorded and simulated (i.e. auralized) sounds.

## 4.1 Measurement of a reference

Traditionally, room acoustical measurements are usually performed with an omnidirectional sound source (a dodecahedron loudspeaker) and an omnidirectional microphone, a figure-of-eight microphone, and a dummy head. These equipment are adequate for capturing impulse responses with which the room acoustical parameters can be determined [4]. To capture spatial impulse responses for auralization more sophisticated equipment are needed.

The measurements of impulse responses are performed with an omnidirectional loudspeaker in a few position on the stage. For auralization purposes, such a source and a few point-like positions are not optimal. A real sound source in a concert hall is an orchestra which occupies the whole stage. In addition, each instrument of an orchestra has a different directivity characteristics. For natural auralization, a sound source should be able to produce directivity patterns similar to musical instruments or a human singer. Current loudspeakers do not have such directivity characteristics, but two research projects are pursuing to build a loudspeaker with adjustable directivity [23, 24]. Such a transducer will finally enable room acoustical measurements with a sound source imitating the directivity of a musical instrument. Then, for example, each single instrument position of an orchestra can be measured with a correct directivity and an orchestra can be represented with a group of individually radiating point sources.

Capturing a spatial impulse response is not a major problem anymore. Binaural responses can be measured with real or dummy heads. Other microphone systems for multichannel reproduction have gained a lot of attention lately, thanks to fast growing business of multichannel audio and home theaters. Farina and Ayalon [25] have presented a survey of different possibilities to capture spatial impulse responses. A recent promising technique, which enables the use of spatial impulse response in convolution process with arbitrary loudspeaker configuration has been presented by Pulkki and Merimaa [26, 27].

High quality auralization needs much better signal-to-noise ratio (SNR) than traditional room acoustical measurement methods provide. The SNR can be maximized with a sweep excitation signal and appropriate signal processing [28]. Good tutorials to measurement process and related signal processing are presented, e.g., by Müller and Massarani [29] and by Merimaa et al. [30].

#### 4.1.1 Recording of anechoic stimulus

The anechoic recording of a single instrument or a human talker is trivial in an anechoic chamber. However, if a concert hall model is auralized, a relevant sound source is the whole symphony orchestra. Recording of a symphony orchestra without acoustics of a recording space is a problem. One possibility is to record each instrument individually in an anechoic chamber, but then ensemble playing (tempo and dynamics) is very hard to control. Another possibility is to record each instrument with a pick-up microphone when the orchestra is playing in an acoustically dry environment. Neither of these two possibilities is a perfect one and currently no high quality, noise-free anechoic orchestra music is available.

## 4.2 Room acoustic modeling

Computational room acoustic modeling has been studied for 40 years and several different methods exist [31, 32]. In principle, the most accurate results are obtained with wave-based techniques such as the finite element method, the boundary element method [33], and finite-difference time-domain methods [34, 35, 36]. They all aim to numerically solve the wave equation in a modeled space, but the current wave-based techniques are far from practical applications. The more pragmatic techniques are ray based, such as ray-tracing [12, 13], beam-tracing [37], and the image-source method [15, 16]. None of the room acoustic modeling techniques is optimal in all geometries and over the whole audible frequency range.

There exist no reflection model (for any of the modeling techniques) which handles specular and diffuse reflections as well as diffraction at the same time [32]. Such a complete model is essentially needed for accurate room acoustic prediction. The image-source method usually operates with pressure signals. It can solve specular reflections and it has been extended to solve also diffraction [38]. Both magnitude and phase can be accurately predicted, but good modeling method for diffuse reflection is not found. In contrast, diffuse reflection modeling is quite simple to implement with ray-tracing, but then predictions are computed with energy signals, not pressure signals which are needed for auralization. Although, e.g., Farina [17] proposed simple method to convert energy responses into pressure responses, the correct phase information is lost when energies are applied in computation. The phase information is essential since in a room, reflections may reach the listener at the same time, thus interfering with each other. The comprehensive reflection model includes also angle dependent material absorption. Its modeling has not yet been solved, although, simplified models have been proposed [39, 40].

### 4.3 Spatial sound reproduction

Spatial sound reproduction is another issue that makes auralization hard. High quality auralization requires good spatial accuracy without any coloration problems.

Binaural reproduction is often mentioned to be good in auralization, but at high frequencies binaural reproduction either with headphones or loudspeakers is almost impossible to realize accurately. Even if individual HRTFs were applied the unidealities of headphones colorate sound and may ruin otherwise high quality auralization. Multichannel reproduction with loudspeakers doesn't have such coloration problems, but other error sources exist. Such problems include, e.g., finding of optimal loudspeaker configuration and best panning method, as well as the elimination of acoustics of the listening room.

For auralization the best sound reproduction system, in our opinion, is a multichannel reproduction in an anechoic chamber. One example of such a system is reported by Pulkki and Hirvonen [26]. In their system 16 loudspeakers are mounted in a 3D layout in an anechoic chamber and vector base amplitude panning (VBAP) [41] is applied to reproduce each simulated reflection from a correct direction.

## 5 Conclusions

Subjective auralization verification is based on comparing sounds measured in a real space with auralized sounds provided by room acoustic modeling. For a valid comparison a comprehensive spatial sound reference is required, but with current equipment such a measurement is difficult. However, with careful design a suitable reference can be measured for case studies enabling meaningful listening tests.

The research community should organize a round robin test for auralizations, in the similar way than room acoustic prediction softwares have been evaluated [42, 43, 44]. Such a common effort of the research community would be necessary in finding the methods and models to implement an auralization system that provides authentic results in any geometry.

## References


[1] M. Kleiner, B.-I. Dalenbäck, and U.P. Svensson. Auralization – an overview. *J. Audio Eng. Soc.*, 41(11):861–875, Nov. 1993.

[2] J.H. Rindel. Evaluation of room acoustic qualities and defects by use of auralization. In *148th Meeting of the Acoustical Society of America*, San Diego, California, 2004.

[3] ArupAcoustics. Arup soundlab, May 2005. http://www.arup.com/project.cfm?pageid=2737.

[4] ISO Standard 3382. *Acoustics – Measurement of the reverberation time of rooms with reference to other acoustical parameters*. International Standards Organization, 1997.

[5] M. Karjalainen. A binaural auditory model for sound quality measurements and spatial hearing studies. In *Proc. 1996 IEEE Int. Conf. Acoust., Speech, and Signal Processing*, volume 2, pages 985–988, Atlanta, GA, USA, May 1996.

[6] N. Zacharov and V.-V. Mattila. GLS – a generalised listener selection procedure. In *the 110th Audio Engineering Society (AES) Convention*, Amsterdam, the Netherlands, May 12-15 2001. preprint no. 5405.

[7] D. Isherwood, G. Lorho, N. Zacharov, and V.-V. Mattila. Augmentation, application and verification of the generalized listener selection procedure. In *the 115th Audio Engineering Society (AES) Convention*, New York, NY, USA, Oct. 10-13 2003. preprint no. 5894.

[8] ITU-R. Recommendation BS.1116-1, Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems. International Telecommunication Union Radiocomminication Assembly, 1997.

[9] E. Parizet, N. Hamzaoui, and G. Sabatié. Comparison of some listening test methods: A case study. *Acta Acustica united with Acustica*, 91(2):356–364, March/April 2005.

[10] M. Kleiner. Speech intelligibility in real and simulated sound fields. *Acustica*, 47(2):55–71, 1981.

[11] H. Kuttruff. Auralization of impulse responses modelled on the basis of ray-tracing results. In *the 91st Audio Engineering Society (AES) Convention*,

New York, NY, USA, Oct. 4-8 1991. preprint no. 3122.

[12] A. Krokstad, S. Strom, and S. Sorsdal. Calculating the acoustical room response by the use of a ray tracing technique. *J. Sound Vib.*, 8(1):118–125, 1968.

[13] A. Kulowski. Algorithmic representation of the ray tracing technique. *Applied Acoustics*, 18(6):449–469, 1985.

[14] W. Pompetzki and J. Blauert. A study on the perceptual authenticity of binaural room simulation. In *Proc. of the Wallace C. Sabine Centennial Symposium*, pages 81–84, June 5-7 1994.

[15] J. B. Allen and D. A. Berkley. Image method for efficiently simulating small-room acoustics. *J. Acoust. Soc. Am.*, 65(4):943–950, 1979.

[16] J. Borish. Extension of the image model to arbitrary polyhedra. *J. Acoust. Soc. Am.*, 75(6):1827–1836, 1984.

[17] A. Farina. Auralization software for the evaluation of the results obtained by a pyramid tracing code: Results of subjective listening tests. In *Proc. 15th Int. Congr. Acoust. (ICA'95)*, pages 441–444, Trondheim, Norway, June 1995.

[18] J.H. Rindel and C.L. Christensen. Room acoustic simulation and auralization - how close can we get to the real room? In *Proceedings of Eighth Western Pacific Acoustics Conference*, Melbourne, Australia, April 7-9 2003. Keynote Lecture; CD-ROM, 8 pp.

[19] J.H. Rindel, C. Lynge, G. Naylor, and K. Rishøj. The use of a digital audio mainframe for room acoustical auralization. In *the 96th Audio Engineering Society (AES) Convention*, Amsterdam, The Netherlands, 1994. preprint no. 3860.

[20] T. Lokki. *Physically-based Auralization - Design, Implementation, and Evaluation*. PhD thesis, Helsinki University of Technology, Telecommunications Software and Multimedia Laboratory, report TML-A5, 2002. Available at http://lib.hut.fi/Diss/2002/isbn9512261588/.

[21] T. Lokki. Subjective comparison of four concert halls based on binaural impulse responses. *Acoustical Science and Technology*, 26(2):200–203, 2005.

[22] T. Lokki. Auralization of simulated impulse responses in slow motion. In *the 118th Audio Engineering Society (AES) Convention*, Barcelona, Spain, May 28-31 2005. paper no. 6500.

[23] C. Ianniello and A. Farina. Design and realisation of a transportable device for the measurement of acoustical spatialisation in theatres and concert halls. Italian research project PRIN-2003 n. 2003090372_004, 2003. http://cofin2003.cineca.it/cgi-bin/relazioni/form-vis_form.pl.

[24] O. Warusfel and N. Misdariis. Sound source radiation syntheses: From performance to domestic rendering. In *the 116th Audio Engineering Society (AES) Convention*, Berlin, Germany, May 8-11 2004. paper no 6018.

[25] A. Farina and R. Ayalon. Recording concert hall acoustics for posterity. In *AES 24th Int. Conf. on Multichannel Audio*, pages 38–51, Banff, Canada, June 26-28 2003.

[26] V. Pulkki, J. Merimaa, and T. Lokki. Reproduction of reverberation with spatial impulse response rendering. In *the 116th Audio Engineering Society (AES) Convention*, Berlin, Germany, May 8-11 2004. paper no. 6057.

[27] V. Pulkki and J. Merimaa. Spatial impulse response rendering: Listening tests and applications to continuous sound. In *the 118th Audio Engineering Society (AES) Convention*, Barcelona, Spain, May 28-31 2005. paper no. 6371.

[28] A. Farina. Simultaneous measurement of impulse response and distortion with a swept-sine technique. In *the 108th Audio Engineering Society (AES) Convention*, Paris, France, Feb. 19-22 2000. preprint no. 5093.

[29] S. Müller and P. Massarani. Transfer function measurement with sweeps. *J. Audio Eng. Soc.*, 49(6):443–471, June 2001.

[30] J. Merimaa, T. Peltonen, and T. Lokki. Concert hall impulse responses - Pori, Finland, reference documentation. Published online at http://www.acoustics.hut.fi/projects/poririrs/, May 2005.

[31] L. Savioja, J. Huopaniemi, T. Lokki, and R. Väänänen. Creating interactive virtual acoustic environments. *J. Audio Eng. Soc.*, 47(9):675–705, 1999.

[32] U.P. Svensson and U.R. Kristiansen. Computational modeling and simulation of acoustic spaces. In *AES 22nd Int. Conf. on Virtual, Synthetic and Entertainment Audio*, pages 11–30, Espoo, Finland, June 15-17 2002.

[33] A. Pietrzyk. Computer modeling of the sound field in small rooms. In *Proc. AES 15th Int. Conf. on Audio, Acoustics & Small Spaces*, pages 24–31, Copenhagen, Denmark, Oct. 31 - Nov. 2 1998.

[34] D. Botteldooren. Finite-difference time-domain simulation of low-frequency room acoustic problems. *J. Acoust. Soc. Am.*, 98(6):3302–3308, 1995.

[35] L. Savioja, J. Backman, A. Järvinen, and T. Takala. Waveguide mesh method for low-frequency simulation of room acoustics. In *Proc. 15th Int. Congr. Acoust. (ICA'95)*, volume 2, pages 637–640, Trondheim, Norway, June 1995.

[36] L. Savioja. *Modeling Techniques for Virtual Acoustics*. PhD thesis, Helsinki University of Technology, Telecommunications Software and Multimedia Laboratory, report TML-A3, 1999.

[37] T.A. Funkhouser, N. Tsingos, I. Carlbom, G. Elko, M. Sondhi, J. West, G. Pingali, P. Min, and A. Ngan. A beam tracing method for interactive architectural acoustics. *J. Acoust. Soc. Am.*, 115(2):739–756, February 2004.

[38] U.P. Svensson, R.I. Fred, and J. Vanderkooy. Analytic secondary source model of edge diffraction impulse responses. *J. Acoust. Soc. Am.*, 106(5):2331–2344, 1999.

[39] J.H. Rindel. Modelling the angle-dependent pressure reflection factor. *Applied Acoustics*, 38(2-4, Special Issue on Computer Modelling and Auralisation of Sound Fields In Rooms):223–234, 1993.

[40] J.H. Rindel. Computer simulation techniques for the acoustical design of rooms - how to treat reflections in sound field simulation. In *Proc. ASVA'97*, pages 201–208, Tokyo, Japan, Apr. 1997.

[41] V. Pulkki. Virtual sound source positioning using vector base amplitude panning. *J. Audio Eng. Soc.*, 45(6):456–466, June 1997.

[42] M. Vorländer. International round robin on room acoustical computer simulations. In *Proc. 15th Int. Congr. Acoust. (ICA'95)*, volume 2, pages 689–692, Trondheim, Norway, June 1995.

[43] I. Bork. A comparison of room simulation software – the 2nd round robin on room acoustical computer simulation. *Acta Acustica united with Acustica*, 86:943–956, 2000.

[44] I. Bork. Simulation and measurement of auditorium acoustics – the round robins on room acoustical computer simulation. In *Proceedings of the Institute of Acoustics*, volume 24, London, England, July 2002.