REVIEW ON AUDIO DENOISING TECHNIQUES

Kodathala Sai Varun
GITAM School of Technology

Email – kodathalasaivarun@gmail.com

Preliminary Analysis:

The audio is preliminarily a signal, since it's a signal any operations on audio is considered to be signal processing. In this report I evaluate the usual signal processing techniques to denoise the noisy audio signal. In the modern electronics, audio is considered to be of two types: stereo audio and mono audio. My analysis starts with configuring the stereo audio to mono audio and mono audio to be as usual for making analysis easier.

The signal/ audio can be characterized as noisy by observing its signal nature either in time domain or frequency domain. The nature can be depicted from the figure below:

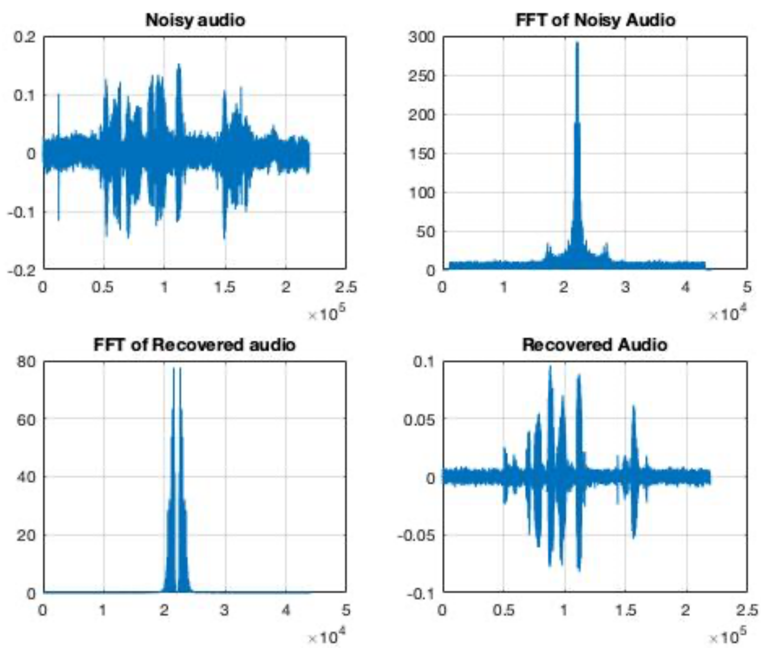


Figure: Signal Analysis

Denoising Techniques:

My approach to the problem is considering the signal as analog signal, when I am working with the analog electronics in my academics, I have learnt the basic nature of signals and its characteristics. From its spatial features I could extract the inference to remove high frequency components or low frequency component or both. In general, a signal is composed of frequency components, frequency is usually the reciprocal of time period. One can characterize those components as high frequency or low frequency based upon cut-off. If I have a signal which is a mixture of 2 frequency components (5Hz and 20 Hz) and the cutoff is 10 Hz then the frequencies below 10 Hz are considered to be low frequency component and frequencies above 20 Hz are considered to be high frequency component. So, accordingly I can filter out either the lower frequency signal, higher frequency signal or both.

This approach is very much similar to problem statement where I need to denoise the signal/ audio. An audio signal composes of both noise signal and informative signal, usually the magnitude of informative signal is more than the noise signal. So, in order to retrieve the informative signal one can, approach two methods either time domain or frequency domain. In time domain one can either do median filtering or thresholding. And in Frequency domain one can approach for filtering techniques.

Time Domain Processing:

1. Median Filtering:

   Median Filtering is usually used in Image Processing to extract the edges and important features of an Image. But in DIP we use 2-D filter, since audio signal is 1-D here I used 1-D filter. The filtering process is follows:

   If $x = (2, 3, 80, 6)$.

   So, the median filtered output signal y will be:

   $y_1 = \text{med}(2, 3, 80) = 3$,
   $y_2 = \text{med}(3, 80, 6) = \text{med}(3, 6, 80) = 6$,
   $y_3 = \text{med}(80, 6, 2) = \text{med}(2, 6, 80) = 6$,
   $y_4 = \text{med}(6, 2, 3) = \text{med}(2, 3, 6) = 3$,

   i.e. $y = (3, 6, 6, 3)$.

   The same approach is applied for whole audio signal and result is as follows:

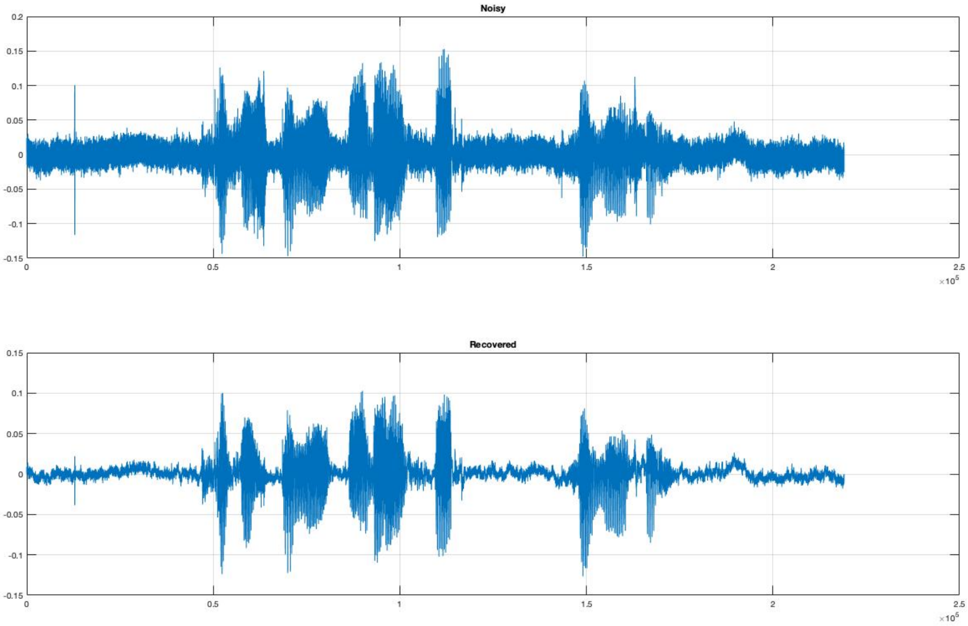


Figure 2: Median Filtering Output

2. Thresholding:

This is formal methodology which lacks the accuracy but looks simple and robust to apply for any kind of signal. From the time analysis of noisy signal it is evident that the informative signal amplitude is higher than magnitude of 0.05 and in negative axis lesser than magnitude of -0.05. So, we need to determine the components in the range of (-0.05,0.05) and make them to zero.

The results are given as follows:

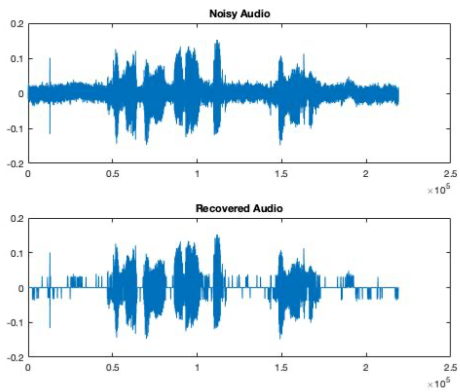


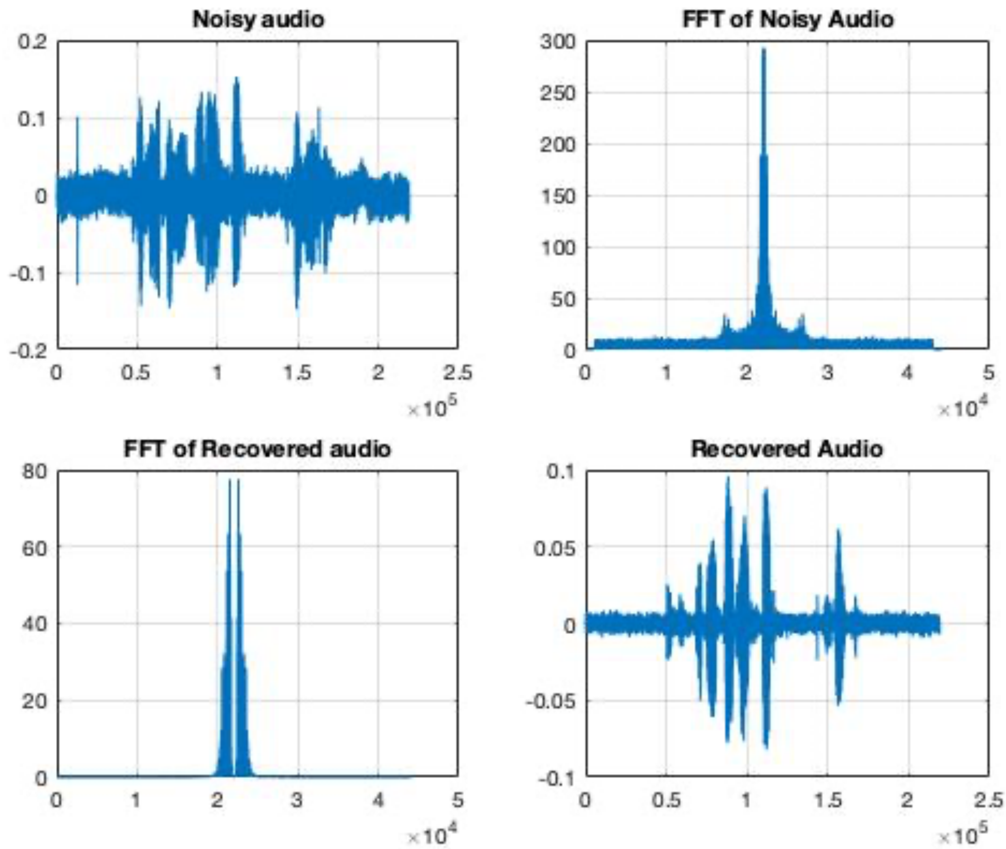
Figure 3: Thresholding Results

Frequency Domain Processing:

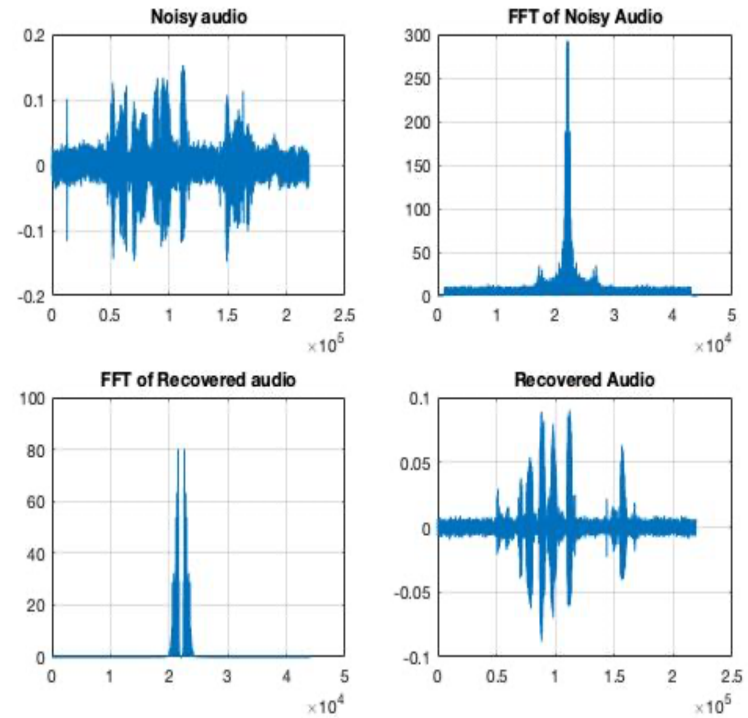
In frequency domain analysis, the signal's frequency components are visualized first. It could be visualized using discrete Fourier transform by applying Fast Fourier Transform. After converting to frequency domain, we have a visualization of signal's magnitude according to frequency, in this model example the amplitude is high is the range between 1 KHz to 3 KHz. So, after the visualization of frequency component one has to retrieve the signal's frequency components between this range. This can be done using digital filters like Butterworth, and there are several filters in practice. For this model one could use a Bandpass filter or combination of both low pass and high pass filters.

And in filters, they can be cascaded for better response and it refers to order of filters. In this model I utilized three different orders i.e. 4, 5 and 6. The analysis is visualized as below:

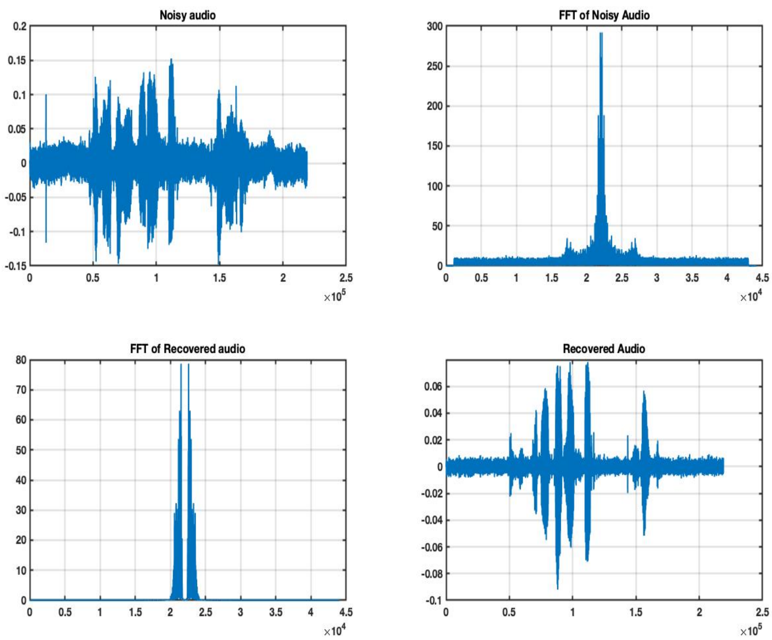
Butterworth – Bandpass Filter with order 4:



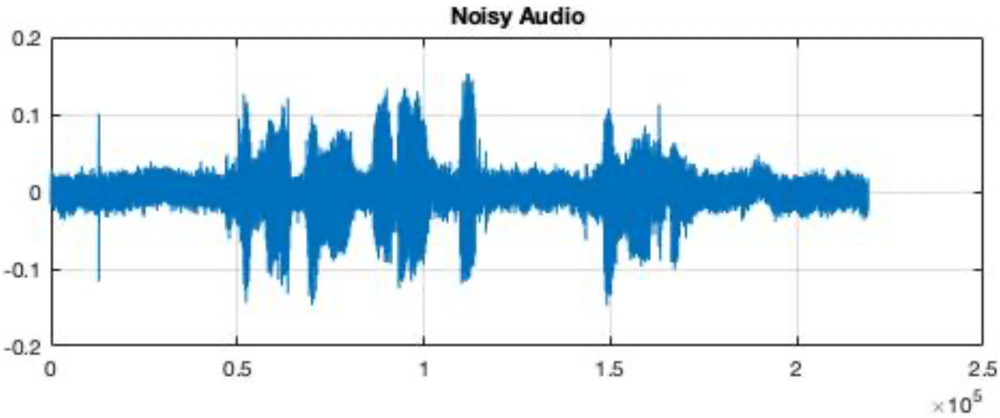
Butterworth – Bandpass Filter Order 5:

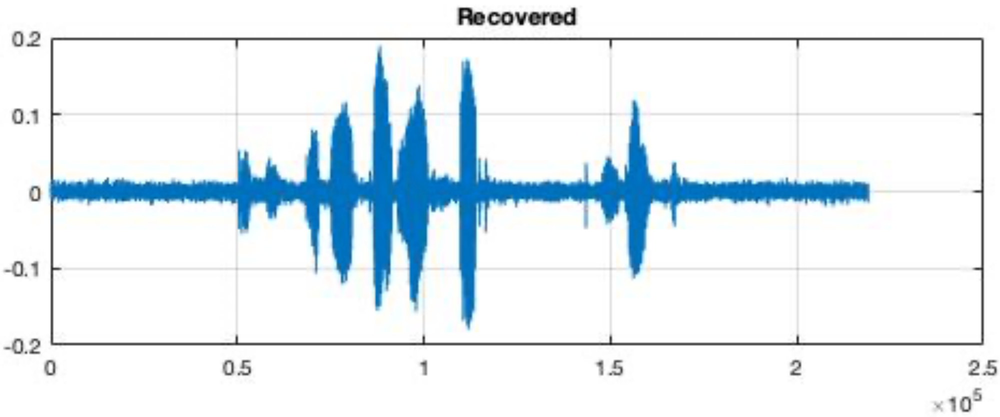


Butterworth Bandpass filter order 6:



Combination of Low Pass and High Pass:





Summary:

The Time domain analysis is completely impractical since its accuracy is very low. One should approach the problem in frequency domain and visualize the frequency components properly. With those frequency components one can filter out noise from informative signal. The analysis in determining the order of filters is that the recovered signal amplitude is reducing if order is too low and too high so, one should choose the order wisely.

The visualizations, programming is done in MATLAB and the same can be inferred using any programming language since the background of the problem is mathematics.

The models, visualizations and results can be collected from:
https://github.com/varunkodathala/audio-denoising