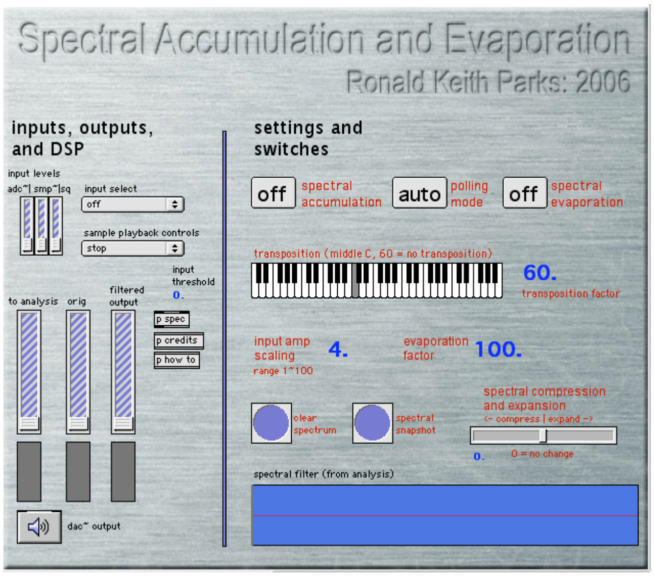
Real-time Spectral Attenuation Based Analysis and Resynthesis, Spectral Modification, Spectral Accumulation, and Spectral Evaporation; Theory, Implementation, and Compositional Implications.

Ronald Keith Parks, Ph.D., Winthrop University (parksr@winthrop.edu)

Abstract: Building upon convolution-based EQ (Settel and Lippe 1997 rev. 2001) spectral analysis data is utilized to attenuate FFT bins (derived from an FFT analysis of noise) to create an FFT/IFFT-based subtractive analysis/resynthesis module. Techniques for modification of analysis data prior to resynthesis, producing a variety of effects, are examined and demonstrated. Methods for retaining information from previous analysis (spectral accumulation) and for systematic data attrition (spectral evaporation) are introduced. A MaxMSP graphic user interface, designed by the author for implementation of the techniques, is discussed and described. Compositional implications are examined and musical examples are utilized to illustrate potential musical applications.

†

1. Spectral Attenuation-Based Analysis and Resynthesis.



Analysis/resynthesis models have historically been oriented toward utilization of Fourier analysis of an audio signal in order to deconstruct the spectra to its component sine waves. Subsequently, the frequency and amplitude information gleaned for each partial from the analysis is distributed to a bank of oscillators for additive-based resynthesis (Lippe, 1996). Once the spectral data is acquired, a variety of modifications may be applied prior to resynthesis (Settel and Lippe, 1994). However, alternate methods of resynthesis may also be employed. This paper describes an analysis/resynthesis technique in which Fourier analysis is combined with FFT/IFFT-based spectral attenuation. Also addressed are some of the intrinsic techniques for modification of analysis data prior to resynthesis.

Building upon the analysis/oscillator bank approach to analysis/resynthesis, attenuation based analysis/resynthesis also employs Fourier analysis of the original audio signal to obtain the frequency and amplitude of the most significant peaks in the harmonic spectrum. In the current implementation analysis is achieved via the MaxMSP fiddle~ object[1] (Puckette, 1998; MSP port by Ted Apel, David Zicarelli). The incoming audio is analyzed and fiddle~ is configured to report the relative amplitude of the thirty-two most significant spectral peaks as determined by the analysis. This information is output from fiddle~ as a list of numbers for each reported spectral peak. The list includes the index number (or partial number), the frequency of the spectral peak in hertz, and the relative amplitude of each spectral peak. At this point in the process, attenuation based analysis/resynthesis departs from previous approaches in that the frequency and amplitude data are stored as sample values at pre-determined locations in a buffer (hereafter referred to as the *spectral index*) instead of being passed on to an oscillator bank. Each sample location in the spectral index corresponds to an FFT frequency bin of a predetermined size. The spectral index address for a given frequency can be determined by f/(sr/FFT-size) where f is the frequency in Hertz, sr is the sampling rate, and FFT-size is the size of the FFT in samples. Resynthesis is achieved by performing a Fourier analysis of white noise, then attenuating each frequency bin of the FFT output by multiplying it by the value reported by the analysis module, and stored in the spectral index, for each frequency bin

[1] The author has developed an FFT-based method for analysis that does not require fiddle~, however, that method is not described in this paper.

location. The FFT/IFFT pair is embedded inside a pfft~ compatible patch to facilitate windowing and a variety of FFT sizes. If no energy was reported at a frequency, then that bin is zeroed (multiplied by zero), and no resynthesis occurs at that frequency location. All bins multiplied by a positive value will output energy at that frequency proportional to the energy at the corresponding spectral location of the analysis source as reported by fiddle~. More succinctly stated, white noise is filtered via a convolution based EQ that utilizes an analysis of the frequency spectra of an incoming signal as the basis for determining each FFT bin's amplitude value. Consecutive analysis/resynthesis, at regular time intervals (i.e. once per FFT), will result in an approximate subtractive-style resynthesis of the analysis source. The accuracy of the resynthesis varies, depending on the complexity of the analysis source and the precision of the analysis data output by the analysis module. As with previous analysis/resynthesis methods, the compositional potential of the attenuation-based analysis/resynthesis method lies not only in the resynthesis of an audio signal, but also in the capability to extract and modify analysis data before resynthesis occurs.

2 Modification of Analysis Data Prior to Resynthesis.

In the current MaxMSP implementation of spectral attenuation-based analysis and resynthesis, the analysis data reported by fiddle~ may be altered before being passed on to the resynthesis module in a variety of ways. These modifications include frequency shifting as well as a variety of spectral-based modifications. The spectral-based modifications include techniques I have designated as *spectral accumulation* (the systematic retention of data from previous analysis), and *spectral evaporation* (the systematic attrition of spectral data subsequent to spectral accumulation or retention).

2.1 Frequency shifting.

Frequency shifting is achieved by proportional displacement of the spectral index addresses for all frequencies output by the analysis module, prior to resynthesis. Given that the spectral index address for a frequency can be determined by f/(sr/FFT-size), frequency shifting is easily achieved by multiplying f by the appropriate transposition factor before calculating the spectral index address for f. The transposition factor is determined by first selecting a base frequency that represents no transposition (C or 261.62558 hertz in the current implementation) then dividing the frequency that represents the desired amount of transposition by the base frequency. For example, given a base frequency (i.e. no transposition) of C, 261.62558 hertz, transposition up by one equal tempered half step is achieved by dividing 277.182617 hertz (or C#) by 261.62558 hertz (C natural) then multiplying the incoming frequency (i.e. the frequency to be transposed) by the result. This calculation is performed before determining the spectral index address for each incoming frequency. For example, to transpose A up by one half step: 440*(277.182617/261.62558) = 466.1637 hertz, or B flat. If all analysis frequencies are multiplied by the same multiplication factor, the resulting resynthesis will be transposed by the interval represented by the distance between the base frequency and the transposition frequency. The current MaxMSP implementation features a graphic interface consisting of a five-octave keyboard icon for selection of the interval of transposition. The user selects the amount of transposition by clicking on the key that represents the desired interval of transposition. This value is output as a MIDI note number then translated into hertz using the MaxMSP MIDI-to-frequency converter object, *mtof.* The MIDI note number is displayed to the right of the keyboard icon and may be changed manually by clicking and dragging on the displayed number, thus allowing transposition by intervals not represented by the equal tempered scale.

It should be noted that simple frequency domain transposition of an incoming signal is easily achieved with the MaxMSP gizmo~ object. Therefore, the allure of the current method lies not in the harmonizer effect so much as in the potential to offset or alter spectral data prior to the application of spectral accumulation or spectral evaporation (discussed below).

2.2 Spectral Compression and Expansion.

It is not necessary to scale all incoming frequencies from the fiddle~ analysis module by the same transposition factor. In practice, some data may not be scaled at all or scaled in a methodical way to create novel effects. A variety of methods such as spectral attrition (deleting some partials while leaving others intact) and other techniques are currently under development by the author.

Among the techniques that have been implemented is spectral compression and expansion (fitting *n* partials into a smaller or larger pitch space than they would normally occupy). This is achieved by subtracting the fundamental (in hertz) from each frequency in the incoming fiddle~ analysis, then multiplying the result by a compression or expansion factor, and then adding this result to the original frequency. This method allows the position of the upper partials to change in relation to the fundamental while the fundamental itself remains unchanged. However, the partials will retain their relative position to one another in the frequency spectra. The current implementation features a slider-based user interface for compression and expansion of spectral data. The range of the slider is calculated so that all partials can be compressed into a pitch space as small as a few hertz above the original fundamental. Conversely, maximum expansion shifts each partial (excluding the fundamental) up one position in the overtone series (i.e. partial 2 at maximum expansion shifts upward to the location in hertz previously occupied by partial 3, 3 maps to 4, and so forth). Perhaps the most musically interesting results are achieved by utilizing settings between these two extremes or in transitions from one compression or expansion factor to another. Future enhancements may include a ratio-based user interface or possibly a function object-based index for time-varying compression-expansion factors.

2.2. Spectral Accumulation.

In the current implementation, analysis data is output from the fiddle~ object once each analysis window, and then sent to the resynthesis module. Spectral accumulation is the systematic retention of that data from previous analyses. In the most recent MaxMSP implementation, spectral accumulation can be switched on and off. When spectral accumulation is off, the spectral index buffer is zeroed at the conclusion of each analysis window. When spectral accumulation is on, the spectral index buffer is not cleared of values received from previous analyses at the conclusion of each analysis window. This results in a type of spectral histogram, with the energy at each frequency bin retaining the most recently received value for that address. For example, if the frequency bin centered at approximately 440 hertz receives a magnitude of *x*, then that magnitude will not be altered until a new value is received from the analysis module, or the user manually clears the data. As a result, the spectral index will contain analysis data from the most recent analysis and from previous analyses. The resynthesized result will bear characteristics of all audio analyzed from the time that spectral accumulation was switched on, or the user last cleared the spectral index buffer.

Several effects can be achieved using spectral accumulation. The incoming audio can be sustained indefinitely at pitch, or, as previously suggested, the analysis data can be shifted along the spectral axis prior to accumulation and resynthesis. Also, infinite reverb type effects can be achieved by gating input so that only signal values above a minimum threshold are resynthesized. Additionally, the fiddle~ object can output analysis data once per window period, or it can be put into ‘poll’ mode and output analysis data only when requested by the user. Using this technique it is possible to take a series of ‘spectral snapshots’, each consisting of only one analysis window, and layer them onto one another in the frequency spectrum. This method is analogous to multiple exposures on the same photographic film, creating composite images taken at different times. As with all collage-like methods in which data is accumulated over time, it is possible to saturate the image, or in this case the spectrum, with too much information. Therefore, a systematic method for data attrition is desirable.

2.3. Spectral Evaporation.

Subsequent to data retention (spectral accumulation) it will, at times, become desirable to thin or completely clear the spectral index buffer. Since the spectral index data is stored as samples in a buffer, it is possible to alter that data after it has been written into the buffer. The indiscriminate deletion of all data is easily achieved by simply clearing the spectral index buffer, thereby erasing all previously accumulated spectral information. However, this method results in an abrupt cessation of sound that is of limited compositional interest. A more systematic approach to clearing the data holds potential for more compositionally germane results.

Spectral evaporation is the systematic time-based zeroing of data in the spectral index buffer. This is achieved by sequentially and randomly selecting a bin addresses and writing a zero to that address, thereby eliminating the energy output from the resynthesis module at that frequency region (if any was present). Currently, a pseudo-random uniform distribution is implemented to select frequency bins for attenuation. The MaxMSP *urn* object is utilized to avoid repetitions of randomly generated numbers. Adjusting the frequency at which the *urn* object produces random numbers regulates the rate of spectral evaporation. In the current implementation the rate of spectral evaporation ranges from 0 to 100, 100 being the fastest possible evaporation and 0 no evaporation.

More systematic and targeted applications for bin attenuation are possible. For example, in the current implementation all frequency bins that have not yet been selected for attenuation are equally likely to be selected (i.e. uniform distribution without repetition of variates). However, other probability distributions may be employed in the spectral evaporation process in order to focus the attenuation on a particular frequency region or regions. Also, alternate methods for bin selection may be of interest. Finally, once selected, bins need not necessarily be attenuated. One possible avenue for future applications of this technique is the random selection of a frequency bin coupled with a randomly selected magnitude for that bin. Both bin selection and the bin value could be linked to some form of user input.

3 Compositional Implications.

To date I have utilized spectral accumulation and evaporation in three interactive computer music compositions, *Afterimage 3* for percussion and MaxMSP and *Afterimage 6* for guitar and MaxMSP, and *Afterimage 7* for flute, violin, cello, piano, percussion, and MaxMSP. Although numerous real-time and non-real-time processes are employed in all three works, the accumulation of spectral data based on real-time analysis of audio input and the time-based spectral evaporation of the accumulated data is featured prominently in each composition.

For example, the opening gesture of *Afterimage 3* consists of a series of sounds produced by sliding one concrete block on top of another. These sounds (basically band-limited noise) are fed into the analysis unit and the analysis data is accumulated for resynthesis (spectral accumulation). Spectral evaporation is then employed to modify the initially accumulated spectral data for approximately the first sixty seconds of the piece. Subsequent applications of spectral accumulation and evaporation include the prolongation of the complex spectra of a sound created when the player taps the concrete block with a large metal bolt.

Spectral accumulation and evaporation appear prominently in the final section of *Afterimage 6*. The guitarist performs a series of staccato chords, each of which is analyzed and the spectra allowed to accumulate. The effect is that of the multiple exposures on the same photographic film as discussed previously. Spectral evaporation is then applied to thin the data. In addition, analysis of subsequent chords, containing a subset of pitches present in the proceeding series of chords, is written to the spectral index, thereby reinforcing some frequency regions while others are allowed to evaporate.

Spectral evaporation with spectral expansion appears prominently throughout *Afterimage 7*. As the work progresses, a specific sonority (that indicates structural boundaries) is subjected to spectral expansion, accumulation, then evaporation. The spectral expansion increases incrementally with each appearance of the sonority, thus gradually decoupling the computer-generated resynthesis component from the acoustic component. Therefore, the distortion of the acoustical spectra of the sonority becomes developmental in nature.

4. Summary and Suggested Possible Future Research Directions.

Real-time spectral attenuation based analysis and resynthesis is a frequency domain based method for collecting data from an audio source then utilizing that analysis data to resynthesize the original sound by mapping that information to an FFT/IFFT based subtractive synthesis module. A variety of methods may be employed to alter the data prior to resynthesis, creating a variety of effects. Additional effects are achieved by retaining data from previous analyses (spectral accumulation) and the systematic attrition of accumulated data (spectral evaporation).

A number of methods for data alteration, both prior to and subsequent to analysis, remain tantalizingly unexplored at this time. Some of the techniques mentioned previously, that are currently being explored by the author, include spectral attrition, weighted probability spectral evaporation, and stochastic alteration of analysis data prior to resynthesis.

The MaxMSP patch described in this paper is available for download at the author's website at http://faculty.winthrop.edu/parksr.

5.Acknowledgements.

The author wishes to acknowledge the work and continuing support of Cort Lippe, without whose mentorship this work would not have been possible. Thanks also to Zack Settel for his invaluable contributions to our field, Miller Puckett and David Zicarelli for MaxMSP, and Ted Apel for the OSX port of fiddle~. Special thanks to the numerous colleagues who aided in the conceptual and programming aspects of this work including Eric Ona, Samuel Hamm, David Kim-Boyle, and James Paul Sain.

References

Settel, Zack and Lippe, Cort (1997, revised 2001). “Forbidden Planet” *MaxMSP examples* included in the standard release of Cycling74's MaxMSP software, www.cycling74.com.

Lippe, Cort (1996). “Music for Piano and Computer: A Description”.

Settel, Zack and Lippe, Cort (1994). “Real-Time Timbral Transformation: FFT-based Resynthesis” Contemporary Music Review: Harwood Academic Publishers.

Puckette, Miller; Apel, Theodore; and Zicarelli, David (1998). “Real-time Audio Analysis Tools for Pd and MSP.” 1998 International Computer Music Conference Proceedings.