Institut für Industrielle Informationstechnik
Hertzstraße 16, Geb. 06.35
76187 Karlsruhe
www.iiit.kit.edu

Bachelor's Thesis, Master's Thesis

Estimation of Optimal Transcription Settings from Audio Data

Motivation

Automatic music transcription systems allow users to generate sheet music or tablature from audio data. In practice, however, the quality of the transcription strongly depends on choosing suitable transcription parameters. In the Klangio Transcription Apps, a transcription wizard is used for this purpose, allowing various settings to be adapted to the specific audio material.

For many users, manually selecting suitable parameters can be time-consuming and requires experience with the system. The goal of this thesis is therefore to develop a method that automatically predicts appropriate transcription settings based on the characteristics of an uploaded audio signal. This should improve ease of use while simultaneously increasing the quality of the resulting transcriptions.

Task Description

As part of this thesis, a system will be developed that estimates suitable transcription parameters for the Klangio Transcription Apps from a given audio signal. First, the existing workflow of the transcription wizard will be analyzed to determine which parameters have the strongest influence on transcription quality.

Subsequently, a dataset will be created or prepared that links audio files with suitable or optimal parameter settings. Based on this data, a neural network will be developed, implemented, and trained to predict appropriate settings directly from audio data. Different audio representations and model architectures will be investigated and compared.

Finally, the developed system will be evaluated using suitable metrics. In addition to the accuracy of parameter estimation, particular attention will be given to determining to what extent the automatically suggested settings improve the quality of the resulting transcriptions and reduce the configuration effort for users.

Prerequisites

- Basic programming skills in Python
- Basic knowledge of Machine Learning or Deep Learning
- Interest in signal processing, audio analysis, or Music Information Retrieval
- Enjoyment of scientific work

Research Area

- Music Information Retrieval
- Signal Processing
- Deep Learning
- Automatic Music Transcription

Degree Program

- [x] Electrical Engineering and Information Tec
- [x] Computer Science
- [x] Music Informatics (HfM Karlsruhe)

Focus

- [x] Signal Analysis
- [x] Modeling
- [x] Implementation

Start

Immediately

Links

Staff Page
Klangio
Project Page

Contact

M. Sc. Sebastian Murgul
Alter Schlachthof 33, 76131 Karlsruhe
sebastian.murgul@klangio.com
Tel.: (0721) 276604 20

