Identification of systematic differences in acoustic properties in categories of laughter in "MULTISIMO" dataset

Mingwei Shi, Master in Computer Science

University of Dublin, Trinity College, 2024

Supervisor: Dr.Carl Vogel

Research into laughter classification is a compelling field that captivates scientists and sociologists seeking to unravel the enigmatic nature of this social signal. This paralinguistic cue possesses a notably intricate acoustic structure. Unveiling its discriminating properties could shed light on the internal acoustic structure of laughter. Previous studies have undertaken experiments to identify these discriminating acoustic properties, presenting a comprehensive pipeline that spans machine learning selection, identification of discriminating properties, and exploring factors influencing them. However, previous research has not released its dataset publicly, and some procedures require enhancement. To construct a more rigorous pipeline and comprehensively analyse discriminating acoustic properties, we compiled our dataset tailored to our research objectives from the "MULTISIMO" raw corpus(Multimodal and Multiparty Social Interactions Modelling), followed by identifying discriminating properties in mirthful and discourse laughter within our constructed dataset, performing regression analysis on the datasets to identify significant features that could explain discourse and mirthful laughter ,and exploring factors influencing discriminative acoustic properties.

The main findings in our work highlight that through a synthesis of the results from the machine learning experiments and regression analysis, we identified five shared discriminating acoustic properties across both experiments and laughter types: fundamental frequency, mel-frequency cepstral coefficient, auditory spectrum, spectral features, and jitter. The first four properties gauge energy-related information in acoustic laughter, while the last describes temporal characteristics. Our findings exhibit both concurrence and disparity with the findings from Tanaka and Campbell(2014),our replicated work, attributable to differences in the acoustic feature set quantity and the total number of utterance instances. Notably, fundamental frequency and spectral features emerge as common discriminating properties in both studies.

This work makes significant contributions both in theory and practice. Theoretically, this research has established a comprehensive pipeline encompassing dataset construction, verification, machine learning design and implementation, identification of acoustic properties, and examination of factors that may influence discriminating properties. This pipeline presents a novel avenue for researchers in audio processing and artificial intelligence. In terms of practical applications, although the study emphasises lies in theory, the developed algorithm shows promise for integration into real-time video systems to assist in laughter classification. It enables dynamic tracking of specific acoustic properties unique to instances of laughter.

Key phrases: Human laughter, Natural language dialogue, Social signals