Master thesis/Research Project

Title: Analysis of voice signals – a signal processing task

Background

Extensive voice usage in the work environment leading to injury of the voice is more and more being considered as an occupational safety and health issue. Earlier studies have shown that preschool teachers is one occupational group where self-reported voice problems are common and several papers has presented research on voice changes in teachers voices. There is a need to develop algorithms that can be used in order to assess voice changes and voice problems.

Research field

Assessing voice by objective measurements as well as research on the actual measurement methods are active research fields. The most well known and used measure when assessing voice is probably average pitch and average sound level followed by jitter, shimmer and harmonic-to-noise ratio. The parameters of average pitch and average sound pressure level have been analyzed in several earlier studies. Jitter and shimmer is the variability of pitch and amplitude respectively. Harmonic-to-noise ratio is the ratio between the harmonic and the noise components in a speech signal .

Thesis

The task of the thesis is to see if conventional speech processing algorithms from e.g. speech recognition or speech coding can be used to detect changes in a voice due to extensive usage of the voice.

The thesis will provide a challenging and inspiring work for an ambitious student with good programming and mathematical skills and an interest for embedded systems.

In this thesis you will implement conventional speech processing algorithms, such as mel-frequency based extraction of feature vectors, which take a block of speech samples as input and give a vector of speech parameters as output. You will use your implemented algorithms to process a data set consisting of recorded voice signals (recordings of teacher voices in a preschool environment). Then you will analyse if there is a change in the speech parameters average value due to the usage of the voice, i.e. if extensive usage leads to e.g. an increase of the average value. In an extension of the thesis, i.e. if there is time, you will use a data set consisting of healthy and injured voices and analyse if the implemented algorithm can distinguished between the healthy and injured voices.

Activities

To meet the above objectives, the thesis is designed as a real industrial development project. The task of the thesis is to design, implement, and evaluate functional algorithms and to present the results in form of a verbal presentation as well as a written report. A standardized model for the master thesis project will be used. In this model the thesis is divided into five parts. A short description of these parts is given below.

Project plan

The student creates a project plan for the project including a detailed time plan.

Information gathering
The student performs a smaller research of literature. This will be the base of the theory part of the report. The students will acquire knowledge of speech parameter estimation techniques and their use in different applications.

Specification
The student specifies the functionality of the algorithms that should be developed.

Evaluation
The specification is first implemented in Matlab, this is done in order to more quickly and easily evaluate the performance of the specified system and reach a functional solution.

Implementation
A real-time implementation is then performed on a Analog Devices Blackfin 522 DSP-card using C.

Report Writing
The theory of pitch shifting, the implemented system, and the results DSP implementation are presented

General

This thesis is conducted within a research collaboration between Umeå University, Dept. of Applied physics and Sahlgrenska Akademin at Göteborg University, Dept of Occupational and Environmental Health, Principal investigators are Fredric Lindstrom, Ph.D., Haibo Li, Ph. D. and Kerstin Person Waye Ph. D.

Contact

Fredric Lindström, Ph. D.

076-1176768

fredric.lindstrom@amm.gu.se