Crosstalk Cancellation Over Loudspeakers Using The RACE Algorithm

Charles Craig Jr.
cjc694@nyu.edu

Christal Jerez
cj713@nyu.edu

Jason Rostowski
jgr266@nyu.edu

Tom Xi
qx244@nyu.edu

ABSTRACT

This paper summarizes the thought process behind our group's choosing to implement a well documented method of reproducing multichannel sound content via stereophonics in Max/MSP: Recursive Ambiophonic Crosstalk Elimination (RACE). As has been discovered, discussed, and validated in numerous scholarly papers, the replication of binaural audio over loudspeakers can be realized through the use of a mixture of adequate crosstalk cancellation algorithms coupled with individualized Head-Related Transfer Functions (HRTFs). [2,3] Because of its ability to recursively cancel cross-talked (XTC) signals to inaudibility without relying upon HRTFs, the RACE algorithm benefits from reduced frequency response errors and the undesirable coloration of tonal areas. [3] A Max/MSP patch which implements XTC over two loudspeakers through the use of the RACE algorithm is introduced and described within the paper.

1. INTRODUCTION

Immersive experiences can be thought of as those indistinguishable from their inspiration. For example, one's reaction to an in-game AI companion's exclamation," Watch out!", through the fixed visual orientation and LR speaker setup of a monitor does not belie the same degree of experiential immersion they'd experience if their friend were to yell a similar phrase during a leisurely stroll through a busy city intersection. Reacting to the AI's warning requires that the listener respond to sonic and visual information emanating from a fixed location. On the contrary, reacting to the friend's warning necessitates a swift visual and aural surveillance of the landscape in order to locate the origin of the warning. Because humans do not experience the world from a fixed point-of-view, our auditory systems come equipped with a slew of physical and neurological behaviors which enable the gathering and processing of important environmental information.

The reproduction of sound through stereophonics (two or more channels) does not provide the fidelity of spatial information required by technologies tasked with the creation of immersive experiences. Although audio signals can have spatial characteristics transposed onto them through convolution (modification through combination) with databased or personalized HRTF's, they cannot be adequately represented over speakers.

In order for each channel of a stereo playback system to present spatialized content, their channel information must not interact with that of the other. Avoiding this interaction is achieved through the use of XTC filters. Essentially, these filters should invert, attenuate and delay each channels signals, then send them to interact with the opposite channels signal. Because a successive order of XTC's is required to make the XTC signal inaudible, the RACE algorithm [3,5] can be used to produce higher order XTC's until the cancellation signal becomes inaudible.

The group was motivated to implement the RACE algorithm in order to gain experience with building XTC filters.

2. METHODOLOGY

In the same manner as that which was described earlier, an out of phase XTC signal was superimposed onto the original signal in order to eliminate crosstalk between the two speakers. The RACE algorithm was utilized because of the need for the crosstalk signal to be recursively cancelled out to avoid adverse spectral coloration. Its use resulted in a wider sonic image, [3] which is evident in the Max/MSP patch.

3. IMPLEMENTATION

Because of it's flexibility and familiarity, the RACE algorithm was implemented in Max/MSP. The patch prioritized the execution of the three main components of the algorithm: 180° signal inversion, attenuation ratio, and the delay length of the left and right channels. The recursive loop was established by feeding each delayed line into the opposite channel, meaning the delayed right signal was added to the left signal, and vice versa. As described in the earlier sections, this recursive loop is at the heart of all cross-talk cancellation systems.

In order to achieve optimal efficiency, the RACE algorithm was written inside of a gen~ object. This allowed for the delay of samples greater than or equal to one. Lastly, the delay length and attenuation ratio were programmed as user controlled variables, allowing users to try out their own individualized customization.

The Max/MSP patch was tested in the NYU research lab. A binaural recording was used as the source audio to better determine the stability of the implementation method. Two Genelec 8010 loudspeakers were placed at approximately one foot apart (the distance was measured from speaker cone to cone).

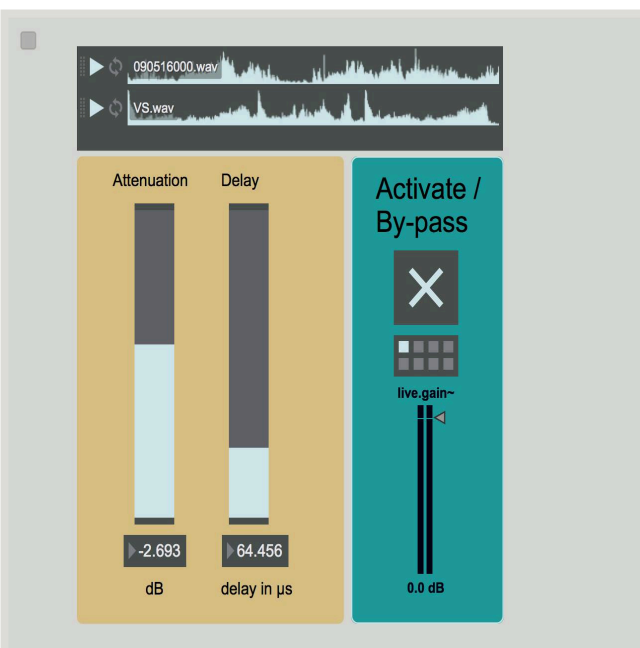

Figure 1 - User Interface

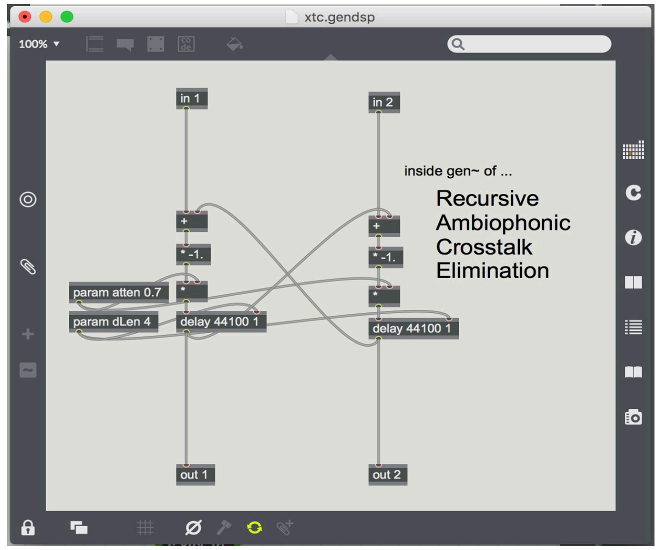

Figure 2 - Inside the RACE Algorithm

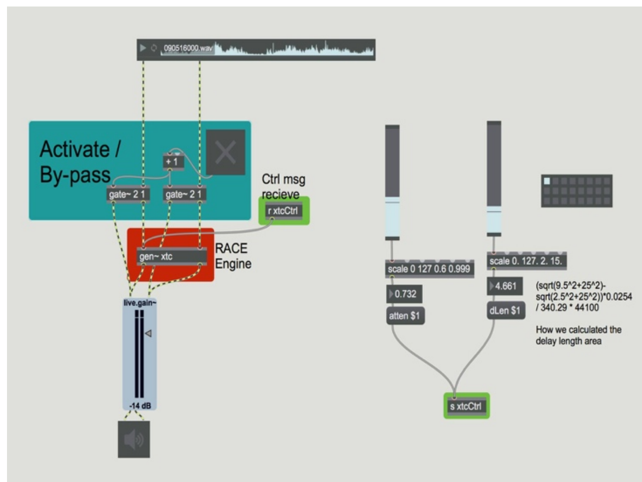

Figure 3 - Max/MSP RACE Algorithm Patch

4. RESULTS

After adjusting the user controlled variables, our implementation of the RACE algorithm resulted in a perceptibly widened soundstage. We were able to hear a correlation between the size of the soundstage as the attenuation and delay amounts were altered, despite Glasgal's contrary opinions.

5. CONCLUSION

In an attempt to further understand XTC systems, as well as attempt to improve upon the existing methods, an alternative RACE algorithm was built in Max/MSP. It combined the Ambiophonic method with existing HRIRs from each loudspeaker position. In this set-up, the HRIRs were convolved with the signal prior to reaching the RACE algorithm in the gen~ sub patch. Although noticeable spatial enhancement was perceived and documented, further research would need to be done for proper implementation.

6. REFERENCES

[1] E.Y. Choueiri, “Optimal Crosstalk Cancellation for Binaural Audio with Two Loudspeakers” Princeton university, pp. 1-24.

[2] M.R. Schroeder, D. Gottlob, K.F. Siebrasse, “Comparative study of European concert halls: correlation of subjective preference with geometric and acoustic parameters*” *Drittes Physikalisches Institut*, Universitat Gottingen, Republic of Germany, 1973, pp. 1195-1201.

[3] R. Glasgal, “360° Localization via 4.x RACE Processing” *Audio Engineering Society Convention Paper 7301*, Presented at the 123rd Convention, New York, NY, 2007, pp. 1-11.

[4] R.B. Schulein, D. Mapes-Riordan, “The Design, Calibration and Validation of a Binaural Recording and Playback System for Headphone and Two-Speaker 3D-Audio Reproduction” *Audio Engineering Society Convention Paper 9130*, Presented at the 137th Convention, Los Angeles, USA, 2014, pp. 1-16.

[5] T.S. Wu, “Listening with Realism: Sound Stage Extension for Laptop Speakers” *Master of Music in Music Technology Thesis Submission*, 2013, pp. 23-25.

[6] T.S. Wu, A. Roginska, R. Glasgal, “Ambidio: Sound Stage Width Extension for Internal Laptop Loudspeakers” *Audio Engineering Society Convention Paper 9077*, Presented at the 136th Convention, Berlin, Germany, 2014, pp. 1-8.