Abstract

Cochlear hearing loss typically leads to a reduction of frequency selectivity, which causes distortions of timbre and pitch in acoustic signals. As a consequence, listeners with cochlear hearing loss have difficulties to discriminate between music instruments and to recognize and follow melodies in complex music. In addition, technical limitations of currently available hearing instruments prohibit an accurate transmission of music and thus cause further perceptual distortions. Therefore, especially cochlear implant (CI) listeners express their dissatisfaction with the quality of music reproduction.

As previous works have found that CI listeners typically prefer music played by isolated instruments over more complex ensemble or orchestra music, in this work we propose strategies for reducing the spectral complexity of music signals and thus for mitigating effects of cochlear hearing loss. To this end, reduced-rank approximations of music signals are investigated, which are obtained by applying dimensionality reduction techniques in the time-frequency domain. Since we mainly address timbre and pitch distortions, the investigations performed in this work are based on non-percussive and purely instrumental classical chamber music with a distinct monophonic leading instrument and a polyphonic accompaniment. Given the lack of suitable instrumental music quality measures which take effects of cochlear hearing loss into account, in this work we also propose metrics which predict changes of auditory distortion in processed music signals and provide estimated music ratings of CI listeners. The reduced-rank approximations are assessed using existing source separation measures and the proposed metrics, and by performing listening tests with normal-hearing (NH) and CI listeners.

A comparison of different spectral transforms and blind dimensionality reduction methods reveals that principal components analysis (PCA) applied in the constant-Q spectral domain of music signals is most effective in terms of spectral complexity reduction while preserving the strongest harmonics of the leading voice. To investigate the benefits of including score information of the leading voice, also partial least squares (PLS) analysis, which is closely related to PCA, and a supervised source separation method are applied to music signals, respectively. Although these methods outperform PCA regarding a high-fidelity reconstruction of the leading voice, the proposed music quality measures predict a considerably higher reduction of auditory distortion and higher CI listener ratings when PCA is used. This result is confirmed by listening tests. While signals processed by PCA are significantly preferred by CI listeners over unprocessed music in up to 75% of all cases, PLS only achieves maximal preference rates of 68%. In contrast, music separated by the supervised source separation method and remixed at higher signal-to-interference ratios is not preferred over the unprocessed case. These results suggest that attenuating the accompaniment and weak harmonics of the leading voice is more effective for improving music appraisal of CI listeners than a strong suppression of the accompaniment and an accurate extraction of the leading voice.