Progress Report PHD

Condensed Version

Manuel Brandner

Supervisor: Alois Sontacchi
Advisor: Robert Höldrich
Assessor: Brian Katz

July 10, 2018

Analysis of Quality and Efficiency in Classical Singing with Objective Parameters

Overview

Although a lot of research interest over the last decades emphasized on investigating the classical singing voice and in the analysis of proficiency in classical singing, yet until now there exists no commonly used assistance system in higher musical education. One main reason is possibly the lack of acoustical descriptors for voice quality and efficiency in singing which can be displayed robust in real-time. Another question arises about the applicability of such descriptors for students and teachers. A possible problem could be, that scientific research software can be too impractical in the use for educational purpose. The main goal of the thesis is to investigate descriptive acoustical parameters known from literature (e.g. used in phoniatrics) which seem to be promising in the use for a real-time feedback system and help to assess and visualize the parameters quality and efficiency in an education friendly software environment.

To do so a closer look at parameters which can be used as global descriptors for all voice types (e.g. soprano, alto etc.) is necessary. Knowledge about voice production in classical singing seems indispensable for correct analysis. Especially, the optimization of the vocal tract configuration deserves a closer look, which is commonly known as formant tuning and the main reason for the possibility to sing for large audiences in an unamplified setting.

Current Work - Directivity in Classical Singing

As a starting point of the thesis the parameter directivity is examined. The question to which extent sound radiation from a singer's mouth changes in regard to different mouth openings is addressed. The influence of the mouth opening has been several times mentioned in literature but not yet investigated. The data of different mouth openings for the same vowels should furthermore allow to distinguish if and how strong influences in the analysis and interpretation of parameters of the singing voice (e.g. glottal source parameters) occur. To accomplish a detailed analysis on the horizontal and vertical plane a measurement setup (double circle microphone array, DCMA) with low setting-up time has been designed and built. Furthermore, a collection of directivity analysis tools have been programmed in Matlab to visualize directivity patterns for impulse repsonse, long-time-average spectrogram (LTAS) data off-line and multichannel

recordings in quasi real-time. The tools have been made available for the IEM and the scientific community via our institutes website (opendata.iem.at).

The first results already show that specific mouth postures (e.g broad or open-mouthed) are more influential than others in the focusing of sound towards the audience. Additional to the fact that a certain singing technique enforces higher levels in the frequency region of the singer's formant, certain mouth postures can further enhance the voice level in forward direction.

Future Work - Vowel Space and Glottal Analysis

After the topic sound radiation is tackled, an investigation of the glottal source parameters, as described in [1] and [2], and its applicability for the classical singing voice and real-time analysis (currently available in the COVAREP repository [1] and restricted to an upper frequency limit of 500 Hz and off-line analysis) is planned. In [1] it is stated that the implementation is more robust and outperforms other current available algorithms.

Furthermore, vowel space analysis for analyzing intelligibility in classical singing will be investigated. The review in [3] discusses the effects on the vowel distribution in singing in comparison to speech and defines the diminished vowel space in classical singing. The change from speech to singing shows a reduction of the span of the vowel space. A parameter analysis of the vowel space can be accomplished by using the definitions stated in [4, 5]. The hypothesis is that a larger span of the vowel space area can be described by a few parameters and is correlated to intelligibility.

An evaluation of such features calculated via customized, self-programmed VST plugins shall be undertaken and used in an educational context. This seems to be promising, because free or inexpensive available digital audio workstations (e.g. Reaper) exist which allow an easy setup and use. Furthermore, recordings of the Franz Schubert and Modern Music competition with 54 participants in the category Lied, which did take place at the University of Music and Performing Arts in Graz in March 2018, are available for scientific analysis.

[1] COVAREP - A Cooperative Voice Analysis Repository for Speech.

Bibliography

[1] Thomas Drugman et al. "Glottal source processing: From analysis to applications". In: *Computer Speech and Language* 28.5 (2014), pp. 1117–1138.

[2] Rubén Fraile and Juan Ignacio Godino-Llorente. "Cepstral peak prominence: A comprehensive analysis". In: *Biomedical Signal Processing and Control* 14.1 (2014), pp. 42–54.

[3] Wencke Ophaug. "The Diminished Vowel Space in Classical Singing and the Tug of War between "Speech-true" and Modified Vowel Qualities". In: *Journal of Singing* 73.3 (2017), p. 293.

[4] Visar Berisha et al. "Characterizing the distribution of the quadrilateral vowel space area". In: *The Journal of the Acoustical Society of America* 135.1 (2014), pp. 421–427.

[5] Stefan Scherer et al. "Reduced Vowel Space is a Robust Indicator of Psychological Distress: A Cross-Corpus Analysis". In: *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2015), pp. 4789–4793.