

Dr. Nir Fink & Yami Thor

Bettcar

Deep Audio Processing (DAP): Eine KI-gestützte Technologie zur Audio-Personalisierung zur Verbesserung der Sprachverständlichkeit und des Hörerlebnisses

Die Sprachverständlichkeit in komplexen akustischen Umgebungen stellt trotz moderner Hörsysteme weiterhin eine große Herausforderung für schwerhörige Menschen dar. Konventionelle Ansätze wie Verstärkung, Störgeräuschunterdrückung und KI-basierte Denoising-Verfahren sind in dynamischen Mehrquellenszenarien oft unzureichend.

Dieser Vortrag stellt Deep Audio Processing (DAP) vor, eine KI-gestützte Technologie zur Audio-Personalisierung, die eine inhaltsbasierte und nutzerspezifische Anpassung ermöglicht. DAP kombiniert Audioszenenanalyse, Quellentrennung sowie eine wahrnehmungsbasierte Neugewichtung von Sprache, Musik und Hintergrundgeräuschen. Ein zentrales Element ist die phonem-basierte, nicht-uniforme Zeitmodifikation (NU-TSM), die eine gezielte zeitliche Anpassung von Sprache erlaubt, um wahrnehmungsrelevante Hinweise zu verstärken, ohne globale Verzerrungen zu erzeugen.

Das System ist für den Echtzeitbetrieb unter niedrigen Latenzanforderungen ausgelegt. Psychoakustische Untersuchungen zeigen eine verbesserte Sprachverständlichkeit, reduzierten Höranstrengungsaufwand sowie ein verbessertes Hörerlebnis.

DAP stellt einen Paradigmenwechsel hin zu wahrnehmungsbasierter Audioverarbeitung dar.