MODELING EARLY REFLECTIONS OF ROOM IMPULSE RESPONSES USING A RADIANCE TRANSFER METHOD

Hequn Bai[1*], *Gaël Richard*[1], *Laurent Daudet*[2]

[1] Institut Mines-Télécom, Télécom ParisTech, CNRS-LTCI, 37-39 rue Dareau, 75013 Paris, France
[2] Institut Langevin, ESPCI ParisTech, Univ. Paris Diderot and CNRS UMR7587,
1 rue Jussieu, 75005 Paris, France

ABSTRACT

In this paper we propose an extension for the Acoustic Radiance Transfer (ART) method for the modeling of room acoustics. The original ART method is very efficient for modeling diffuse reflections and the late reverberation but does not well represent the early echoes. We then propose, in this paper, an extension of the ART method which allows to model the early part while keeping the advantages of the original method for the late reverberation simulation. The experimental results confirm that the proposed method gives more accurate reconstruction of the early reflections than the traditional ART method in average and that comparable accuracy can be obtained at lower complexity and memory requirements than the traditional ART method.

Index Terms— Room impulse response, radiance transfer method

1. INTRODUCTION

Room acoustics reconstruction considers the problem of modeling the sounds received at the listener's position in a sound scene. Multiple sound sources can be considered as well as their reflections and diffractions from surface materials and objects in order to simulate a realistic sound. The sound propagation effect is commonly divided into two parts: Early Reflections (ER) and Late Reverberation (LR). ER helps in localization, while LR gives an impression of the size of the environment and level of furnishing and absorptivity [1].

Numerous methods have been proposed for modeling the Room Impulse Response (RIR). Amongst geometric methods, the most widely used are Image Source Method (ISM) [5] and Ray tracing method [6]. However, since the computation time of both methods becomes very large if high reflection orders are needed, geometric methods are only suitable to model low-order specular reflections.

The late reverberation tail of RIR usually presents a distinct noise-like character due to the isotropic addition of a multitude of late acoustics reflections [9]. This motivated the deployment of statistical method for modeling the LR [2, 7]. Nevertheless since the statistical methods do not take into account the detailed room geometry and the precise positions of sound sources and listeners, the prediction accuracy is thus not guaranteed.

In [4], an Acoustic Radiance Transfer method was presented which can handle both diffuse and nondiffuse reflections. Since the high-order reflections computation is precomputed and decoupled from the run-time computation procedure, this method is well suited for dynamic applications with moving sources. A case study has shown that the simulated RIR using this method can predict several acoustic parameters with acceptable accuracy, and thus gives a good estimation of the RIR decay curve, especially for the late reverberation. However, due to the discretization of the room geometry, the accuracy of the early reflections is not guaranteed. Hybrid methods were proposed in [3, 10] where ART method is used to model diffuse reflections and late reverberation while the geometric methods are used to model the more important early reflections. Although it is common to split ER and LR, it would be desirable to have a single generic method that would be accurate for both parts. To that aim, we propose in this paper an extension of the ART method which allows us to model the early part while keeping the advantages of the original method for late reverberation simulation.

The paper is organised as follows. In section 2 we recall the basics of the ART method and the main computation procedure. In section 3 we highlight the inaccuracy of the ART method for early echoes and introduce a new proposal. The experimental results are given in section 4 and some conclusions are suggested in section 5.

2. RADIANCE TRANSFER METHOD

The ART method models the sound energy received at the listener's location by the contribution of the outgoing sound energy flux on the surface in each direction. Let $L(x,\Omega,t)$ be the time-dependent outgoing sound energy flux at point x along direction Ω. Then the Acoustic Rendering Equation can be described as

$$L(x,\Omega,t) = L_0(x,\Omega,t) + \int_s R(x,x',\Omega,t)L(x',\frac{x-x'}{|x-x'|},t)\mathrm{d}x',$$

where L_0 is the initial emitted sound energy and L is the total outgoing energy. $R(x,x',\Omega,t)$ is the reflection kernel which describes how the outgoing energy from point x' influences the outgoing energy at point x in direction Ω [4].

In order to solve the Acoustic Rendering Equation by numerical simulation, the room surface s is discretized into N patches. The Bidirectional Reflectance Distribution Function (BRDF), which defines how the incoming sound flux is reflected at a surface, is discretized by dividing the hemisphere into solid angles. The hemisphere is devided into m azimuth angles and n elevation angles.

The outgoing energy resulting from the n^{th} reflection can be written in discretized form as

$$\hat{L}_n^{\bar{i}} = \sum_{\bar{j}=1}^{\infty} F_{\bar{j},\bar{i}}\hat{L}_{n-1}^{\bar{j}}. \quad (1)$$

with $\bar{i} = \{k,p,s\}$ denoting the radiance transfer element in patch k, azimuth p and elevation s, and $\bar{j} = \{l,q,t\}$ denoting the radiance transfer element in patch l, azimuth q and elevation t. $F_{\bar{j},\bar{i}}$ is

*This work was partly supported by the European Commission under contract FP7-ICT-287723 - REVERIE.

978-1-4799-0972-8/13/$31.00 ©2013IEEE

an energy transfer response which models the reflected energy flux from element $\bar{i}$ by the contribution of the reflected or emitted energy flux from element $\bar{j}$. $\hat{L}_{n-1}^{\bar{j}}$ is the outgoing energy for element $\bar{j}$ in the $n-1^{th}$ reflection, and $\hat{L}_n^{\bar{i}}$ is the outgoing energy for element $\bar{i}$ in the n^{th} reflection. If $\hat{L}_n^{\bar{i}}$ is are put into vector form ϕ_n and $F_{\bar{j},\bar{i}}$s into matrix F,

$$\phi_n = \begin{pmatrix} \hat{L}_n^1 \\ \vdots \\ \hat{L}_n^N \end{pmatrix}, F = \begin{pmatrix} F_{1,1} & \cdots & F_{1,N} \\ \vdots & \ddots & \vdots \\ F_{N,1} & \cdots & F_{N,N} \end{pmatrix}, F_{\bar{i},\bar{i}} = 0, \quad (2)$$

then the total outgoing energy from the surfaces can be written in discretized form as

$$\phi = \sum_{n=1}^{\infty} \phi_n = \phi_0 + \sum_{n=1}^{\infty} F^n \phi_0 = (I + \sum_{n=1}^{\infty} F^n)\phi_0. \quad (3)$$

where F^n is the n^{th} convolution of the transfer matrix F.

The numerical simulation can be divided into an off-line computation and a run-time computation. During the off-line computation, the transfer matrix F is computed. We use energy-based ray tracing to compute the element to element responses $F_{\bar{i},\bar{j}}$. Rays are emitted uniformly within the solid angle of element i over the patch. When each ray encounters a geometric primitive, it is specularly reflected, according to the pre-defined BRDF. The reflected energy flux in each outgoing direction is calculated to form the one-order element to element response $F_{\bar{i},\bar{j}}$. The transfer matrix F is convolved n times to get high order responses.

Run-time computation can be divided into three stages: initial shooting, energy propagation and final gathering. At the initial shooting stage, rays are uniformly emitted from the sound source. These rays are reflected and collected in each reflected angle to form the initial condition ϕ_0. Energy propagation consists of convolving the initial energy condition ϕ_0 with the high order element-to-element transfer matrix $\sum_{n=1}^{\infty} F^n$. During the final gathering stage, the energy responses from all visible radiance transfer elements are accumulated at the listener's location [4]. As for the initial shooting stage, we use ray tracing to calculate the amount of energy received by the listener from each element.

3. FINAL GATHERING SCHEMES

In [3, 8], rays are uniformly sent from the listener and collected at each patch to calculate the energy response at the final gathering stage, using a reciprocity principle. Since these works only use ART method to model diffuse reflections, the effect of the directional discretization in the final gathering stage were not discussed. For simulation of early echoes, it is important to accurately model the reflection angle and the energy amplitude for each solid angle. Using such final gathering method tends to result poor accuracy in early echoes. Discretizing the geometry and directions introduces errors in two ways. Firstly, the outgoing energy flux in a certain direction of one patch is considered constant. In fact, it can have considerable variation, especially for the patches close to the source. Secondly, because the sound sources and listeners are free to move in the acoustic space, there exist positions for listeners where the final gathering energy responses are largely under- or over- estimated.

More natural way is to send ray from the center of each patch and to accumulate them at the listener's position, as what they did for modeling diffuse radiosity. The rays are collected by a receiving

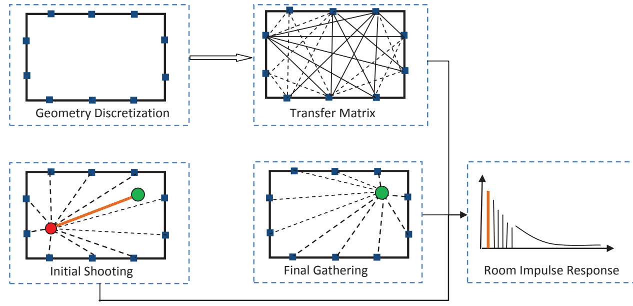

Figure 1: One-pass numerical simulation. The red spot denotes the sound source and the green spot denotes the listener. The orange line indicates the direct sound.

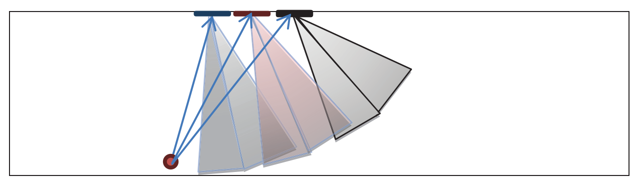

Figure 2: Original scheme: final gathering scheme by emitting rays from the centers of patches.

volume at the listener's position. The receiving volume is adjustable as in the ray tracing method. The received energy responses are convolved with the initial shooting matrix and the transfer matrix to get the early echoes as well as the late reverberation. The direct sound is calculated at the initial shooting stage. Together with the direct sound, the whole RIR is generated in a single computation pass, instead of splitting the RIR into ER and LR and modeling them separately. The computation procedure is illustrated in Fig. 1. However, this method still leads to large areas where the listener receives no sound, and some other areas where he receives overlapped sounds from adjacent patches, as one can see on Fig. 2.

To make the reflected sound more uniformly distributed and to better cover the acoustic space, we propose to emit rays uniformly over each patch along the solid angle, instead of emitting all rays from the center of the patch. For each solid angle of a patch, we randomly choose 20 points on that patch. We send one ray from each point along a random direction within the solid angle. These rays are accumulated at the receiving volume to generate the final gathering responses. The uniform final gather scheme is illustrated in Fig. 3.

In the next section, we compare the performance for early reflections simulation of our approach (called herein ART-proposed) to the ART-traditional method.

4. EXPERIMENTAL RESULTS

4.1. Experiment

The test environment is a rectangular room with dimensions $4m \times 3m \times 2m$. The room surface is divided into 104 triangles. For each patch, the hemisphere is divided into 144 directions, 24 in azimuth and 6 in elevation, with 15^o resolution in each dimension. The sound source is fixed at $(1, 2, 1)$, while 40 listeners' positions are uniformly selected on the plane at height $z = 1$. Energy responses

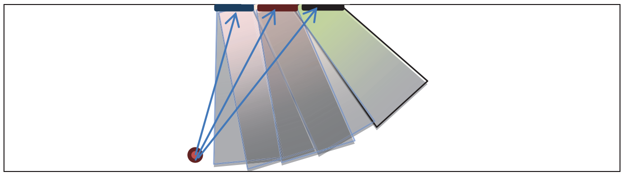

Figure 3: Proposed scheme: final gathering scheme by emitting rays uniformly over the patch.

are generated using each method and are evaluated by comparison with reference signals obtained by ray tracing. The evaluation of the energy responses at these 40 locations is presented on Fig. 6 as a 2D map. The energy responses at two locations, one picked at the central of the room and the other near the surface, are given on Fig. 4 and Fig. 5 to show the reconstruction details.

In another set of experiments we study the effect of discretization of the room surface and directions. The number of patches varies from 32 to 416, and the solid angle resolution varies from 15^o to 30^o. Their combined patterns are shown in Table 1. For each discretization pattern, 40 energy responses are simulated at the locations designed above, and their reconstruction accuracy as well as the average computation time are shown in Table 1.

4.2. Evaluation Criterion

The ray tracing method is used as the reference method to evaluate the accuracy of both methods. Since the performance of the ART method in the reconstruction of diffuse reflections and the late reverberation has been studied in [3, 4], we mainly focus on the accuracy in early echoes simulation. We thus consider the simulated energy responses up to the fourth order reflections. The comparison with more accurate and computational efficient methods, such as the beam tracing method, will be left to the future work.

As a first evaluation criterion, we use normalized cross-correlation to assess the distance between the simulated energy impulse responses simulated with the different ART methods, and the baseline ray tracing method. The formula is given in (4) as

$$\Delta = \frac{f(t) \star g(t)}{||f(t)||_2||g(t)||_2} \tag{4}$$

where $f(t)$ and $g(t)$ are the considered signals, $\star$ denotes correlation, and $||\cdot||_2$ denotes the l_2 norm of $\cdot$.

Since the individual impulses in the reconstructed energy response using ART methods can have slight time shift due to surface discretization, the normalized cross-correlation is not sufficiently informative to assess the results. Thus, the difference between the accumulated energy is also used to evaluate the performance.

Let $G(t) = \int_0^t g(t)\,\mathrm{d}t$ and $F(t) = \int_0^t f(t)\,\mathrm{d}t$, then the error between the accumulated energy is

$$\varepsilon = \int_0^{T_M} |G(t) - F(t)|\,\mathrm{d}t \tag{5}$$

where $[0, T_M]$ is the time interval considered which in our case roughly corresponds to include the first four orders reflections.

4.3. Results

Fig. 4 and Fig. 5 display the reconstructed early echoes at listener's location $(2, 1.5, 1)$ and $(1, 0.5, 1)$ (case A and B in Fig. 6). In

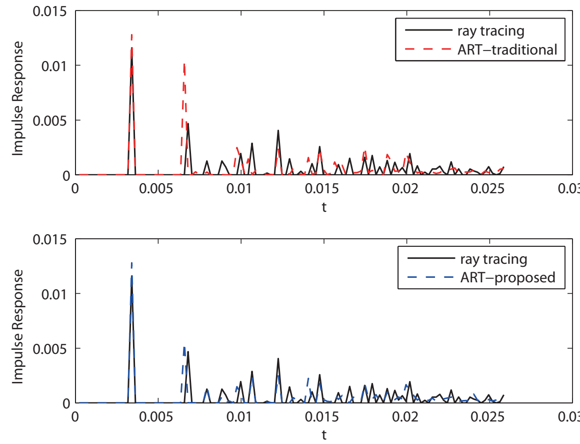

Figure 4: Case A: energy response obtained by the ART methods and the ray tracing method. The sound source is at $(1, 2, 1)$ and the listener is at $(2, 1.5, 1)$.

Fig. 4, the proposed method shows a better reconstruction. Note that the first echo is largely over-estimated by the ART-traditional method, but more accurately modelled by the proposed method. In Fig. 5 the first echo is totally missed by the ART-traditional method, but is well preserved using the proposed method. The improvement is due to the averaging effect of the uniform distribution of the emission points over each patch. Note that, as shown on Fig. 5, there still exist a few impulses that are not well modeled by any of the two methods. In this case, the listener is very close to the surface, which leads to strong surface quantization effects.

Fig. 6 shows how the error of accumulated energy ε varies at different locations in the room. The darker color in the error map indicates smaller accumulated energy error, and thus has better reconstruction accuracy. The error map shows that the reconstructed responses using ART-proposed are in general more accurate than the traditional method (noted ART-traditional) at almost every position. The errors at some locations such as at $(1, 0.5, 1)$ and $(2.5, 2, 1)$ are effectively reduced by the proposed method.

Combined with high order reflections, the proposed method is able to effectively model the whole RIR. In order to check that the proposed method did not affect the late reverberation, we computed the reverberation time (T_{30}) of both methods, and found no significant difference. The comparison using cross-correlation as criteria gives similar conclusion, and hence is not presented here.

Table 1 compares the performance of the two methods under different discretization patterns. In this table, N denotes the number of patches, m the number of discretization in Azimuth, and n the number of discretization in elevation. We use $\bar{\varepsilon}$ for the average error of accumulated energy, $max.\varepsilon$ for the maximum error of accumulated energy, $\bar{\Delta}$ for the average correlation, and $min.\Delta$ for the minimum correlation. t is the computation time in minutes.

It is interesting to observe from Table 1 that dividing the hemisphere finer does not necessarily bring improvement to the accuracy of the early echoes. On the contrary, the performance is slightly degraded in some tests, although not in a systematic way. However, this observation indicates that when more computation power and memory is available, first trying to divide the surface into finer patches can be a better option.

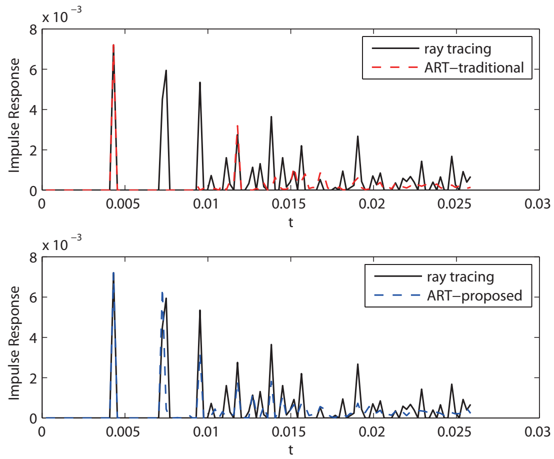

Figure 5: Case B: energy response obtained by the ART methods and the ray tracing method. The sound source is at $(1, 2, 1)$ and the listener is at $(1, 0.5, 1)$.

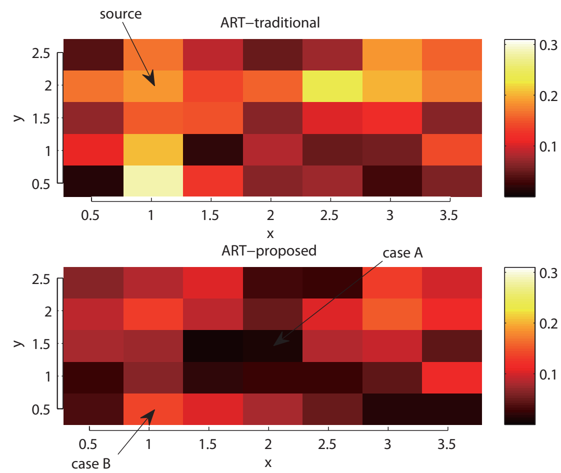

Figure 6: Error map of accumulated energy (lower value indicates better result). The sound source is at $(1, 2, 1)$ and the listeners are located uniformly on the plane at height $z = 1$.

Table 1: Performance comparison for the traditional (Trad.) and proposed (Prop.) methods under different discretization patterns.

N	m	n	Algo.	$\bar{\varepsilon}$	$max.\varepsilon$	Δ	$min.\Delta$	t
32	12	3	Trad.	0.11	0.37	0.67	0.38	0.56
			Prop.	0.09	0.18	0.69	0.43	0.52
32	24	6	Trad.	0.13	0.67	0.58	0.25	0.77
			Prop.	0.08	0.15	0.65	0.46	0.91
104	12	3	Trad.	0.11	0.26	0.73	0.59	1.99
			Prop.	0.07	0.15	0.79	0.64	1.95
104	24	6	Trad.	0.12	0.28	0.71	0.55	2.53
			Prop.	0.07	0.18	0.79	0.65	2.62
416	24	6	Trad.	0.08	0.22	0.78	0.68	15.1
			Prop.	0.05	0.12	0.81	0.71	15.6

Comparing the two methods under the same discretization pattern, it is clear that the proposed ART technique outperforms the traditional ART method, at a very mild cost in terms of computational requirements. As seen in the average values in the last two discretization patterns, the proposed method can sometimes achieve a performance similar to the traditional method used with finer discretization. In other words, the proposed method is a competitive alternative to using smaller patches, when high precision is needed with limited computational resources.

5. CONCLUSION

In this paper we investigated the use of the radiance transfer method in modeling early reflections of room impulse responses. We proposed a novel energy final gathering method, which outperformed the traditional scheme. Using the new scheme, a single generic method can efficiently model both the early and the late reverberation of RIRs. Although radiance transfer method are not guaranteed to reconstruct the RIRs as accurately as ray tracing method, the fact that they decouple the source and listener positions from the bulk of computation makes them an appealing choice for real-time acoustic rendering.

6. REFERENCES

[1] H. Kuttruff, *Room Acoustics*. Elsevier Science Publishing Ltd., 1991.

[2] J. Garas, *Adaptive 3D sound systems*. Kluwer Academic Publishers Norwell, 2000.

[3] L. Antani, A. Chandak, M. Taylor, and D. Manocha "Direct-to-indirect acoustic radiance transfer," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 2, pp. 261–269, 2012.

[4] S. Siltanen, T. Lokki, and L. Savioja, "The room acoustic rendering equation," *The Journal of the Acoustical Society of America*, vol. 122, no. 3, pp. 1624–1635, 2007.

[5] J. Allen, and D. Berkley, "Image method for efficiently simulating small-room acoustics," *The Journal of the Acoustical Society of America*, vol. 65, no. 4, pp. 943–950, 1979.

[6] A. Krokstad, S. Strom, S. Sorsdal, "Calculating the acoustical room response by the use of a ray tracing technique," *Journal of Sound and Vibrations*, vol. 8, no. 1, pp. 118–125, 1979.

[7] J. Moorer, "About this reverberation business," *Computer music journal*, vol. 3, no. 2, pp. 13–28, 1979.

[8] L. Antani, A. Chandak, L. Savioja, and D. Manocha, "Interactive sound propagation using compact acoustic transfer operators," *ACM Transactions on Graphics*, vol. 31, no. 1, 2012.

[9] E. A. Lehmann and A. M. Johansson, "Diffuse reverberation model for efficient image source simulation of room impulse responses," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 6, 2010.

[10] T. Lewers, "A combined beam tracing and radiant exchange computer model of room acoustics," *Applied Acoustics*, vol. 38, pp. 161–178, 1993.