SYNTHESIZING THE DYNAMIC SPECTRA OF THE DIDGERIDOO

Lydia Ayers

Tower 14, Flat 7A, HKUST
Clear Water Bay, Kowloon
HONG KONG

Andrew Horner

Department of Computer Science
HKUST
Clear Water Bay, KLN, HK

ABSTRACT

This paper describes a timbre morphing design that simulates the dynamic spectra of the didgeridoo. The design captures many of the instrument's subtle timbral and expressive characteristics based on an underlying additive synthesis model in Csound. The design is very flexible, and allows vocal sounds, harmonic sweeps and flutter tonguing.

Figure 1. Ash Dargan Playing the Didgeridoo.

1. INTRODUCTION

Researchers have previously studied the acoustic properties of the didgeridoo (see Figure 1) [6, 14] and morphed the sound of a double bass and a didgeridoo [10]. In this project, we have modelled the didgeridoo using timbre morphing in Csound [13], which captures many of its subtle timbral and expressive characteristics. This paper gives some background for the didgeridoo and then describes the source material we used in this research. We discuss the features of the timbre morphing design which we use to model the dynamic spectra of the didgeridoo, and, finally, we examine musical expression on the didgeridoo, including vocal sounds and harmonic sweeps.

2. THE DIDGERIDOO

The didgeridoo is an Australian buzzed lip instrument traditionally made from a tree trunk that has been hollowed out by termites [9]. Every tree is uniquely shaped, and the meandering termite pathways in the walls of the instrument help determine the character of its sound. Some instruments have a large bell at the end of the tube where the tree root flares out which adds a rich color to the sound. The instrument is played with the lips buzzing into one open end of the tube. The lips are much looser than for the brass instruments, making the sound quite rich and deep. The didgeridoo plays a drone tone colored with vocal sounds and harmonics. This paper examines synthesis of representative sounds produced on a didgeridoo with a tone pitched at B1 (62.5 Hz).

3. THE SOURCE MATERIAL

We worked from CD recordings by Ash Dargan (Larrikia tribe, Darwin, NT, Australia), *Ash, Dust & Dirt* [3] and *Didgeridoo Made Easy: A Beginner's Guide* [4]. (We obtained his permission to analyze timbres from his CDs to create the designs for this research. We have also corresponded with him regarding the results of this research.) Because we wanted to use the resulting design in a composition, we started with "Wangurra" (sleep) [3], one of the pieces that used a didgeridoo in B1, a key that fit the composition. We then chose a collection of interesting sounds to analyze for the instrument model.

4. SYNTHESIZING THE DIDGERIDOO

The additive synthesis model from our previous work [2] was a good starting point, but since the didgeridoo plays basically a single fundamental drone tone and its harmonics, colored by constantly-changing vowel sounds, we found that we could not use the average spectrum to model the didgeridoo as effectively as we have modeled other wind instruments. Instead, we focused on morphing together a string of timbres, rather than slurring a group of pitches.

The following discussion will consider a 2.344 second group of three sounds using the voice to produce "ah-ee-oo too wah," from the beginning of "Wangurra" [3]. A spectrogram [8] of the segment shows a harmonic sweep at the beginning, which sounds like the vowels "ah-ee-oo," followed by a percussive "t" at the beginning of "too" and then a "wah" (see Figure 2).

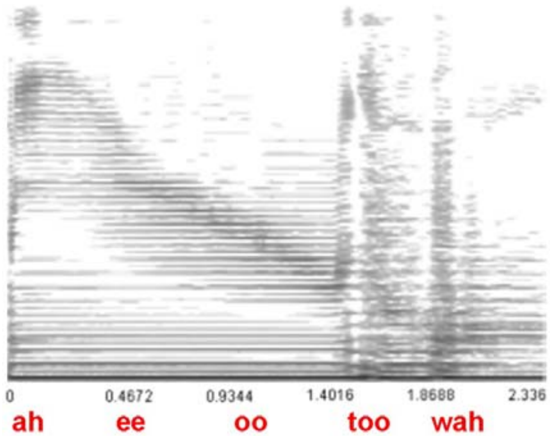

Figure 2. Spectrogram of Original Didgeridoo B1 "ah-ee-oo too wah" Sound Segment.

Figure 3 shows a time-varying phase vocoder analysis [5, 15] of the segment. The harmonics sweep from the high harmonics at the lower left to the lower harmonics at about time 1.5, when the noisy "t" interrupts the sound, followed by the bright spectrum of the "w" in the "wah."

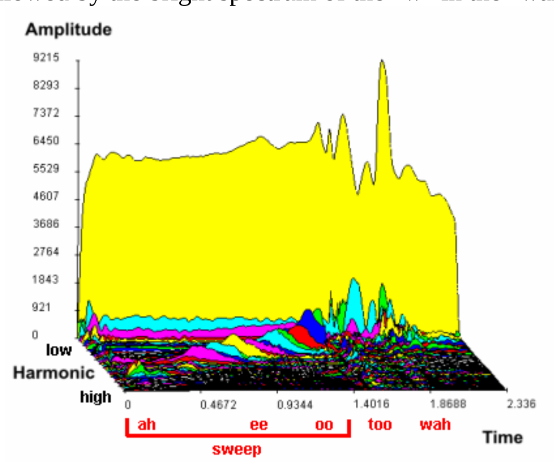

Figure 3. Phase Vocoder Analysis of Original Didgeridoo B1 "ah-ee-oo too wah" Sound Segment.

We divided this sound segment into three sounds, a harmonic sweep (the changing vowel sound) and the syllables "too" and "wah." We describe the two syllables in the following sections, as they provide good illustrations for how we modelled the didgeridoo. We modelled the harmonic sweep as an expressive technique and describe it in Section 5, *Musical Expression on the Didgeridoo.*

4.1. Modelling the "Too" Sound

The vocal "too" syllable has two phonemes, the noisy "t" in the attack and the "oo" vowel, which is most of the sound. We took two snapshots of this sound (see Figure 4), one for the "t" at time .035 and one for the "oo" at time .267. We chose times from the amplitude peaks in the sections of the waveform that represented the phonemes.

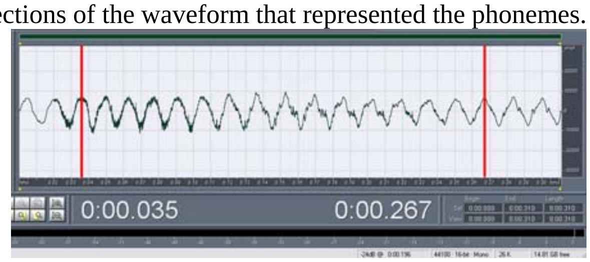

Figure 4. Snapshot Selection from Didgeridoo B1 "Too."

4.2. The Spectral Snapshots

After choosing the snapshot times, we examined the phase vocoder analysis graphs (Figures 5-6).

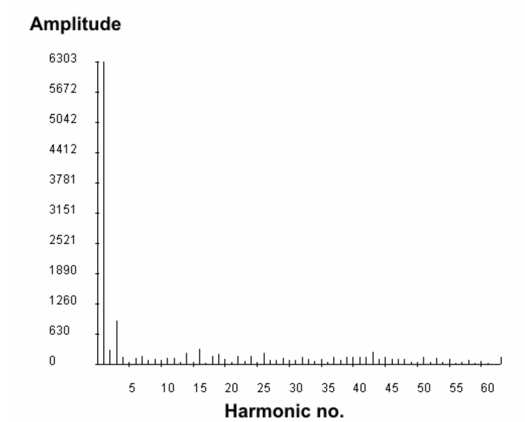

Figure 5. Spectral Snapshot of "t" at Time .035.

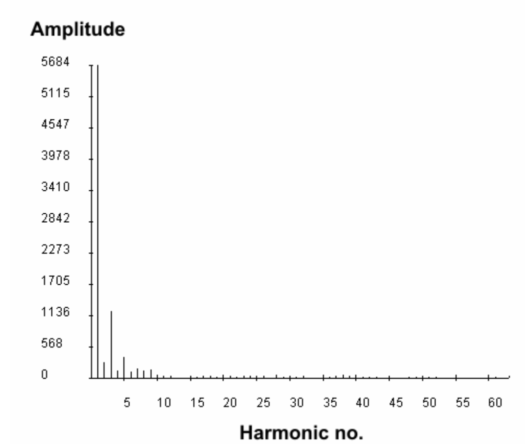

Figure 6. Spectral Snapshot of "oo" at Time .267.

We stored the amplitude data from the graphs in two Csound score function tables. We reserved the number at index 0 in the table for an amplitude at an inharmonic frequency (see the next section, *The Wah Sound*).

F51 0 128 -2 h0(v) h1 h2 h3 …
F101 0 128 -2 h0(v) h1 h2 h3 …

Randomization in the additive synthesis orchestra ensures that repeated performances of the same sound have varied spectra in order to make the performance expressive. Using this method with one spectral snapshot is unlikely to capture the full quality of the noise in the “t” sound, but the result is sufficient for our musical purposes, and the “oo” part does sound reasonably like the original sound.

4.3. The Wah Sound

The “wah” sound contains two phonemes, a short “oo” for the “w” and an “ah.” We chose the snapshot time for the “w” in the unstable beginning at time .049 seconds (see Figure 7), and the time for the “ah” near the end at time .659 seconds (see Figure 8) because this part of the sound becomes richer as it begins to morph into the next sound, another harmonic sweep.

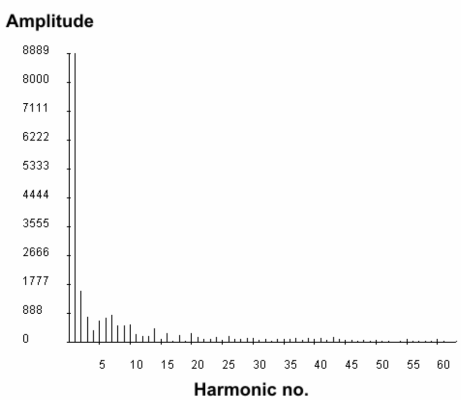

Figure 7. Spectral Snapshot of "w" at Time .049.

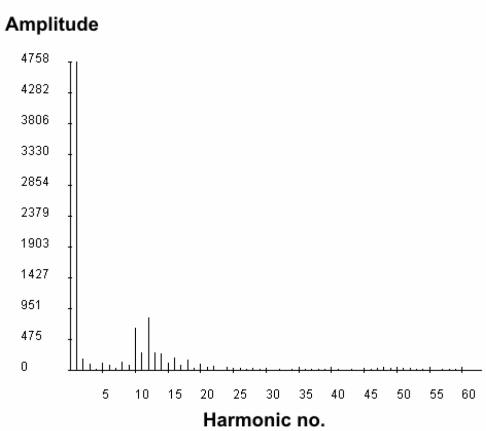

Figure 8. Spectral Snapshot of "ah" at Time .659.

The waveform analysis [7] summary showed that this sound has an extra inharmonic frequency of 174.5 Hz, which is 13 Hz flatter than the third harmonic of 187.5 Hz (see Figure 9).

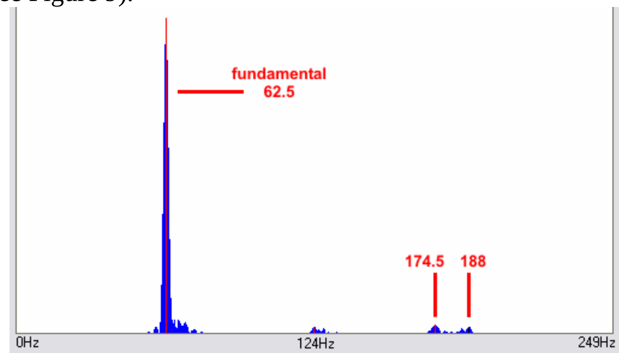

Figure 9. Waveform Analysis of Original Didgeridoo B1 "wah."

Interestingly, however, although the summary doesn’t show it, the waveform analysis graph shows the expected harmonic at about 188 Hz and its amplitude is only slightly weaker than the amplitude of the harmonic at 174.5. We added the extra harmonic into the function table at index 0, after calculating its amplitude in relation to the amplitude of the fundamental. We set a slightly varying frequency of about 170-175 for this harmonic to model the natural frequency variations of a performer’s voice. The synthesized “wah” sounds more like the original sound than the synthesized “too.”

5. MUSICAL EXPRESSION ON THE DIDGERIDOO

Didgeridoo players use a variety of vocal imitations of animal sounds in their performances [4, 9, 12]. The imitation of the buffalo is a particularly gritty sound which uses a combination of vocalization and flutter tongue (typically accomplished by gargling in the back of the throat). The amplitude modulation that we used in the previous wind instrument designs produces a close approximation to this sound [1]. This section describes vocal sounds and harmonic sweeps, expressive techniques that increase the musical potential of the design.

5.1. The Harmonic Sweep

The sound segment begins with a vocal sound similar to a smooth owl-like “who” descending from about B4 to E4. The didgeridoo resonance pulls the voice toward the didgeridoo harmonics similar to the stepped effect a sweeping band-pass filter produces. The resulting sound is more or less similar to the phonemes, “ah,” “ee” and “oo.” We divided the sound into one pair of slices for each of these phonemes. After producing a tone by morphing the six sounds, we added a band-pass filter sweeping from the

40th harmonic to the 20th harmonic. This method successfully captures this effect.

5.2. Morphing the "Ah Ee Oo Too Wah" Segment

Morphing the sounds together produced the combined original segment, "ah-ee-oo too wah." In our previous additive synthesis instruments, line segments connected the note parameters for slurs and morphed the transitions. In the case of the didgeridoo, the code for the slur feature is useful in connecting sound segments of different timbres. Each harmonic is one component signal for the spectrum, with its own frequency and amplitude line segments for the group of morphed notes.

5.3. Multiple Harmonic Sweep

Another harmonic sweep moves up and down twice (see Figure 10). For musical use of the sweeping band-pass filter effect, it was enough to set the envelopes determining the harmonic sweeping without concern for the exact original vowel sounds. That is, the effect was pretty similar no matter which original sound we applied it to. We incorporated the sweeping band-pass filter with up to 5 (mostly) randomly-determined moving segments.

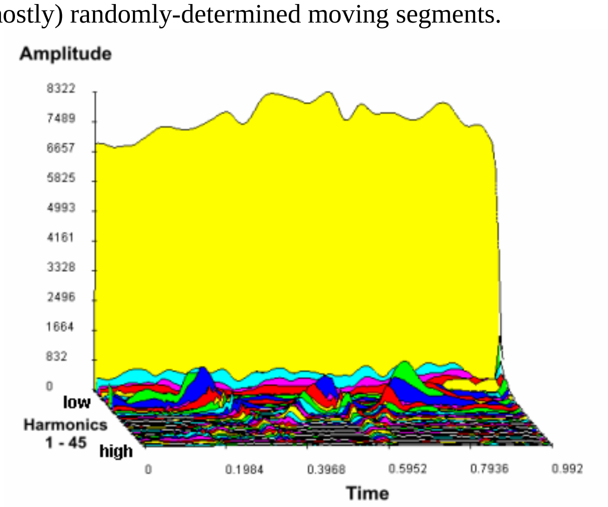

Figure 10. Time-Varying Phase Vocoder Analysis of Multiple Harmonic Sweep on Original Didgeridoo B1 Sound.

6. CONCLUSION AND FUTURE WORK

We used timbre morphing based on additive synthesis to model the dynamic spectra and musical expression of the didgeridoo. The design captures the subtle characteristics of many of its timbres, including vocal sounds, harmonic sweeps and flutter tonguing. Future work includes extending the model to synthesize more vocal sounds combined with didgeridoo sounds, along with modelling didgeridoos in other keys.

7. ACKNOWLEDGEMENTS

We would like to express our appreciation to Ash Dargan for giving permission to use his photo and to analyze timbres from his didgeridoo recordings to create the designs for this research. We also appreciate his input in personal correspondence regarding the results of this research. This work was supported by the RGC Research Grant 613508.

8. REFERENCES

[1] Ayers, L. "Synthesizing Timbre Tremolos and Flutter Tonguing on Wind Instruments," in *Proceedings of the International Computer Music Conference*, Miami, USA, 2004.

[2] Ayers, L. and Horner, A. "Synthesizing the Chinese Suona," in *Medi(t)ations, Australasian Computer Music Conference* 2006, Adelaide, Australia, pp. 15-22.

[3] Dargan, A. "Wangurra" (Sleep) on *Ash, Dust & Dirt*. Indigenous Australia, 1999.

[4] Dargan, A. *Didgeridoo Made Easy: A Beginner's Guide*. Indigenous Australia. 2001.

[5] Dolson, M. 1986. "The Phase Vocoder: A Tutorial," *Computer Music Journal*, 1986, 10:4.

[6] Fletcher, N.H. "Acoustics of the Australian Didgeridoo," Australian Aboriginal Studies, 1.

[7] Hibbert, B. *Bell Waveform Analysis Program*. Great Brookham, Surrey, UK. www.hibberts.co.uk.

[8] Horne, R. *Spectrogram*. Visualization Software. www.visualizationsoftware.com/gram.html.

[9] Hudson, D. *Making and Playing the Didgeridoo*. Indigenous Australia. 1A2021DVD.

[10] Kuchera-Morin, J. *Paleo* for Double Bass and Computer Generated Tape, Discordia Music, Chicago, Illionois.

[11] Neuenfeldt, K. (ed.) *The Didgeridoo: From Arnhem Land to Internet*. John Libbey & Co., Sydney, NSW, Australia.

[12] Schelberg, D. *Didgeridoo: Ritual Origins and Playing Techniques*. Binkey Kok Publications, Diever, Holland.

[13] Vercoe, B. Csound: A Manual for the Audio Processing System and Supporting Programs with Tutorials. MIT Media Lab, Cambridge, MA. 1992.

[14] Wiggins, G.C. "The physics of the didgeridoo," *Physics Bulletin*, 39. 1988.

[15] Wun, S. *PVan*. www.cs.ust.hk/~simon/pvan. 2000.