Lab 5: 3D Audio

In this lab we will learn how to create 3D sounds for headphone playback. We will make use of simple filters and HRTFs to create static and moving sources. Use the three sounds `fly.wav`, `helicopter.wav`, and `crumble.wav` in the lab archive as sources for the 3D recording that you will create.

Part 1. Static sources using ITD/ILD cues.

Assume the following source locations:

a) Straight ahead

b) 45 degrees to the left

c) 80 degrees to the right

d) 160 degrees to the left

For each location find the source's delay between the two ears (assume a source distance of 2 meters), and design two filters that will simulate that ITD and ILD features (feel free to round the IDT delays to an integer sample size). Assume that when sounds come from the side of the head the attenuation at the contralateral ear is by a factor of 0.7. From sounds coming medial plane (between the ears) there will be no attenuation due to the head. For positions moving from the medial plane towards the sides you can interpolate between no attenuation and a factor of 0.7. Design and plot the filters that correspond to the locations shown above and use them to make 3D sounds with the given source material. Listen to them through headphones and verify that they sound somewhat localized (it won't sound perfect, but it should be believable).

Part 2. Static sources using HRTFs

In the lab archive you will also find a directory with HRTF files. In that directory you will also find code for the function `load_hrtf` which returns the left and right HRTF filters given as input a source's azimuth and elevation. These filters will be much better than the ITD/ILD filters for localizing sounds.

Apply the HRTFs on the given sources and create 3D sounds that correspond to the locations given above. For each source, you will beed to convolve it with the left and right HRTF of the desired position and generate two sounds, one for each channel. Verify that they sound correct using headphones; are they better than before? What differences do you observe?

Part 3. Dynamic sources

In this part you will need to make a moving sound source. In order to do so we will make use of a fast convolution routine based on your STFT code from lab 1. In order to perform fast convolution we can perform an STFT of the sound to use, multiply each time frame of this transform with the DFT of the filter that we want to impose and then use overlap add to transform back to the time domain.

Start by taking each source from above, and apply your STFT on it. Make sure that the size of the transform is the same as the HRTF's filter length. The hop size should be the same as the DFT size and you will need to zero pad by as much as the DFT size in order to avoid performing circular convolution. Do not use an analysis/synthesis window.

Once you compute this STFT, go through its every time frame and element-wise multiply it with the desired HRTF filter to generate the STFT of the left and right sounds. Figure out which HRTF angle to multiply each frame with so that by the end of the sound you will have made it go around your head.

Make sure that you don't fumble the convolution. The input has been zero-padded, so the filter needs to be zero-padded at the same amount when you take its DFT. Also remember that in the STFT you keep only the first N/2+1 frequencies, so you will need to do the same for the filter DFT otherwise you will not be multiplying the corresponding frequencies.

Once you perform these operations you will have generated two STFT matrices, one for the left channel and one for the right. Use your inverse STFT routine and play the stereo sound through your headphones. You should hear a convincing rendering of the original sounds circling around your head.

Extra credit

You can use extra credit assignments to add a few more points to your grade, they are optional but will help your final grade. Use the file `dry.wav` to make a short story. This is a 7-channel file of a scene from a really bad B-movie. Each channel contains a different sound. If you play that sound you probably won't hear most of the content since you won't have a 7 speaker setup, import it in a multi-channel editor such as Audacity and you will get a sense of what's in there. Since it sounds so boring you need to add some reverb and 3D-locate the sounds so that it sounds more exciting. Try to make it sound better using what we have done so far. Keep in mind that you can add different amounts of reverb for each sound, and you can dynamically 3D place them as well.