BME 680 / EE 682

Winter 2018

Due: January 31st

Project 1 Sound Localization

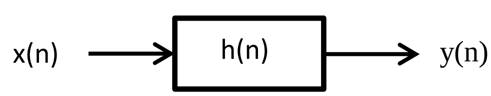
Introduction

In this project, we will look more closely at convolution, correlation, and system identification techniques. We will examine these topics within several application areas including (1) using convolution to develop a system that can perform sound localization, and (2) understand the difference between convolution and correlation by trying to identify a mystery file.

Skills learned include:

- System identification methods
- How to implement a convolution to position sounds at different locations
- How to modify the convolution algorithm to assess correlations between unknown data sets

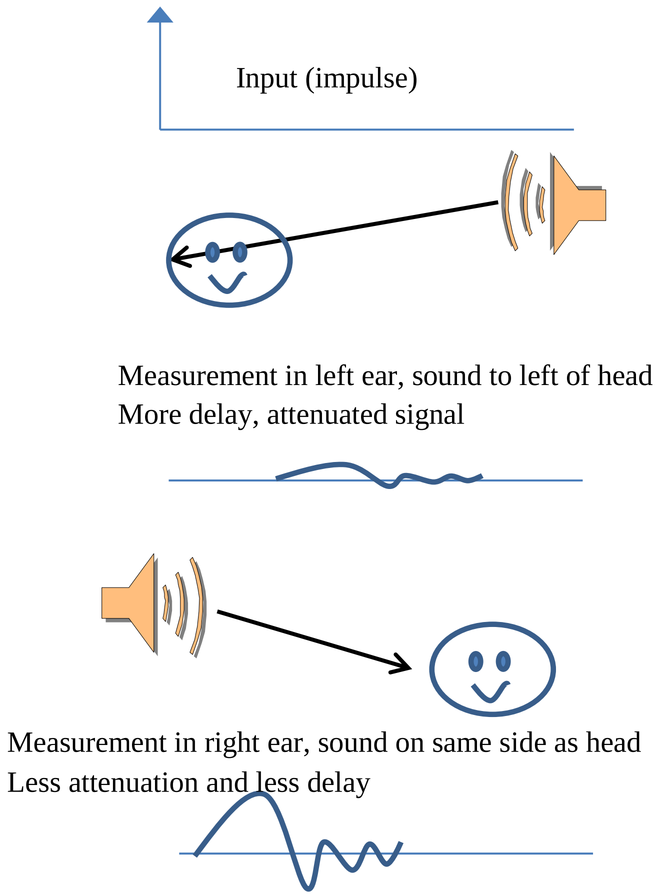
A discrete linear time-invariant system such as the one shown below can be characterized by its impulse response.



The impulse response of the system represents how the system responds to an instantaneous input. This means that if you present the system with an instantaneous input signal at time n (i.e. the Kroenecker delta function $\delta(n)$), the response that you measure in response to that (y(n)) is an estimate of the system's impulse response or its transfer function - h(n). Once you know a system's transfer function, you can use it to estimate how the system can respond to other types of inputs by convolving the impulse response with the input to get the output.

$$y(n) = x(n)*h(n)$$

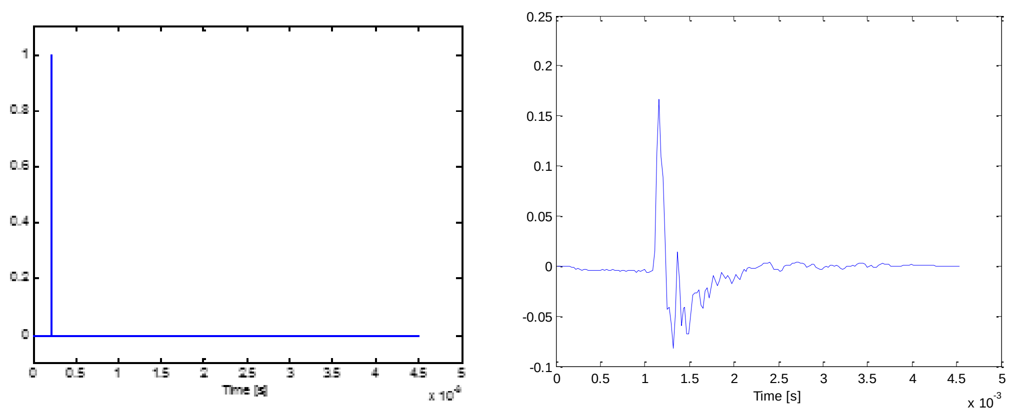
An example of a system that you might want to characterize is a person's head and how the head influences the reception of a sound in either ear of the head. The input to the system is a sound in a room coming from a location relative to the head. If a sound is coming from the opposite side of the head as the ear in which the measurement is being done, this ear will receive the sound later and at a lower amplitude than the ear on the same side as the sound. In this way, the system (the head) results in attenuating and delaying the sound signal. The impulse response for a sound coming from the opposite side of the head as the ear in which a measurement is made would impart this delay and this attenuation on the signal. On the other hand, if the measurement was made in the same side of the head from where the sound is originating, the measurement in the ear would be less attenuated than less delayed than the previous test scenario. Therefore the impulse response for this location of sound would impart less attenuation and less delay on the signal. See the cartoon representation of this below.



Our brains use these attenuations and delay differences between the two ears to know where sounds are located in the environment. These cues are called interaural time difference (ITD) to represent the delays between the ears and interaural level difference (ILD) to represent the amplitude or level differences between the ears. By knowing the left and right ear impulse response for sounds at different locations relative to the head ($h_l(n)$, $h_r(n)$), we can spatially position a sound $x(n)$ anywhere in the room by convolving the sound with the left and right ear impulse response for a given location and then presenting these convolved signals $y_l(n)$, $y_r(n)$, to the left and right ears respectively.

We have provided you with five impulse response pairs, each pair includes a measurement made in the left ear and a measurement made in the right ear in response to a click. The files include a 'Right' for right ear or a 'Left' for left ear. These impulse responses were recorded using microphones placed within the ears of a manikin called KEMAR. During recordings of the sounds, the manikin had microphones placed within each of its ears, one in the left and one in the right. Broadband sounds (i.e. clicks, swept tones, or Golay sequence) were presented at five different locations around the room and the impulse response was measured in each ear by recording the microphone measurement in the ear in response to the broadband sound.

A click is a broadband signal, meaning it includes many frequency components. A click can be generated by having a sound go from a low value to a high value instantaneously. A click is shown in the figure 1(a) below. The microphone measurement will not look like the click below, but rather will be delayed, attenuated, and distorted as shown in Figure 1 (b). This is due to the distortions partially imposed by the head system. Therefore, we call this the head-related transfer function.



(a) (b)

Figure 1: (a) Click used to evoke an impulse response. (b) Measurement in the ear of a KEMAR manikin in response to a click.

Exercises and Report

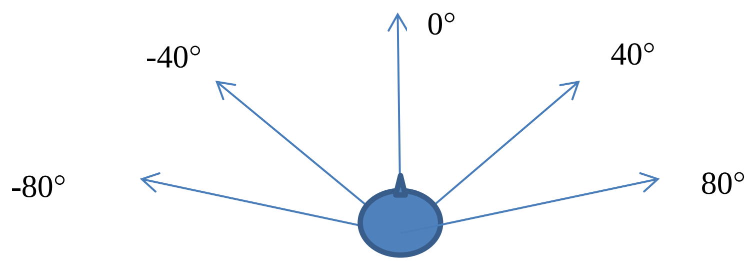
In this project, you are given 10 files. Each file represents an impulse response for a sound presented at a specific location in the room in a given ear. There are five different locations where sounds were emitted and an impulse was recorded in each ear, therefore there are 5 pairs of sounds for each location. Your task is to determine where the sound originated from for each of the impulse response pairs. We have provided you with a sample sound file in .wav format (sentence.wav). You can use this sample file to “spatially move the sound around the room” using the impulse responses. This will involve convolving the sentence with the impulse responses provided and then listening to the sounds through headphones. I have also provided you with some examples of signals that were used to elicit the impulse response signals.

For your report, you will want to do the following tasks and answer the included questions.

1. Write a convolution algorithm in a language of your choice (e.g. Matlab). The structure of your convolution algorithm will look something like this:

   Conv_signal = doConvolve(input1,input2);

2. Use your convolution algorithm to identify from which direction relative to the head each impulse response pair represents (i.e. where is the sound coming from for each impulse response given). The five possible directions are shown below. i.e. identify pair 1 = -40° pair2 = 80° etc.



Pair	Location (e.g. -80, -40, 0, 40, or 80)
1	
2	
3	
4	
5	

3. Make a one or two-line change to your convolution algorithm and convert it to a correlation algorithm.
4. There is an advantage to using a pair of Golay sequences instead of a click impulse response because the out-of-phase autocorrelation coefficients of the sequences sum to zero. Golay sequences are oftentimes used instead of a delta function like a click because it helps reduce measurement noise. A simple example is the Golay sequence of length 2, which is (1,1) and (1,-1). Show that the out-of-phase coefficients of the autocorrelation of these two sequences sums to zero and explain why you might want to use two Golay sequences instead of a click to evoke an impulse response when noise is present in the room.
5. You will be given a mystery file. The mystery file contains one of the impulse responses that you have worked on in steps 1 and 2, however, random noise has been added to the signal. Determine which of the impulse responses the mystery file contains – indicate both the ear and the pair number or the direction. Write a brief explanation describing how you found your answer. Include any source code or plots used to verify your answer.

Additional notes:

The .wav files are all sampled at 44100 samples/sec. If using Matlab, uou will likely find it useful to use the wavread and wavwrite functions to read and write data from/to audio file format. Remember that when you do a wavwrite or wavread, you should specify the sample rate. See the following example:

```
% Write a stereo .wav audio file that has data to be sent to the left ear and the right ear.
% Data is sampled at 44100 samples/sec with a file name called sample_file.wav
wavwrite([left_data_in_column right_data_in_column],44100,'sample_file.wav');
```