



Invitation to PhD Defense

by Amir Hossein Poorjam



Partly online session due to COVID-19 - sign up and questions: Please sign up via doodle to participate at the defense ([doodle link](#)) and/or at the reception in the lunch room ([doodle link](#)). It will also be possible to follow the defense via ZOOM. If you wish to participate in the defense via ZOOM, please send an email to secretary Kristina Wagner Røjen (kwro@create.aau.dk), and she will invite you to the session. Please also contact Kristina with any further questions regarding the defense.

Place

Seminar Room No. 4.517 & 4.521
Rendsburggade 14, Aalborg University
(and ZOOM)

Date

August 16, 2021

Program

11:00 – 11:05

Welcome by moderator Kasper Rodil

11:05 – 11:50

Presentation by
Amir Hossein Poorjam

11:50 – 12:30

Lunch & Coffee break

12:30 – 14:30 (latest)

Questions

14:30 – 15:00

Assessment

15:30

Reception and Announcement from
the committee

Title

**Automatic Quality Control in Speech Data Collection
with Application to Remote Pathological Voice Analysis**

Assessment committee

Professor Elmar Nöth

Fredrich-Alexander-Universität, Erlangen–Nürnberg, Germany

Professor Junichi Yamagishi

National Institute of Informatics, Japan

Associate Professor Kasper Rodil

Department of Architecture, Design and Media Technology,
Aalborg University, Denmark

Supervisors

Professor Mads Græsbøll Christensen

Department of Architecture, Design and Media Technology,
Aalborg University

Associate Professor Jesper Rindom Jensen

Department of Architecture, Design and Media Technology,
Aalborg University

Host

Department of Architecture, Design and Media Technology
Aalborg University

Abstract

In this thesis, we develop accurate algorithms and methods to automatically control the quality of speech recordings in a data set by 1) recognizing the signals or segments of the signals that are degraded or that do not comply with the context of the data set, 2) identifying the type of degradation, and 3) informing the choice of appropriate enhancement algorithms. In the first step, by analyzing the distribution of mel-frequency cepstral coefficients (MFCCs), we demonstrate that degradation in speech signals foreseeably alters the mean and the covariance of the MFCCs, and the amount of this modification is correlated with the level of degradation. This attribute of MFCCs makes them a good candidate for identifying different quality patterns in speech signals. Next, shifting the focus to pathological voices, we propose a variety of algorithms to identify the presence and the type of short-term and long-term anomalies and degradations in speech signals. In the third step, to address the importance of quality control in remote speech-based applications, we analyze the performance of the pathological voice detection system under mismatched acoustic conditions, and we show how integrating the proposed quality control algorithms with appropriate enhancement techniques can effectively compensate for this mismatch and improve the recognition accuracy. Finally, we demonstrate the effectiveness of the MFCC features in other quality-control-related applications by proposing a novel speech enhancement algorithm based on detecting clean speech components of a noisy observation in the MFCC domain.

This study is a step toward the development of robust speech-based applications in general, and remote pathological voice analysis systems in particular, that are capable of operating in a variety of acoustic environments. Moreover, the applications of the methods presented in this thesis extend well beyond speech-based systems. They can be used for controlling the quality of other sensor modalities in different applications such as telemedicine or the monitoring of engineering systems.