Neural Echo Cancellation

Implementing echo and noise cancellation using a neural network on an embedded system

Students

Selvin Blöchlinger

Robin Eberle

Introduction: The target of this thesis was the development and implementation of a neural network capable of acoustic echo cancellation on a microcontroller. Noise and echo cancellation are vital for passenger announcement systems in trains and other public transportation, which experience a lot of noise and strong acoustic feedback over the speakers. The project was approached from two sides: On one side, capable neural network architectures were identified and corresponding models were trained. On the other hand, the implementation of a neural network on a microcontroller was analyzed and implemented.

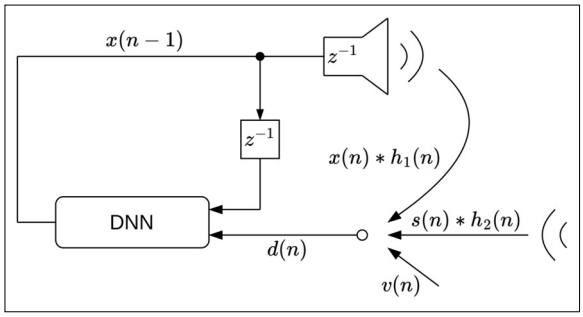
Approach: After extensive research on AI models capable of acoustic echo cancellation, four promising models from the ICASSP acoustic echo cancellation challenge were selected for further investigation. The models were trained on a custom data pipeline tailored to the problem at hand, and then evaluated using a wide range of performance metrics, which were themselves examined for their meaningfulness.

An investigation into methods and tools for implementing neural networks on a microcontroller was conducted and a possible approach developed. This was then tested by benchmarking the performance of different neural network architectures and layer types. For model inference on real-time audio data, an embedded software application was developed to capture audio from a microphone, process it through a neural network and then output it to either a PC or a speaker.

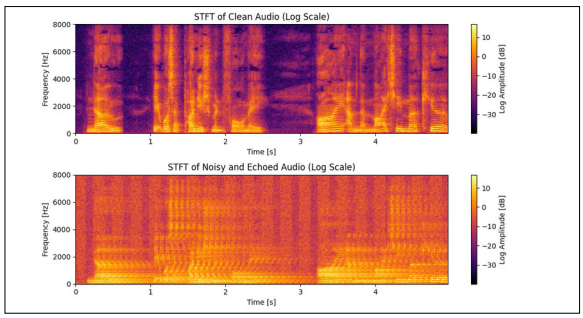
Result: The work on this thesis revealed several problems. The models examined, although showing solid improvement while training and reasonable scores on test data, produce disappointing results when listening to their audio output. The model performance analysis on the other hand showed that the state-of-the-art architectures are far from being small enough to be implemented on a microcontroller with a Cortex-M7 processor.

The obtained results suggest that acoustic echo cancellation using a neural network implemented on low-cost hardware is not very promising. It would therefore be necessary to either invest a lot more into finding an efficient model architecture capable of attaining the desired results, or to use a different approach altogether, like the more classical adaptive filters.

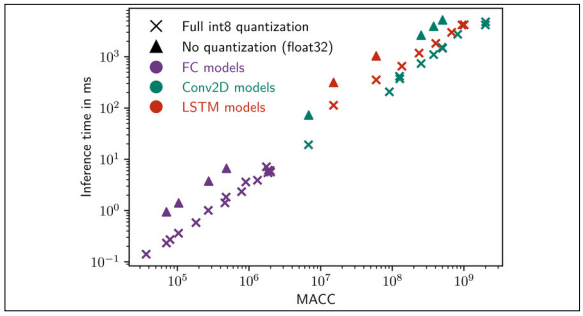
Diagram of the acoustic situation from a signal processing perspective.
Own presentment



The STFT of a clean (top) and corrupted (bottom) audio signal, visualized on a logarithmic scale.
Own presentment



Performance analysis of various neural networks with standard layers, executed on the STM32H753ZI microcontroller.
Own presentment



Advisor
Hannes Badertscher

Subject Area
Data Science, Electrical Engineering