

### Problem Set #3, SP part

- Issue date: Nov. 20, 2022; Deadline: 23:59, Dec. 4, 2022
- use `[y, Fs]=audioread('xxx.wav')` to load signal;
- use `audiowrite('xxx.wav', y, Fs)` to save signal;
- Please use `figure`, `subplot(x, x, x)` and `title('')` to make your output beautiful.
- Please use matlab to output “. Pdf” and upload to Gradescope.
- **Good luck and have fun!**

Our sense of hearing provides us rich information about our environment with respect to the locations and characteristics of sound producing objects. For example, we can effortlessly assimilate the sounds of birds twittering outside the window and traffic moving in the distance while following the lyrics of a song over the radio sung with multi-instrument accompaniment. In many cases we need to record the audio, transmit and preprocess the audio signal in electronic way. For example, when we are in online learning, we need our voice to be heard by each other. Besides, we may also find it interesting to transfer the style of a recorded voice. All of those fall into the range of audio signal processing.

#### 1. Denoising and Filtering

In this section, you should try to generate a segment of noised speech signal with echo and restore the signal by filtering.

- a) The input files are given as “speech.wav” and “noise.wav”. You should use function **audioread** to read the time-domain signal and the sampling frequency  $F_s$ . You are required to write a .mlx file named “sp\_3\_solution.mlx” where you should perform the following steps **IN GIVEN ORDER**. (4 Points)
- b) Add noise to the speech signal. Plot the original speech signal and noised speech signal in time domain as a waveform and frequency domain as its amplitude spectrum respectively. (8 Points)

Hint: The x-axis of the time domain waveform should be in seconds (s), which means that you need to calculate the corresponding time of each sampling point on the timeline first. For the frequency domain, you should give the amplitude as the y-axis and concrete frequency (Hz) from 0 to  $F_s/2$  as the x-axis. You can use function `fft()` and `abs()` to get the N-point DFT of the signal.

- c) The sound will be reflected back when it encounters an object in the transmission, and the auditory effect is like the echo of your shout in the valley. It is not difficult to understand that it is caused by the mixture of the original signal and the reflected signal. Therefore, the model of echo mixing can be expressed as follows:

$$y(t) = x(t) + a * x(t - T_0)$$

$$y(n) = x(n) + a * x(n - N_0)$$

Try to add echo to the noised speech signal from b). The echo is added after 1s, and the attenuation coefficient of the echo is 0.2 ( $T_0 = 1, \alpha = 0.2, N_0 = T_0 * F_s$ ). Plot the noised speech signal with echo in time domain as a waveform and frequency domain as its amplitude spectrum respectively. (8 Points)

Hint: You can add zeros before the start of the signal to get the echo. At the same time, the length of the signal will change by adding echo.

- d) Remove the echo by inverse filtering. The Z-transform converts a discrete-time signal into a complex frequency-domain.

$$y(n) = x(n) + \alpha * x(n - N_0)$$

$$y(n) \Leftrightarrow Y(z) \quad x(n) \Leftrightarrow X(z)$$

$$Y(z) = (1 + \alpha * z^{-N_0})X(z) = H(z) * X(z)$$

$$H(z) = 1 + \alpha * z^{-N_0}$$

Therefore, echo mixing can be regarded as the generation of signal  $x(n)$  after passing through  $H(z)$ . Then removing the echo can be realized by inverse filtering, and the inverse system of  $H(z)$  is

$$H^{-1}(z) = \frac{1}{1 + \alpha * z^{-N_0}}$$

You can use function **filter** to implement inverse filtering as follow:

`b = [1]; a = [1, zeros(1, N0 - 1), α];`

`x = filter(b, a, y);`

Try to implement inverse filtering to the noised speech signal with echo from c) to remove the echo. Plot the signal before and after the inverse filtering in time domain as a waveform respectively. (8 Points)

Hint: You can also use function **sound** to play the speech signal to verify whether the elimination of echo is successful

- e) Design a FIR filter use `fdatool` command and filter the noised speech signal without echo from d) with function `filter` to remove the noise. You are required to declare the filter type you used and briefly explain the design parameters you choose. You are then required to plot: (1) The magnitude response of the filter in frequency domain; (2) The noised speech signal without echo from d) and the filtered signal in time domain as a waveform and frequency domain as its amplitude spectrum respectively. (12 Points)

Hint: You can also use function `sound` to play the speech signal to verify whether the elimination of noise is successful

## 2. Audio Signal Processing

In this section, you should try to realize several modulation methods for a given signal.

- a) The input file is given as “handel.wav”. You should use function **audioread** to read the time-domain signal and the sampling frequency  $F_s$ . You are required to write a .mlx file named “sp\_3\_solution.mlx” where you should perform the following steps **IN GIVEN ORDER**. (2 Points)
- b) Plot the signal in time domain as a waveform and frequency domain as its amplitude spectrum. The x-axis of the time domain waveform should be in seconds (s). For the frequency domain, you should give the amplitude as the y-axis and concrete frequency (Hz) from 0 to  $F_s/2$  as the x-axis. You can first use function **fft()** and **abs()** to get the N-point DFT of the signal. (6 Points)
- c) Make a simple equalizer. The equalizer allocates different weights various different frequency bandwidth. This can be realized by a couple of band-pass filters. In this case, eight filters are adopted for 32, 64, 125, 250, 500, 1k, 2k and 4kHz. Design the filters as you prefer and splice the signals filtered by your filters. Before filtering, consider the following questions. What is the filter type you choose? Is it time-domain or frequency domain? Is it IIR or FIR filter? What is the reason that you choose it? After that, Plot the spectrum of the difference of the original signal and filtered signal. (The difference should be minimized as far as possible, which is your task.) (24 Points)
- d) Assign weights for each filter to make the voice of human clearer. The values of weights can range from -12dB to 12dB. Plot the spectrum of the filtered signals again and save it through function **audiowrite**. (Hint: dB is a parameter describing gain  $G(\omega) = 20 \log_{10}(|H(e^{j\omega})|)$  . (18 Points)
- e) Resample the signal and apply the function **sound(y,fs)** with previous sampling frequency  $f_s$ . The inconformity of sampling frequency leads to the change of voice. Use function **resample** to try it. However, there may be two problems. When we upsample the audio, we need to drop out weak signals. When we downsample the audio, we need to padding to maintain the length of the audio. Plot the time-domain sequence and Fourier spectrum before and after the voice transferring. (10 Points)