

Echtzeitfähiger Nachhallschätzer für virtuelle Umgebungen

Dirk Schröder, Philipp Dross, Michael Vorländer

Institut für Technische Akustik, 52066 Aachen, Deutschland, Email: dsc@akustik.rwth-aachen.de

Einleitung

Eine raumakustische Simulation in Echtzeit stellt eine Erweiterung multimodaler virtueller Umgebungen dar. Gerade bei der Darstellung von virtuellen Räumen, wie z.B. einem Theater, einer Kirche oder einem Konzertsaal (Abb. 1), unterstützt der auditive Stimulus den räumlichen Eindruck des visuell Dargebotenen. Im Gegensatz zur Qualität der visuellen Komponente fehlt es der raumakustischen Simulation allerdings immer noch an einer realistischen und zumindest perceptiv korrekten Darstellung, sofern die Bedingungen der Echtzeit erfüllt werden sollen.



Abbildung 1: Beispiel für ein virtuelles Modell eines Konzertsaals.

Eine der markantesten akustischen Größen eines Raumes ist sein Nachhall, da er ihm seine individuelle Klangcharakteristik verleiht. Theoretisch sollte der späte Nachhall einen sehr hohen Diffusitätsgrad aufweisen. In Echtzeit-Systemen wird daher, aufgrund der begrenzten Rechenzeit, der späte Nachhall kostengünstig durch einen exponentiell abklingenden Nachhallschwanz simuliert, wobei beim Bau der binauralen Filter von einer directionalen Gleichverteilung der empfangenen Schallenergie ausgegangen wird. Die Praxis zeigt allerdings, dass die Annahme eines völlig diffusen Schallfelds in einer Vielzahl von Raumtypen, z.B. Flachräume, Korridore, L-Form, etc., nicht gültig ist. Möglich auftretende Flutter-Echos oder verzerrte Nachhallkurven würden daher unerkannt bleiben[1].

Vorgestellt wird die Implementierung eines echtzeitfähigen Nachhallschätzers, der in ein bestehendes Echtzeit 3D-Audio Rendering System[2] für die Auralisation virtueller Umgebungen integriert ist. Zeitliche und spektrale Informationen über den empfangenen Schall werden dabei zur Laufzeit mit Hilfe eines schnellen Ray Tracers gesammelt und in den momentanen Filterbau einbezogen, wodurch eine realistischere Auralisation der virtuellen Umgebung ermöglicht wird.

Ray Tracing

Für die Berechnung des diffusen Anteils der Raumimpulsantwort wird ein stochastisches Ray Tracing-Verfahren verwendet. Der Ray Tracer simuliert die Ausbreitung eines Energie-Impulses mit Hilfe einer finiten Anzahl von Energiepartikeln, die zum Zeitpunkt t_0 von der Position der zu simulierenden Schallquelle aus zufallsverteilt in den Raum abgeschossen und solange verfolgt werden bis ein vordefiniertes Abbruchkriterien erfüllt wird (z.B. Laufzeitüberschreitung, zu geringe Restenergie des Partikels). Die Energiepartikel breiten sich dabei mit Schallgeschwindigkeit geradlinig im Raum aus, wobei die Partikel aufgrund der Luftabsorption stetig Energie verlieren. Kollidiert ein Partikel mit einem Raumobjekt (z.B. Wände, im Raum befindliche Gegenstände) wird abhängig vom Absorptionsgrad des getroffenen Objektes weitere Energie absorbiert. Anschliessend wird das Energiepartikel entweder an dem getroffenen Objekt spiegelnd reflektiert oder diffus in den Raum gestreut. Als Entscheidungskriterium für die Reflexionsart wird der materialabhängige Streugrad s verwendet, der das Verhältnis zwischen nicht-gestreuter Energie und reflektierter Gesamtenergie für das getroffene Material darstellt[3].

Aufgrund der frequenzabhängigen Absorptions- und Streukoeffizienten wird der zu simulierende Frequenzbereich in einzelne Frequenzbänder unterteilt (Terz- oder Oktavauflösung), die unabhängig voneinander mit der gleichen Partikelanzahl simuliert werden. Die Energie-Impulsantwort an der gewünschten Messposition wird mittels einer Detektorkugel bestimmt, welche in n Azimuth-Gruppen und m Elevations-Gruppen, also $n \cdot m$ Richtungsgruppen, unterteilt ist. Während der Simulation wird für jede dieser Richtungsgruppen die Energie und Laufzeit durchfliegender Energiepartikel aufgezeichnet, wodurch für jede Gruppe ein drei-dimensionales Energie-Histogramm erstellt wird, mit k Zeitintervallen der Länge Δt auf der Ordinate, den einzelnen Frequenzbänder f_b auf der Abszisse und die zeit- und frequenzabhängige Energie $E_{m,n}(t_k, f_b)$ auf der Applikate (Abb. 2(a)). Die Ergebnisse können auf zweierlei Weise interpretiert werden, zum einen als die zeitliche Energieverteilung $E_{m,n,f_b}(t_k)$ eines Frequenzbands und zum anderen als die spektrale Energiedichte $E_{m,n,t_k}(f_b)$ eines Zeitintervalls Δt .

Filterbau

Die zeitliche Auflösung der Energie-Histogramme ist üblicherweise wesentlich geringer als die Abtastrate, weshalb für die binaurale Raumimpulsantwort eine zeitliche Struktur konstruiert werden muss. Da die einzelnen Ener-

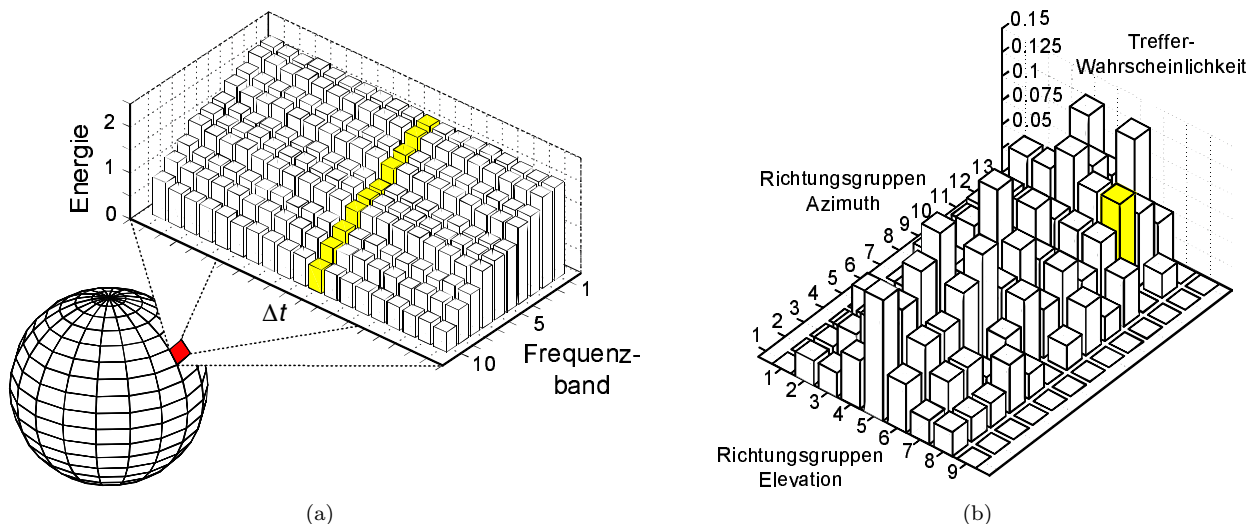


Abbildung 2: a) Für jede Richtungsgruppe der Detektorkugel wird während der Ray Tracing Simulation ein drei dimensionales Energie-Histogramm erstellt. b) Trefferwahrscheinlichkeiten der Richtungsgruppen für ein Zeitintervall Δt .

giebeiträge des diffusen Schallfelds aufgrund ihrer zeitlichen Überlappung im späten Teil der Impulsantwort nicht mehr als getrennte Rückwürfe wahrgenommen werden können, ist es möglich diese zeitliche Struktur mit Hilfe einer Poisson-verteilten Dirac-Impulsfolge nachzubilden. Die Wahrscheinlichkeit $w_n(t)$ mit der genau n Ereignisse innerhalb eines Zeitintervalls Δt auftreten ist bei einer Poisson-Verteilung beschrieben durch

$$w_n(t) = \frac{(\mu t)^n}{n!} e^{-\mu t} \quad n = 0, 1, 2, \dots, \mu > 0, t \geq 0 \quad (1)$$

wobei μ die mittlere Ereignishäufigkeit darstellt.

Die einzelnen Dirac-Impulse werden den entsprechenden Zeitintervallen zugeordnet und anschließend mit den zu diesem Zeitintervall wahrscheinlichsten kopfbezogenen Impulsantworten gefaltet. Dies wird anhand der Trefferwahrscheinlichkeiten der einzelnen Richtungsgruppen zum Zeitpunkt Δt entschieden, die für jedes Zeitintervall aus der entsprechenden über alle Frequenzbändern gemittelten Energieverteilung auf der Detektorkugel bestimmt wird (Abb. 2(a), Abb. 2(b)).

Um aus den beiden so generierte Rauschprozessen $r_i(t)$ die binaurale Impulsantwort zu erhalten, muss $r_i(t)$ noch ein zeitabhängiges spektrales Verhalten aufgeprägt werden. Hierfür kann die Proportionalität zwischen dem Energiedichtespektrum $E(f)$ und dem Quadrat des Betragsspektrums $|S(f)|$ des Amplitudendichtespektrums $S(f)$ eines Zeitsignals $s(t)$ ausgenutzt werden:

$$E(f) \sim |S(f)|^2; \quad s(t) \circ \bullet S(f) = |S(f)| e^{j\phi(f)} \quad (2)$$

Da die statistischen Energiemittelwerte pro Zeitintervall jeweils nur in Frequenzbändern von Terz- oder Oktavbreite vorliegen, ergeben sich die einzelnen Betragsspektren durch Radizieren der über alle Richtungsgruppen aufsummierten Kurzzeit-Energiedichtespektren $E_{t_k}(f_b)$,

$$E_{t_k}(f_b) = \sum_n \sum_m E_{m,n,f_b}(t_k) \quad (3)$$

die anschließend noch abhängig von der Abtastrate und Zeitintervallbreite interpoliert werden.

Für eine effiziente Gewichtung der Rauschprozesse $r_i(t)$ werden diese in k Segmente Hamming-gefenstert (zur Vermeidung von Diskontinuitäten) in den Frequenzbereich transformiert und dort mit den entsprechenden interpolierten Betragsspektren multipliziert. Die binaurale Impulsantwort erhält man nun durch Rücktransformation der einzelnen Segmente in den Zeitbereich und Einsortierung in das Gesamtfilter.

Validierung & Performanz

Eine Präsentation von Validierungs- und Performanztests würde den Rahmen dieses Manuskriptes weit überschreiten. Daher sei an dieser Stelle auf [2, 4] verwiesen.

Literatur

- [1] Rindel, J.H., *Evaluation of room acoustic qualities and defects by use of auralization*, 148th Meeting of the Acoustical Society of America, San Diego, CA, USA, 2004.
- [2] Lentz, T., Schröder, D., Vorländer, M., Assenmacher, I., *Virtual reality system with integrated sound field simulation and reproduction*, EURASIP Journal on Applied Signal Processing, Special Issue on Spatial Sound and Virtual Acoustics, 2007.
- [3] Cox, T. J., Dalenbäck, B.-I. Dalenbäck, Antonio, P. D., Embrechts, J. J., Jeon, J. Y., Mommertz, E., Vorländer, M., *A tutorial on scattering and diffusion coefficients for room acoustics*, Acta Acustica united with ACUSTICA, 92:1-15, 2006.
- [4] Schröder, D., Dross, P., Vorländer, M. *A fast reverberation estimator for virtual environments*, Proceedings of the 30th AES International Conference, Saariselkä, Finland, 2007.