

Vortragender: Dipl.-Ing. Aleksej Chinaev (Fachgebiet Nachrichtentechnik)

Thema: Bayes'scher Postprozessor für die Schätzung der spektralen Rauschleistungsdichte und seine Leistungsfähigkeit in der Sprachsignalverarbeitung

Abstract:

Die Qualitätsverbesserung gestörter Sprachsignale und robuste Spracherkennung sind große und interessante Forschungsfelder. Häufig wird hier von additiven Störungen ausgegangen, die sich aus den verrauschten Sprachsignalen besonders gut im Frequenzbereich mittels Spektraler Subtraktion entfernen lassen. Die Schätzung der spektralen Rauschleistungsdichte (SRLD) spielt dabei eine entscheidende Rolle.

Besonders herausfordernd ist die SRLD-Schätzung in Gegenwart von instationären Hintergrundstörungen, wie Lärm einer Menschenmenge, Maschinengeräusche, Straßenlärm, Geschirrgeklapper oder Bürolärm. Hier verfolgen konventionelle Verfahren die zeitvariante SRLD nur während der Sprachpausen und halten den SRLD-Schätzwert konstant bei Sprachaktivität, während der die Störung vom Sprachsignal 'verdeckt' und die SRLD-Schätzung inakzeptabel fehlerbehaftet wird.

Im Fachgebiet Nachrichtentechnik haben wir uns der Herausforderung gestellt, die SRLD während der Sprachaktivität zu schätzen. Als Ergebnis haben wir ein Bayes'sches Schätzverfahren entwickelt, das auf den Namen Maximum a-posteriori basierter Postprozessor getauft wurde. Die Funktionsweise des Postprozessors und seine Leistungsfähigkeit in der SRLD-Schätzung, in der Qualitätsverbesserung gestörter Sprachsignale und in der robusten Spracherkennung werden im Rahmen des Vortrags ausführlich diskutiert.