Spatialisation and Panning for Headphones

Sally-anne Kellaway S430029204
Digital Audio Systems 2013 DESC9115
Submission for Lab Report 2

Product proposition

Technology pushes to deliver audio visual material to be as "real" as possible – delivering experiences which mimic the way we see and hear. Common executions of spatial audio for delivery on headphones will apply Inter-aural Time Differences (ITD), Shoulder Reflection (SHR) and Pinna Reflections (PR) to the direct mono input in two channels (left and right) and will then sum the two as a stereo file.

These techniques offer value to the realm of conference telephone calls. Calls involving many people may sound 'cluttered', and without spatialisation, multiple people talking at the same time may result in any of the participants' words being unintelligible. Selective digital signal processing on the signals being summed at the receiver's phone may alleviate this issue. This project proposes spatially separating each participant and providing a way to "rank" the sources to imply a "host" or "curator", and a "participant" status.

Product Goal

This product aims to provide the listener the ability to detect spatial difference between two or more sources. The contained MATLAB script executes a room reverberation function, an ITD function and Shoulder Reflection functions in sequence.

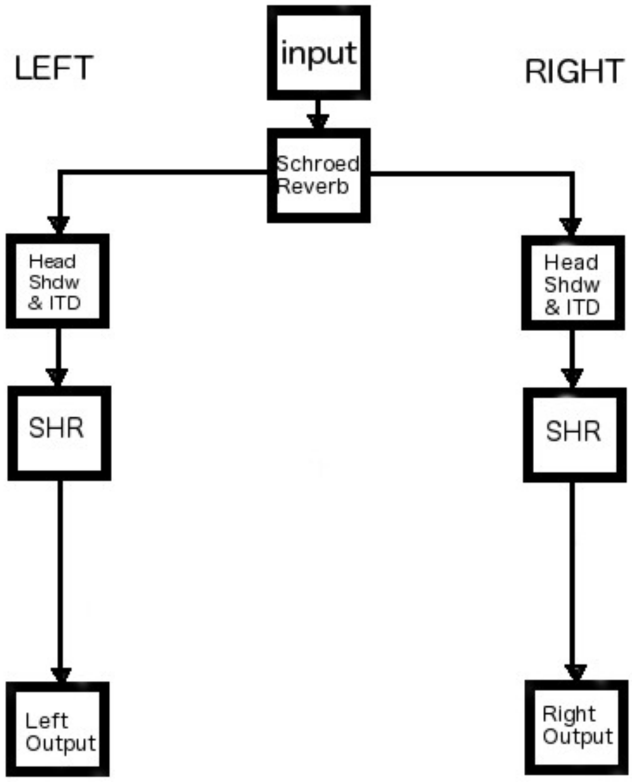
Execution

Resources Provided

- Script.m Executes the sequence of functions
- ITD_HS.m Applies Head Shadowing and Interaural Time Differences
- Shoulder_reflect.m Applies Shoulder Reflections
- Schroeder.m Applies Reverberation
  - fbcomb.m Generates Comb Filters for reverberation
  - allpass.m Generates All Pass Filters for reverberation
  - tdl.m Generates the Tapped Delay Line for reverberation
- overhere_input_mono.wav Input file
- Output Files - Spatialisation_and_panning_demo_t-90_e0_r2.wav, Spatialisation_and_panning_demo_t-45_e45_r2.wav,

Spatialisation_and_panning_demo_t35_e45_r1.wav,
Spatialisation_and_panning_demo_t135_e0_r1

Description of Processing



Script.m

Defines constants - Theta, elevation, rank and input file
This script applies Reverberation, Interaural Time Difference, Head Shadowing and Shoulder reflections to a mono input signal and writes the result to a wav file in the selected directory.

Schroeder.m

This reverb was selected as it offers fairly realistic reverberation with a relatively low cost on memory. Avoiding convolution reverbs will potentially avoid issues with large numbers of participants requiring large amounts of memory. This reverb uses a series of comb filters and tap delays to process an input signal with reverberation. Although the number of comb filters and tap delays, as well as their durations and gain factors are exposed within the function itself, two ranges of values are defined, with the switch variable “rank” exposed in the Script.m file.

ITD_HS.m

Calculates Interaural Time Delay and Head Shadowing effect from the angle of the sound.
Head Shadowing is executed first, both for the ipsi and contralateral ear. The variable theta is the angle the sound should be positioned at, and to calculate the contralateral ear, the negative value of theta is used. The following equation develops the

coefficient α (θ), which is used along with the angular frequency ω (speed of sound divided by radius of head), to develop the head shadowing filter Hhs.

$$\alpha(\theta) = 1.05 + 0.95 \cos\left(\frac{\theta}{150^\circ}\ 180^\circ\right)$$

$$Hhs = \frac{(\omega_\circ + \alpha Fs) + (\omega_\circ - \alpha Fs)z^{-1}}{(\omega_\circ + Fs) + (\omega_\circ - Fs)z^{-1}}$$

The ITD is then calculated from the theta θ value in radians. From this, ω is multiplied by the sum of θ rads and sin(θ rads) to obtain the ITD in seconds. Since milliseconds are the unit of time used throughout the sequence of functions, this conversion is made.

The final execution is the summation of the delayed and undelayed signals. The extra samples required for the delay are calculated by creating a vector of zeros the length of the ITD. This delay is then sequenced with the outputs of the Head Shadowing filter two ways - one output channel has the delay sequenced first and the other has it sequenced after. The two outputs are then concatenated to a stereo file if testing is required.

Shoulder_reflect.m

Calculates Shoulder reflections from the angle theta (horizontal) and angle of elevation

The Shoulder Reflection function will result in a single echo per channel, which is dependent on theta θ and elevation **ϕ** (phi). The following equation is used.

$$T_{sh} = 1.2 \times \frac{180^\circ - \theta}{180^\circ} \times \left(1 - 0.00004\left((\phi - 80^\circ)\ \frac{180^\circ}{180^\circ + \theta}\right)^2\right)$$

As can be noted, the use of simplified figures decreases the ability for this solution to accurately represent a wide range of individuals' shoulder echoes (for example, variations in neck length and shoulder width are not accounted for). However, this provides a good start point for experimentation.

At this point, a sequence is introduced which adjusts the gain of the working channels SHR_L and _R. Following this, the ITD function outputs are tested for consistent length before adding the shoulder reflection. It was found that when the angle θ was above 90° or -90° that the SHR and ITD overall lengths were different, which caused errors in the summation of the signals. The length (in samples) for both ITD and SHR are assigned to variables. A test for the larger of the ITD_size and SHR_size is executed, with the longer file being subtracted from the shorter to result in the difference – a number of samples that must be added to the shorter file for the summation to be carried out within MATLAB. The two signals are then added, normalized and concatenated to form a stereo output for stage based testing.

Within the script, the sound is then played at the sampling frequency, and written to a wav file for future listening.

Evaluation of Result

This project has developed a product, which achieves the goals laid out in the introduction. It has also laid clear a path for continued development to deepen the results achieved.

Strong elements of the product are its ability to laterally place a sound source in the 3D field. Comparing the four provided output files, it can be noted that for rank one files, the 3D panning is very precise, although externalisation from the headphones can be more pronounced. Differences in elevation and front-to-back shadowing could be more well developed. Rank two files provide a lesser sense of position – the feeling of "definite left" from the -90° theta angle should be much stronger (as in the theta 135° and theta 35° files.

The reverb allows an extremely flexible method for homogenising multiple input sounds, and its wide range of variables allow for a high level of flexibility to the development team when determining ranks. These are still available to the user to edit too, which is another positive element to this choice of reverberation technique.

Areas for development include a deeper conveyance of the elevation of the input sound source, as well as a higher level of shadowing when the sound source is behind the listener. It could be argued that these elements are not critical for the project, as professional group conversations are most often laid out as a round table setting, which does not usually include sourced behind or above the listener. However, a greater need for externalisation and distance attenuation would be of great benefit to the product, which would be a by-product of a pinna reflections function added to the program.

Future Developments

Within scope of project as it stands, to assist with the achievement of the goal of the product, the development of a small number of elements for the project is foreseeable. As previously mentioned, the development of a Pinna Reflections module will assist in deepening the elevation and front-to-back panning in the program. This element would be executed through the development of a pole-only filter for the ipsilateral ear to approximate the deep notches which occur due to the interaction of the pinna with incoming sounds. The contralateral ear will not require this layer of filtering as the effect of the pinna is not as great on the already filtered sounds arriving at this ear are not direct to the structure of the pinna. This element would be achievable within a three-week period, inclusive of assimilation in to the existing program.

Beyond the scope of this project (as a future iteration), it would be interesting to offer not just rank variables, but additionally to offer room variables to give the flexibility of smaller or larger virtual meetings. This element would develop a longer period to develop, as provision of multiple "rooms" would require the creation of ranks per room.

References

Rocchesso, D 2002, 'Spatial Effects', in U Zolzer (ed.), *DAFX – Digital Audio Effects,* John Wiley & Sons Ltd, Stafford, QLD.

Cassidy B and O'Neil L, ND, '3-D Spatialization and Localization and Simulated Surround Sound with Headphones'. University of Victoria, VIC. Available at http://web.uvic.ca/~loneil/elec484/project/LucasONeil_BrendanCassidy_Project.pdf

Miranda Jofre, LA (ed) 2013, 'Schroeder.m', 'tdl.m', 'fbcomb.m', 'allpass.m', Week 8 Digital Audio Systems Class Materials, University of Sydney, NSW.