

Theoretischer und subjektiver Einfluss des Aufnahmeraumes auf den Wiedergaberaum

Andreas Haeussler, Steven van de Par

Carl von Ossietzky Universität, Institut für Medizinische Physik Acoustics Group, D-26111 Oldenburg,
andreas.haeussler@uni-oldenburg.de

Einleitung

Heutzutage ist es üblich sich Musik- oder Sprachaufnahmen aus einem Raum (Aufnahmeraum) in einem anderen Raum (Wiedergaberaum) erneut anzuhören. Beispielhaft für das Szenario ist eine Konzertaufnahme bzw. eine Studioaufnahme oder die Übertragung einer Audiokonferenz. Diese Wiedergabekette ist schematisch in Abbildung 1 dargestellt und wird als „Raum-in-Raum“ (RinR)-Wiedergabe [1] bezeichnet. Im folgenden wird diskutiert inwiefern der zweite Raum in der Wiedergabekette zu einer grundlegenden Änderung des Verlaufs der Raumimpulsantwort (RIR) und dessen resultierenden Spektrums führt.

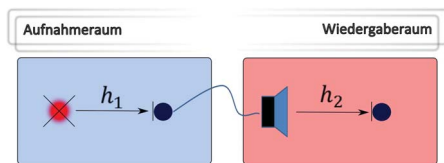


Abbildung 1: RinR-Wiedergabe mono: Eine Quelle spielt ein Signal s im Aufnahmeraum ab, welches über ein Mikrofon aufgenommen und im Wiedergaberaum über einen Lautsprecher abgespielt wird. Anstatt eines Zuhörers befindet sich im Wiedergaberaum ein Mikrofon. Die RIR h_1 und h_2 beschreiben jeweils die Sender-Empfänger-Strecke

Grundlagen

Im Aufnahmeraum sendet eine Quelle ein Signal s aus, welches von einem Mikrofon empfangen wird. Das aufgenommene Signal x kann mathematisch als Faltung der RIR h_1 des Aufnahmeraumes mit dem Signal s betrachtet werden ($x = h_1 * s$). Das verhaltene Signal x wird nun im Wiedergaberaum über einen Lautsprecher abgespielt und erneut mit einem Mikrofon an der Abhörposition aufgenommen, was mathematisch mit einer Faltung der RIR h_2 beschrieben wird ($y = h_2 * x$). Bei der Zusammenfassung der Wiedergabekette entsteht eine Faltung beider RIR:

$$y = s * (h_1 * h_2) \quad (1)$$

Exemplarisch ist die Faltung zweier Impulsantworten h_1 und h_2 in Abbildung 2 logarithmisch dargestellt¹. In der linken Spalte ist zu erkennen, dass die Faltung den Energieabfall der RinR-Impulsantwort verlängert. Die oberste RIR (Raum 1) hat eine Nachhallzeit von $T_{60} = 600$ ms und die mittlere (Raum 2) eine Nachhallzeit von $T_{60} = 1200$ ms. Durch die Faltung entsteht eine resultierende Nachhallzeit von $T_{60} = 1500$ ms. Die

¹Zur deutlicheren Darstellung, sind die logarithmischen RIRs mit einer Zeitkonstante $\tau = 2.5$ ms geglättet worden

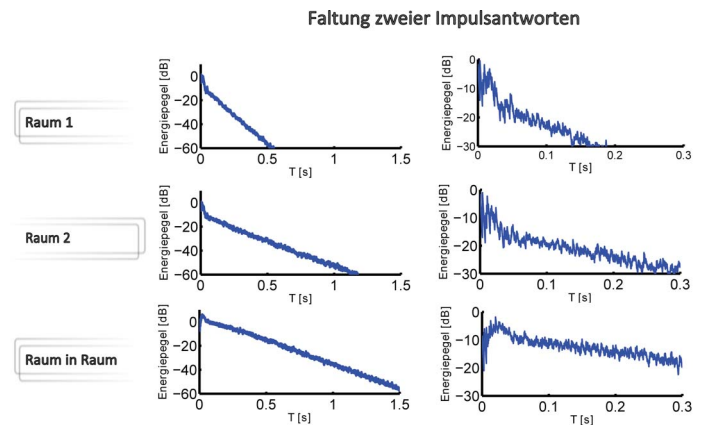


Abbildung 2: Simulierte RIRs von Raum 1 (oben) und Raum 2 (mitte) und RinR (unten) als normalisierter Energiepegel. Die linke Spalte zeigt die RIR mit einem Abfall auf -60 dB, die rechte Spalte einen vergrößerten Ausschnitt der ersten 300 ms bis -30 dB.

RinR-Impulsantwort fällt im Logarithmischen weiterhin annähernd linear ab, allerdings ist in der rechten Spalte ein Einschwingen von ca. $5 - 7$ dB erkennen. Dies deutet auch darauf hin, dass die Faltung zu mehr Signalenergie in den ersten 30 – 50 ms führt.

In Abbildung 3 sind die Spektren H_1 und H_2 der einzelnen Räume (oben/mitte) und das resultierende Spektrum einer RinR-Impulsantwort (unten) dargestellt. Für die weitere Betrachtung wird die Schwankungsstärke über die Standardabweichung (STD) klassifiziert. Oberhalb der Schroederfrequenz kann das logarithmische Betragsspektrum als log-Rayleigh-verteilt angenommen werden und hat eine STD von ca. 5.6 dB [2]. Dieser Wert ist gültig, wenn der Mittelwert über mehrere Oktaven gleich ist. Für das RinR-Spektrum ergibt sich eine erhöhte STD von ca. 8 dB. Im Frequenzbereich entspricht dies einer Multiplikation der komplexen Spektren oder einer Addition der logarithmischen Spektren:

$$Y = H_1 \cdot H_2 \rightarrow \log(Y) = \log(H_1) + \log(H_2) \quad (2)$$

Somit addiert sich die STD der einzelnen logarithmischen Spektren. Für unkorrelierte Spektren ergibt sich:

$$STD_{\text{RinR}} = \sqrt{STD_1^2 + STD_2^2} \approx \sqrt{2} \cdot 5.6 \text{ dB} \quad (3)$$

Sobald die Spektren eine Korrelation untereinander aufweisen, ist der Korrelationskoeffizient der Spektren mitzuberücksichtigen.

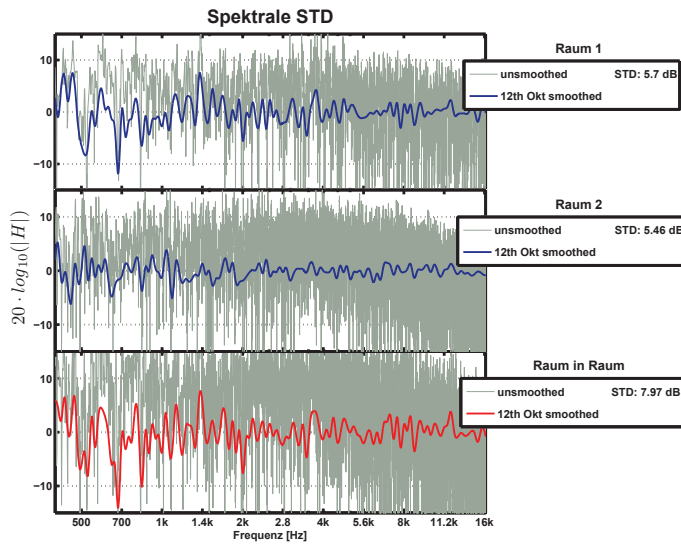


Abbildung 3: Spektren zweier Einzelräume (Raum 1 und Raum 2), sowie das resultierende Spektrum der Faltung dieser beiden Räume. Die graue Linie zeigt das ungeglättete (unsmoothed) Spektrum (log) der RIR, in blau ist eine Glättung über $\frac{1}{12}$ -Oktaven zu sehen. Gemessene STDs der Spektren der Einzelräume und des RinR-Spektrums sind in der Legende angegeben.

Färbung

Wie im vorhergegangenen Kapitel zu sehen ist, entsteht bei der RinR-Wiedergabe eine erhöhte STD im Spektrum. Zu diesem Zweck wurde ein Hörversuch durchgeführt, mit dem Ziel, die Färbungsstärke einer einzelnen RIR auf die Färbungsstärke einer RinR-Impulsantwort anzupassen. In Abbildung 4 ist das Vorverarbeitungsschema der RIR abgebildet. Es werden die frühen Reflexionen extrahiert und ein FIR Filter generiert, welcher die spektralen Eigenschaften, allerdings keine zeitliche Information der RIR besitzt. Dieser Filter wird mit einem Klaviersignal gefaltet. In dem Hörversuch sollten die Versuchspersonen die spektrale Färbungsstärke eines einzelnen Raumes über den Faktor k auf die Färbungsstärke einer RinR-Wiedergabe (Faktor $k = 1$) anpassen.

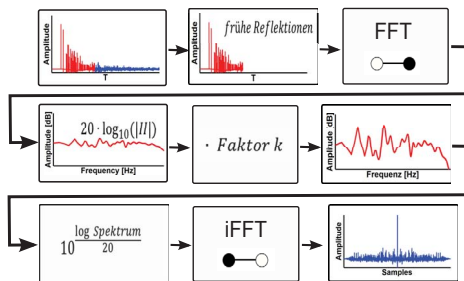


Abbildung 4: Blockschaltbild für die variable Anpassung des Spektrums

Die Ergebnisse der resultierenden STDs der Spektren sind in Abbildung 5 dargestellt. Es ist zu erkennen, dass die meisten RinR-Konditionen zu einer erhöhten Färbungsstärke hinleiten und dass die STD kein guter Schätzer für die subjektive Färbungsstärke ist.

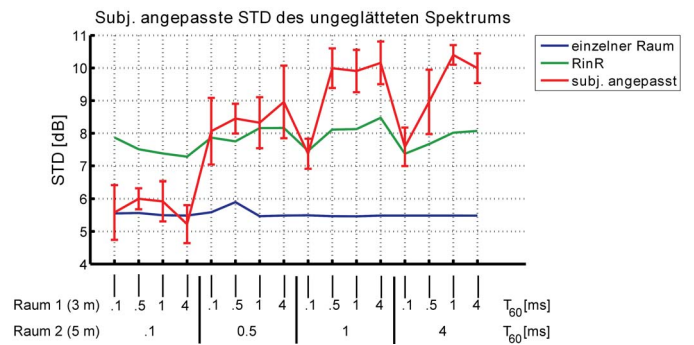


Abbildung 5: Ergebnisse der subj. Bewertung der Färbungsstärke einer RinR-Wiedergabe im Vergleich zu einem einzelnen Raum. Auf der x-Achse sind die Nachhallzeiten der einzelnen Räume zu sehen und die Distanzen von Quelle-Empfänger. In Blau ist die STD des Spektrums für die verschiedenen Nachhallzeitkombinationen dargestellt. Die grünen Datenpunkte zeigen die STD der RinR-Wiedergabe. In Rot ist die STD des subj. angepassten Spektrums auf die RinR-Wiedergabe zu sehen.

Durch die Glättung der Spektren über eine Oktave und die erneute Berechnung der STD ist in Abbildung 6 zu erkennen, dass die subjektiven Daten durch die Glättung sehr gut vorhersagbar sind.

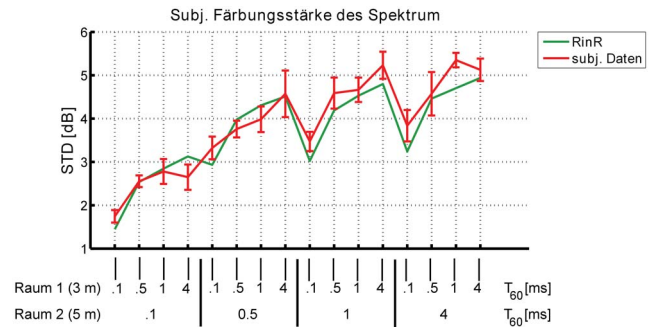


Abbildung 6: Ergebnisse der RinR-Wiedergabe (grün) und der subj. angepassten Spektrums (rot) aus Abbildung 5 mit der Glättung der Spektren.

Zusammenfassung

Die resultierende Form der RinR-Impulsantwort unterscheidet sich von den einzelnen RIR durch eine verlängerte Nachhallzeit und durch ein verstärktes Einschwingverhalten. Durch die Erhöhung der spektralen STD, kommt es zu einem Anstieg der gehörten Färbung. Als gutes Maß zur Vorhersage der Färbungsstärke einer RinR-Impulsantwort hat sich die Glättung des Spektrums über eine Oktave mit erneuter Berechnung der STD erwiesen.

Literatur

- [1] C.C.J.M. Hak, R.H.C. Wenmaekers: *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] M. Schroeder: *Die statistischen Parameter der Frequenzkurven von grossen Räumen* Acustica 4 (1954) 594-600