2021 ASAS Homework 4: Head related transfer functions and filtering via FFT

Prof. Yi-Wen Liu

Due Sunday April 11, 2021.

Assigned reading material for this week's seminar: L. Hausmann, M. von Campenhausen, F. Endler, M. Singheiser, and H. Wagner, "Improvements of sound localization abilities by the facial ruff of the barn owl (Tyto alba) as demonstrated by virtual ruff removal," PLoS One 4(11) e7721, Nov. 2009.

Homework description

In this homework, you will learn to apply HRTFs to mono sound, make a two-channel output, and forge the perception of sounds arrival from a certain spatial direction (θ, ϕ) when listened through a stereo headphone. These HRTFs were measured from Prof. Liu's head many years ago.

Please run the starter code first, and the head-related impulse responses for the left and right ear will be stored in `lhrir` and `rhrir`, respectively. To begin block-wise processing, please try using either the rectangular window or the Hann window. Somewhat awkwardly, these two lines

```
win_hann = hann(L+1);
win_hann = win_hann(1:end-1);
```

creates a Hann window which satisfies the constant overlap-add (COLA) condition for a hop size of L/2. Create a "windowed frame" by performing sample-wise multiplication of a block of the input mono signal with `win_hann.`

After this, every windowed frame needs to be convolved with the filters `lhrir` and `rhrir`, respectively. The result becomes stereo. You are required to do it in the frequency domain by completing the following steps:

Step 1: Zero-pad both the windowed frame and the HRIRs to a sufficient FFT length N_zp > L+M-1, where L denotes the block length, and M denotes the filter length. The FFT of HRIRs can be regarded as the HRTFs sampled at equally-spaced frequencies.

Step 2: Multiply the FFT of the windowed frame with the left and right HRTFs.

Step 3: Apply inverse FFT to the result (which has two channels), and the outcome should be an array of length N_zp for the left and right channel, respectively.

Step 4: Shift the time index by L/2, repeat steps 1-3, and perform *overlap and add*.

When everything is done correctly, the results should be numerically identical to what you can obtain by directly using `conv()` to perform FIR filtering. You should check it. Anything other than an identical result might cause audible frame-rate error.

Basic Activities for writing up the report (80%):

1. Listen to the result of mono-to-stereo conversion and compare the perceived quality with or without zero-padding. Again, let's emphasize that it is wrong to do it without zero-padding in step 1. The purpose here is to train yourself to tell how things can go wrong by hearing.
2. Is it possible that imposing a delay between two channels can produce a similar virtual effect? Please shift one of the two channels by up to K samples and describe what you hear. Here, please choose K such that $K\Delta t$ is approximately the maximal possible *time difference of arrival* between your ears, where $\Delta t = 1/fs$ denotes the sampling period.

Bonus for writing up the report (20%)

I have a few different suggestions this week. Please do at least one of the following.

A. You may study how HRIRs and HRTFs vary systematically when we change (θ, ϕ). You may attempt to estimate ILD and ITD from the HRIRs.

B. Implement block-wise FFT-based FIR filtering to record mono from a microphone and play stereo to the headphone *in real-time*. You may find the following functions helpful: `audioDeviceReader(), audioDeviceWriter().`

C. If you are less inclined to explore the technical details of this homework further, you may choose to conduct blind-tests with your group partners to see if each of you can distinguish small differences in the direction of arrival (DoA) such as +5 vs. -5 degrees (left vs. right), 85 vs. 95 degrees (front vs. back) and so on. Find out the minimum azimuth difference for you to consistently notice the difference. I'd be curious to know whether people whose head has a similar shape to mine perform better? Also, change the input signal to any piece of music, and you may also compare if sensitivity to the difference in DoA depends on the type of music being played.