HRTF-Based Mixing Plugin

YUTONG HAN 440409724

Digital Audio System, DESC9115, Semester 1 2015 Graduate Program in Audio and Acoustics Faculty of Architecture, Design and Planning, The University of Sydney

ABSTRACT

This report is to provide an implementation of HRTF-Based spatial mixing plugin system, the plugin contains limiter, compressor, and HRTFs processing component which using head related impulse response (HRIR). It is designed for earphones play back system, which the aim of this project is providing the operable mixing plugin for digital audio workstation, such as Pro-Tools, Audition, Cubase or other software.

1. INTRODUCTION

From gramophone to iPod, from mono to stereo to 5.1 even 9.1 speakers system, people increasingly require higher quality and greater performance of sound. In resent years the spatial sound effects have become a popular technology, which aims to provide good quality of spatial hearing, such as Dolby Atmos.

Although spatial sound effects of loudspeakers system is almost a mature technology, however the binaural spatial sound effects still in developing, it is hard to find a commercial spatial sound effects plugin especially for binaural play back system. This report is to implement spatial sound effects with mixer system, which is based on head related impulse response (HRIR). The spatial sound process is based on related HRTFs.

2. RELATED TECHNOLOGY

2.1 Limiter

Limiter is used for control over the high peaks of input signal and reduces the level of volume in dB as little as possible. The basic diagram is shown below.
2.2 compressor and expander

Compressor and expander are based on RMS measurement. Compressors attenuate strong signals; expanders attenuate weak signals. Compressors are used mainly to decrease the dynamic range of audio signals so that they fit into the dynamic range of the playback or broadcast system. Expanders are used for increasing the dynamic range of signals, for noise reduction.

2.3 Head Related Transfer function and Head Related Impulse Responses

Spatial localization is based on lateralization, which sound come to left and right ears are different. There are two aspects, interaural time difference (ITD) and interaural level difference (ILD). Because of extra distance sound travels to the contralateral ear, it can cause time difference between two ears which people can depend on ITD to locate the sound source. In addition, due to the acoustic shadow from pinna, skull, shoulder and other part of the body that affects the sound level difference between two ears. The shadowing is significant at frequencies above 2 kHz, however it does not exist when the frequency is below about 800 Hz.

However, one of the most common ways to obtain a binaural data from source in space by using measured Head Related Impulse Responses (HRIR). HRIR refers to measure an impulse response from a source at a specific location around the head by using a real or dummy head with microphones in the ear. HRTF refers to convolve the HRIR data by using the convolution definition to calculate, but it is an easy way to use the convolution function in MATLAB.

\[ y[n] = f[n] * g[n] = \sum_{m=0}^{M-1} f[n - m]g[m] \]

Dummy head can stimulate a reliable Head Related Impulse Responses (HRIR), For example, in the acoustic lab, the dummy head with 45-degree and 1 meter from source, and 90-degree and 1 meter from source. From the impulse response figure, the ITD in 45-degree and 90-degree are roughly 0.354ms and 0.771ms. In terms of ILD, according to octave band figures of 45-degree dummy head, the sound pressure level is roughly same in low frequency, however in high frequency the sound pressure level is around 10dB difference. It is good to match the theoretical expectations.
3. IMPLEMENTATION

This HRTF-based mixing plugin is designed for earphones play back system, which the aim of this project is providing the operable mixing plugin for digital audio workstation. According to CIPIC HRTFs data (http://interface.cipic.ucdavis.edu/), the mixing system can process the sound depending on different people. Therefore the sound could be variety, for example, music industries could depend on different people’s spatial interests to process different music by using different HRTFs.

Basically, this plugin is consisted of five components: input the audio files, set the parameters of Limiter, set the parameters of compressor, load the HRIR data which depend on different size of ears, and set the total gain value. Figure 1 shows the diagram of this plugin system. Figure 2 shows the plugin looks like.

![Diagram of the plugin system](image1)

**Figure 1**
4. SIGNAL PROCESSING

There are three major signal-processing components and two main signal path, each of those component can switch on or off in this plugin: limiter is used for limit the high volume of signal, which is based on peak level measurement, the attack time should be very small, typically less than a millisecond, so that the reaction is very fast. Compression processing is used to reduce the dynamics of the input signal.

HRTF is a function, which based on HRIR data from measurement. At first, select the different shape of ear from the CIPIC HRTF database, such as large pinna or small pinna. Then, obtaining HRTFs from HRIRs by using convolution function. For given HRIR data of large and small pinna, the source was moved in the horizontal plane clockwise around the head. The vertical-polar azimuth goes from 0°in steps of 5° to 355°. The input angle of HRTF is based on index of azimuth, for example index number is 1=0°, 10=45°etc.

5. EVALUATION

As results, through the binaural analysis by AARAE, the ITD of 0 degree, 45 degree and 90 degree are 0.002268, 0.4059 and 0.7234 (ms) respectively. It is very close to the results which I did in acoustic lab. The ITD in different angles are also matching the theoretical expectations. Therefore depend on those results, this plugin is reliable to do a spatial sound process.
6. DISCUSSION AND CONCLUSION

There are some issues for this plugin that the main problem is that the HRIR data is not perfectly matched everyone’s ear. In further work, it would be more concentration on flexible using, which it could instantaneous process the music for people. Second problem is that when input large sound data or many sound tracks to process with this plugin, it will take long time.

### Attachment Structure

<table>
<thead>
<tr>
<th>files name</th>
<th>final_review.pdf</th>
</tr>
</thead>
<tbody>
<tr>
<td>classical_drum.m</td>
<td>demo</td>
</tr>
<tr>
<td>comp_expan.m</td>
<td>compressor function</td>
</tr>
<tr>
<td>my_fft_conv.m</td>
<td>convolution function</td>
</tr>
<tr>
<td>limiter.m</td>
<td>limiter function</td>
</tr>
<tr>
<td>large_pinna_final.mat</td>
<td>HRIR Data for small and large ear</td>
</tr>
<tr>
<td>small_pinna_final.mat</td>
<td></td>
</tr>
<tr>
<td>Evaluation.m</td>
<td>Evaluation demo</td>
</tr>
<tr>
<td>anechoic_voice.wav</td>
<td></td>
</tr>
</tbody>
</table>

### Reference

DESC 9115 Tutorial WEEK6: Head Related Transfer Function and Impulse Response”, Sydney University, William L. Martens