

Music Signal Processing for the Reduction of Auditory Distortions in Hearing-impaired Listeners

(Musiksignalverarbeitung zur Reduktion auditorischer Verzerrungen bei Schwerhörigen)

von Anil M. Nagathil

Kurzfassung

Ein sensorisch bedingter Hörverlust führt typischerweise zu einer Reduktion der Frequenzselektivität, die Klangfarben- und Tonhöhenverzerrungen in akustischen Signalen hervorruft. Dies hat zur Folge, dass Schwerhörige Schwierigkeiten bei der Unterscheidung von Musikinstrumenten und der Erkennung von Melodien in komplexen Musikstücken haben. Hinzu kommt, dass derzeit verfügbare Hörhilfen aufgrund technischer Einschränkungen keine exakte Musikübertragung ermöglichen und dadurch zusätzliche wahrnehmungsbezogene Verzerrungen erzeugen. Daher äußern insbesondere Cochleaimplantat (CI) Träger ihre Unzufriedenheit über die Qualität der Musikwiedergabe.

Da CI-Träger in früheren Arbeiten eine Präferenz für Soloinstrumentalmusik gegenüber komplexerer Ensemble- und Orchestermusik gezeigt haben, schlagen wir in dieser Arbeit Strategien zur Reduktion der spektralen Komplexität von Musiksignalen und damit zur Linderung von Effekten eines sensorisch bedingten Hörverlusts vor. Dazu werden Niedrig-Rang Approximationen von Musiksignalen untersucht, die durch Anwendung von Dimensionsreduktionsverfahren im Zeit-Frequenz-Bereich bestimmt werden. Da wir uns hauptsächlich mit Klangfarben- und Tonhöhenverzerrungen befassen, erfolgen die Untersuchungen in dieser Arbeit auf der Grundlage von nicht-perkussiver und rein instrumentaler klassischer Kammermusik mit einer klar definierten monophonen führenden Stimme und einer polyphonen Begleitung. In Anbetracht eines Mangels an passenden Bewertungsmaßen, die Effekte eines sensorisch bedingten Hörverlusts berücksichtigen, werden in dieser Arbeit zusätzliche Maße vorgeschlagen, die Änderungen der auditorischen Verzerrungen bei prozessierten Musiksignalen sowie Musikbewertungen durch CI-Träger vorhersagen. Die Niedrig-Rang Approximationen werden mit Hilfe von bereits existierenden Quellentrennungsmaßen, den vorgeschlagenen Maßen und auf der Grundlage von Hörversuchen mit Normalhörenden (NH) und CI-Trägern bewertet.

Ein Vergleich verschiedener Spektraltransformationen und blinder Dimensionsreduktionsverfahren zeigt, dass eine im Constant-Q Spektralbereich angewendete Principal Component Analysis (PCA) die effektivste spektrale Komplexitätsreduktion bewirkt, unter gleichzeitiger Beibehaltung der stärksten Harmonischen der führenden Stimme. Um Vorteile durch eine Zunahme von Noteninformationen der führenden Stimme zu untersuchen, werden außerdem das Partial Least Squares (PLS) Verfahren, das mit der PCA verwandt ist, sowie ein überwachtes Quellentrennungsverfahren auf Musiksignale angewandt. Obwohl diese Methoden der PCA hinsichtlich der Wiedergabetreue der führenden Stimme überlegen sind, sagen die vorgeschlagenen Maße bei Verwendung einer PCA eine beträchtlich höhere Reduktion von auditorischen Verzerrungen und höhere Bewertungen durch CI-Träger voraus. Dieses Ergebnis wird durch Hörversuche bestätigt. Während CI-Träger die durch eine PCA prozessierten Signale in 75% der Fälle signifikant gegenüber unprozessierten Signalen bevorzugen, erreicht das PLS-Verfahren nur eine maximale Präferenz von 68%. Musik, die dagegen durch das überwachte Quellentrennungsverfahren zerlegt und dann bei höheren Signal-zu-Interferenz Verhältnissen neu gemischt wird, wird nicht bevorzugt gegenüber dem nicht prozessierten Fall. Diese Ergebnisse deuten darauf hin, dass eine Bedämpfung der Begleitung sowie von schwachen Harmonischen der führenden Stimme für die Verbesserung der Musikwahrnehmung von CI-Trägern effektiver ist als eine starke Unterdrückung der Begleitung und eine wiedergabetreue Extraktion der führenden Stimme.