Applicant: Athanasios Lykartsis, Audio Communication Group, TU Berlin

Einsteinufer 17, 10587 Berlin, Germany. +493031475772, athanasios.lykartsis@tu-berlin.de

Using Music Information Retrieval Methods for Speech Rhythm Analysis – Results, Challenges and Perspectives

Abstract

The creation of an automatic language identification (LID) system based on speech rhythm characteristics has been shown to be a promising challenge. Speech rhythm is difficult to quantify precisely, as speech lacks the comparably rigid metrical framework of music (Patel, 2008). Therefore, a research was initiated which sought to answer the following questions: Which acoustic features of the speech signal can best be used to quantify speech rhythm for LID? And can the gap between high-level speech rhythm perception and low-level rhythmic features to be extracted from the speech be bridged? To resolve these issues, we conducted experiments extracting relevant features given by acoustic correlates and derived from duration, loudness, fundamental frequency and spectral qualities (Lykartsis & Weinzierl, 2015). Having achieved some satisfactory results, we are interested in the challenges and perspectives which lie ahead: Can the features be adapted using results from speech perception and cognition, so that LID performance can become even higher? How can we take advantage of the differences and similarities between musical and speech rhythm in order to produce better features for the quantification of speech rhythm? We plan on performing listening tests to determine the relevant dimensions of speech rhythm and applying more current MIR methods for LID. The talk will give an overview of the planned experiments, further future work and focus on finding a common ground in the cognition of speech and musical rhythm.

References (in alphabetical order)

Lykartsis, Athanasios, and S. Weinzierl. "Using the Beat Histogram for Rhythm Description and Language Identification," *INTESPEECH 2015,* Dresden, September 2015.

Patel, Anirrudh D. Rhythm. In *Music, Language and the Brain*, Oxford University Press, New York, 2008.