Assignment: Homework 7

What to Hand In

1. Your code with appropriate comments and a **readme file**. You are responsible for making us understand your code and know how to run it.
2. A report with answers, figures, and explanations.
3. Any sound files requested in the problems.

How to Hand It In

1. Put all of your files in one folder. Compress this folder and name it <firstname>_<lastname>_HW7.zip. For example, "Zhiyao_Duan_HW7.zip".
2. Submit via blackboard to the appropriate entry**.**

When to Hand It In

It is due immediately at the beginning of the day (i.e., 0:00AM) specified on the course calendar. **Late assignments will not be graded.**

Problems (10 points in total)

In this homework we will be implementing three methods of generating spatial audio effects for stereo headphones. Taking a **single-channel** audio signal as input (**for stereo input, we need to mix the two channels first**), we will convert it to a two-channel (stereo) signal to be played back by the headphone. By manipulating the signals going to the two ears, we will "place" the audio somewhere on the horizontal plane of our ears. The first method will be the simple amplitude panning. The second method will additionally consider phase/time difference between the two ears. The third method will further consider the filtering effect of the head modeled by the Head-Related Transfer Functions (HRTF).

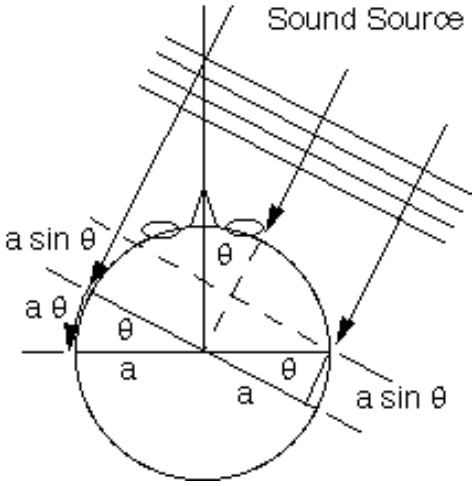
1. (3 points) Amplitude panning. In amplitude panning, the signals that go to the two channels are identical except for their amplitudes. When the signal is panned to the left (right), the left (right) channel equals to the input signal while the right (left) channel is silent. When the signal is panned to the middle, the left and right channels are the same. There are different formulae for amplitude panning. Here is one:

$$\mathrm{y[left]} = \mathrm{x}\cos\left(\lambda\,\frac{\pi}{2}\right),$$
$$\mathrm{y[right]} = \mathrm{x}\sin\left(\lambda\frac{\pi}{2}\right),$$

   where x is the single-channel input signal, y is the two-channel output signal, λ is the pan parameter ranging between 0 and 1. The left-right signal ratio varies between $\pm\infty$ accordingly. The good thing about this formula is that the summation of the energy of the two channels is constant. Implement this amplitude panning with a

slider to control the pan parameter in real time on the OMAP-L138 board. Listen to the effect, do you think it is effective? You might feel that the signal is placed somewhere in your head.
Note: If your slider doesn't work, you will receive up to 2 points.

2. (3 points) IID and ITD. Amplitude panning is simple but it does not consider the interaural time/phase difference. In this problem we will be implementing both the IID and ITD to place the signal outside of the head. Refer to the figure on the right and use the following formulae:



$$\text{IID (linear amplitude)} = \frac{1-\frac{1}{3}\sin\theta}{1+\frac{1}{3}\sin\theta},$$

$$\text{ITD} = \frac{a}{c}(\theta + \sin\theta),$$

where θ is the azimuth, ranging from $-\frac{\pi}{2}$ (left) to $\frac{\pi}{2}$ (right); a is half of the distance between the two ears; c is the speed of sound. From these equations, we can see that the IID ranges from 2 to $\frac{1}{2}$, or 6 dB to -6 dB. Implement this effect on the OMAP-L138 board with a slider to control the azimuth. Choose appropriate values of a and c in your implementation.
Note: If your slider doesn't work, you will receive up to 2 points.

3. (4 points) In this problem we will use HRTFs to make the spatial effect even more realistic.
    a. (1 point) Use MATLAB to load the HRIR data. You will find two structs: "l_hrir_S" and "r_hrir_S", corresponding to the head-related impulse responses of the two ears. In each struct, "content_m" contains the impulse responses. Each row corresponds to one position (azimuth and elevation) described in "azim_v" and "elev_v", respectively. As we will only use the horizontal plane (i.e., elevation 0 degree, corresponding to Rows 73-96 of the array), you can discard impulse responses corresponding to other positions. Plot the left-ear impulse response when the azimuth is 90° and 270°, respectively. Which one is on the left side of the dummy head? You can see that both impulse responses really only start around 240 samples. This delay corresponds to the distance between the loudspeaker and the dummy head. Using the sampling rate and the speed of sound, you can calculate this distance. What is it roughly?

b. (1 point) Now you can import the impulse responses at azimuth=120° into OMAP L-138 and convolve them with your input signal to create stereo output. This will place your input signal to the 120° azimuth location on the horizontal plane.
   Hint 1: As only the relative (instead of the absolute) time delay matters, you can discard the first 240 samples of all impulse responses. This saves memory. In addition, as the entire impulse response may be too long to fit into the memory, you may use just a portion of it. Apparently, the longer the better. Report the length you finally use.
   Hint 2: to calculate the convolution in real time, you basically view each incoming sample of the input signal as an impulse. The HRTF filter thus has a response to it, which is a scaled HRIR by the input signal sample. This response will affect current and future output signal samples. You can simply maintain a buffer for the future output signal and update it every time you receive an input sample.
c. (2 points) Implement a slider to control the azimuth ranging from 0° to 360°. Convolve the signal with the impulse responses at the closest position. (You do not have to interpolate the impulse responses for non-registered positions.) Again, you may just use a portion of each impulse response to save memory and computation. What do you think about this spatial effect?
   Hint: The implementation procedure is the same as b, with the only difference that here the HRIRs change over time, determined by the slider.