A New Spectral-Geometric Optimized Architecture for Distortion-Free Binaural Speech Intelligibility

Shen, Elizabeth (School: Jesuit High School)

Commercial hearing aids struggle with poor speech clarity in noise, notably “cocktail party” settings. Many devices also distort the spatial cues critical for sound localization. As a result, over a fifth of hearing aid users have discontinued usage due to discomfort. To address the fundamental speech intelligibility and spatial fidelity tradeoff, this project presents a new binaural signal-processing architecture inspired by biological auditory processing. Implemented in MATLAB and Python, the system processes signals from binaural microphones. A lightweight Conv-TasNet model classifies input sound and computes energy masks for extracted speeches. The identified speech signals feed a spatially revised GCC-PHAT function, performing direction-of-arrival (DOA) estimation to track speech in multi-speaker environments. The architecture then uses DOA information to optimize an adaptive multi-constraint LCMV beamformer, improving upon the single-constraint beamformers used in traditional hearing aids. In the LCMV, target steering is derived from user orientation, spatial constraints from localized speech sources, and a covariance matrix from exponentially weighted recursion, enabling distortion-free speech enhancement. This real-time architecture was evaluated on over 12,000 open-source recordings via the Short-Time Objective Intelligibility metric, which assesses speech intelligibility on a 0-1 scale. Results demonstrate the algorithm improved speech clarity by an average of 20% compared to leading ML filtering and 54% compared to classical noise suppression methods such as spectral subtraction or Wiener filtering. Furthermore, realistic interaural time difference cues are preserved for a spatially natural listening experience.

Awards Won: