ECE 340: Discrete Time Signals & Systems, Fall 2021

Lab. 4
Spectrum Estimation of Multimedia Signals

1. In this question, you will read an audio file using MATLAB and find information about the sampling and quantization parameters about the digital audio signal.

   (a) Download the music file `love_mono22.wav` from the Lab website and save it in your working directory. Play the audio signal using the system media player and listen to it using your headphone. Write down the first few words of the song.

   (b) Read the audio file in MATLAB using the `audioread` function. Represent the audio signal using the matrix `x`. Find the matrix size and the sampling rate. Having 8 bits for each sample, calculate the bit-rate (bits/sec) and the duration of the audio signal.

2. In this question, you will estimate the spectrum of the audio signal using the MATLAB `fft` function. Note that the `fft` function calculates the discrete Fourier transform (DFT) of an *N*-point signal $x[k]$, which is defined as follows.

$$X[r] = \sum_{k=0}^{N-1} x[k] e^{-j(2\pi kr/M)} \qquad\qquad 0 \le r \le N-1$$

   The DFT can be considered as an approximation of the discrete-time Fourier transform (DTFT) of $x[k]$. Note that a signal with N samples will have N DFT coefficients. The m-th coefficient $X[m]$ represents the strength of the *m*-th basis function present in the signal. If N is even, the first (N/2+1) DFT coefficients will represent the positive frequencies, and the last (N/2-1) coefficients will represent the negative frequencies. Suppose the discrete audio signal is obtained from a continuous-time audio signal with a sampling frequency F_s. The frequency (in Hz) corresponding to the *m*-th coefficient $X[m]$ is given by

$$f_m = \frac{F_s}{N} \times m, \qquad m = 0, 1, 2, ...(N/2)$$

   (a) Using the `fft` function in MATLAB, calculate the discrete Fourier transform of the audio signal (stored in matrix `x`) obtained in Q1(b).

   (b) Note down the values of $X[0]$, $X[1]$, and $X[2]$.

   (c) Scale the coefficients $X[r]$ as follows:

$$X'[r] = \frac{X[r]}{\sqrt{N}}$$

   (d) Plot the magnitude spectrum of $x[k]$ by plotting the absolute values of $X'[r]$. Scale the x-axis appropriately so that the frequency is represented in KHz. Plot the magnitude of $X'[r]$ (in y-axis) using decibel (dB) scale. Note that

$$|X'(r)| \ (in\ dB) = 20 \log_{10} |X'(r)|.$$

   (e) Comment on the magnitude spectrum plot.

3. In this question, you will learn how to do spectrum estimation using the MATLAB *pwelch* function.

   Many of the random phenomena that occur in nature are functions of time. For example, the thermal noise voltage generated in the resistor of an electronic device such as a radio receiver is a function of time. An audio signal that is transmitted over a telephone line is also a function of time. These are examples of stochastic (random) processes.

   In statistical signal processing and physics, the **spectral density** or **power spectral density** is a positive real function of a frequency variable associated with a stationary stochastic process, or a deterministic function of time, which has dimensions of power per Hz. It is often simply called the *spectrum* of the signal. Intuitively, the spectral density captures the frequency content of a stochastic process.

   Note that human speech is a highly nonstationary signal, and the audio signal characteristics change almost every 30-40 msec. Therefore, you can not directly calculate the Fourier transform of the whole signal in one go. This is something beyond the lecture materials. Generally, the power spectrum is calculated as follows: the audio signal is divided into overlapped blocks. The Fourier transform is calculated on each block, and the resulting FT's are averaged to get an estimate of the frequency content of the signal. The MATLAB function *pwelch* will do all this for you.

   (a) Using the following code, estimate the power spectral density of the signal "love_mono22.wav", and plot the spectrum.

   ```
   N = 512;
    [Px, F]=pwelch(x, N, [], N, Fs);
   plot(F/1000, 10*log10(Px)); %Plots the power spectrum
   %scaling F by 1000 will represent frequency in kHz
     xlabel('Frequency (kHz) ');
     ylabel('Power Spectral Density (in dB) ');
   ```

   Since the sampling rate of the audio signal is 22.05 kHz, the maximum frequency in the plot corresponds to 11.025 kHz. The frequency axis has been scaled to show the frequencies in KHz.

   (b) From the power spectrum, note down the frequency range in which the signal has most energy (e.g. the top 20 dB with respect to the DC (zero-frequency) power).

   (c) Note down the frequency of the (annoying) tonal noise present in the signal.

4. In this question, you will learn to calculate the power spectrum of a 2-D image.

   (a) Download the `ayantika.tif` and `Q4.m` from the Lab website.

   (b) Display the image. Comment on the quality of the image - do you observe any annoying artifacts?

   (c) Run the MATLAB program Q4.m, which will calculate and plot the spectrum of the 2-D image. Note that unlike the audio which is 1-D, the image is a 2-D signal. Therefore, the Fourier spectrum has two frequency axes – horizontal and vertical. Note down the normalized range of the horizontal and vertical frequencies shown in the plot.

   (d) The main lobe at the centre (0,0) represents the spectrum of the actual image signal. There are a few other peaks around the main lobe, which are due to the noise (the noise appears as horizontal and vertical lines in the image) in the image. Note down the (2-D) frequencies of these noise peaks.