

Theoretischer und perceptiver Einfluss einer Raum-in-Raum-Wiedergabe

Andreas Haeussler, Steven van de Par

Carl von Ossietzky University Oldenburg, Acoustics Group Cluster of Excellence Hearing4all

Email: andreas.haeussler@uni-oldenburg.de

Einleitung

Heutzutage ist es üblich, sich Musik- oder Sprachaufnahmen, welche in einem Raum (Aufnahmeraum) aufgenommen sind, in einem anderen Raum (Wiedergaberaum) wiederzugeben. Einige Beispiele hierfür sind Konzertaufnahmen, Studioaufnahmen, allerdings auch Audio-/Videokonferenzen oder die Anwendung von künstlichem Hall in einer Indoor-Konzertsituation. Deshalb stellt sich die Frage, inwiefern sich diese „Raum-in-Raum“ (R*R) Wiedergabeform [1] von einem einzelnen Raum unterscheidet. In der vorliegenden Studie wird eine Bass-Drum in einem Raum aufgenommen und in einem anderen Raum wiedergegeben (siehe Abbildung 1). Dabei vermischen sich die räumliche Information der beiden Räume. Für eine plausible Wiedergabe ist die präzise Wiedergabe eines Transienten Instruments sehr wichtig. Deshalb soll der Einfluss auf die Knackigkeit einer Bassdrum in einem einzelnen Raum mit der R*R-Wiedergabe verglichen werden.

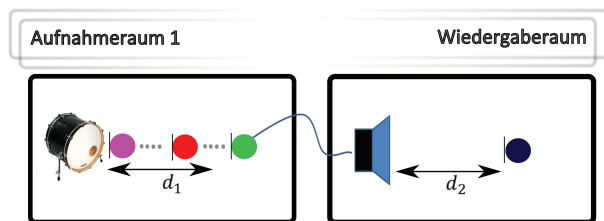


Abbildung 1: Schematische Darstellung des Knackigkeitsexperiments

Vorhergegangene Studien

Technisch lässt sich die Wiedergabekette von Aufnahmeraum und Wiedergaberaum als Faltung der Raumimpulsantworten (RIRs) beider Räume beschreiben. Aus [2] ist bekannt, dass die R*R-Impulsantwort (R*R-IR) andere statistische Eigenschaften als die zugrundeliegenden einzelnen RIRs hat. Die Faltung der RIRs führt zu einem verlängerten Energieabfall im Vergleich zu den zugrundeliegenden RIRs und der Abfall des diffusen Nachhalls ist nicht mehr konsequent linear [3]. Die zeitliche Struktur der Pulse der R*R-IR verdichtet sich im Vergleich zu den einzelnen RIRs [2], ebenso verschiebt sich die Energie zeitlich nach hinten [3]. Im Vergleich zu dem Direktschall kumuliert die Energie zu den frühen Reflexionen [3]. Die Eigenschaften der R*R-IR können auch im Frequenzbereich analysiert werden.

Aus der Faltung der RIRs im Zeitbereich wird im Frequenzbereich eine Multiplikation der komplexen Spek-

tren. In logarithmischer Auftragung der Betragsspektren folgt eine Addition. Schroeder klassifizierte Schwankungen über die Frequenzen des Spektrums über die Standardabweichung (σ_{spec}) und stellte fest, dass für einen einzelnen Raum eine σ_{spec} von 5.6 dB gilt [4]. Die Faltung zweier RIR führt zu einer generellen Erhöhung von σ_{spec} um $\sqrt{2}$ [3]. Eine Untersuchung zur subjektiv empfundenen Färbung ergibt, dass dieser Anstieg von σ_{spec} mit einer Erhöhung der subjektiv wahrgenommener Färbung einhergeht [3].

In den dieser Studie werden die zeitlichen Eigenschaften der R*R-IR diskutiert und Anhand eines Hörexperiment zur Knackigkeit die Einflüsse des Einschwingvorgangs der R*R-IR im Vergleich zu einzelnen Räumen untersucht.

Hörexperiment: Knackigkeit

Methode

In einem Hörversuch ist der Einfluss einer R*R-Wiedergabe auf die subjektive Empfindung der Knackigkeit des Klanges im Vergleich zu einzelnen Räumen mit einem Rating-Experiment (1 = wenig knackig, 7 = sehr knackig) beurteilt worden. Im Aufnahmeraum findet eine Aufnahme einer Bass-Drum in fünf verschiedenen Distanzen statt (siehe Abbildung 1). Das aufgenommene Signal wird daraufhin im Wiedergaberaum abgespielt und an der Abhörposition von einem Mikrofon aufgenommen.

Tabelle 1: Physikalische Maße und T_{30} -Zeiten der einzelnen Räume für die Bewertung der Knackigkeit.

Aufnahmeraum 1	Breite = 6 m, Länge = 7 m, Höhe = 3 m $V = 126 \text{ m}^3$ $d_1 = [0.2 \ 0.5 \ 1 \ 2 \ 4] \text{ m}$ $\text{Ø}T_{30} = 272 \text{ ms}$
Aufnahmeraum 2	Dimensionen, siehe Aufnahmeraum 1 $\text{Ø}T_{30} = 860 \text{ ms}$
Wiedergaberaum	Breite = 5 m, Länge = 7 m, Höhe = 3 m $V = 105 \text{ m}^3$, $d_2 = 3 \text{ m}$ $T_{30} = 500 \text{ ms}$

Testsignale

Die Räume sind virtuell mit einem Raumsimulationsprogramm [5] generiert worden, Details sind Tabelle 1 zu entnehmen. Über die Faltung konnte somit einem trockenen Bass-Drum Signal der Raumeindruck der simulierten Räume bzw. der R*R-IR aufgeprägt werden.

Aufbau des Hörexperiments

Der Hörversuch ist in einer einwandigen Hörkabine an der Universität Oldenburg durchgeführt und über Kopfhörer (Sennheiser HD650) Versuchspersonen präsentiert worden. An dem Hörexperiment nahmen elf Versuchspersonen (Vp) mit einem Durchschnittsalter von 29 Jahren teil. Jede Vp hat pro Kondition eine Bewertung abgegeben (kein Retest), bei einem RMS-Pegel von 74 dB SPL. Knackigkeit steht für die Plausibilität des Transienten und wurde nach der Definition von [6] den Versuchspersonen erklärt.

Ergebnisse

Die Ergebnisse des Hörversuchs sind in Abbildung 2 dargestellt. Auf der Y-Achse ist die Bewertung der Knackigkeit aufgetragen, wobei das Urteil 7 für sehr knackig und das Urteil 1 für wenig knackig steht. Auf der X-Achse sind die Distanzen d_1 (Aufnahmeraum) und d_2 (Wiedergaberaum) aufgetragen. Über die Vp ist der Mittelwert und der Standardfehler ermittelt.

Die höchste wahrgenommene Knackigkeit besitzt die Bassdrum aus dem Aufnahmeraum 1 bei einer Distanz von 20 cm. Es ist zu sehen, dass die Knackigkeit der Aufnahme im Aufnahmeraum mit anwachsender Distanz kleiner wird. Der Wiedergaberaums im Mittel eine Bewertung etwas kleiner als 5.

Die R*R-Wiedergabe der Aufnahmen aus Aufnahmeraum 1 hat eine deutlich niedrigere Bewertung und für größer werdende Distanzen im Aufnahmeraum sinkt die Knackigkeit.

Zu diesen Ergebnissen ist noch die Bewertung der Knackigkeit von Aufnahmeraum 2 dargestellt. Die Knackigkeit fällt hier im Vergleich zu Aufnahmeraum 1 steiler über die Distanz ab.

Die R*R-Wiedergabe wird im Vergleich über die Distanzen im Mittel weniger Knackig bewertet. Nicht selbsterklärend ist allerdings der Fakt, das Aufnahmeraum 2, trotz der höheren $\overline{OT}_{30} = 860$ ms niedriger Bewertet wird als die R*R-Wiedergabe ($\overline{OT}_{30} = 610$ ms).

Diskussion

Die Knackigkeit einer Wiedergabe für einen einzelnen Raum scheint von der Distanz zwischen Quelle und Empfänger abzuhängen, was für beide Aufnahmeräume zu erkennen ist. Deutlich wird die Abnahme der Knackigkeit sobald sich das Mikrofon außerhalb des Hallradiuses befindet. Die Knackigkeit wird auch durch die Nachhallzeit beeinflusst, deutlich zu erkennen, da der trockenere Aufnahmeraum im direkten Vergleich bei selber Distanz knackiger bewertet wurde.

Beim Vergleich der R*R-Wiedergabe mit Aufnahmeraum 2 fällt allerdings auf, dass die R*R-Wiedergabe trotz geringerer durchschnittlicher Nachhallzeit ($\overline{OT}_{30} = 610$ ms, siehe Tabelle 2) niedriger bewertet wurde. Interessant ist auch der Vergleich der einzelnen Nachhallzeiten der R*R-IR miteinander, dargestellt in Tabelle 2 (unten). Bei einer theoretischen JND von 10% [1] liegen alle Nachhallzeiten

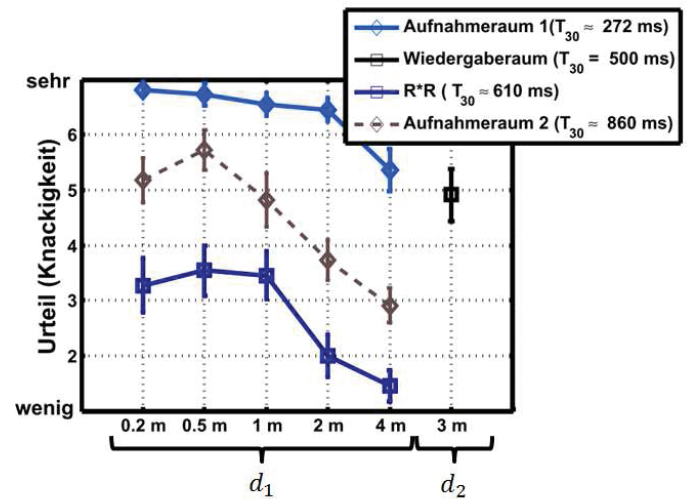


Abbildung 2: Ergebnisse des Knackigkeitsexperiments, für Aufnahmeraum 1, dem Wiedergaberaum, der Wiedergabekette von Aufnahmeraum 1 und dem Wiedergaberaum, sowie der Bewertung für einen Aufnahmeraum 2 mit einer erhöhten T_{30} -Zeit im Vergleich zu Aufnahmeraum 1

innerhalb dieser JND, was den Rückschluss zulässt, dass die Nachhallzeit hier keinen hörbaren Einfluss auf die Bewertung der R*R-Wiedergabe über die Distanzen hat.

Tabelle 2: T_{30} -Zeiten für Aufnahmeraum 1 und 2 und der R*R-Wiedergabe in Abhängigkeit der Distanzen

	\emptyset	0.2 m	0.5 m	1 m	2 m	4 m
Aufn 1						
T_{30} (ms):	272	253	254	257	292	308
Aufn 2						
T_{30} (ms):	860	822	835	843	891	912
R*R						
T_{30} (ms):	610	592	599	598	624	636

Exemplarisch sind in Abbildung 3 drei der fünf R*R-IR dargestellt mit den dazugehörigen Spektren. Die Impulsantworten sind zueinander verschoben geplottet. Bei einer kleinen Distanz von 20 cm im Aufnahmeraum, ist zunächst kein besonderer visueller Effekt sichtbar. Sobald sich die Distanz im Aufnahmeraum erhöht, wird zunächst ein kleines Plateau bei 1 m sichtbar, bei 4 m ein deutlicher Einschwingvorgang der R*R-IR, auf welchen durch [1,2,3] aufmerksam gemacht wird. Für die resultierenden Spektren ist ebenso kein bestechender Abfall von hohen Frequenzen zu sehen, welcher die Knackigkeit beeinflussen könnte.

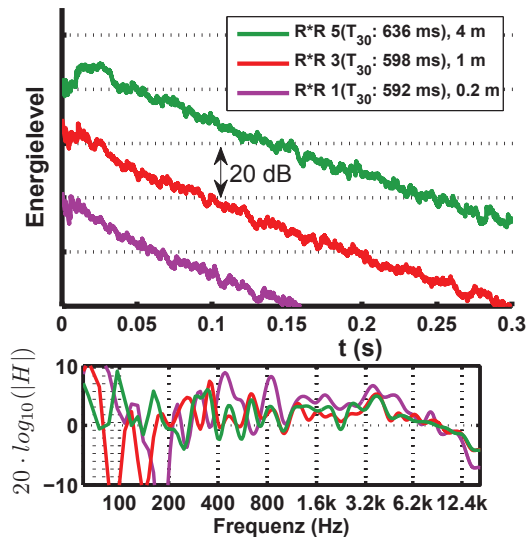


Abbildung 3: Der Kurvenverlauf der R*R-IR (relativ zueinander verschoben) in Abhängigkeit der Distanz von 0.2 m, 1 m, 4 m im Aufnahmerraum (oben) und die resultierenden $\frac{1}{3}$ -Oktave geglätteten Spektren (unten).

Der Rückschluss auf den Einschwingvorgang der R*R-IR, welcher hier den Unterschied in der Knackigkeit zwischen Aufnahmerraum 2 und der R*R-Wiedergabe darstellt, ist somit möglich. Ebenso sollte der Einbruch der Knackigkeit der R*R-Wiedergabe über die Distanzen auf den Einschwingvorgang zurückzuführen sein.

Schlussfolgerung

Eine Faltung zweier Impulsantworten verändert die wahrgenommenen zeitlichen und spektralen Eigenschaften (Knackigkeit/Färbung). Dabei sind zeitliche Eigenschaften, wie der Transient der RIR ein hervorzuhebendes Problem. Ein Hörversuch konnte zeigen, dass die R*R-Wiedergabe für eine Bass-Drum generell weniger Knackig bewertet wurde, als der Aufnahmerraum, bzw. der Wiedergaberaum alleine. Dabei hat schon eine Distanz im Aufnahmerraum von 20 cm Einfluss auf die wahrgenommene Knackigkeit. Die Faltung der RIRs führt zu einer Verschmierung des Transienten. Die Energie in den frühen Reflexionen prägt sich im Vergleich zum Direktschall immer stärker aus und es konnte gezeigt werden, dass dies zur Verringerung der Knackigkeit führt.

Die spektralen Eigenschaften einer Aufnahme können durch lineare Manipulation des Signals aus dem Aufnahmerraum (z. B. durch Equalising) korrigiert werden [7]. Temporäre Eigenschaften, wie der Transient scheinen in dieser Hinsicht schwerer zu korrigieren, da man dazu wahrscheinlich Zeitvariante, nichtlineare Signaltransformation handelt (z. B. Gate, Transient Shaper), welche zu hörbaren Artefakten führen können.

Literatur

- [1] Hak C.C.J.M., Wenmaekers R.H.C.: *The Impact of Sound Control Room Acoustics on the Perceived Acoustics of a Diffuse Field Recording*. WSEAS Transactions on Signal Processing 6.4 (2010): 175-185.
- [2] Vogel P., de Vries D.: *Electroacoustic system response in a hall: a convolution of impulse sequences*. Journal of the Audio Engineering Society 42.9 (1994): 684-690.
- [3] Haeussler A., van de Par S.: *Psychoakustische Untersuchung zum Einfluss des Aufnahmerraums auf den Wiedergaberaum*. , Fortschritte der Akustik. DAGA 2014, 40. Jahrestagung für Akustik
- [4] Schroeder M. : *Die statistischen Parameter der Frequenzkurven von grossen Räumen* Acustica 4 (1954) 594-600
- [5] Wendt T., van de Par, Steven; Ewert, Stephan D.: *A Computationally-Efficient and Perceptually-Plausible Algorithm for Binaural Room Impulse Response Simulation*, JAES Volume 62 Issue 11 pp. 748-766; November 2014
- [6] Lindau A. et al. *A Spatial Audio Quality Inventory (SAQI)*. Acta Acustica united with Acustica 100.5 (2014): 984-994.
- [7] Grosse J., van de Par, S., *Perceptually accurate reproduction of recorded sound fields in a reverberant room using spatially distributed loudspeakers*, Selected Topics in Signal Processing, IEEE Journal of , vol.PP, no.99, pp.1,1 (2014)