Proceedings of 20th International Congress on Acoustics, ICA 2010

23-27 August 2010, Sydney, Australia

Frequency- and Time-dependent Geometry for Real-time Auralizations

Sönke Pelzer (1) and Michael Vorländer (1)

(1) Institute of Technical Acoustics, RWTH Aachen University, Aachen, Germany

PACS: 43.55.Ka

ABSTRACT

The quality of present-day room acoustic simulations stands and falls by the quality of the underlying CAD room models. A ‘high-quality’ room model does not implicate that it has to be highly detailed with a lot of small objects and ornamentation. High accuracy of a model and its auralization is only achieved when basic acoustic principles are regarded. This means for geometrically based simulations that wavelengths from 1.7cm till 17m have to be handled with regard to their reflection/scattering pattern at room or object surfaces. As this spans the dimensions of the majority of known objects, walls etc., only an adapting room model that changes its level of detail accordingly to the incident sound wave frequency can ensure correct results, e.g. for low-frequency specular reflections. When it comes to real-time auralizations, which are used in sophisticated virtual reality systems, another emerging aspect is how to deal with very complex room geometries in limited computation time. Using active frequency dependent geometry has great advantages in this discipline due to much faster simulations when simple geometries with a low polygon count are involved. Lower frequencies typically travel much longer than higher ones in common rooms and thus they produce the majority of computation load which is now significantly reduced, yielding a great speed-up potential. Going on, the introduction of a temporal discretization reducing room model details step by step over the duration of the room impulse response can save valuable computation resources as well. This technique makes use of perceptual characteristics of the human ear, which is not able to distinguish fine structures in the late part of the impulse response. Furthermore, most of the energy in this late part originates from diffuse reflections for which the exact geometry does not matter. Thus, the active geometry model switches to simpler structures for late reflections. In this contribution, the newly developed active geometry model, which uses a frequency and time dependent level of detail, will be presented as well as results from comparative listening tests. These could point out the necessary complexity for the highest detail step as well as the maximally allowed simplification based on human perception.

INTRODUCTION

Room acoustical simulation software is a popular tool today. It is used to predict and improve the acoustics of existing buildings and those to be planned, respectively. The application range is not limited to performance spaces such as concert halls, but implies a broad range of areas including lecture halls and schools, offices and factories, railway stations and airport terminals, to mention just some of them. The requirements for all these applications differ significantly from one to the other, e.g. the simulation of a concert hall should deliver a very high quality of the generated impulse responses (IRs) while the difficulties for airport terminals or office buildings with several floors and many rooms lie in their size and complexity.

In the history of room acoustical simulations until today, the input CAD models have almost always been a simplification of architectural models or a new design where room boundaries were approximated very roughly by an acoustic expert. The simplifications are necessary due to laws of physics and also due to the computational constraints. The physics behind geometrical acoustics (GA) makes use of a wave-particle dualism where the sound waves that travel through a space are represented by particles or sound rays, which are reflected on surfaces following simple geometric rules. But when objects or surfaces are roughly of the same size or smaller compared to the wavelength of the incident sound wave, those geometric reflections would be in contradiction to the acoustic effects that occur in such situations, and those modeled reflections would simply be wrong. Consequently, the acoustic CAD model has to be matched to all sound frequencies of interest covering three decades of human hearing, which is not possible while using just one model.

Usually, a compromise is made concerning the degree of detail in the acoustic model - but without any verification. So far it is not known how much simplification is allowed without being audible or how much accuracy is needed when modeling fine details of the room structure. A highly detailed room model runs the risk of bad accuracy for mid and low frequencies due to wrongly modeled reflection patterns and especially very high unnecessary computational effort. On the other hand a very low detailed model might be too simple, missing out on important acoustic characteristics.

By running listening tests, it is possible to discover some guidelines for the development of acoustic models. Comparative tests using different degrees of spatial resolution can unveil the abilities of the human ear to detect changes of the surrounding geometry. For such listening tests, it is necessary to produce a set of models of the same room with graduated level of detail (LOD), sorted by the size of the smallest structure that appears in the model. By comparing all models among each other in a listening test, a value in *meters* can be obtained that marks an audibility threshold for structure sizes. This design limit can be referenced in future when modeling rooms for auralization purposes, a further elaboration of details would not be audible. The definition of a smallest occurring structure size will implicitly define a cut-off for low frequencies when using GA simulations on this model.

The idea is now to use all these models with graduated levels of detail at the same time in one simulation. When a sound wave is reflected at a wall, each frequency component of this wave will see a different wall surface which is chosen from the set of detail-graduated CAD models and matched to the wavelength.

In the course of this paper, a very effective and efficient algorithm will be presented that is able to implement this idea. An acoustic model will be decomposed into different LOD steps, and during the simulation it will adapt to the frequency of incident sound waves, so that it always presents a reflection surface that supports the physics behind GA simulations. Furthermore, the algorithm will reduce the average number of polygons that have to be taken into account for intersection tests compared to a simulation that would use the highest-detailed model only. This will save valuable computation time while modeling physically more accurate reflections at the same time. An additional speed-up will be achieved by using the simplified models also for higher frequencies in the late part of the impulse response, which has lower amplitudes and is energetically dominated by diffuse reflections that are less dependent on surface orientations.

AURALIZATION AND VIRTUAL REALITY

The main reason for speeding up simulations is found in the field of auralizations (Vorländer 2007). Static offline auralizations that use fix positions for sound sources and the listener, and especially fix view directions for these, will never be able to deliver a good representation of the room acoustics that is to be observed. It is a typical human behaviour to move around and especially use head motion and rotation to closely listen to sound sources and localize them. Actually, the capability of human hearing is highly increased by head movements (Blauert 1997).

Thus, auralizations become a much richer experience and also much more meaningful when they are able to account for and react on listener movements, with head rotations at capital importance. Such interactive real-time auralizations are used in Virtual Reality (VR) systems, where they are combined with 3D visual reproduction to fully immerse a user into a virtual world, forming today's state-of-the-art tools in many development and analysis processes.

Even after recent major leaps in the architecture of hardware and software, that helped VR systems to reach outstanding qualities and abilities, real-time auralizations with physically based soundfield rendering have significant computational complexity that can only be handled by distributed computation on PC-clusters (Schlütter 2009). Thus, any speed improvements that do not affect the simulation quality are very welcome.

PREVIOUS WORK AND BACKGROUND

The core techniques of present-day room acoustical simulations comprise image source (IS) (Allen & Berkley 1979) and ray tracing (RT) (Vorländer 1988) algorithms, or combinations of both. These are GA based methods that work well in large spaces. For small rooms, wave effects such as room modes become very important, thus FEM/BEM simulations offer good results. Anyway, last-mentioned are not suitable for real-time applications due to their heavy computational load.

Important milestones in GA simulations were the successful integration of a forward audibility test of image sources (Vorländer 1989, Funkhouser et al. 1998) scattering (Heinz 1993), sound transmission (Thaden 2005, Schröder 2007) and diffraction (Svensson 1999, Stephenson 2004, Schröder & Pohl 2009).

The main task in all GA based algorithms is to intersect the resulting sound rays (beams, cones etc.) with the room geometry, which is done by nominating the polygon that is hit and applying a reflection pattern to the ray's travel line. In general, this substantial task is linearly dependent with *O(n)* on the number of polygons *n* in the scene. Figure 1 shows an example of real simulations with an acceleration technique based on Binary Space Partitioning (BSP), which reduces the computational effort to *O(log(n))* (Schröder & Lentz 2006).

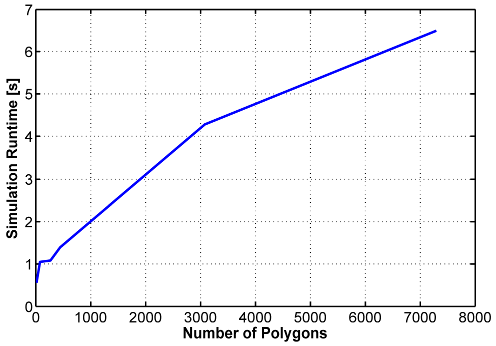
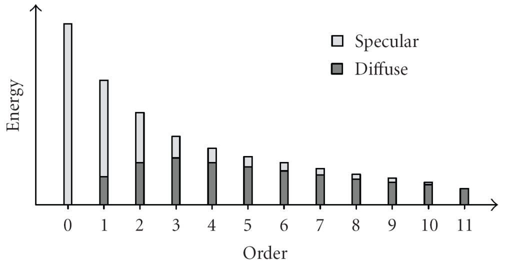


Figure 1. Increasing simulation runtimes with higher polygon counts using a BSP-based acceleration technique.

Nevertheless, the number of IS that have to be generated at a first stage grows significantly with the desired reflection order and the number of polygons. Hence, the polygon count has to be kept as low as possible and the IS model itself is only applicable for a limited number of reflections.

After Kuttruff (1995), however, these scattered reflections become the energetically dominant part in the room IR already after three reflections, as shown in Figure 2.

In any implementation of GA based simulations, the performance strongly depends on the number of polygons in the scene, so again this number has to be as low as possible - but still high enough to ensure a satisfactory soundfield reproduction.



Source: (Kuttruff, 1995)

Figure 2. Scattered energy becomes the energetically dominant part of an impulse response already after 3 reflections.

Other studies on the number of polygons and necessary complexity of CAD models were carried out by (Pompetzki 1993), who used an image source algorithm to build auralization filters that were compared to measured impulse responses afterwards. This work, using a model of a lecture hall in three different spatial resolutions (48, 126 and 238 polygons), has identified and properly targeted key aspects of acoustic room modeling, from which important insights can be derived: the accuracy of a simulation does not necessarily increase with the degree of detail of the model. A proper approximation is sufficient. From today's point of view, two important aspects are missing in this work, namely the involvement of scattering and the frequency-dependency of room structures.

Siltanen, Lokki and Savioja (Siltanen et al. 2006) also worked on simplification of room geometries. They implemented a modified Marching Cubes algorithm that could automatically reduce the level of detail of a complex CAD model, yielding a seamless set of models with arbitrary number of polygons. Unfortunately, the automatic geometry reduction is not yet reliable enough, and appropriate acoustic models can still only be obtained manually by an experienced acoustician. The automatically generated models contained structural artifacts and were not able to prove that important room acoustical parameters remain unchanged between different detail steps.

SIMULATION SOFTWARE RAVEN

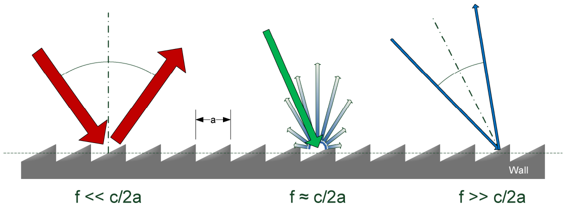
The room acoutics simulation software *Raven* (Schröder et al. 2010), that is currently developed at the Institute of Technical Acoustics at RWTH Aachen University in Aachen, Germany, uses a hybrid geometric approach which combines advantages of the IS technique for modeling all occuring low-order specular reflections with precise time resolution and RT technique for modeling specular and scattered reflections to reproduce a diffuse late reverberation field.

Most present-day knowledge of Virtual Acoustics is aggregated in *Raven*'s algorithms, enabling a physically accurate simulation of sound propagation in complex environments, including important wave effects such as sound scattering (Schröder 2007), airborne sound insulation between rooms (Wefers 2007) and sound diffraction (Schröder & Pohl 2009). Freely movable sound sources and receivers are supported, and also modifications and manipulations of the geometry itself (Schröder & Ryba 2010).

In the current version, *Raven* is now able to handle different versions of the same scenario in parallel, featuring a geometry LOD controller that is described in the following sections.

PHYSICS OF SURFACE REFLECTION

When sound waves hit the surface of an object or wall, they are reflected with a certain pattern. For glossy surfaces, the reflection is mainly specular while rough surfaces scatter sound in various directions. Especially small structures on surfaces cause very high scattering coefficients. The reflection pattern is strongly dependent on the angle and frequency of the incident sound waves, as shown in Figure 3.



Source: (Vorländer, 2007)

Figure 3. Frequency-dependent reflection patterns at rough surfaces with structures of dimension *a*. *f* denotes the frequency of the incident wave and *c* the sound speed. Low frequencies are not able to see structural fine details.

This illustrates the challenge that was already pointed out in the Introduction. Human hearing is able to detect frequencies ranging from 20Hz up to 20kHz, which results in wavelengths from 1.7cm to 17m in air at room temperature. On the left hand in Figure 3, it is shown how low frequencies are not able to see structural details on surfaces. Their reflection pattern is geared to the underlying superordinate structure.

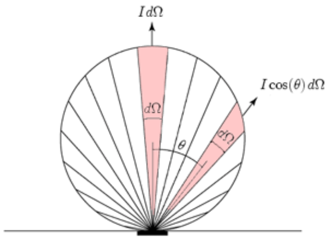


Figure 4. Scattering at rough surfaces can be approximated by Lambert's cosine law.

For wavelengths that are roughly the same size as the structural details, the sound will be scattered without preferred direction but with an angle dependent scattering pattern. This can be approximated by Lambert's cosine law, see Figure 4, or by other more specific characteristics.

It is obvious that running a simulation for low frequencies with a model with high details leads to physically wrong results. Furthermore the computation time is strongly dependent on the polygon count and thus unnecassarily high.

Simulating very low frequencies with GA based methods is generally unfavorable. They rely on approximating spherical waves by plane waves which is only valid for long distances between sources, receivers and walls. Furthermore they neglect phase interference effects which is only valid for cases where sound waves and especially room modes are densely overlapping in random directions. This is given above the *Schroeder Frequency* (Schroeder 1962):

$$f_S \gg 1200\sqrt{\frac{T}{V}} \tag{1}$$

V denotes the room volume and *T* the reverberation time.

Thus, GA is often used and most precise in large spaces that lead to a very large number of polygons to be handled. Especially for use in real-time applications, it would be very advantageous to run the simulation for lower frequencies on simpler geometries for speeding up the simulation and to comply with the laws of physics. In addition, the absorption in common rooms is low for deep frequencies causing high reflection orders, so that they carry great potential for a speed-up and quality improvement.

LISTENING TESTS

In listening tests that were carried out in co-operation with the Audio Communications Group at TU Berlin, some guidelines for the development of acoustic models could be discovered. The method of constant stimuli following a 3AFC-paradigm was applied, eight models with graduated LOD as shown in Figure 7 were presented. The listening tests unveiled the abilities of the human ear to detect changes of the surrounding geometry.

All structurally graduated models were pairwisely compared to the model with the highest detail level, and the subjects were asked to pick the different auralization from a list of three. In this test, 33% correct answers (67% wrong answers resp.) indicate no discriminability and values close to 100% indicate a high discriminability.

The listening test included several source and receiver positions as well as various dry input signals. Figure 5 shows the histogram of the empirical relative number of wrong answers as well as the matched normally distributed probability density function. The structural detail sizes (in *cm*) are shown logarithmized to be perceptually adequate, and they are mirrored at the ordinate to help fitting the Gaussian curve. Thus observations are restricted to the relevant half of the probability density and distribution, yielding a threshold of noticeable simplification at 75% in the cumulative distribution function that represents the collective psychometric function of all subjects (Gelfand 2004), see Figure 6.

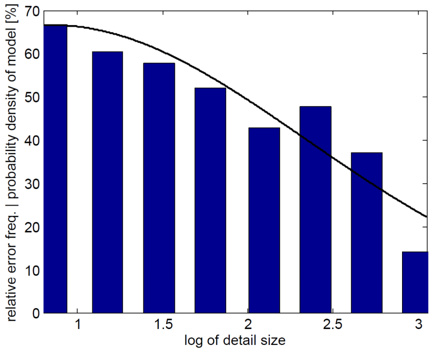


Figure 5. Histogram showing the relative number of wrong answers in a 3-AFC listening test and a matched Gaussian probability density.

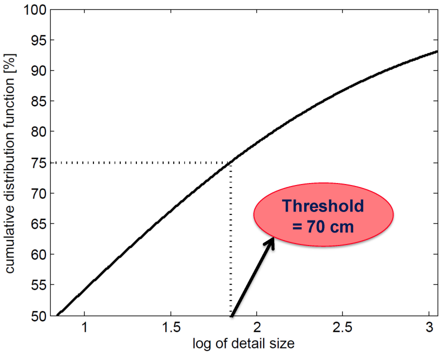


Figure 6. The cumulative distribution function for the density as shown in Figure 6, and the resulting threshold at 75%.

The threshold of noticeable simplification yields to *70cm*, defining a structure size which is not needed to go beyond when modeling fine details, because further elaboration would not be audible. This result is valid only for the presented room type at this stage.

LEVEL-OF-DETAIL MODELING

We propose to apply a LOD graduation for the room models which is known from computer graphics. The frequency of incident sound waves will decide which geometry the particles will 'see' when they hit a surface. As the RT simulation is run separately for different frequency bands, the geometry can easily be exchanged in between.

For the IS generation, a decision has to be made if the simulation's purpose is offline calculation or real-time rendering.

Offline. For offline calculations, there are no time constraints, so the ISs can be generated for every model of the LOD set. This will end up in very high memory consumption but ensures physically correct specular reflections for the whole frequency range.

Real-time. For online auralization, the LOD method is very beneficial. On a sliding scale between the different degrees of detail, the memory consumption and audibilty test latency can be adjusted to the available processing power of the hardware. Adjacent rooms can run simple models while e.g. at the listener's location full details are provided.

The RT algorithm that is used in *Raven* (and the IS audibility test as well) is accelerated by a BSP tree. The BSP data structure divides the scene consecutively in two parts by defining cutting planes. The result is a tree of depth $log_2(P)$ where P is the number of polygons (which are well suited to define cutting plane positions). The search for an intersection point, which is the basic task in RT, is then reduced from $O(P)$ to $O(log_2(P))$. As the root of a BSP tree is simply the first cutting plane, followed by the complete geometry information as branches, this root can be seamlessly exchanged in the intersection algorithm depending on the frequency of the sound ray that is to be traced.

The LOD model set generation starts with a highly detailed room model which is then structurally simplified step by step. This is done by excluding structural details that fall below a certain minimun size and replacing them with plain surfaces, with attention to the scattering coefficients that have to be matched to the material's acoustic properties appropriately. This should be independent from the fine structure and their simplification, as the highest detailed model should account for the correct scattering coefficients on all materials already. Nevertheless, when fine structures are neglected, appropriate scattering coefficients must be ensured.

After the simplification process, the smallest remaining surface size defines the minimum frequency that this model can be used for, see Figure 3.

Ideally, the total number of models should be matched to the desired frequency resolution of the simulation. For real-time auralizations, it is common to limit the frequency range to 10 octave bands. During the listening tests, it became apparent that differences of structural details below the size of 70 cm become more and more indistinguishable. This minimum size of 70cm qualifies the model to run simulations as from 500Hz. Thus, for octave resolution we propose a set of five models, as shown in Table 1.

Table 1. Parameters for the proposed set of LOD models, with minimum structural sizes and resulting valid frequencies.

Minimal structure size	>[m]	0.7	1.4	2.8	5.6	11.2
Minimal valid frequency	>[Hz]	500	250	125	63	31

Figure 8 illustrates a possible simplification that uses three different models.

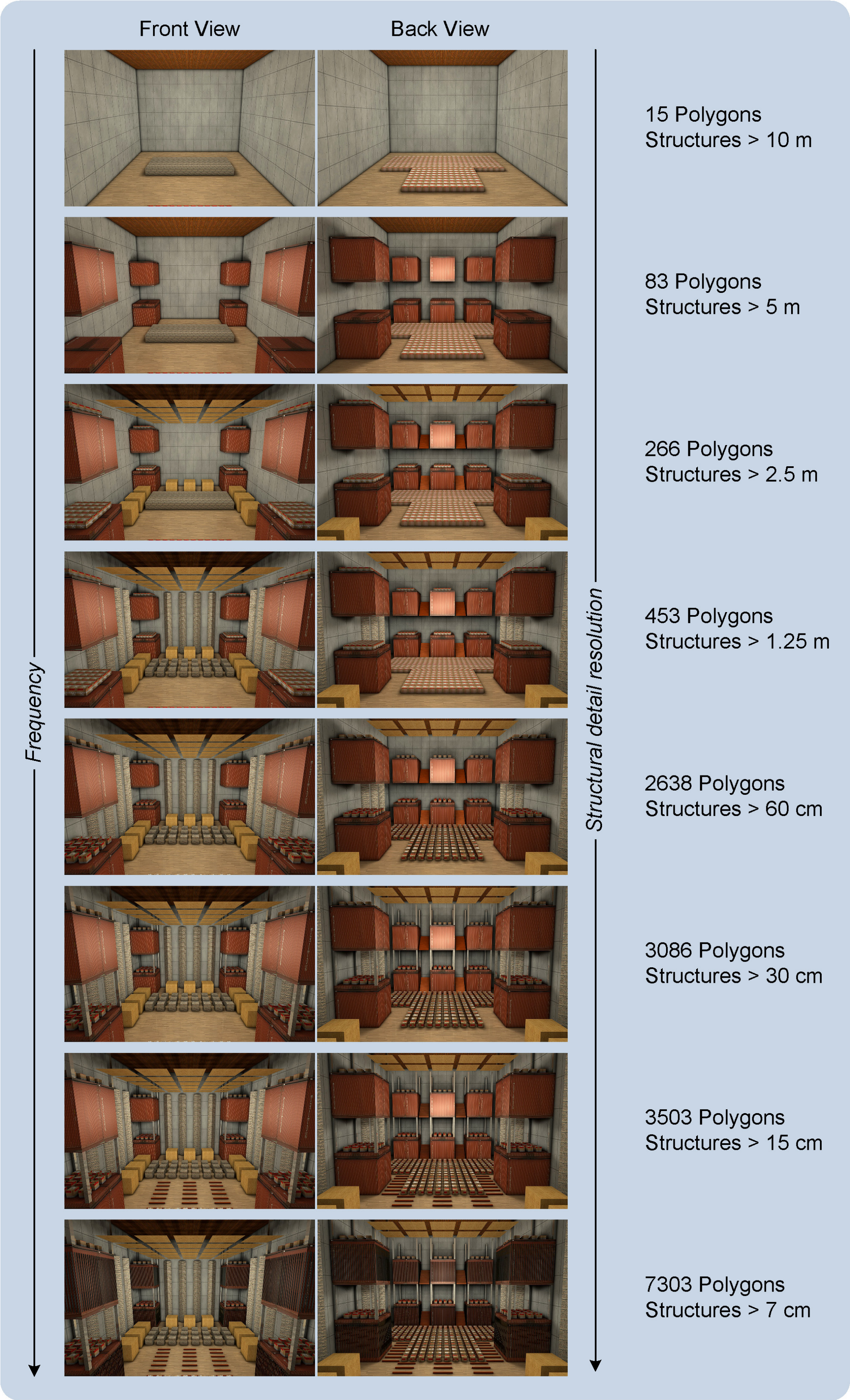


Figure 7. Example of an acoustic LOD model set with eight different levels of detail.

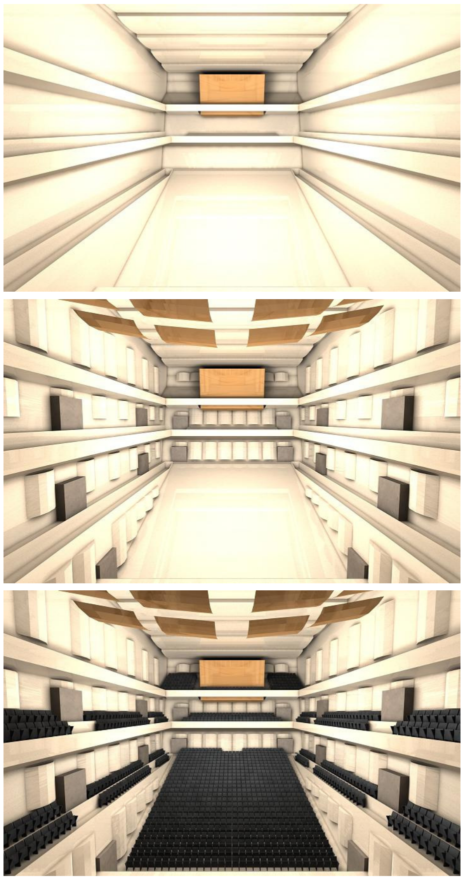


Figure 8. Example of an acoustic LOD model set (Konzerthaus Dortmund) with low detail (top) medium detail (middle) and high detail (bottom).

TIME-DEPENDENT GEOMETRY

Once a set of models is created, it can further be used to improve the speed of the ray tracing engine. Therefore, in addition to the frequency-dependent selection of the model, a time-dependent LOD control sequence will be added. Here, the model details decrease with increasing time in the IR. This is based on psychoacoustic masking effects of the late part of the reverberation as well as the fact that late reflections are majorly diffuse, so that structural fine details are not that relevant anymore.

Beginning from the model that was selected with respect to the current frequency, all simplified versions of this model with less detail will be loaded in parallel. The ray tracing engine can now access and jump between all different models without any loss of performance by using the BSP structure.

After a certain number of reflections, that is freely definable, the geometry can be completely exchanged while the particle figuratively flies through the air. On the next surface collision, the sound rays see a simplified version of the model. As a favourable effect, the intersection test that has to find this next collision point is accelerated due to the lower number of polygons in the simplified model. By continuing this scheme, the geometry is simplified step by step in the temporal course of the resulting IR.

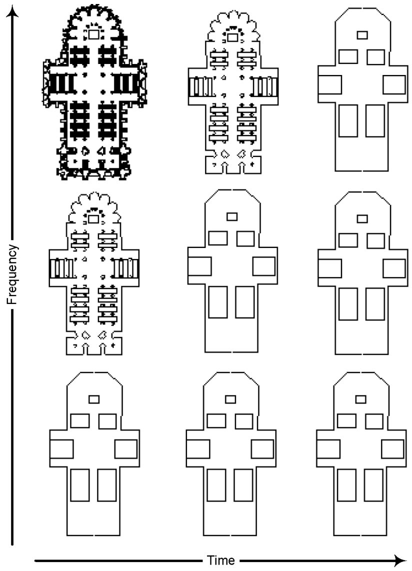


Figure 9. Frequency- and Time-dependent LOD model substitution strategy, example of Cologne Cathedral.

RESULTS

The performance of the whole simulation process is mainly dominated by the RT algorithm, which consumes much more computation power compared to the audibility test of the IS. The audibility test runtime is not affected by the LOD control, unless the physically more precise combination of IS of all LOD models is desired. Figure 10 shows how the RT runtime as the most time consuming part of the simulation is highly accelerated by the approach of frequency-matched geometry, yielding an overall speed-up of more than factor *2x* in a test environment that used 8 different LOD, see Figures 7 and 10.

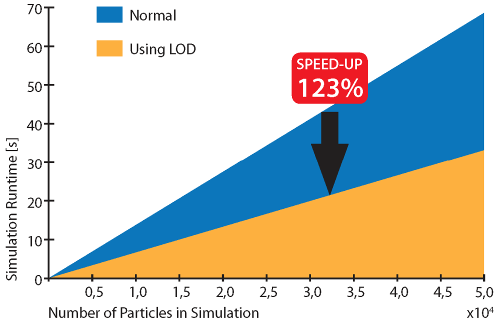


Figure 10. Simulation runtimes for the highest detailed model compared to the frequency-dependent LOD model set for various number of RT particles.

Figure 11 shows the further and even higher acceleration potential for the RT engine when using time-dependent geometry substitution. The speed-up depends on the number of reflections before switching to the next simpler model of the LOD chain. As proposed by (Kuttruff 1995), diffusely reflected sound becomes the dominant part of the IR already after three reflections. As the exact direction of sound rays after wall reflections is not of high importance in a diffuse sound field and especially as scattered wall reflections are less dependent on the exact surface orientation, it is feasible to switch to simplified geometries for the simulation of later parts in a room IR.

Assuming a diffusely developing sound field after three reflections, it is suggested to start the simplification at this stage after the third reflection. The current model is then continuously substituted by the next level of detail in direction of reduced detail after every third reflection. With this strategy, the overall performance in a broadband simulation that already uses frequency-matched models can still be accelerated almost by an additional factor *3x*, as shown in Figure 11. This figure also shows possible speed-ups for more conservative approaches with higher reflection order intervals (‘Triggerinterval’) between model substitutions.

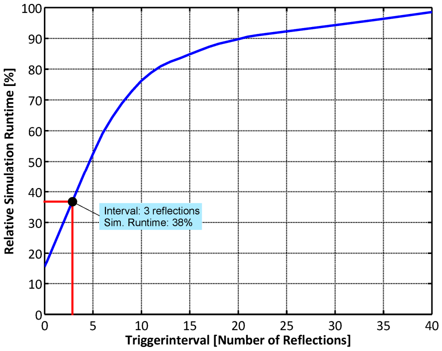


Figure 11. Reduction of simulation runtimes by simplifying geometry details for later parts of the impulse response after every x reflections (‘Triggerinterval’).

The combination of frequency- and time-dependent geometry in a simulation yields a speed-up factor of about *6 times*.

CONCLUSION AND OUTLOOK

The approach of graduated CAD models is convincing in consideration of the physics behind geometrically based simulation algorithms. Integrated into GA simulations, this higher precision comes together with about 2 times faster RT calculations. If this approach is combined with continuously decreasing LOD throughout the temporal progress in the IR, the simulation runs 6 times faster without audible negative effects on the IRs. This is very appealing for physically based auralization softwares that are under real-time constraints, such as *Raven*.

The availibility of a set of models with graduated complexity in such real-time auralizations allows adjusting the simulation accuracy and speed to the current needs in large scenarios and to the available computation power. In addition to the advantages for the RT, it also enables the IS engine to select a model from this set that is appropriate in terms of accuracy, memory and speed.

So far, the graduated models have to be constructed manually. With improvements of the automated simplification algorithms that can be expected in the future, the presented approach can be fully integrated in the simulation software or at least provide preliminary work for the dismissal of fine structures.

Another important topic for future research lies in the field of scattering patterns, which are strongly dependent on the structural fine details and their geometric orientation and thus have to be extensively investigated to improve the LOD approach.

ACKNOWLEDGEMENT

The author would like to thank Professor Michael Vorländer for his support and discussion, all participants of the listening tests and the german research foundation (DFG) for supporting this project.

REFERENCES

Allen & Berkley (1979), “Image method for efficiently simulating small-room acoustics” The Journal of the Acoustical Society of America, **64(4)**

Blauert (1997), “Spatial Hearing” MIT Press Cambridge MA, 2^{nd} edition

Funkhouser et al. (1998) “A beam tracing approach to acoustic modelling for interactive virtual environments” Computer Graphics, SIGGRAPH ’98, **21**

Gelfand (2004), *Hearing. An Introduction to psychological and physiological acoustics* (Dekker, 4th Ed., New York) p.251

Heinz (1993), “Binaural Room Simulation based on the Image Source Model with Addition of Statistical Methods to include the Diffuse Sound Scattering of Walls and to Predict the Reverberant Tail” Applied Acoustics **38**

Kuttruff (1995) “A Simple Iteration Scheme for the Computation of Decay Constants in Enclosures with Diffusely Reflection Boundaries” The Journal of the Acoustical Society of America **98**

Pompetzki (1993) “Psychoacoustic verifcation of computer models for binaural room simulation” Doct diss, Ruhr-Univ. Bochum

Schlütter (2009) “Parallelisierung von Algorithmen für die Echtzeit-Raumakustiksimulation“ Master thesis, RWTH Aachen University

Schroeder (1962) “Digital computers in room acoustics” Proc. 4th ICA Copenhagen

Schröder & Lentz (2006) “Real-time processing of image sources using binary space partitioning” Journal of the Audio Engineering Society (JAES), **54** Nr. (7/8)

Schröder (2007) “A Fast Reverberation Estimator for Virtual Environments” AES 30th International Conference, Saariselkä, Finland

Schröder (2007) “Hybrid Method for Room Acoustic Simulation in Real-time” 19th International Congress on Acoustics (ICA), Madrid, Spain

Schröder & Pohl (2009) “Real-Time Hybrid Simulation Method Including Edge Diffraction” EAA Symposium on Auralization, Espoo, Finland

Schröder & Ryba (2010) “Real-time auralization of dynamically changing virtual environments” Submitted to 20th International Congress on Acoustics (ICA)

Schröder et al. (2010) “Virtual Reality System at RWTH Aachen University” Submitted to International Symposium on Room Acoustics (ISRA)

Siltanen et al. (2006) “Geometry Reduction in Room Acoustics Modelin” Proc. of the Institute of Acoustics, **28** Pt. 2

Stephenson (2004) “Beugungssimulation ohne Rechenzeit-Explosion: Die Methode der quantisierten Pyramidenstrahlen- ein neues Berechnungsverfahren für Raumakustik und Lärmimmissionsprognose“ Doct. Diss., RWTH Aachen University

Svensson (1999) “An Analytic Secondary Source Model of Edge Diffraction Impulse Responses” Journal of the Acoustical Society of America, **106(5)**

Thaden (2005) “Auralisation in Building Acoustics” Doct. Diss., RWTH Aachen University

Vorländer (1988) “Die Genauigkeit von Berechnungen mit dem raumakustischen Schallteilchenmodell und ihre Abhängigkeit von der Rechenzeit“ ACUSTICA **66**

Vorländer (1989) “Simulation of the transient and steady state sound propagation in rooms using a new combined sound particle – image source algorithm” J. Acoust. Soc. Am. **86, 172**

Vorländer, M (2007) *Auralization* (Springer, 1^{st} Edition)

Wefers (2007) “Real-time Auralization of Coupled Rooms” EAA Symposium on Auralization, Espoo, Finland