PRODUCT PROPOSAL
SAMPLE LIBRARY PREPARATION TOOL

Ben Jancovich 460383886

Final Review, Digital Audio Systems, 2016
Graduate Program in Audio and Acoustics
Faculty of Architecture, Design and Planning, University of Sydney

PRODUCT OUTLINE

The modern audio production process is highly dependent on efficient workflows and pre-preparation of content. Fast turnaround is crucial. For this reason, it is standard practice for producers and sound designers to maintain libraries of pre recorded sounds. Whether those sounds have been purchased as sample libraries or recorded by the end user, it may be necessary that they undergo various types of signal processing to minimize any additional mixing required once they are loaded into to a production. A common example is that of a commercially available drum sample library. Often the hi-hat samples are monaural, or feature a very narrow stereo image, and can also contain unnecessary low frequency information. These issues can be remedied in the production session once the samples have been added, but this interruption to workflow can limit creativity and reduce operational efficiency. This product has been developed as a solution to this problem. It is intended to batch process large numbers of audio files to prepare them for later use.

The signal processing modules are implemented in what is known as a multiband processing network, which consists of several sets of parallel signal processing chains that each work on a different part of the frequency spectrum. This allows for a higher degree of control than processing the entire broadband signal. The central concept is that what is considered a desirable sonic characteristic for one part of a signal's spectrum, may not be desirable for another. An exaggerated stereo width effect for example, may be desirable in the high-mid frequency range of a bass synthesizer sound, and this can be created using an inter-aural time difference. In the low range however, the phasing introduced by this delay between channels will result in an unintended loss of low frequency information. Multiband processing eliminates this problem.

RECOMMENDED GUIDELINES FOR OPERATION

It is recommended that the system be used for batches of similar files to allow for the best possible parameters to be used. For example, processing an entire drum kit worth of sounds in one batch would not produce ideal results, as different settings are appropriate for each sound in the drum kit. In the case of that example, best results would be seen if all the hi-hat samples were processed together as a batch, followed by a batch for snare drums and so on. Suggested generalized module settings are as follows:

- High band ITD time: 10ms
- Mid band ITD time: 0-10ms (transient blurring may occur)
- Low-Mid Crossover: 120Hz
- Mid-High crossover: 1000Hz
- Auto-fade time: 50 samples

MULTIBAND PROCESSING BACKGROUND

Central to the system's architecture is a linear phase band splitter. The band splitter modules are the first modules in the signal chain. They use Infinite Impulse Response filters to spectrally split the sound into three frequency bands for separate processing by subsequent effects modules.

Phase distortion is an inherent issue with IIR filter designs, and in multiband applications this can pose particular problems. Each filter will produce a different amount of phase distortion, and individual filers may feature different phase distortion across different frequencies (different group delay), and while this may not cause problems while listening to each of the bands individually, once their signals are summed together, the result will be audible comb filtering effects caused by the phase differences between them. It is critical that this be avoided. Therefore a linear phase solution is necessary.

In other words, all filters must have the same phase distortion for all frequencies. According to Linkwitz, (1976, p.6) for a two band system to accomplish this, the z-plane representation of the two filters must have identical poles, and the high pass must have zeros at complex frequency = 0. There must also be an even number of zeros to ensure there is no phase difference between the two bands at the crossover frequency.

An additional condition for the filter design for this application is that the filters must also sum to unity gain, with no peak or notch at the crossovers. Linkwitz (1976, p.6) states that the -6dB crossover amplitude needed for this flat summed frequency response necessitates that the poles be double poles. He points out that two cascaded Butterworth filters will satisfy all these requirements.

Therefore, each filter in this system is constructed from two 4^{th} order Butterworth filters that are cascaded to create what is known as a Linkwitz Riley filter, in this case of the 8^{th} order. The first Butterworth introduces phase distortion, and then signal is run through the second filter in reverse, which cancels the phase distortion of the first, resulting in a linear phase response but with a small delay introduced.

The result is a computationally efficient filter with a steep 48dB per octave roll-off, linear phase response and zero passband and stopband ripple.

The effectiveness of this system was demonstrated in the proof of concept prototype, which consisted of a two-band split. The frequency response of these two bands and their summed output is shown in Figure 1. The phase response of the prototype's low pass filter is shown in figure 2.

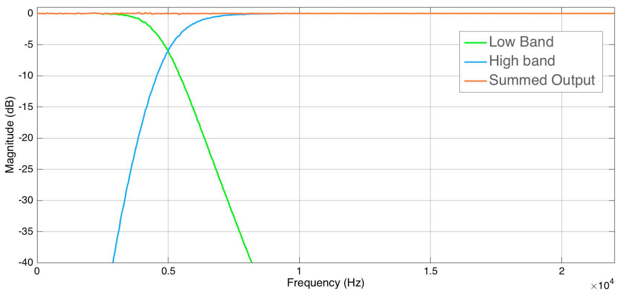

Figure 1 - Frequency response of prototype band splitter

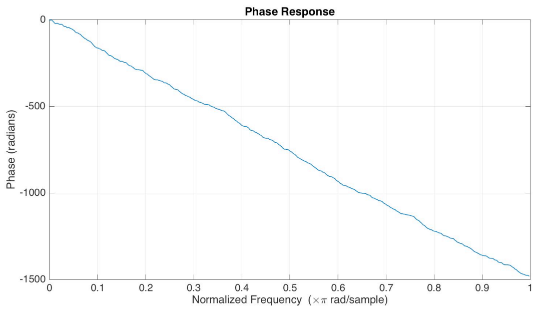

Figure 2 - Phase response of prototype low pass filter

The band splitting system in the final product will follow this same design, but will consist of three parallel blocks, each receiving the broadband signal as its input; a low-pass filter, a high-pass filter cascaded to a low-pass filter, and a high pass filter.

Since the mid band is being filtered twice, an additional delay will be introduced, however the length of this delay can be easily calculated, and a corresponding delay will be added to the low and high pass bands. Again, this ensures phase coherence between the bands.

After the band splitting stage, the resulting three bands will then be run through the rest of the signal chain shown in the system block diagram shown in figure 3.

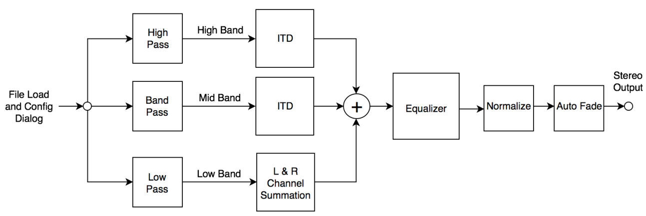

Figure 3 - System block diagram

MID & HIGH BAND – INTERAURAL TIME DIFFERENCE MODULE

The ITD modules take user defined ITD times and delays one channel of the input signal to create an inter-aural time difference. The psychoacoustic effect of this is a sense of width. This is a well-known technique for creating pseudo-stereo effects, and is based on the precedence effect demonstrated by Haas (1972). Figures 4 and 5 show visualization of the stereo image of a test signal before and after processing with ITD.

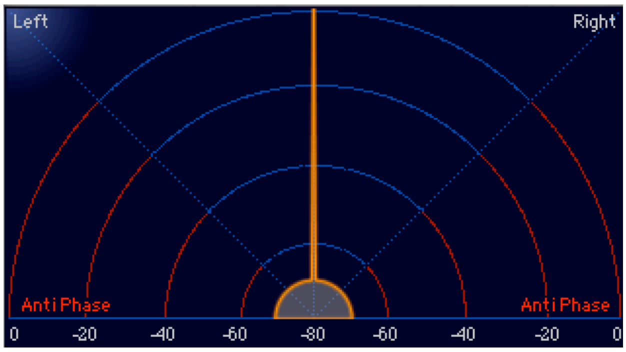

Figure 4 - Test Signal Before ITD Processing

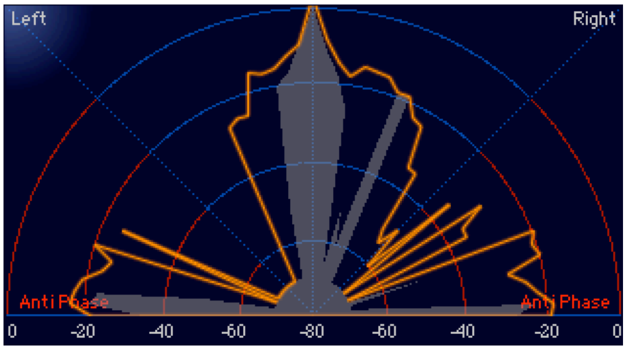

Figure 5 - Test Signal After ITD Processing

LOW BAND – CHANNEL SUMMATION MODULE

The channel summation module ensures that the information in both channels of the low frequency band is identical. This module operates in one of two ways. If the original input signal is monaural, the single channel is copied to channel two. If the signal is stereo, the two channels are summed and then gain scaled by a factor of 0.5. This gain scaling ensures that the amplitude of the signal remains at unity. The resulting signal is then copied to both channels of the module's output signal.

BAND SUMMATION

After the ITD and low-band channel summation modules, the resulting three signals are summed into the broadband signal. As there have been changes to the length of the signals introduced by the delays in the ITD modules, the ends of shorter signals are padded with zeros to ensure that all three are the same length.

BROADBAND EQUALIZER MODULE

This module has not yet been implemented, but it is intended to function as a sort of master cleanup equalizer. To remove unwanted high or low frequency information, thereby increasing amplitude headroom. It will consist of two Butterworth filters, a high pass and low pass, each with user definable slope from and cutoff frequency. The filter designs are identical to those used in the band splitter, so implementation requires a simple adaptation of existing code, with the addition of a user-defined slope selection. Each filter may be bypassed from the signal chain.

NORMALIZATION MODULE

This module that ensures that the absolute maximum amplitude of the signal is equal to 1, thereby ensuring that all output signals are as high in amplitude as possible, without clipping.

AUTOFADE MODULE

This module is simply designed to prevent output signals from having non-zero endings. It takes a user-defined fade-length in samples, *n*, and uses a recursive operation to calculate a decay curve *n* samples long, and ending at zero. The last *n* samples of the input signal are then multiplied by the values of the decay curve to scale their amplitude.

Figures 6 and 7 show the amplitude of a test signal with a non-zero ending before and after auto-fade processing.

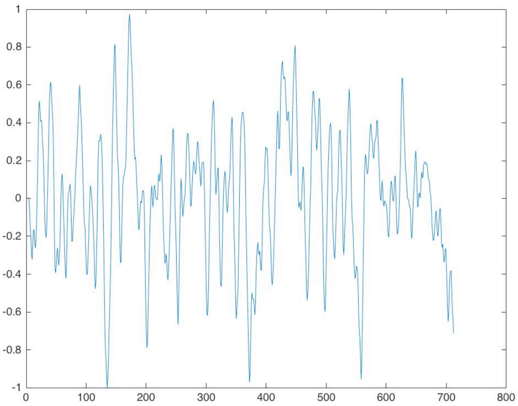

Figure 6 - Test signal with non-zero ending

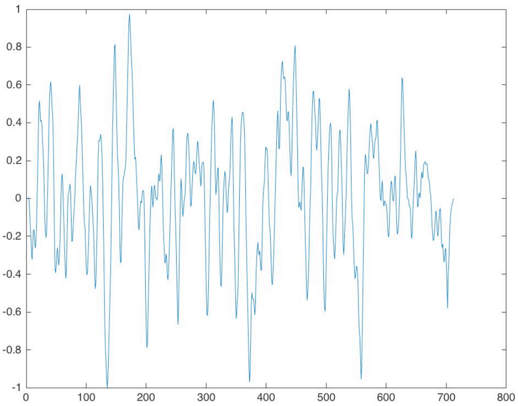

Figure 7 - Test signal after auto-fade processing

SUBJECTIVE TESTING

Subjective listening tests are proposed as a way to evaluate the effectiveness of the stereo image processing component of the system and to more accurately determine the range of ideal values for ITD. Testing would take place in an acoustically treated listening room. Conditions need not be anechoic, as this would not reflect normal listening conditions, however control of background noise, reflections and room modes would increase validity of the results. Particular attention should be paid to reverberation characteristics of the testing environment, as reflections would increase diffusivity of a monaural source, thereby giving it an artificially wide stereo image. Ideally, participants would be “expert listeners”, possessing some critical listening experience. Music and audio production professionals would be suitable, as they make up the target market for this product.

Listeners are to be seated at a computer, positioned at the focal point of two loudspeakers, and played a randomized sequence of test signals, some processed with the system, some not. The test signals would consist of various musical sounds, single snare drum hits and single piano notes for example, as well as broadband noise signals. To eliminate visual bias, the loudspeakers would be hidden from view behind an acoustically transparent curtain.

The first set of tests would consist of a absolute estimation exercise where listeners would be asked to indicate the perceived width of the low, mid and high components of each signal. The subject’s source width reports would be input by way of a software GUI featuring a diagram like the one shown in figure 8, and on-screen buttons to widen or narrow the dotted lines indicating where they perceive the boundaries of the source width to be.

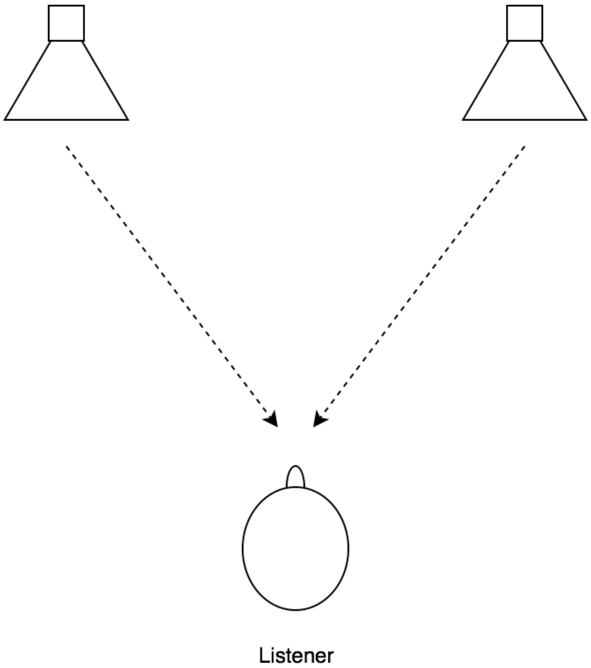

Figure 8 - Listening Test Input GUI

A second set of tests would also be conducted with the same testing conditions, but listeners would be played mono test stimulus and asked to click on-screen buttons to widen the mid and high band components of the signal until they find a setting for that stimulus that they find most pleasing.

REFERENCES

Linkwitz, S. H. (1976). Active Crossover Networks for Noncoincident Drivers. *Journal of the Audio Engineering Society*, *24*(1), 2–8.

Haas, H. (1972). The Influence of a Single Echo on the Audibility of Speech. *J. Audio Eng. Soc*, *20*(2), 146–159. Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=2093