

STEREO SPREADER REVIEW

Arturo Fernandez 420164911

Digital Audio Systems, DESC9115, Semester 1 2013
Post Graduate Program in Audio and Acoustics
Faculty of Architecture, Design and Planning, The University of Sydney

ABSTRACT

A stereo spreader is a very useful tool when trying to blend sounds of the same or similar frequency content and make them appear as if they were spread through the left and right speakers. The spreader function that is going to be described in this paper will vary a signal's time and pitch characteristics resulting in a stereo spread signal, it is going to point out the problems the code may have and how it can be improved.

1. INTRODUCTION

A stereo spreader effect also known as stereo widener make the stereo image of a sound appear wider within two speakers, therefore giving it more space within a mix of different sounds. This paper is going to explain the approach given to this effect and how it can be implemented, analyse the given results and suggest starting values for the possible user. Another algorithm will be analysed taking consideration of possible further improvements.

2. STEREO SPREADER

The input values of this function are delay L (ms), delay R (ms), pitch L, pitch R, and mix. A block diagram is provided so that the basic signal flow of the effect can be explained:

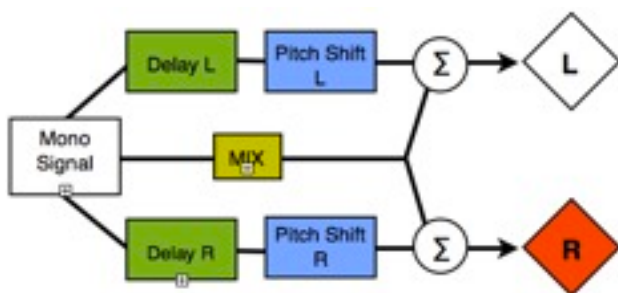


Fig. 1 Block Diagram of spread function

As shown in the diagram, the signal to be processed needs to be mono. This signal will be sent to the left and right channels assigning the specified delay to it. Then, it will go to the pitch shift and be processed with the input values, then the dry mono signal will be processed through the mixer and this will determine the amount of dry signal to be summed with the left and right channels, the output signal will be a stereo file. This effect was tested with different audio signals, a 1kHz sine wave, an acoustic guitar, vocal track, saxophone and a vocal track recorded through a telephone. The conclusions drawn to that was that in the case of harmonic instruments, the input values for the function would need to remain relatively low, as the pitch and time differences from the original signal and the

processed one would start creating phasing problems, not necessarily canceling the signal, but would create comb filtering as well as perceptual differences from where the sound was coming from, in the specific case of the acoustic guitar, the sound would feel like it was jumping from one speaker to the other, and never really settle in a specific space over time. In the other hand, melodic instruments have higher tolerance for high values and when the widening effect threshold is breached, the effect acquires harmonising qualities, which can be used to create vocal harmonies, of even making an instrument deeper, for example a snare that cannot cut through a mix can be processed in a way that there is a slap-back delay with a lower tuned snare.

Some of the starting values to get good results out of the effect might be setting the left and right delays somewhere around 15 and 30 milliseconds, then for the pitch values a good start would be -2 cents all the way up to -8 cents or +2 to +8, if the result of this is to fatten the sound then those values would be fine, otherwise calculating the actual pitch of the song in relation to the pitch shift could be used to create some sort of harmonic coherence.

3. CONSIDERATIONS FOR FURTHER IMPROVEMENT OF THE CODE

According to the analysis done in the audio files, some important considerations need to be done with this effect. As mentioned before, depending on the desired product the effect's threshold varies, and basically works better in small values for some sounds and works well in higher values for different sounds. Considering the effects this effect has on the phase alignment of the audio file, and how this changes the perception dramatically then I decided to analyse a similar effect and point out some of the improvements based on this other algorithm.

3.1 Virtual Sound Algorithm for Wide Stereo Sound Stage

The output of this effect is the widening of a stereo signal for use in two dimensional layouts. The main design implements three different FIR Filters; binaural synthesis through FIR filter convolution, crosstalk canceler in a form of high-order IIR Filter to maximise cancellation performance and a direct filter. The output of this algorithm is a 2x2 matrix (stereo signal).

To preserve timbre and centre image, the original signal is played back through the direct filter, this filter matches the time delay and output level. All three filters are later simplified into a low order FIR filter (panorama filter) in a 2 by 2 FIR filter matrix giving easy computation and a simple structure. (Kim, S. 2004.)

Following is an explanation of how each filter is structured, the math behind it and how they all sum together.

Binaural Synthesis:

HRTF's are convolved with the input audio data recreating a virtual sound image. This filter would result in two widely equally spread sources in front of the listener, left channel of binaural signal is synthesised and sent to the left speaker, and the process is repeated for the right speaker. This is represented by the equation in fig. 1

$$\begin{bmatrix} L_L(z) & R_L(z) \\ L_R(z) & R_R(z) \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^N L_{Li}(z) & \sum_{i=1}^N R_{Li}(z) \\ \sum_{i=1}^N L_{Ri}(z) & \sum_{i=1}^N R_{Ri}(z) \end{bmatrix}$$

Fig. 1 Binaural Synthesis Mathematical Representation

$L_{Li}(z)$ and $L_{Ri}(z)$ represent the HRTF's from the designated virtual source on the left side to left and right ear. To simplify the computation, instead of doing different processes (four filters and four convolutions), the layout is built within a 2 by 2 matrix which will distribute the values in a more efficient way. (Kim, S. 2004.)

Crosstalk Canceler:

The widening effect is played at its full capabilities when using headphones, in the downside if the signal is played back through loudspeakers, the crosstalk phenomena changes the results. The crosstalk phenomena refers to the audio signals coming from the left speaker delaying in the listener's head and arriving late to the right ear, this happens equally to the left ear. This late signals give spatial cues to the main signal, therefore giving a false representation of what the resulting sound should be. In order to solve the crosstalk phenomena a crosstalk cancellation filter is applied to the algorithm;

$$C(z) \cong -\frac{\sum_{i=1}^R \sum_{k=1}^{\delta} \frac{c_i}{b_i^k} z^{-(\delta-k)}}{D_{\min}(z)} \begin{bmatrix} H_{22}(z) & -H_{12}(z) \\ -H_{21}(z) & H_{11}(z) \end{bmatrix}$$

Fig.2 Crosstalk Canceller equation

The design of this function will be explained next. The direct inversion method is used. This method is represented by the following matrix:

$$C(z) = \frac{1}{D(z)} \begin{bmatrix} H_{22}(z) & -H_{12}(z) \\ -H_{21}(z) & H_{11}(z) \end{bmatrix}$$

Fig. 3 Cross-cancellation matrix

In this matrix, the direct sound coming from the left and right speaker is represented by $H_{11}(z)$ and $H_{22}(z)$ accordingly. $D(z)$ would be the determinant (fig.3), it defines what signal is being cancelled in the left and right sources.

$$D(z) = H_{11}(z)H_{22}(z) - H_{12}(z)H_{21}(z)$$

Fig. 4 Determinant

In order to simplify the inverse process, the determinant is separated into delay by using the root-finding algorithm and

extracting the maximum phase subsystem($D_{\max}(z)$), doing this by comparing the absolute values of the roots to 1. (Kim, S. 2004.)

Direct Filter Design

This filter helps the original signal preserve the timbre and centre channel image. This filter is represented by the following equation:

$$A(z) = az^{-\alpha}$$

Fig. 5 Direct Filter

The complex variable is represented by z , α is the time delay and a the output level.

After all the filters are in place there is a minimisation technique used to convert the three different filters into low order FIR filters. What this minimisation technique does is adding the binaural synthesis and the crosstalk canceller creating a widening filter (Fig.6) and then adding this filter with the direct filter creating a "Panorama Filter" (Fig. 7) (Kim, S. 2004.)

$$W(z) = \begin{bmatrix} C_{11}(z) & C_{12}(z) \\ C_{21}(z) & C_{22}(z) \end{bmatrix} \begin{bmatrix} L_L(z) & R_L(z) \\ L_R(z) & R_R(z) \end{bmatrix}$$

Fig.6 Widening Filter

$$P(z) = \begin{bmatrix} W_{11}(z) + A(z) & W_{12}(z) \\ W_{21}(z) & W_{22}(z) + A(z) \end{bmatrix}$$

Fig. 7 Panorama Filter

The output stereo signal is represented by y_L and y_R :

$$y_L = P_{11}(z)L + P_{12}(z)R$$

$$y_R = P_{21}(z)L + P_{22}(z)R$$

Fig. 8 Output stereo signal

4.

DISCUSSION

The only main issue in the original spreader function is the phase coherence. The other differences such as the pitch shift for harmonics is a good parameter to have so that the effect can have drastic changes but at the same time keep phase coherence. In order to achieve this, a good implementation of the "Virtual Sound Algorithm" is to use the direct filter in the original design of the stereo spread function, this way allowing the audio to be have the phase correction needed and avoiding the unusual discrepancies in the spatial placement of sounds caused by this.

5.

CONCLUSION

Given the results of the first spreader function, a straightforward way of achieving the effect desired is clearly pointed out in the discussion, by adding the direct filter to the original design of

this function. This direct filter gives stability to the phase coherence of the code, but at the same time keeping the ability to acquire the harmonic qualities to create different layers of pitched up/down sounds.

6. REFERENCES & BIBLIOGRAPHY

Kim, S. 2004. Virtual Sound Algorithm for Wide Stereo Sound Stage. In Audio Engineering Society. San Francisco, CA, October 28-31, 2004. New York: JAES. 1-6.

Smith, S. 2013. Stuart Smith . [ONLINE] Available at:<http://www.cs.uml.edu/~stu/>. [Accessed 27 May 2013].

Zölzer, U, 2011. DAFX: Digital Audio Effects. 2nd ed. West Sussex, United Kingdom: Wiley & Sons.