

## Advances in Blind Multichannel Wiener Filtering of Noisy Speech

Das mehrkanalige Wiener-Filter (engl. *multichannel Wiener filter*, MWF) ist ein Signalverarbeitungsverfahren zum Minimieren des mittleren quadratischen Fehlers (engl. *mean-square error*, MSE) zwischen seinem Ausgangssignal und einem gegebenen Zielsignal. In Anwendungen der Sprachsignalverarbeitung, in denen mehrere Mikrofone ein Zielsignal gestört durch additive Umgebungsgeräusche und Nachhall aufnehmen, könnte das MWF somit eine MSE-optimale Schätzung des sauberen Sprachsignals liefern. Es würde dadurch sowohl Störgeräuschunterdrückung als auch Enthüllung leisten – beides Schlüsselaufgaben in der Sprachsignalverbesserung. Die Anwendung des MWF benötigt jedoch *a priori* Kenntnis der akustischen Übertragungsfunktionen (engl. *acoustical transfer function*, ATF) zwischen der Signalquelle und den Mikrofonen sowie des Leistungsdichtespektrums (engl. *power spectral density*, PSD) der Sprache und der Kovarianzmatrix der Störung.

Die vorliegende Arbeit unterbreitet daher Vorschläge, wie die Leistungsfähigkeit des MWF blind angenähert, d.h., die benötigte Menge an *A-priori*-Information so gering wie möglich gehalten werden kann. Zu diesem Zweck wird die bekannte Zerlegung des MWF in einen räumlichen Verarbeitungsschritt in Form eines MVDR-Beamformers (engl. *minimum-variance distortionless-response*, MVDR) und eine spektralen Nachbearbeitung im Sinne eines einkanaligen Wiener-Filters betrachtet. In dieser Arbeit werden blinde Kanalschätztechniken benutzt, um geeignete ATF-Schätzungen zu erhalten und so eine blinde Anwendung des MVDR-Beamformers zu ermöglichen. Die räumlichen Eigenschaften des Störschallfeldes in Form einer normierten Kovarianzmatrix werden als im Vorfeld messbar und somit als *a priori* bekannt angenommen.

Den Kern dieser Arbeit bildet ein neuartiges, zweistufiges Schätzverfahren für die Spektralverbesserung im MWF, das sogenannte Postfilter. Im ersten Schritt wird ein einfaches statistisches Modell der Mikrofonensignale vorgestellt und *maximum-likelihood* (ML) optimale Schätzer für die Sprach- und Störsignal-PSD werden hergeleitet. Aus diesen ML-optimale PSD-Schätzungen wird anschließend ein *ML-basiertes* Postfilter konstruiert. Im zweiten Schritt wird das ML-basierte Postfilter selbst als statistische Größe aufgefasst. Die Modellierung seiner Verteilung in Form einer Postfilter-Likelihood und das Einführen einer *A-priori*-Verteilung des idealen Wiener-Postfilters ermöglichen eine Bayes'sche Korrektur der geschätzten Postfilter-Werte. Diese *Bayes-Verfeinerung* wird als Erweiterung des ML-basierten Postfilters in Form einer vorberechneten Nachschlagetabelle umgesetzt.

Der blinde MVDR-Beamformer und die ML-basierten und Bayes-verfeinerten Postfilter werden in verschiedenen akustischen Konfigurationen in Computersimulationen sowie auf Basis der CHiME-4-Sprachdatenbank untersucht. Die vorgestellten Postfilter zeigen gute Störgeräuschunterdrückung, gemessen in segmenteller SNR- und PESQ-Verbesserung, und reduzieren auch in begrenztem Umfang späten Nachhall. Das Bayes-verfeinerte Postfilter bietet dabei eine konsistente Verbesserung des ML-basierten Postfilters. Diese Ergebnisse bestätigen die Perspektive für ein blindes MWF basierend auf den verfügbaren Mikrofonensignalen bei konsequenter Nutzung einer begrenzten Menge an *A-priori*-Information.