Project Plan

Neural networks for speech separation

Qingzhuo Zhang

ID: 29830885

The written report will be submitted at the end of week 3. This should contain:

= a copy of the presented presentation

= a layman's summary of the project (1 page)

= a trainings needs analysis

= a gantt chart with Milestones

= a record of meetings with supervisor (with feedback)

A copy of the presented presentation

（will have）

**1 Introduction**

The speech separation is to separate monaural sounds which contain more than one voice. In this project, neural network and machine learning will be used as a tool to solve the problem. Basically, the task can be described as: to separate a monaural sound record, which mixed two (or more) people’s voices, into two (or more) independent voice records for different people.

**2 Background and Application**

The speech separation problem is originally beginning with “cocktail party problem”. When people in a cocktail party, it is easy for them to distinguish the voice from different people. Speakers are effortless to focus on the voice of the people who they are talking and ignore other people’s voices, background music, and noise. However, for computer, this is very difficult to achieve so. Especially, the monaural sound is harder because of no position information. Before the popularity of machine learning, this problem has been tried by many methods but all disappointed. While neural network seems very good at dealing with those problems with help of deep learning and training, the performance of the separated results still lower than people expected. There still have a large distance to achieve the goal that separating mixed speeches clear enough to be recognition by machine.

Speech separations have a wide range of applications, including denoising, mobile telecommunication, signal processing, speech recognition and so on. The performance of speech separation is crucial for the continuing work of the signals or voices. With more and more accurately the machine can understand human languages, speech separation becomes an indispensable skill of artificial intelligence.

**3 Methodology**

Machine learning has showed a prosperous scenario by solving many problems which are easy for human but very difficult for machine, such as image identification, speech recognition, etc. Speech separation is one of those problems as well.

In this project, deep neural network(DNN), convolutional neural network(CNN) and recurrent neural network(RNN) will be used to solve this problem separately. Then the results by each neural network will be compared. It is also possible to combine different neural networks to achieve better performances.

The basic idea is turn the voice signal into spectrogram and separate the spectrogram then turn back to voice signals. The separated results will be checked with the original voice. So this is a supervised learning model.

**4 Plan**

The plan

This project plan is a kind of prediction of what will happen. It is quite hard to plan this project for the reason that the coding work more or less unpredictable. If the task has been settled before did not finish in time, then the work after would be impossible to continue