Name:Parvez Rahman

IIIT Guwahati,Assam

Musical Notes Identification

**Background**

A soundtrack involves a mixture of various entities such as various musical instruments like piano, guitar, drum, etc. Or human voices. The most important feature of a musical soundtrack is the notes present in it. In many applications, it is important to know where the notes in a sound-track begin and end. It finds application in musical education, composing , recording, etc.

**Problem Statement**

This is a mini-project which I performed in my 1st semester of MTech. Here, I have taken 3 different kinds of soundtracks. There is an operation which involves concatenating them together to produce a composite sound-track. The task is to identify the notes in the composite soundtrack by performing various signal processing operations.

**Data**

I have used 3 different sound clips which are in wav format-

cartoon001.wav- 50 samples present (<https://github.com/Parvez318/Musical-Notes-Identification/blob/master/cartoon001.wav>)

cartoon049.wav- 46 samples present (<https://github.com/Parvez318/Musical-Notes-Identification/blob/master/cartoon049.wav>)

cartoon058.wav- 28 samples present (<https://github.com/Parvez318/Musical-Notes-Identification/blob/master/cartoon058.wav>)

**Approach**

* I started by reading each of the soundtracks onto three different objects using the **‘audioread’** function so that they are stored permanently
* I concatenated them to produce a composite track stored in ‘y’
* I performed **thresholding** operation using the **‘wthresh’** function which makes the samples in the soundtrack more separated and identifiable
* Then, I applied **averaging** using a self-defined averaging function
* The final form that I got was such that we can easily count from the plot, the number of samples present in the soundtrack.

**Solution**

Link to code on github----https://github.com/Parvez318/Musical-Notes-Identification/blob/master/Musical\_notes\_iden.m

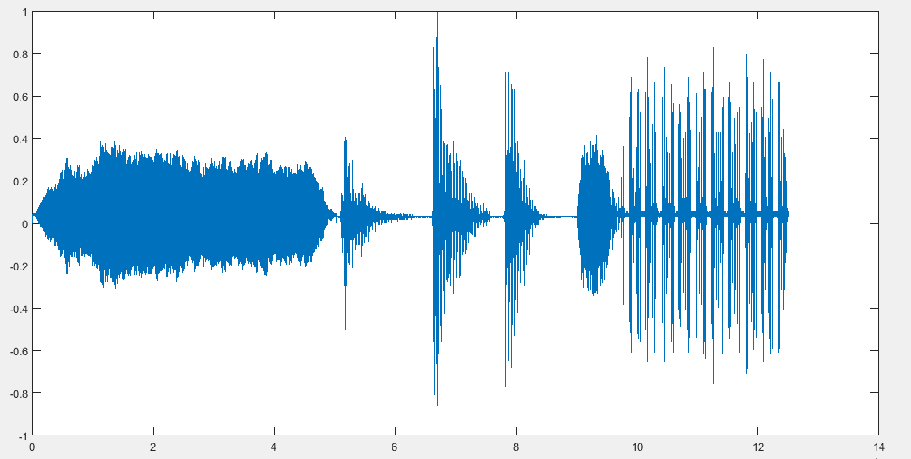
****

Figure 1:The composite signal ‘y’

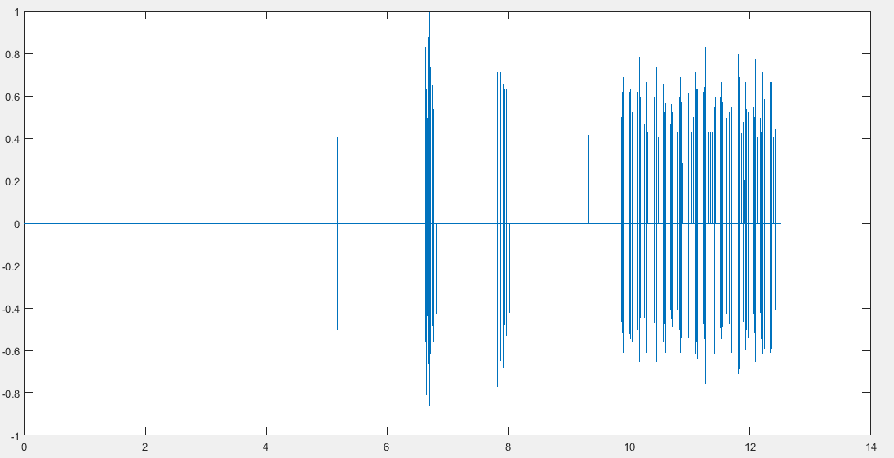


Figure 2:The signal after thresholding

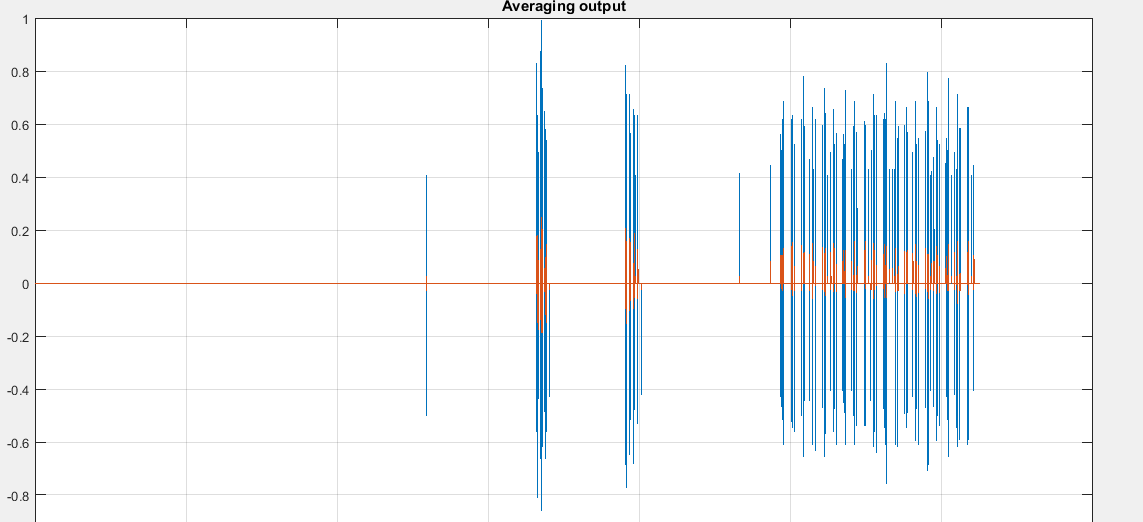


Figure 3:The final output form after averaging

**Result**

In this experiment, I have used thresholding and averaging operations to optimise the soundtrack so that we can easily count the notes, which are the spikes that stand out clearly in the final form. The task which I could not complete is to identify the time of occurrence of each note separately. This can be carried forward using Wavelet Decomposition techniques and Time Frequency analysis, which I will try in near future.