

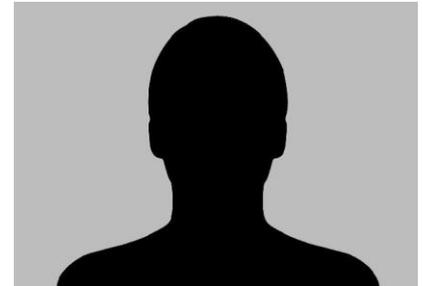
Signalverarbeitungsmethoden zur Enthaltung von Audiosignalen

Diese Arbeit untersucht die Enthaltung und künstliche Verhallung von Audiosignalen. Hall kann die Verständlichkeit der Sprache beeinträchtigen. Trockene Audiosignale werden mit Raumimpulsantworten (RIR) verhallt und anschliessend mit unterschiedlichen Algorithmen enthüllt. Durch die Enthaltung soll die verhallte Sprache für Menschen verständlicher werden.

Mit der Sinus-Sweep-Methode und der Maximum-Length Sequence-Methode (MLS) wurden unterschiedliche RIR gemessen. Dabei stellte sich heraus, dass die Sinus-Sweep-Methode ein besseres Signal-Rausch-Verhältnis (SNR) als die MLS-Methode aufweist. Zur Generierung von RIR wurden die Spiegelquellenmethode und die Finite-Difference-Time-Domain-Methode (FDTD) verwendet. In dieser Arbeit wurde zusätzlich die effiziente Hybrid-Methode verwendet, welche eine Kombination der Spiegelquellen- und der FDTD-Methode ist. Ein mit der generierten RIR verhalltes Sprachsignal sollte dabei möglichst natürlich klingen. Dazu wurden die Spektrogramme der generierten RIR und der gemessenen RIR visuell verglichen. Zusätzlich wurde ein trockenes Sprachsignal mit der generierten RIR und der gemessenen RIR verhallt und akustisch überprüft. Für einen natürlichen Klang mussten die generierten RIR nachbearbeitet werden. Unter Berücksichtigung der Absorption von Schallwellen in der Luft wird eine gemessene RIR sehr gut approximiert.

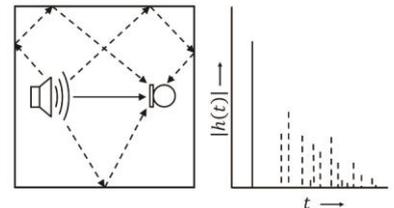
Als Enthaltungs-Algorithmen wurden Weighted Prediction Error (WPE) mit einer Online- und Offline-Version, Cepstrum Mean Subtraction (CMS), Cepstrum Liftering und Spectral Subtraction erarbeitet. Die Arbeit beinhaltet ausschliesslich Single-Channel-Methoden zur Enthaltung.

Die Algorithmen wurden mit diversen objektiven Gütekriterien beurteilt und mit Punkten bewertet. Zusätzlich wurde eine Umfrage mit Hörbeispielen durchgeführt. Die objektive und die subjektive Bewertung zeigten dabei kaum Übereinstimmungen. Der CMS-Algorithmus und der Cepstrum-Liftering-Algorithmus führten für Menschen störende Artefakte ein. Der WPE-Algorithmus in der Online-Version hatte weder objektiv noch subjektiv einen feststellbaren Enthaltungseffekt. Die Offline-Version des WPE-Algorithmus enthalte das Audiosignal bei allen Kriterien sehr gut. Dieser Algorithmus ist jedoch rechenintensiv. Die Spectral Subtraction erreichte bei fast allen objektiven Kriterien schlechte Resultate. Subjektiv betrachtet wurde der Algorithmus allerdings sehr gut bewertet und ist zudem sehr recheneffizient.

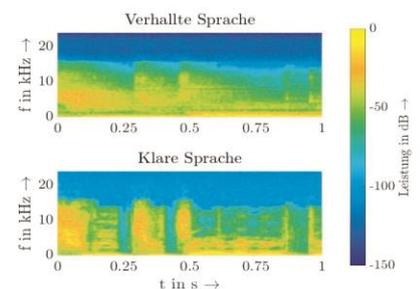


Diplomierende
Moritz Oppliger
Hannes Stoll

Dozent
Sigisbert Wyrsch



Links: Schematischer Hall in einem Raum. Rechts: Schematische Raumimpulsantwort (RIR). Direktsignal durchgezogen, Reflexionen gestrichelt dargestellt.



Leistung eines Sprachsignals in Abhängigkeit der Zeit t (x-Achse) und der Frequenz f (y-Achse). Die normierten Leistungspegel sind mit Farbe dargestellt.