ANALYSIS-RESYNTHESIS OF MUSICAL TONES USING SPECTRAL MODELLING TECHNIQUE

[1]Ong Bee Suan & [2]Minni Ang
Music Department
Faculty of Human Ecology
Universiti Putra Malaysia
43400 UPM Serdang
tel. 03-9486101 ext 1738
Fax. 03-9161689
[1]beesuan@music.upm.edu.my
[2]minni@music.upm.edu.my

ABSTRACT

Analysis-resynthesis is a modelling technique for the generation of musical tones. It looks at the spectrum of a real musical instrument and tries to recreate it through the application of additive and subtractive resynthesis algorithms. It encompasses different techniques that have a three-step process in common: firstly, a recorded sound is analysed, next, the musician modifies the analysis data, and finally, the modified data is used in resynthesising the altered sound. The sound of a musical instrument is contributed primarily by the addition of the harmonic partials that occur in the sound waveform. The analysis of a recorded sound is performed through the estimation of its power spectrum. The input waveform is windowed, with window widths based on the amplitude envelope parameters, obtained through analysis of the amplitude of the waveform with respect to time. The Fast Fourier Transform [FFT] is then applied to these discrete sections to obtain a spectral model of the sound signal. This data is then used for the resynthesis of the original sound. The modelling technique is tested using digitally generated waveforms with known spectral components and amplitude envelopes. Results indicate that the method developed is able to return the results expected, and may be used for the analysis and resynthesis of real musical instrument sounds.

Keywords: *Music synthesis; digital tone generation; computer music; digital music; spectral modelling; music synthesis techniques.*

1.0 Introduction

Digital sound synthesis is a type of sound signal generated using computers. Digital sound synthesis generates a stream of numbers representing the samples of the sound signal waveform. The first experiments in digital sound synthesis began in 1957 by researchers at Bell Telephone Laboratories in Murray Hill, New Jersey [1,2,3]. The concept of *unit generators* has been used in digital sound synthesis. Unit generators are signal-processing modules such as oscillators, filters, and amplifiers which can be interconnected to form synthesis instruments or patches that generate sound signals.

There are many different general methods for the synthesis of musical instrument sounds. Some of the techniques have been built by emulating the function of the acoustics of musical instruments, rather than of the sounds themselves. Sometimes, these techniques are also based on the mechanics of a natural sound production process. Some techniques use samples of the instrument. Some other techniques look at the spectrum of a real instrument, as the harmonic overtones present in the spectrum are caused by the modal vibrations of the instrument's vibrating system. These modes affect the tone quality or timbre of the instrument's sound.

Analysis-resynthesis is a modelling technique, which looks at the spectrum of a real instrument and tries to recreate it. It encompasses different techniques that have a three-step process in common. These are:

(i) A recorded sound is analysed.
(ii) The musician modifies the analysis data.
(iii) The modified data is used in resynthesising the altered sound

This modelling technique is based on either additive or subtractive resynthesis. Besides that it also can based on the combination of these two techniques. [4,5].

The general overview of analysis-resynthesis is shown in Figure 1 below.

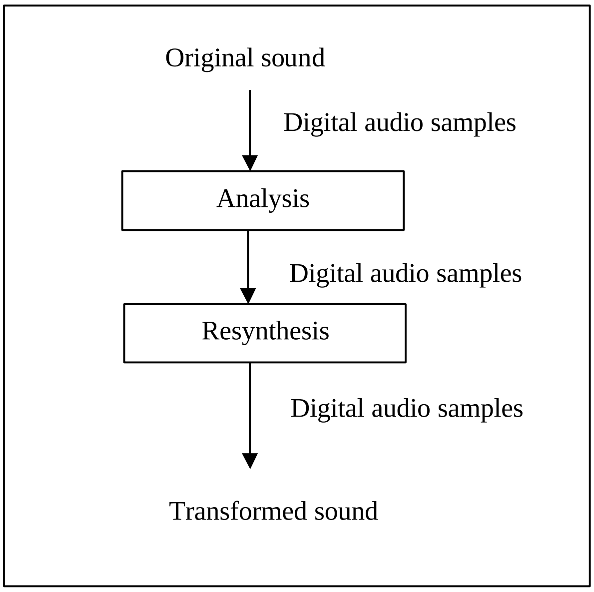


Figure 1. Overview of the Analysis-Resynthesis Technique.

This paper presents the results of the testing of the techniques involved in the analysis stage, which produces the data used to create the model of the sound spectrum used for resynthesis purposes. It focuses on the approach used to extract the data pertaining to the harmonic partials of the spectrum. Test waveforms with known harmonic partial components are digitally generated and analysed using the specified approach. The analysis data is used to regenerate the tone. This verifies the usability and reliability of the methodology, for the analysis of unknown spectra of real musical instrument sounds. A sample real musical instrument tone is then analysed and resynthesised using the approach tested.

2.0 Literature Review

Xavier Serra and Julius Smith developed spectral modelling synthesis, in 1990. [6]. The spectral modelling synthesis technique is a set of techniques and software implementations for the analysis, transformation and synthesis of musical sounds. This technique aims to get general and musically meaningful sound representations based on analysis, from which musical parameters might be manipulated while maintaining high quality sound. [7]. These techniques employ parameters that describe the sound spectrum, regardless of the acoustic mechanisms that may have produced them. Spectral modelling synthesis technique is developed through Fourier analysis, which considers a pitched sound to be made up of various sinusoidal components, where the frequencies of the higher components are integral multiples of the frequency of the lowest component. [8]. The advantage of these techniques compared with plain sampling is that musicians can manipulate these coefficients in a variety of ways, in order to create new sounds such as sound morphing, which can be achieved by varying the coefficient. [8].

2.1 Spectrum Analysis

Spectrum analysis is important for spectral modelling, as samples alone do not provide information about the spectral constituents of a sampled sound. There are two categories of spectrum analysis. These are harmonic analysis and formant analysis. Harmonic analysis focuses on the identification of the frequencies and amplitudes of the spectrum components, while formant analysis uses the estimation of the overall shape of the spectrum's amplitude envelope.

2.2 Additive Synthesis

Additive synthesis is one of the oldest and most heavily researched synthesis techniques, based on the summation of elementary waveforms to create more complex waveforms. This technique is accepted as the most powerful and flexible spectral modelling technique. It was among the first synthesis techniques in computer music [9]. It was described extensively in the very first article of the very first issue of the Computer Music Journal [10]. Additive synthesis technique assumes that any periodic waveform can be modelled as a sum of sinusoids at various amplitude envelopes and time-varying frequencies. Basically, it puts together a number of different wave components together, which can be partials or harmonics, to arrive at a particular sound [11]. Figure 2 shows the additive synthesis function by summing up sinusoids in order to form specific waveforms. Additive synthesis allows more control than any other kind of synthesis, as it permits fine control over individual frequency components [12]. Moreover, a synthesis may interpolate between the frequency spectra from two or more different sounds [13].

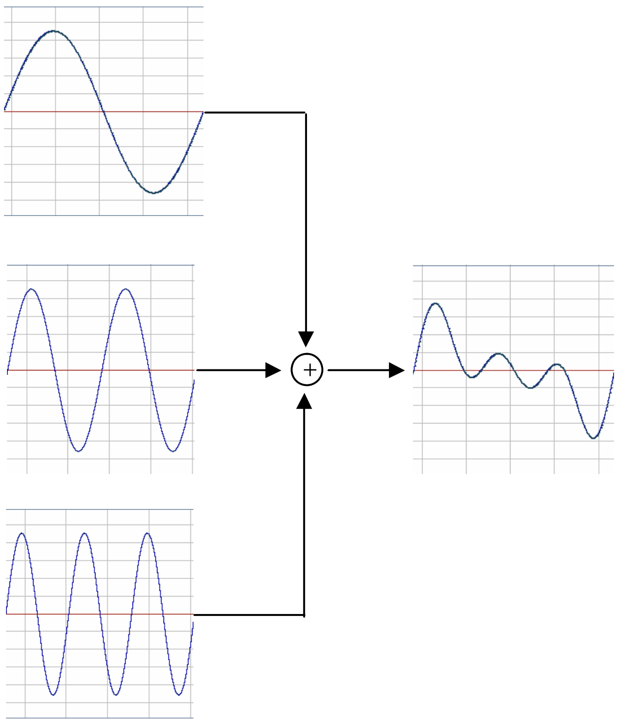


Figure 2. Additive synthesis function by summing up sinusoids in order to form specific waveform.

2.3 Subtractive synthesis

Subtractive synthesis implies the use of filters to shape the spectrum of a sound source by subtracting unwanted partials of its spectrum, while favouring the resonation of others. As the source signal passes through a filter, the filter boosts or attenuates selected regions of the frequency spectrum. If the original source is spectrally rich and the filter is flexible, subtractive synthesis can shape close approximations of many natural sounds, as well as a wide variety of new and unclassified timbres. [2]. This technique has been used successfully to model percussion-like instruments and the human voice. [8].

3.0 Methodology

The analysis of a sound, to identify the harmonics that occur in the sound signal, is performed through the estimation of its power spectrum. Samples of musical tones are analysed using the Short-Time Fourier Analysis, to determine the time-varying frequency characteristics. The analysis is carried out on short segments of the input signal, through a technique called windowing, with window widths based on the amplitude envelope parameters, obtained through analysis of the amplitude of the waveform with respect to time. After that, each windowed segment is sent through a bank of narrow bandpass filters, where every filter is tuned to a specific centre frequency. The Fast Fourier Transform (FFT) is then applied to these discrete sections to obtain a spectral model of the sound signal. This data is then used for the resynthesis of the original sound. The modelling technique is tested using digitally generated waveforms with known spectral components and amplitude envelopes.

3.1 Testing an artificial digitally generated tone

Firstly, a mathematical formula, for example,

$$y = a_1 \sin(\omega_1 t) + a_2 \sin(\omega_2 t) + \ldots + a_n \sin(\omega_n t)$$

is designed. With a sampling rate of 44100 Hz and 16-bit resolution (which gives CD quality sound), this formula is programmed into a computer to generate a tone waveform. The Fast Fourier Transform (FFT) is applied to the waveform to estimate the power spectrum. The frequency components of y obtainable in this way is up to a maximum of the Nyquist frequency (half of the sampling rate frequency). In order to obtain a_1, a_2,…., a_n as in the formula above, the power spectral density level (in decibel units) is converted back into relative amplitudes of the component frequencies in the waveform. Through this method, it is shown that this technique is reliable.

3.2 Analysing a real musical instrument tone

A real musical instrumental's tone is recorded in an acoustically dry room, with the use of a digital recorder and a condenser microphone. The waveform of the musical tone is transferred into a computer, and the background noise is filtered off. After that, the waveform is divided into four portions based on its ADSR amplitude envelope.

3.2.1 ADSR amplitude envelope

A simple ADSR amplitude envelope is shown in Figure 3 below. It shows the changes in the amplitude of a note over its duration.

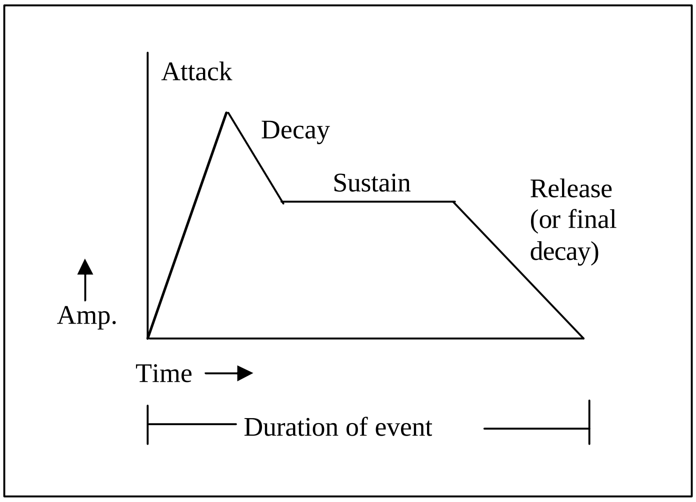


Figure 3. A simple ADSR amplitude envelope

The ADSR amplitude envelope has 4 portions. These are the attack, during which time the amplitude rises from zero to a peak value; the decay, when the amplitude falls off from the peak value to the sustain level; the sustain; and the release, during which time the amplitude falls from the sustain level to zero level.

Each portion is divided into even smaller segments, a technique called *windowing.* The Fast Fourier Transform (FFT) is applied to each of the segments to obtain each frequency component's power spectral density. Finally, the result is converted to obtain the relative amplitude of each frequency component.

4.0 Findings

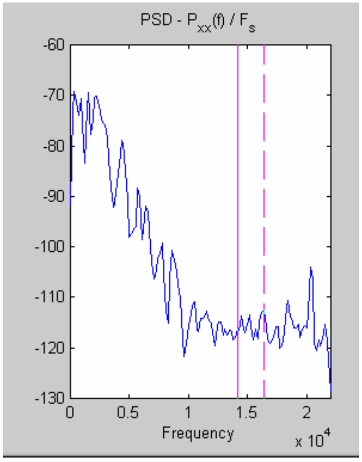
Through the use of digitally generated tones, it is shown that the phase spectrum has no effect on the timbre of the sound. This is proved through interchanging some sine functions from the formula above with cosine functions of the same relative amplitude values. Visually, these two waveforms are different, but aurally it makes no difference to the sound. As the difference between a cosine function and a sine function is just the phase ($\pi/2$ radian), the phase spectrum has no effect on the timbre of the sound.

Table 1 below shows the occurrence of partials at various times of a real-life violin sample. It is seen that even through a small change in time, such as from T1 = 0.00 to 0.02 seconds and T4 = 0.06 to 0.08 seconds, the number of partials present and their relative strengths (as indicated by the power spectral density value) does change.

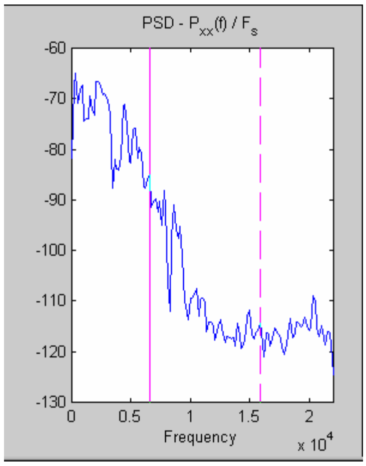
Table 1. Partials present at different times, T1 to T4

T1 = 0.00 - 0.02 s		T2 = 0.02 - 0.04 s		T3 = 0.04 - 0.06s		T4 = 0.06 - 0.08s	
Partial	Power Spectral Density	Partial	Power Spectral Density	Partial	Power Spectral Density	Partial	Power Spectral Density
1	-69.1373	1	-65.0844	1.5	-63.1140	1.5	-56.7533
2.5	-70.6058	2.5	-67.3060	2.5	-56.5064	2.5	-52.2947
4.5	-69.3875	3.5	-73.8939	4.0	-62.5164	4.0	-60.0379
6.5	-70.1501	4.5	-69.4108	5.0	-69.1281	5.0	-66.2307
12.5	-78.8482	6.5	-66.5868	6.5	-56.2584	6.5	-51.4873
16.5	-88.4921	8.0	-68.8855	9.0	-66.0217	9.5	-64.7921
18.5	-91.8721	10.5	-81.9172	11.5	-70.1657	11.5	-82.5957
22.5	-99.1548	13.0	-71.0968	13.0	-69.6122	13.0	-67.9170
25.0	-100.6654	15.5	-75.8836	15.5	-75.2744	16.5	-73.0206
30.5	-110.8789	16.5	-79.5662	16.5	-75.9718	18.0	-77.4029
32	-114.1680	19.0	-85.1778	18.0	-84.9477	19.0	-75.6573
		20.5	-89.8884	19.0	-78.6361	20.5	-75.9249
		21.5	-90.1700	20.5	-80.0141	21.5	-78.2444
		22.5	-88.0939	21.5	-84.9975	23.0	-81.1998
		25.0	-91.0499	23.0	-81.9226	24.5	-86.2425
		26.5	-95.1761	25.5	-85.2931	25.5	-84.3226
		30.5	-107.4556	29.0	-104.5531	28.0	-108.4071
				30.5	-100.652	29.5	-101.7282
						30.5	-97.0943
						32.0	-97.7667
						34.0	-107.5029
						37.0	-100.5544

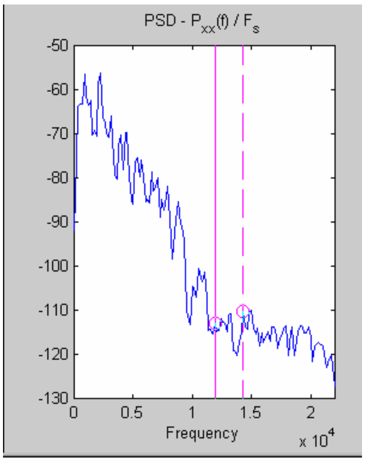
These results indicate that the waveform spectrum varies with time even over very small segments of less than 0.1 seconds (Figure 4). The FFT is also shown to be a viable method of obtaining the partials present in a real-instrument tone. This data can then be used for resynthesis purposes.



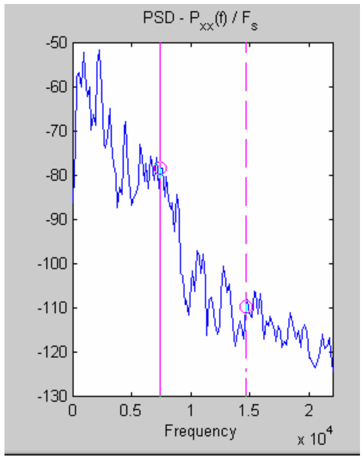
T1=0.00 – 0.02 s



T1=0.02 – 0.04 s



T1=0.04 – 0.06 s



T1=0.06 – 0.08 s

Figure 4. Changes in Power Spectrum Density and Partials versus Time.

5.0 Discussion

There are some non-harmonic partials identified in the violin tone under study. This is possibly the contribution of residual noise and transient sound from the instrument, such as sound from the scraping of the string of the violin, the bow and others. These non-harmonic partials were identified using the Fast Fourier Transform applied to the waveform. These non-harmonic partials occur simultaneously with harmonic partials. When reproducing the waveform using only harmonic data without including the non-harmonic partials, the sound quality of the tones generated resembles the original sound but lacks authenticity. At the same time, without considering the contribution of the time varying partial components present in the original waveform, the sound which is reproduced does not sound real.

6.0 Conclusions

The FFT applied to real-life musical instrument tones is able to provide detailed analysis of partials present. These partials, both harmonic and non-harmonic, are shown to vary in time. The results of the analysis may be used for resynthesising the original waveform, or through the application of transformation functions, for the creation of new sounds.

References

[1] David, E., M. Mathews, and H. McDonald. 1958. "Description and the result s of experiments with speech using digital computer simulation." *Proceeding of the 1958 National Electronics Conference.* New York: Institute of Radio Engineers. Pp. 766-775. Roads, C. (1998). "The Computer Music Tutorial". Cambridge: The MIT Press.

[2] Wood. (1991). "Recollections with John Robincon Pierce." *Computer Music Journal* 15(4): 17-28.

[3] Roads, C. 1980. "Interview with Max Mathews." *Computer Music Journal* 4(4): 15-22. Reprinted in C. Roads, ed. 1989. *The Music Machine*. Cambridge, Massachussetts: The MIT Press. Pp. 5-12.

[4] Serra. (1989). "A system for sound analysis/transformation/synthesis based on a deterministic plus stochastic decomposition." Stanford: Center for Computer Research in Music and Acoustics, Department of Music, Stanford University.

[5] Serra, X. & Smith (1990). "Spectral Modelling synthesis: a sound analysis/synthesis system based on a deterministic plus stochastic decomposition." *Computer Music Journal* 14(4): 12-24.

[6] Vaggione. (1996). "Articulating Microtime". *Computer Music Journal*. 20(2): Pp.33-38

[7] Serra, X. (1998). *Spectral Modelling Synthesis.* [Online]. Available: http://www.iua.upf.es/eng/recerca/mit/sms/docs/spectral-models/section5.html

[8] Miranda, E.R. (1998). "Computer Sound Synthesis for the Electronic Musician". Oxford: Focal Press.

[9] Serra, X. (1997). *Musical Sound Modeling with Sinusoids plus Noise.* [Online]. Available: http : // www. iua. upf. es/ eng/ recerca/ mit/ sms/ docs/ msm/ background. html

[10] Moorer, J. A. (1977). "Signal Processing Aspects of Computer Music-A Survey." *Computer Music Journal.* 1(1): Pp. 4-37.

[11] Gould, C. (1997). *Beginner's Synthesizer Faq*. [Online]. Available: http://tilt.largo.fl. us/faq/synthfaq.html

[12] Fitz, K., and L. Haken. (1996). "Sinusoidal Modeling and Manipulation Using Lemur". *Computer Music Journal*. 20(4): Pp. 44-59.

[13] Dodge, C., and T. Jerse. (1985). *Computer Music*. New York: Schirmer Books.

[14] Pellman, S. "An Introduction to the Creation of Electroacoustic Music". Wadsworth Publishing Company.