Audio Enhancement with Fourier Transform

I213E 2022 - NGUYEN Dinh Mau

I213E 2022 - NGUYEN Dinh Mau
s2210403 - maudn@jaist.ac.jp

Outlines

1. Introduction - Problem
2. Objectives
3. Background
4. Methodology
5. Results
6. Discussion
7. References
8. Appendix

1. Introduction

Audio enhancement is the process of improving the quality of audio signals to make them more listenable, clearer, and more suitable for a particular application. One of the fundamental techniques in audio processing is the Fourier transform, which provides a mathematical representation of audio signals in the frequency domain.

NHK World is the international arm of Japan's public broadcaster, the Japan Broadcasting Corporation (NHK). It provides news and information to an international audience, including a 24-hour English-language news channel, as well as content in Japanese, Spanish, and Arabic. The content covers a wide range of topics including international news, business, culture, and technology. The services are available via cable TV, satellite TV, and online streaming.

1. Introduction - Problem

JAIST is in mountainous are, therefore radio signal is sometimes distorted

Audio collected from students in JAIST who are curious of how Japanese people react to the international events

I213E 2022 - NGUYEN Dinh Mau
s2210403 - maudn@jaist.ac.jp

*Pictures shown are for illustration purpose only

2. Objectives

- To study and analyze the various audio enhancement techniques and the Fourier Transform method.
- To design and implement a software tool that uses the Fourier Transform to enhance audio signals.
- To compare the performance of the Fourier Transform-based enhancement with existing audio enhancement techniques.
- To evaluate the effectiveness of the Fourier Transform-based enhancement in terms of clarity, noise reduction, and overall audio quality.

3. Background

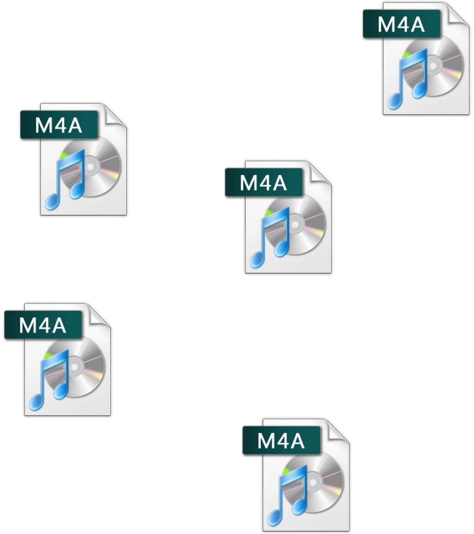
- **Audio signals**: Time & frequency domain representation
- **Fourier transform**: Convert time-domain signals to frequency-domain
- **Audio enhancement**: Improve audio quality (clarity, loudness, noise reduction)
- **Fourier transform in audio enhancement**: Analyze and modify signals in frequency domain.
- **Bandpass filter**: A bandpass filter is a type of electronic filter that passes signals within a specific frequency range and attenuates (reduces) signals outside that range.

I213E 2022 - NGUYEN Dinh Mau
s2210403 - maudn@jaist.ac.jp

4. Methodology

1. Obtain audio signals
2. Convert signals to frequency domain using Fourier transform
3. Analyze frequency components of the signals
4. Apply bandpass filter in frequency domain
5. Convert back to time domain using inverse Fourier transform
6. Evaluate the performance of the enhancement system.

4.1 Obtain audio signals - m4a files



Audio collected from students in JAIST who are curious of how Japanese people react to the international events

I213E 2022 - NGUYEN Dinh Mau
s2210403 - maudn@jaist.ac.jp

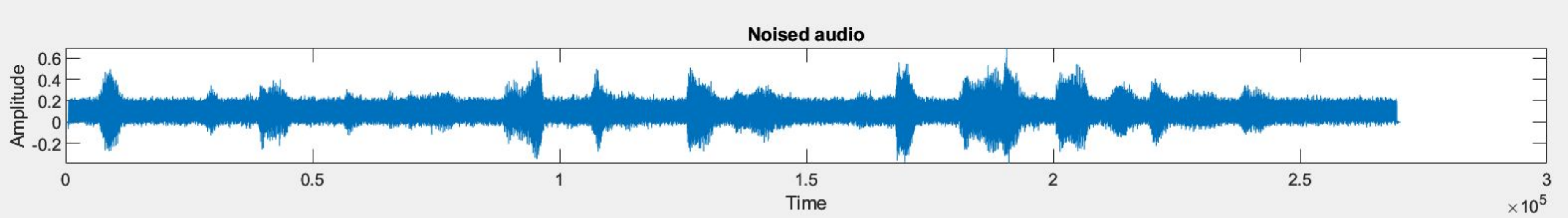
*Pictures shown are for illustration purpose only

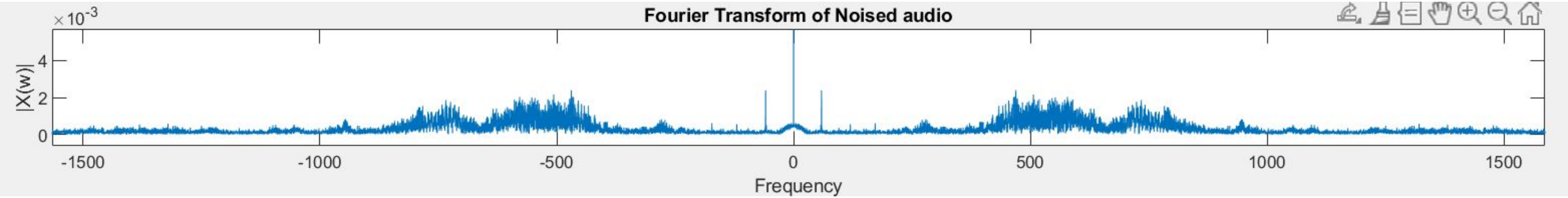
4.2 Convert signals to frequency domain using Fourier Transform

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi kn/N}$$

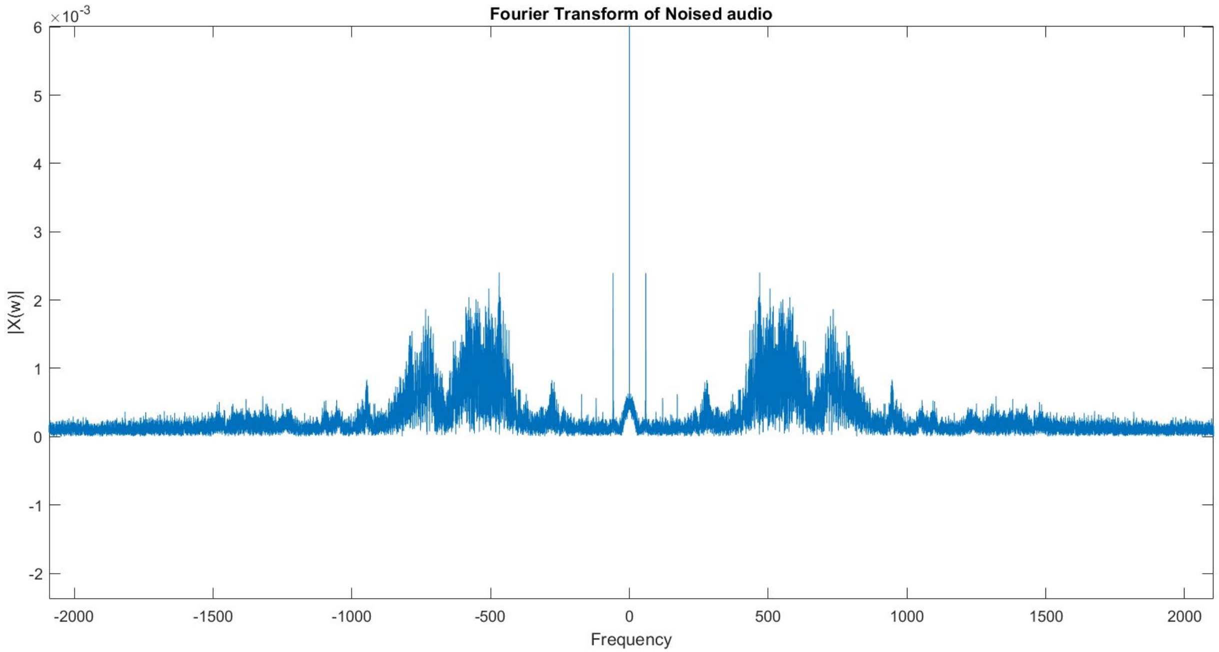
I213E 2022 - NGUYEN Dinh Mau
s2210403 - maudn@jaist.ac.jp

4.2 Convert signals to frequency domain using Fourier Transform





4.3 Analyze frequency components of the signals

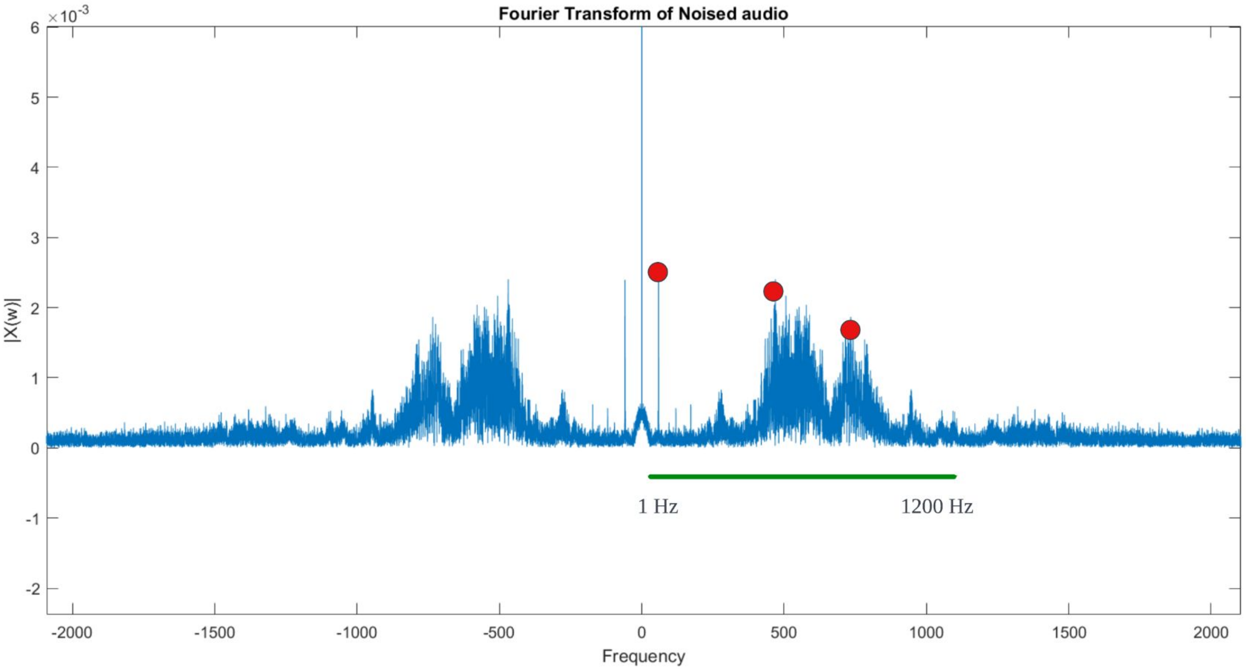


I213E 2022 - NGUYEN Dinh Mau
s2210403 - maudn@jaist.ac.jp

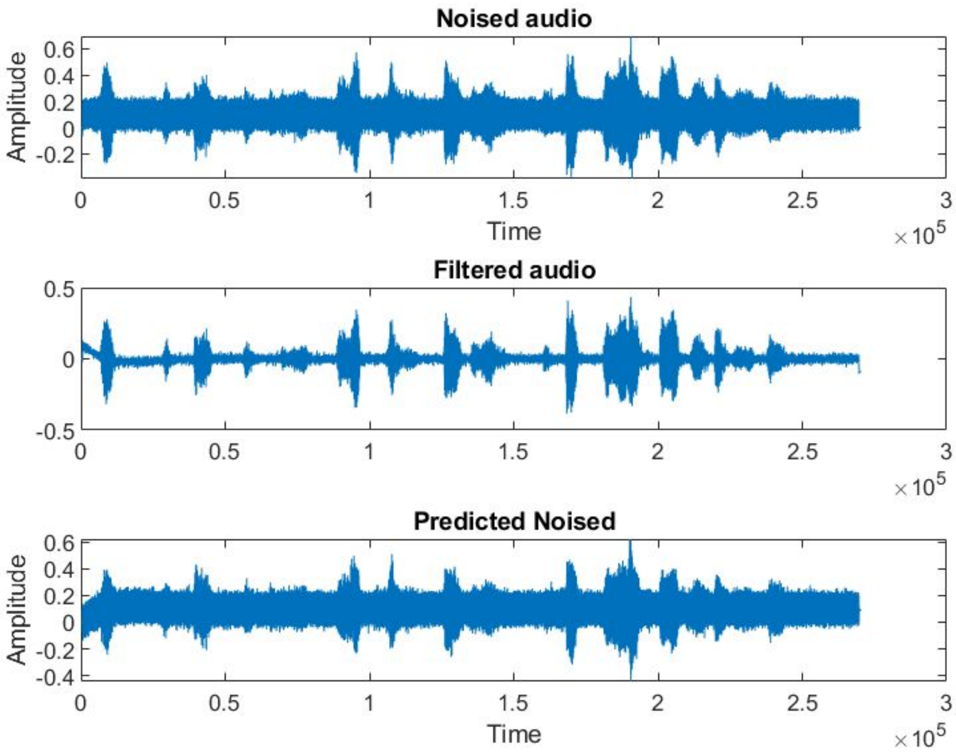
4.3 Analyze frequency components of the signals

During a conversation, the fundamental frequency of a typical adult man ranges from **80 to 180 Hz** and that of a typical adult woman from **165 to 255 Hz** [1]

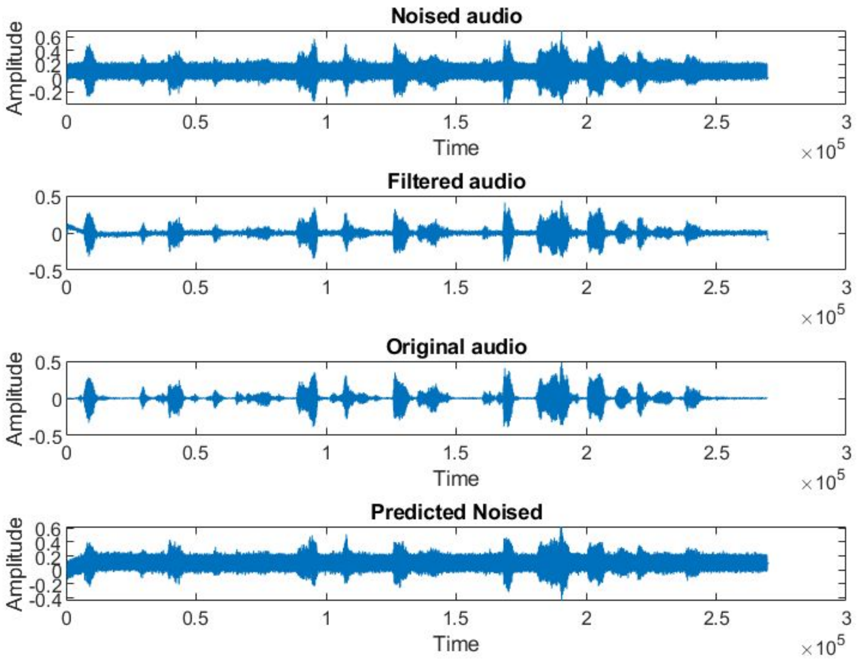
4.4 Apply Bandpass Filter (1 Hz - 1200 Hz)



4.5 Convert back to time domain (inverse Fourier Transform)



4.6 Comparison



I213E 2022 - NGUYEN Dinh Mau
s2210403 - maudn@jaist.ac.jp

5. Result

Signal-to-Noise ratio

- SNR: ratio between signal power and noise power
- Definition:

$$SNR = \frac{P_{signal}}{P_{noise}} = \left(\frac{A_{signal}}{A_{noise}}\right)^2$$

$$SNR(dB) = 10\,{}^{10}\log\left(\frac{P_{signal}}{P_{noise}}\right) = 20\,{}^{10}\log\left(\frac{A_{signal}}{A_{noise}}\right)$$

$$\text{MSE} = \overbrace{\boxed{\frac{1}{n}\sum_{i=1}^{n}}}^{\text{Mean}} \big(\overbrace{\boxed{Y_i - \hat{Y}_i}}^{\text{Error}}\big)^{\overbrace{\boxed{2}}^{\text{Squared}}}$$

I213E 2022 - NGUYEN Dinh Mau
s2210403 - maudn@jaist.ac.jp

5. Result

- Signal to Noise Ratio (SNR): measures the ratio of the power of the original clean signal to the power of the residual noise after processing.
- Mean Square Error (MSE): measures the average difference between the original clean signal and the processed signal.
- Students are happy :D

	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
SNR (dB)	-6.4812	-12.366	-10.2102	-10.031	-9.8805
MSE	0.0023	0.00068	0.0011	0.0012	0.001

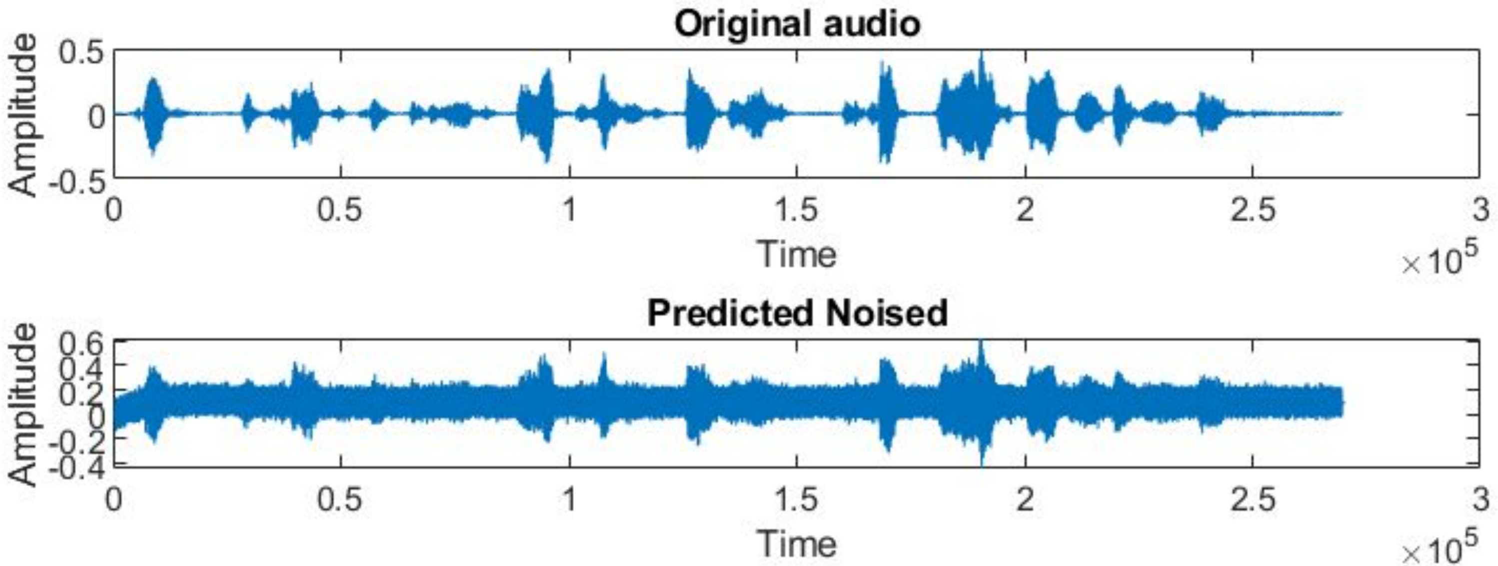
6. Discussion & Improvements

- The processed signal has less noise as checked intuitively
- Frequency range needs removed depends on specific signal
- Original audio patterns seem still say in the noise, there are rooms for improvement
- Other enhancement methods may be used to investigate more: Apply looping, more complicated filters

I213E 2022 - NGUYEN Dinh Mau
s2210403 - maudn@jaist.ac.jp

6. Discussion & Improvements

- Some of the original signal are still in noise, need some freq selection improvements



8. Appendix

- Source code: https://github.com/NgDMau/audio-enhancement-matlab
- Dataset: https://github.com/NgDMau/audio-enhancement-matlab/tree/main/Distorted%20audios

ขอบคุณ

谢谢

شکریہ

Thank you

Terima kasih

ありがとう

Cảm ơn

NGUYEN Dinh Mau
maudn@jaist.ac.jp