"WARM CHORUS": RE-THINKING THE CHORUS EFFECT USING AN ORCHESTRAL SECTION MODEL

Richard Dudas
Hanyang University College of Music
222 Wangsimni-ro, Seongdong-gu
133-791 Seoul, South Korea

ABSTRACT

The "warm chorus" effect described in this paper was initially designed for a string quartet concert in which a simulated string orchestra effect was requested by the performers. Due to the fact that the various "industry-standard" chorus effects which were available to us at the time left a lot to be desired from a musical standpoint, it was decided to create a tailor-made effect for the concert, using the Max/MSP environment, which could be used subsequently as a generic real-time chorusing tool. In order to overcome the mechanical quality inherent in many generic chorus algorithms, the "warm chorus" uses a model based on the performer arrangement in an orchestral string section. This novel approach inspired by human behaviour is further enhanced with additional signal processing techniques to help reduce the perceived "phasiness" of the processed sound, resulting in a chorus effect that is more natural than traditional off-the-shelf chorusing effects.

1. INTRODUCTION

The quest for a more realistic chorusing effect appropriate for instrumental (rather than exclusively vocal) sound sources was initiated while collaborating with the New Asia String Quartet in 2007 [11]. The quartet was performing George Crumb's *Black Angels* alongside Haydn's *Seven Last Words*, Op.51, which additionally exists in a prior orchestral version and a later choral version by the composer. Since we were already amplifying and adding artificial reverb to the instruments for the performance of the Crumb piece, the members of the quartet thought it would be appropriate to also process the Haydn work in order to create a larger, ensemble-like sound, thereby alluding to the other versions of the piece.

Our unanimous dissatisfaction, however, with the credibility of the chorusing effects available to us in both hardware and software, coupled with the apparent lack of recent research into chorusing *instrumental* sound, led the author to embark on the creation of a dedicated string orchestra simulation effect for the concert. Since the objective was to create a rich string orchestra sound, the author (also an amateur string player) decided to use the player arrangement of an orchestral string section as an architectural model for the algorithm. The original version of the resulting "warm chorus" was first used in concert in the Spring of 2007, and has since undergone several updates, modifications and improvements.

2. THE TRADITIONAL CHORUS EFFECT

Artificial chorusing effects can be traced back to the early days of electronic music [1] [8]. These effects were implemented electronically and can still be found today in many analog guitar pedals and external effect units, due to their having henceforth furnished a quintessential (and sometimes clichéd) "sound" within many musical genres.

2.1. The Basic Chorus

In the realm of audio digital signal processing, the most basic chorus effects imitate their analog predecessors and are thus implemented with a series of modulated taps on a single delay line summed together to for a chorused output (see Figure 1) [3] [10]. Although this processing scenario is computationally efficient, it has a several drawbacks including unwanted notch filtering throughout the spectrum and a perceived mechanical regularity to the resultant output sound, due to the Low Frequency Oscillators (LFOs) used for modulation.

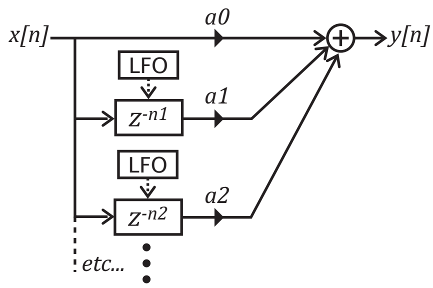

Figure 1. Block diagram of a traditional chorus effect.

2.2. The Industry Standard Chorus

The "industry standard" version of this effect attempts to compensate for the unwanted comb filtering by basing the chorus, paradoxically, on a comb filter — albeit one that includes separate delay-line taps for its feedforward and feedback loops (see Figure 2) [4]. This important modification, along with using the inverse of the signal in the feedback loop, both improves the thickness of the sound and mitigates to some extent the perceived "phasiness" in the upper end of the spectrum. Two or more of these units may be used in parallel for added sound density, although the result still has a mechanical affect more closely resembling its analog forebears rather than a genuine ensemble performance.

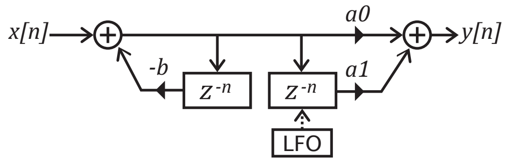

Figure 2. The "industry standard" chorus.

2.3. Pitch Synchronous Overlap Add Techniques

More recently, chorusing algorithms based on Pitch Synchronous Overlap Add (PSOLA) techniques have been implemented [9]. Using pitch-synchronous techniques tends to reduce the unwanted artefacts found in traditional chorusing algorithms, though they are not eliminated completely. Some PSOLA-based chorusing algorithms require pre-calculation of the input sound, limiting their use to non-real-time situations, but the main drawback from a user's perspective — even in the real-time implementations — is that, due to the pitch-synchronous nature of the algorithm used, they require a monophonic input source for proper pitch analysis and waveform segmentation.

2.4. Spectral Techniques

Some spectral-domain techniques for chorusing have also been explored in recent years. The most successful of these use cross-synthesis techniques to combine a solo voice and the sound of a pre-recorded chorus [2]. Alas, the spectral-domain techniques have hitherto been primarily limited to the modification of vocal sounds; to extend them to the processing of instrumental sounds would otherwise require pre-recording an appropriate actual instrumental ensemble, something which is not particularly difficult, but not always practically feasible.

3. AN ORCHESTRAL SECTION MODEL

In contrast to traditional chorusing algorithms, which for the most part attempt to design the effect on computational terms, the design of the "warm chorus" is fundamentally based on a human-inspired model: that of the placement and comportment of performers in an orchestral string section. The idea is that, in a string section, the first stand will probably have performers who are the best players in the section; they are the leaders for those sitting behind them and will likely play with the best intonation, have more accurate timing and, being more confident performers, will play slightly more prominently. Additionally, the sound of these players will be the least delayed and subjected to air absorption since they are seated closest to the audience. The players in the back of the section may not be such proficient players as those in front of them, so we presume that with each subsequent stand, their transposition from ideal intonation will be farther than that of the players in front of them, and their timing may be slightly more lax. Since these players are also likely to be more timid or uncertain about the accuracy of their playing, their volume will be consequently lower on the average, but have a larger range of variation than those players ahead of them. Finally, the sound of each stand will be slightly filtered and delayed due to air absorption and distance from the audience.

Although this is a gross over-simplification of the make-up of an orchestral section, and does justice neither to performers, nor to conductors and artistic directors (neither of whom would congregate all of the weakest players in the back of the section), it does, notwithstanding, provide a simple and effective human-based model upon which to design a chorus effect in much the same way that the general description of the dimensions and materials of a concert hall may offer a simple, effective model for a reverberation algorithm.

4. DESGNING THE "WARM CHORUS"

Instead of being built upon the traditional modulated delay line, the "warm chorus" uses a set of time-domain granular harmonizers with optional input delay, whose transpositions are set to different minute (generally in the range of 2 to 40 cents) values. Although this is significantly more computationally expensive than traditional algorithms, it more accurately models a multiple-player scenario. Using such a time-domain transposition method is considerably cheaper than using spectral-domain pitch shifting, and the somewhat "blurry" quality of the pitch-*asynchronous* overlap-add of the harmonizers actually helps achieve a thicker overall sound with fewer transpositions (and less of a regular "woozy" vibrato-like feeling arising from sinusoidally modulated delay taps). The granular harmonizers used for the first version of the "warm chorus" were simply taken from the standard Max/MSP harmonizer example, *transposer~*[1], but subsequently a new 4-overlap granular harmonizer with built-in allpass filtering was designed for the purpose.

4.1. Harmonizer Structure

The modified harmonizer used for the chorus effect is designed with four sequential overlapping sinusoidal windows each 90 degrees out of phase from the previous. These windows are controlled by a single phasor~ object[2] which simultaneously synchronizes the read locations of the four delay taps and the windowing functions, and acts as a trigger signal for four sample-and-hold units. The delay line read locations have some degree of random variance that is held constant for the duration of the window; this helps create irregular phase cancellation and reinforcement in the upper end of the spectrum. A set of four 8th order allpass filters inserted before summing the overlapping windows further contributes to shifting phases throughout the spectrum, so any beating that occurs after summing the closely-

[1]The MSP transposer~ example is an implementation of the earliest and most basic granular digital pitch shifters which, in turn, were modeled on earlier analog rotary tape head devices See [1] and [8] for a detailed history.

[2] The phasor~ object in Max/MSP generates a periodic ramping function from 0 to 1.

spaced transpositions will not have an obvious regularity. Using a lower-order allpass filters saves only minimal CPU, and did not appear to be as effective upon repeated subjective listening.

This harmonizer unit, diagrammed for the sake of clarity in Figure 3, is not very useful for the purposes of general audio transposition (where 2-overlap seems to be both more practical and produce better results). Nonetheless, for the purposes of a chorusing algorithm it works quite well: the transposed sound is stable within the -100 to +100 cent transposition range and contains some necessary degree of spectral complexity resulting from irregular delicate amplitude modulation in different regions of the spectrum.

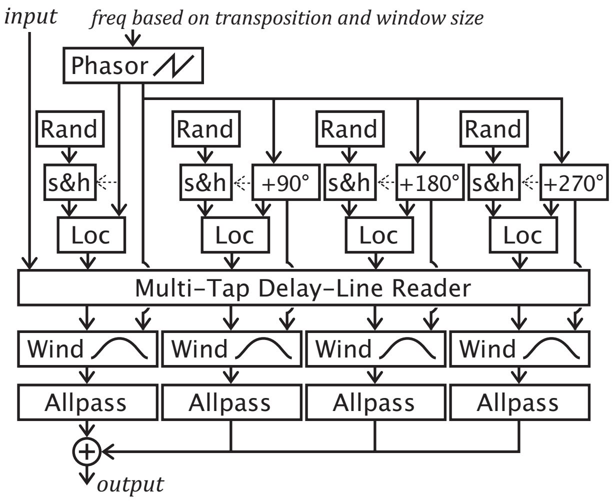

Figure 3. Chorus Harmonizer Block Diagram

4.2. Harmonizer Intonation

It is worth mentioning that, quite unlike an LFO-modulated delay tap —which produces alternating sharp and flat transpositions of the input sound — all of the transposition values used for the harmonizers are positive values (i.e. sharp). Young musicians in student orchestras are often told that sharply tuned notes are perceived to be less "out of tune" than flat ones. Whether or not this is true may be open to debate, but research on intonation has shown that professional musicians have an overall tendency toward higher-pitched intonation for the majority of both equally-tempered and just intervals [6], so this sharp-tuning scenario has been used in the "warm chorus" effect.

4.3. Additional Time-Domain Processing

The output of the more greatly detuned harmonizers is subjected to slight lowpass filtering to simulate air absorption in an exaggerated manner. This also mitigates eventual brightening of the upper end of the spectrum when multiple transpositions are mixed. The volume of each of the transpositions is then attenuated with a slow, randomly varying ramp function. Lower amplitudes and wider variance are given to transpositions that have been more greatly detuned. This amplitude variance helps to keep any one transposition from audibly dominating the chorused "scene," and to keep the chorusing effect from having an obvious regular pulse. Additional frequency-shifting and allpass filtering of each of the harmonizers' outputs were also considered and tested, but both were deemed to be of limited or inconsequential benefit.

4.4. Magic Numbers

The relationships of the transposition values, as well as the relationships of delay times, window sizes and random variance of the various parameters, are generated using numbers from the prime number series. This was intuitively decided upon in order to further help reduce perceptually regular beating, modulation and phasing between multiple closely-spaced transpositions, delays and window durations. It seems to be fairly effective. The parameters for the allpass filters in the harmonizer units are randomly set, empirically. For the moment this works satisfactorily, although it is certainly an area for future research and improvement.

4.5. Diffusion and a Second Stage

Before summing the harmonized output, a second stage of transpositions is effectuated upon the first. This effectively means that each of the 8 signals is transposed again 8 times, resulting in a total of 64 individual harmonizations of the original input. Instead of simply summing the signals, and putting them through another shared multiple-tap delay line, a Fast Walsh-Hadamard Transform is used for summation and diffusion. This unitary transform allows the outputs of the first stage to be combined in different sum and difference combinations each retaining a constant output volume. Although this transform is often used for data reduction and encryption, solving differential equations and removing crosstalk, it has also been shown to be an effective means of diffusion for feedback loops in reverberation algorithms [5]. The addition of the transform precludes using a shared delay line for the harmonizers in the second stage, but the richness of the resulting chorused sound is greatly improved.

4.6. Spectral-Domain Processing

After summing and attenuating the eight outputs of the second stage of transpositions, the original and chorused sounds are fed into a Short Term Fourier Transform (STFT) in order to eliminate some of the unwanted brightness, "phasiness," and amplitude modulation resulting from the time-domain part of the algorithm. First, the phase spectrum of the chorused sound is separated from its amplitude spectrum, and the phases are "locked" using a simple summing technique between adjacent frequency bins. This technique has been used effectively when reconstructing phases within a phase vocoder algorithm [7], and in this context also seems to help reduce brightness and unwanted phasing in the upper partials of the chorused spectrum.

The amplitudes of the original sound are then used to stabilize disconcerting amplitude modulation in the lower end of the spectrum. This is achieved by non-linearly cross-fading the original and chorused amplitudes across the frequency spectrum. Smoothing the amplitudes from one analysis frame to the next was also tried, but proved to impose an unnatural quality to the resulting sound. The processed amplitudes and phases are then recombined to produce a final chorused output. A diagram of the entire "warm chorus" algorithm is shown in Figure 4.

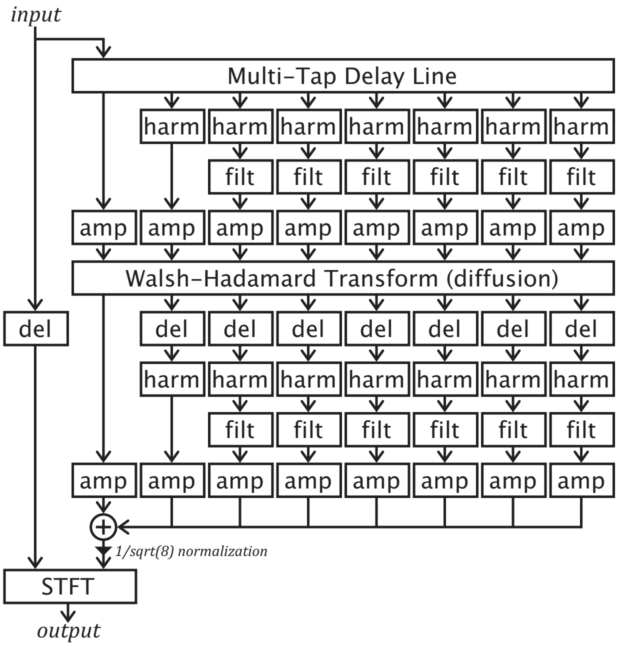

Figure 4. General Structure of the "Warm Chorus."

5. ADVANTAGES AND DRAWBACKS

The resulting effect is aurally convincing, though computationally expensive. Processing can be reduced, and the result is still passable if only 4x4 (instead of 8x8) transpositions are used. It is possible that 3 stages of 4 harmonizers could be used to achieve 64 transpositions, but this has yet to be tested. The other major drawback is a fair amount of input/output latency due to the inclusion of the STFT.

On the positive side, the principal advantage of the algorithm is that can accept any kind of input sound — monophonic, polyphonic, noisy or reverberant — and the processed result is equally effective. The "warm chorus" is also easy to use: it is fitted with a few simple user controls to adjust the intonation, sloppiness and fullness as appropriate for different types of audio input.

6. CONCLUSION

There is further work to be done refining this algorithm, but it is nonetheless musically effective in its current form. Above all, the author believes its novelty lies in the underlying orchestral section and human behavioural model. This will hopefully provide inspiration for re-thinking the design of other traditional audio effects.

7. ACKNOWLEDGMENTS

The author would like to thank violinist Eui-Myung Kim for instigating the project, as well as composer and colleague Jongwoo Yim for his help and advice during the initial stages of conception and development.

8. REFERENCES

[1] Bode, H., "History of electronic sound modification," *Journal of the Audio Engineering Society*, Vol. 32, No. 10, New York, NY, USA, 1984.

[2] Bonada, J., "Voice Solo to Unison Choir Transformation," *Proceedings of the Audio Engineering Society 118th Convention*, Barcelona, Spain, 2005.

[3] Chamberlin, H., *Musical Applications of Microprocessors*, Second Edition, Hayden Books, Rochelle Park, NJ, USA, 1985.

[4] Dattorro, J., "Effect Design, Part 2 : Delay-Line Modulation and Chorus," *Journal of the Audio Engineering Society*, Vol. 45, No. 10, New York, NY, USA 1997.

[5] Jot, J.-M., "Efficient models for reverberation and distance rendering in computer music and virtual audio reality," *Proceedings of the International Computer Music Conference*, Thessaloniki, Hellas, 1997.

[6] Kopiez, R, "Intonation of Harmonic Intervals: Adaptability of Expert Musicians to Equal Temperament and Just Intonation," *Music Perception*, Vol. 20 No. 4, University of California Press, Berkeley, CA, USA, 2003.

[7] Puckette, M., "Phase-Locked Vocoder," *Proceedings of the 1995 IEEE ASSP Conference on Applications of Signal Processing to Audio and Acoustics*, Mohonk, NY, USA, 1995.

[8] Roads, C., *The Computer Music Tutorial*, MIT Press, Cambridge, MA, USA, 1995.

[9] Schnell, N., Peeters, G., et al., "Synthesizing a Choir in Real-Time Using Pitch Synchronous Overlap Add (PSOLA)," *Proceedings of the International Computer Music Conference*, Berlin, Germany, 2000.

[10] Smith, J. O., *Physical Audio Signal Processing for Virtual Musical Instruments and Digital Audio Effects*, W3K Publishing, USA, 2010.

[11] Yim, J., Dudas, R., "New Music Media at Hanyang University: Studio Report," *Proceedings of the International Computer Music Conference*, Copenhagen, Denmark, 2007.