**ELG 5376 Digital Signal Processing**

**Term Project, Winter 2019**

1. Overall description

To practice some of the filtering concepts described in class, in this year's project you will have to implement a 3-D sound rendering system using binaural cues and Head Related Transfer Functions (HRTFs), and producing a stereo sound file to be played through headphones. Your system should be able to reproduce the effect of a sound source moving around the head of the listener. The implementation must be done in Python (NumPy), i.e., Matlab code or other languages will not be accepted.

How does this relate to the course?

HRTFs or their time domain equivalent head related impulse responses (HRIRs) are typically FIR filters and impulse responses. To use them in producing 3-D sounds, we need to compute a filtering or convolution operation, which are concepts at the very core of the course. Computational complexity and filter design are also factors to be considered, and different implementations corresponding to material covered in class are to be investigated.

Theory of 3-D audio sound rendering using HRTFs:

1) An overview of the concept of head related transfer functions and binaural sounds can be found here: https://en.wikipedia.org/wiki/Head-related_transfer_function.

2) For a more detailed description and additional background on 3-D sounds, the classic 1994 textbook "3-D Sound For Virtual Reality and Multimedia" by D.R. Begault is quite accessible for people new to the field:
https://www.researchgate.net/publication/2571893_3-D_Sound_For_Virtual_Reality_and_Multimedia (click download).
A (legal) scanned copy of the same textbook can be found here: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20010044352.pdf.

3) The 2011 paper "Headphone-Based Spatial Sound" by V.R. Algazi and R.O. Duda is also a good tutorial on the topic:
https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5670436.

Will the 3-D sound rendering be perfect?

To be perfect the rendering system would require HRTFs or HRIRs to be measured individually on each person, and it would also require to know the characteristics (frequency response) of each type of headphone used in the rendering system, to remove (compensate for) the impact of each headphone. In this project we will not aim for such accuracy. However, it is still possible to reproduce overall 3-D sound effects with the approaches described in this project, even though there may be artifacts such as sounds which seem to be moving inside one's head rather than outside, or front-back ambiguities in some cases.

Where to find sets of HRTFs (or HRIRs)?

It is easy to find public domain sets of HRTFs from the web. For example, the following page lists a few sources: https://en.wikipedia.org/wiki/Head-related_transfer_function.

CIPIC HRTF Database
https://www.ece.ucdavis.edu/cipic/spatial-sound/hrtf-data/

LISTEN HRTF DATABASE
http://recherche.ircam.fr/equipes/salles/listen/index.html

ITA HRTF-database
http://www.akustik.rwth-aachen.de/go/id/lsly

Aachen Impulse Response Database
http://www.iks.rwth-aachen.de/en/research/tools-downloads/databases/aachen-impulse-response-database/

Spatial audio impulse response compilation captured at the WDR broadcast studios
http://audiogroup.web.th-koeln.de/wdr_irc.html

Spherical Far-Field HRIR Compilation of the Neumann KU100
http://audiogroup.web.th-koeln.de/ku100hrir.html

Spherical Near-Field HRIR Compilation of the Neumann KU100
http://audiogroup.web.th-koeln.de/ku100nfhrir.html

MIT Media Lab
http://sound.media.mit.edu/resources/KEMAR.html

ARI HRTF Database
https://www.kfs.oeaw.ac.at/index.php?option=com_content&view=article&id=608:ari-hrtf-database&catid=158:resources-items&Itemid=606&lang=en

2. More detailed project specifications

For the filtering method to be implemented in Section 3a), two different stereo (2 channel, binaural) output .wav sound files of 20 seconds need to be produced, where the sound source should be perceived as moving:

- (1st stereo binaural file) on a horizontal a plane at the level of the ears, i.e., sound moving in circle from the front of the head, to the left of the head, to the back of the head, to the right of the head, and to the front of the head;
- (2nd stereo binaural file) on a vertical plane crossing the two ears, i.e., sound moving on a half-circle from the left of the head, to the top of the head (above), to the right of the head.

For the filtering method to be implemented in Section 3b), only the first scenario above needs to be considered (sound moving on horizontal a plane at the level of the ears).

For the input sound file to be filtered and rendered as a moving sound source in the headsets, use a mono (1 channel, monaural) input sound file with a sampling rate of 44.1 kHz or 48 kHz. Make sure that it corresponds to a sound file originally recorded or produced at 44.1 kHz or 48 kHz, which will include high frequency content up to 20 kHz (as opposed to a sound file originally sampled at 8 kHz and then upsampled to 48 kHz, which would only have sound content up to 4 kHz). Because all frequencies can affect the binaural perception of a sound source direction, it is a good idea to use an input sound file which has a wide range of frequencies in it. You may want to try your code with a few different input sounds, as the 3-D effect can be better with some sounds than other. Several web sites such as https://freesound.org can be used to find sound files.

3. Different filtering methods to be implemented:

   a) Using experimentally measured HRIRs (downloaded from the web, see section 1)
      i) Time domain convolution equation for FIR causal filters.
      ii) Frequency domain FIR filtering using the "overlap and add" block processing method.
   b) Using HRIR filters designed from a basic synthetic model of HRTFs:
      Use the "frequency sampling" FIR filter design method (to be covered in class, but not explicitly described in the reference textbook) to obtain HRIR FIR filters at different angles, from the frequency response of the HRTF model described on p. 997 (bottom of right column) of the 1993 paper "Modeling head related transfer functions" from R.O. Duda:
      https://www.ece.ucdavis.edu/cipic/files/2015/04/cipic_Duda1993ModelingHRTFs.pdf.
      Then perform the filtering operation using the time domain convolution equation with the HRIR FIR causal filters previously designed.

**Project evaluation**

Projects are to be done in teams of two students. On an exceptional basis, projects can also be done individually, but permission needs to be asked for this. Teams of three students will not be accepted.

In your report, it is required to present the mathematical description that you implement. For the coding part, you should only make use of standard NumPy libraries, and not make use of additional public domain software packages, libraries, codes, etc.

The final report should be in the form of a journal paper, with a maximum of twelve pages. Use the format of the IEEE Institute, for which Word or LaTex templates can be found at:

(Word) https://ieeeauthor.wpengine.com/wp-content/uploads/Transactions-template-and-instructions-on-how-to-create-your-article.doc
(Latex) https://ieeeauthor.wpengine.com/wp-content/uploads/WIN-or-MAC-LaTeX2e-Transactions-Style-File.zip

You must write the reports yourself; plagiarism can lead to a failing mark, see the leaflet "Beware of Plagiarism" https://www.uottawa.ca/about/sites/www.uottawa.ca.about/files/plagiarism.pdf for guidelines. This link provides several examples to explain what is acceptable and what is not acceptable when citing, paraphrasing or using previous work.

The following criteria will be considered when grading the reports:

- the technical soundness of the material presented;
- the clarity and quality of the document, including the presentation of results;
- in addition to the report, you are also required to submit sound files corresponding to the outputs of the different methods that you implement, as well as your source code which can be executed by the Corrector. The quality of the sound files and the correct execution of the source code are factors to be considered by the Corrector.

Reports, sound files and source codes are to be submitted using the Virtual Campus platform. To keep things simple, please provide only one submission per team (submission to be done with the account of only one team member), and indicate clearly the names of the two team members on the report.

**Deadline:**

Last day of classes: April 5th. A penalty of 10% per day can be applied for late submissions.