Abstract

The thesis focuses on the usage of convolutional-recurrent neural networks in the field of audio quality analysis: automatic artefacts classification for a real music content. The key contribution in this approach, compared to the existing research, is that the examined model is evaluated in terms of detecting acoustic anomalies without the comparison to the known reference signal, since it is often unavailable at the time of validation. Currently, non-reference audio quality assessment methods are mainly developed for speech only. However, real music signals are more complex to analyze, since they may include instrumental sounds, speech, singing voice as well as various audio effects.

The thesis describes the research conducted by the author, as the process of developing the prototype model for automatic audio artefacts classification. First, we describe the most important information and problems related to the manual and automatic audio quality assessment. Then, we present all steps related to designing and implementing the artificial neural network model, which was evaluated in terms of various architectures and input parameters. The last section includes a discussion about the obtained results, and conclusions.

The integral part of the performed research is also the creation of the custom database, which was used for the process of model evaluation. The database is a set of signals with artefacts selected based on ITU Recommendation BS.1284-2. A publicly available dataset with clear (unprocessed) music signals MUSDB18 was used as a basis.

The main purpose of this thesis was to improve as much as possible the expensive and time-consuming process of manual listening tests. These are still the most popular way of assessing the audio quality, because of their high effectiveness for real music signals. The examined prototype model is not proposed to be a replacement for them – at this point, even the best known reference methods do not provide such accurate results. However, limiting the scope of manual tests cases by classifying the selected artefacts first, can speed up the process of manual testing and fixing issues in the examined audio processing chain.

To the best of the author's knowledge, there is no existing recommended method of automatic audio quality assessment for real music content without the comparison to the known reference signal, and the prototype model examined in this research is the first published usage of convolutional-recurrent neural networks for the task of automatic non-reference artefacts classification in real music signals.

Kamila Organiściak