Continuous and exchangeable directivity patterns in room acoustic simulation

Sönke Pelzer, Martin Pollow, Michael Vorländer

Institute of Technical Acoustics, RWTH Aachen University, 52066 Aachen, E-Mail:spe@akustik.rwth-aachen.de

Introduction

Room acoustic simulation by using geometrical acoustics is usually implemented with binaural receivers and a given sound source directivity. This way, however, the simulated signals are restricted to a specific set of HRTFs and a fixed radiation pattern of the simulated source. Adjustments such as individualization of the HRTFs or modification of the source directivity pattern cannot be performed after the simulation is finished.

Using a general interface between transfer path and the directivities of both sender and receiver instead, the RIRs can be freely composed of the transfer paths and the directivity patterns of sender and receiver in order to obtain the desired result. To do so, the simulation is performed for a set of basic radiation patterns, called spherical harmonics.

Geometrical Room Acoustics Simulation

The room acoustics simulation software *Raven* [1] that is currently developed at ITA Aachen comprises a hybrid method based on geometrical acoustics using image source and ray tracing. Besides room acoustical parameter prediction this program is also used for auralizations, usually generating binaural room impulse responses (BRIRs) from an HRTF database. The program was extended to represent sources and receivers using spherical harmonics (SH) base functions, so that arbitrary source and receiver directivities can be applied after the simulation is finished. Due to its flexibility, the SH receiver representation allows e.g. subsequent individualization of HRTFs for headphone reproduction or Ambisonics decoding on loudspeaker setups.

Immersive Auralizations

To achieve immersive auralizations, a natural sounding 3-D sound scape must be reproduced with dynamic components, many sound sources and the possibility for user interaction. The calculation and convolution of many room impulse responses (RIRs) is very demanding in terms of computation power, so that this seems to be feasible only on fast compute clusters.

However, the most common and natural user interaction while listening to auralizations is turning ones head to listen to the room and localize sound sources. When using tracked headphones and HRTF data in the SH domain, the system can react on these head movements by re-orientating the HRTF dataset, implemented by a multiplication with the Wigner-D rotation matrix [2] as shown in Eq. 1. This operation is lightweight and can be done in real-time, providing at least a minimum of interaction at very low cost.

$$h_{nm,\mathrm{rot}}(f) = D^{m}_{k,n}(\alpha,\beta,\gamma) \cdot h_{nm}(f) \tag{1}$$

General Acoustic Transfer Path

If both the sender and the receiver are to be kept flexible using a SH representation, the RIR has to be calculated for any combination of the spherical base functions. The result is a big SH to SH matrix with dimensions N x M for a source with SH order of M and a receiver of order N. Figure 1 illustrates the generalized transfer path including the possible rotation matrices.

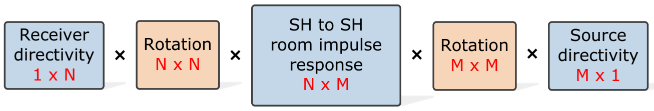

Figure 1: Generalized transfer path structure with separated source and receiver representations in SH domain and a matrix of RIRs for a maximum of flexibility.

Source Directivities

Directivities in geometrical simulations are usually implemented with directional filtering. Radiation patterns of loudspeakers and also musical instruments are available, e.g. in the Common Loudspeaker Format (CLF) or OpenDAFF [3]. Following an idea, initially proposed by Weinreich in 1980 [4], it is possible to decompose the source radiation into orthonormal base functions on the sphere (SH), deriving a spatially continuous solution for the measured radiation pattern. As of today, measurements of musical instruments were done with up to 64 microphones, so that SH coefficients can be derived approximately up to the 7^{th} order.

Receivers

In common auralization tools, receivers are implemented binaurally using HRTF datasets. After the simulation, these HRTFs cannot be changed anymore. With the flexible SH representation, arbitrary HRTFs can be multiplied in SH domain on finished simulation results. This allows finding a matching dummy head dataset or using the own individual HRTFs, if available. Multiplying the directivities of common microphone setups, such as XY, M/S, ORTF, is possible as well as Ambisonics decoding for reproduction on loudspeaker systems. Furthermore, when using HRTF in SH domain, there is no need to implement interpolations, as the data is continuously defined for any angle.

So far, it not known yet up to which order HRTFs has to be encoded into SH coefficients to cover the human ear's maximum localization abilities. The continuous definition can be of advantage here, but this has yet to be proven. The magnitude of SH coefficients in higher orders can also be reduced if a suitable center point is chosen, e.g. closer to the ear than to the center of the head [5].

Implementation

With the presented method, local disk space consumption can become an issue. The pre-processing of the SH RIR matrix for an receiver of order 30 and convolution with anechoic source signals takes already approx. 5GB per minute (16Bit/44.1kHz) of disk space. For complex scenes these impulse responses have to be rendered for each sound source, so that in scenes with many sources it can be advisable to superpose all those sources. The ability to change the source signal content is lost, but an unlimited number of sources can be superposed.

The convolution of SH RIRs with HRTFs in SH domain is usually real-time capable due to the short filter length of HRTFs with typically only 256 samples. But a receiver of order 30 streams nearly 1000 channels of data non-stop and simultaneously. In a test scenario this 1000-channel convolution used 12ms CPU time per 256-sample frame (6ms), so that continuous playback would not be possible. A parallelized convolution on a multi-core CPU could solve the problem, but another solution is to convolve only the coefficients of the first 21 orders (484 channels). This scalability enables real-time capable playback again, using a reduced spatial resolution. The HRTFs are still continuously defined, so that the minimum order as from which a negative effect on source localization might be detectable is not yet known and still has to be investigated.

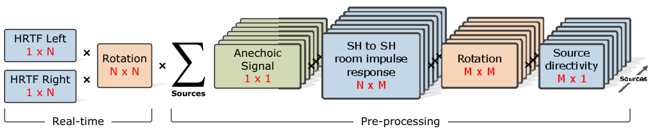

Figure 2: Immersive real-time auralization of multiple sound sources using head-tracking and a generalized transfer path of SH impulse responses, source directivities and HRTFs.

Real-time performance

A possible real-time auralization implementation using the presented method is illustrated in Figure 2 and benchmark results on different CPUs dependent on the receiver order are shown in Figure 3. The convolution was implemented as a single-threaded serial operation without multi-core support. As a result, the tested CPUs were capable of real-time reproduction with a receiver of approx. order 20.

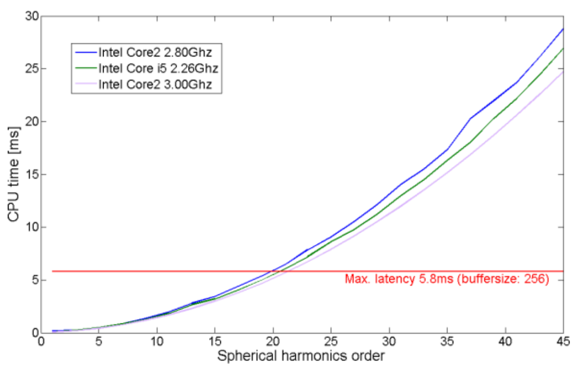

Figure 3: Single-threaded convolution performance of SH receivers with varied orders on three different CPUs.

Applications

The ability to vary source and receiver characteristics in finished simulations can be used as presented for individualization of HRTFs and immersive real-time auralizations, but also a broad field of other applications is possible. Simulations are often used for the design of sound reinforcement systems, different loudspeakers can now be quickly exchanged and objective measures used for comparison as well as immediate auralizations.

The flexibility of source characteristics can be used to implement time-varying directivities, as seen in some musical instruments, especially when player movements are considered. But also machines with any kind of rotating equipment usually show different directional radiation dependent on the operating conditions.

Conclusions

The presented method implements a room acoustic simulation with a generalized transfer path with separated and flexible elements. Sources and receivers are modeled using spherical harmonics, and the room impulse response becomes a SH to SH matrix. This enables to vary the spatial characteristics of sources and receivers in finished simulation results. Application of directivity data and source/receiver rotations are cheap matrix multiplications, so that these operations can be done in real-time, e.g. for immersive real-time auralizations with a tracked headphone.

The method was implemented into ITA's room acoustics simulation framework Raven and can be used now as an effective model base for listening tests. It still has to be investigated which SH order is needed for sources and receivers to appropriately model the radiation of real sources and the characteristics of HRTFs. Therefore future research is planned using the presented technique.

References

[1] Schröder, D.: Physically-based real-time auralization of interactive virutal environments. PhD thesis, RWTH Aachen University, Aachen, Germany, 2011

[2] Pendleton, J. D.: Euler angle geometry, helicity basis vectors, and the Wigner D-function addition theorem. American Journal of Physics, vol. 71, p. 1280, 2003

[3] Wefers, F.: OpenDAFF - Ein freies quell-offenes Software-Paket für richtungsabhängige Audiodaten. Fortschritte der Akustik, DAGA, 2010

[4] Weinreich, G., Arnold, E. B.: Method for measuring acoustic radiation fields. J. Acoust. Soc. Am. 68:2, 404-411, 1980

[5] Ben Hagai, I., Pollow, M., Vorländer, M., Rafaely, B.: Acoustic centering of sources measured by surrounding spherical microphone arrays. J. Acoust. Soc. Am. Volume 130, Issue 4, pp. 2003-2015, 2011