CHROMAX, THE OTHER SIDE OF THE SPECTRAL DELAY BETWEEN SIGNAL PROCESSING AND COMPOSITION

Arshia Cont, Carlo Laurenzi
Ircam
1 place Stravinsky
75004 Paris

Arshia.Cont@ircam.fr
Carlo.Laurenzi@ircam.fr

Marco Stroppa
University of Music and Performing Arts
Urbanstr. 25
70182 Stuttgart, Germany

stroppa@mh-stuttgart.de

ABSTRACT

Spectral delays have been used for a long time as a way to colour and shape spectral characteristics of sound.
Most of available software is controlled by drawing an envelope on a window that represents spectral bins, and by setting a maximum delay time. Despite its comfort, such a simplistic approach does not imply any methods for allowing symbolic manipulations on spectral data that are often required by composers and sound designers.
Chromax proposes an alternative dynamic parameterization of spectral delays, allowing fine and complex compositional manipulations. It implements a bin-synchronous spectral processing using the new Gen~ technology available in Max6 [1], and provides algorithms to dynamically specify a filter, a delay and a feedback level for each bin of a processed sound.

1. INTRODUCTION

Spectral delays may be considered as an extension of traditional delays. An incoming sound is subdivided into frequency bands and each band can be assigned a different delay and played back. In addition a feedback can be added.
Even if alternative implementations have been proposed [2], the frequency subdivision is usually performed via a Fast Fourier Transform (FFT). Spectral delays may become computationally demanding if the number of bands is very high, raising efficiency issues and accordingly reducing control over parameters. Moreover, since the user must set delay times and feedback values for each frequency band, a large amount of data needs to be specified; this raises issues of interface design.
The currently available spectral delays are either standalone applications (such as iZotope *Spectron* [3], or the now discontinued *NI Spektral Delay*, [4]), or special objects for Max or Pure Data. The interface is usually conceived as a time vs. frequency window (time is the Y-axis). The user draws an envelope over frequencies and sets a value for the maximum delay. A second window may be used to control feedback levels.

1.1. Control strategies

The need for better control strategies than those available on commercial software was already mentioned in a paper by Kim-Boyle [5]. He implements a spectral delay as a Max object using pfft~, and provides some macros to manipulate the frequency window, such as "increment or decrement by a small amount the entire contents of the buffer" or "write a value to a specific set of bins" [5, p 43]. He also mentions the automatic computation of delay values by "performing an FFT analysis of a control signal, which may not be the same signal as that processed" [5, p. 43].
From a purely compositional standpoint, however, these controls remain quite basic. For instance, it would be hard with them to compose a precise rhythm.
Gibson [6], drawing examples from his own works, points out that a spectral delay can be also used as a compositional tool and that delay times can be set at once for many bands in the form of a Max message. His paper, however, does not delve into the details of how high-level compositional data are mapped onto low-level spectral parameters and the main control still seems to be window-oriented.

1.2. Time issues

A further complexity stems from the need of accurate timing and constraints imposed by Short-Time Fourier Transforms (STFT), where the smallest time unit is the hop-size, that is typically ¼ of the analysis window if a Hanning window is used, which depends on the sampling rate (SR) and the window size (WS).
For instance, for a SR of 96KHz and a WS of 8192 points, the time resolution is 21.3 msec. Although this value seems to be very small from the point of view of a composition (it corresponds to a little more than a hemidemisemiquaver of a sextuplet with a MM of 120!), when planning rhythmical patterns with a large amount of feedback, round-off errors may rapidly lead to unsynchronized sloppy rhythms that composers would reject.
Designing musically satisfying, albeit computationally reasonable time scaling algorithms, while giving the maximum freedom to the composer, is a challenge and implies a trade-off between scientific precision and aesthetical needs.
Kim-Boyle [5] also mentions another problem of temporal limitations, when spectral delays are used with stereo files for spatial purposes at the resolution of inter-aural time differences (few msec).
Finally, when dynamic delays are allowed, the process becomes further complicated, since such delays must correspond to the timing constraints of the STFT so that feedback values remain rhythmically coherent. A common side-effect when missing this consideration is timing offsets after each feedback cycle.

2. SPECTRAL GENERATOR

Chromax is a spectral generator that provides separate templates for the delay and the feedback windows and a large set of parameters. *Chromax* achieves high-level generation of low-level spectral templates using two base generators, spectral pattern handling and temporal dynamic controls provided to users as described below.

2.1. Base Generators

Spectral templates generated by *Chromax* consist of aggregated *atoms* over the frequency domain. For the time being, *Chromax* allows two types of base atoms: *Partial atoms*, used for harmonic spectrum generation, and *Formant atoms*, defined by centre frequency, amplitude, bandwidth and skirt width [12]. Since, there is currently more support for the *Partial* class than for the *Formant* class, we base our discussions in this paper on *Partial atoms*.

For *Partial* template generation, the commonly used method applies the following equation derived from McAdams [7] to a base frequency bin domain:

$$f_n = nf_0 + (n+shift)f_0 + n^s f_0 \quad (1)$$

where n is the partial number, $f0$ the fundamental frequency, *shift* the amount of $f0$ by which each partial is shifted, and s is the stretch factor defined as:

$$s = log(b) / log(2) \quad (2)$$

The same method is used to compute frequency bin locations for additive synthesis in the *OMChroma* environment [8] [9]. Equation (1) permits to generate a wide range of partial relations beyond pure harmonic structures.

To make it more suitable to spectral delays, each "partial" also possesses a bandwidth (called *variance* in *Chromax*), which defines its spectral thickness, expressed either in Hz or in cents (Figure 1). For a composer this difference is essential: a value in cents produces spectra whose bandwidths increase as a function of frequency and therefore tend to sound more clustery in the upper range, whereas values in Hz are more effective in lower frequencies.

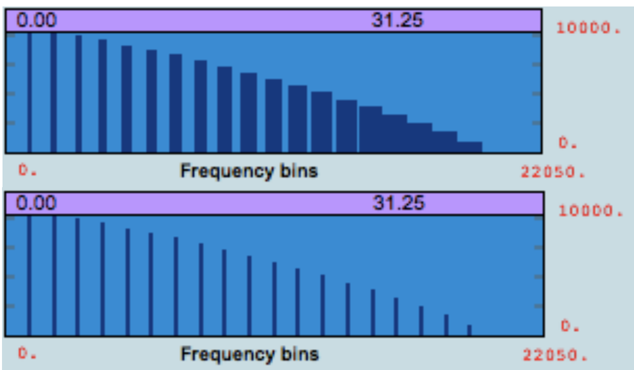

Figure 1: *Harmonic template containing 19 partials with a f0 of 1000 Hz and bandwidth of 50 cents (above) or 50 Hz (below).*

When excited by a noise burst, the response of a spectral delay whose window is a simple contiguous frequency band (Figure 2) consists in a frequency-limited delayed burst.

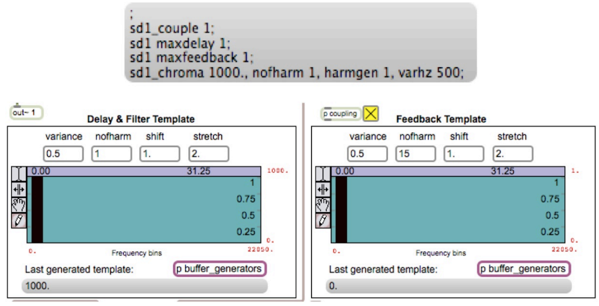

Figure 2: *Chromax template created by the message box above.*

With a feedback of 100%, and in a sample-synchronous setup, the burst should infinitely repeat itself, as in Figure 3, without any attenuation. Most existing developments, however, fail to do this, due to scheduling problems of the host platform (as is the case of *pfft~* in Max). We will get back to this point in Section 2.3.

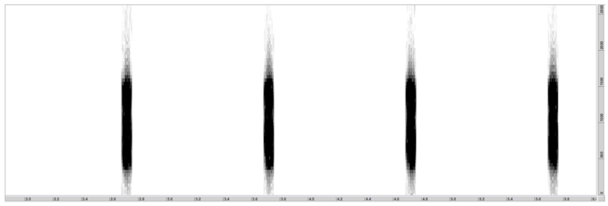

Figure 3: *Sonogram of a noise burst processed by a spectral delay with the template shown in Figure 2.*

2.2. Atom Shapes: time point vs. glissando

Working with composers such as Marco Stroppa made it clear that a more refined control of the shapes and characteristics of each *atom* was required. This lead us to designing two shapes.

With a rectangular shape (as in Figure 1), the spectral range is delayed simultaneously, and can therefore be considered as a single point in time. With a Gaussian shape, the external bins will smoothly precede the centre frequency, while the bin located in the centre of the curve will have the maximum delay (Figure 4).

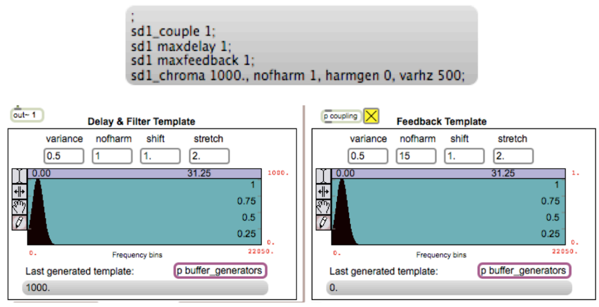

Figure 4: *Chromax template with Gaussian shape.*

This produces a double glissando, ascending and descending toward the value of the central bin (eventually repeated if the feedback is activated, as in Figure 5). The glissando is granulated by the spectral size of each bin and by the time resolution.

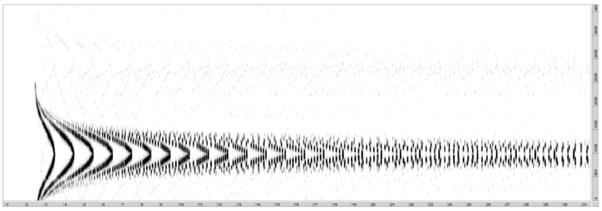

Figure 5: *Sonogram corresponding to the template of Figure 4.*

A further temporal refinement consists in changing the slope of the spectral envelope, in order to produce, for instance, descending spectra (as in Figure 1). When excited by a noise burst, a sequence of descending narrow-frequency bursts is generated. Depending on the slope, the sequence may accelerate or decelerate. If the band is Gaussian, an additional, fast, double glissando will spread around the central bin of each burst, as seen above; with feedback this can result in progressively increasing densities, whose time fabric can be finely controlled.
This version of the spectral delay was first used in *…of Silence* by Marco Stroppa [10].

2.3. Composed rhythmic patterns

Our preliminary experiences with *Chromax* showed that the control of time examined in Section 2.2 was still not accurate enough, especially when using a long feedback. Better time strategies had to be devised.
Our goal was to achieve an as strict as possible temporal precision in spite of the STFT's temporal constraints, so as to match rhythms written in a symbolic score in a way that a composer would consider to be adequate and accept.
In order to comply with such temporal patterns, the generator has to rhythmically quantize the requested delays, so as to:

1. Keep them as integer multiples of the STFT's time-steps (or hop-size) set by user; and
2. Keep the rhythmic results coherent and as close as possible to the written pattern.

This constraint is achieved in *Chromax* by dynamically quantizing the requested rhythms using their greatest common divisor and the underlying analysis/re-synthesis time-step. This allows for the timing correctness of the delayed bins with respect to the written patterns using feedback. The situation becomes more complicated with poly-rhythmic structures that would require segmentation on top of quantization. This issue will be tackled in future work.
Following this solution, further parameters were introduced (called *harmlist* and *harmweights*) to allow the specification of individual partial numbers and weights (that is, time shapers).
In this way, setting *harmweights* of rectangular bands to values corresponding to symbolic rhythms (Figure 6) allows to generate specific patterns, where each frequency band produces a "note" every 1/n time units (a time unit is the maximum delay, after normalization of the weights).

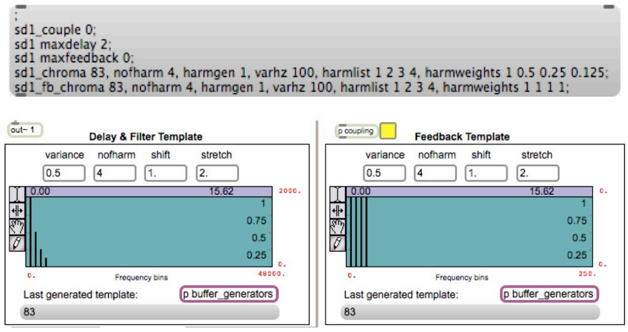

Figure 6: *Chromax template generating a symbolic rhythm.*

For instance, for a spectrum of 4 frequency bands with a weight of, respectively, 1, 0.5, 0.25 and 0.125 and a maximum delay of 2 sec, the first band will be delayed by 2 sec, the second by 1 sec (2*0.5), and so on. When excited by a noise burst, this template generates a band-based arpeggio corresponding to the rhythm shown in Figure 7 (at a metronome of quaver = 60).

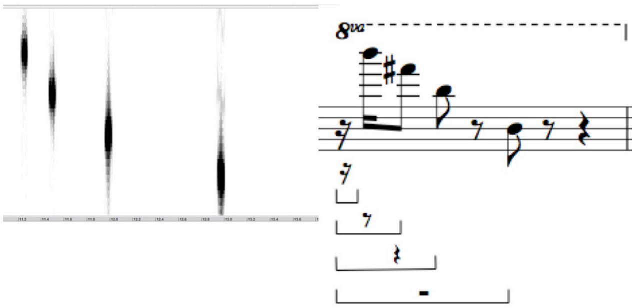

Figure 7: *Sonogram and rhythmic pattern generated by a noise burst processed by the Chromax template shown in Figure 6.*

2.4. Special cases: between sound processing and synthesis

For his opera *Re Orso* [11] Marco Stroppa required special rhythmical templates that simulated the sonic result of a particular wave-table synthesis technique: a cluster made of oscillators reading the same harmonic wave table and whose fundamental frequency is slightly and evenly shifted with respect to one another, generates a typical pattern if all the oscillators start with the same phase (Figure 8). This effect is sometimes called "zero phasing" in the world of computer music composition.

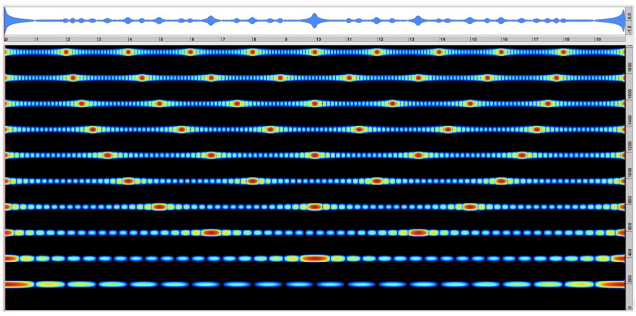

Figure 8: *Sonogram of a cluster of phase-locked oscillators, reading a wave table made of 10 harmonics of equal amplitude. The f0 of the lowest oscillator is 200 Hz, the other f0's are shifted by 0.05 Hz with respect to the previous one.*

Using a long feedback and setting *Chromax* values in specific ways, similar patterns can be produced with spectral delays with more freedom of control than in the wave-table synthesis. For instance, Figure 9 shows an inverted pattern with respect to the one displayed in Figure 8 (the fastest repetition is the lowest, rather than the highest frequency).

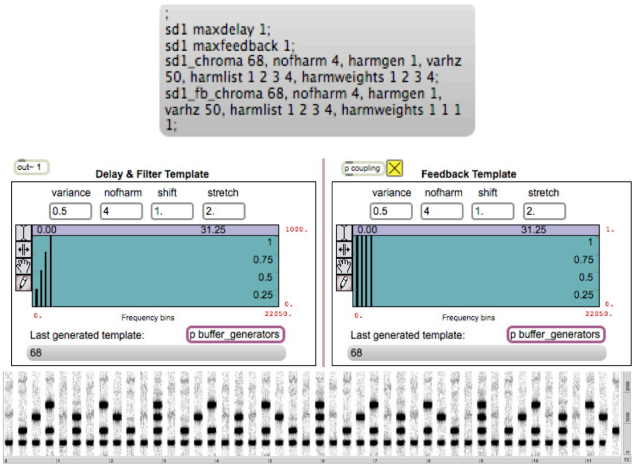

Figure 9: *Chromax message box, template and sonogram in response to a noise burst simulating an inversed "zero-phasing" rhythmical pattern.*

This usage of the *Chromax* template opened rich compositional perspectives and became one of the main processing tools used in *Re Orso*.

Combining particular settings of *harmweights* with non linear spectral slopes, Marco Stroppa created progressively accelerating or decelerating "zero-phasing" patterns, with predictable, and, therefore, composable, time features (Figure 10).

2.5. Dynamic control

Besides the static template generation shown above, the computational scheme allows to produce a dynamic and time-varying control of all the parameters mentioned above, as well as further features. In this case, the *Chromax* template evolves over time while generating complex rhythmical structures synchronous to the DSP scheme imposed by a STFT, that are probably very hard to express using the traditional symbolic notation.

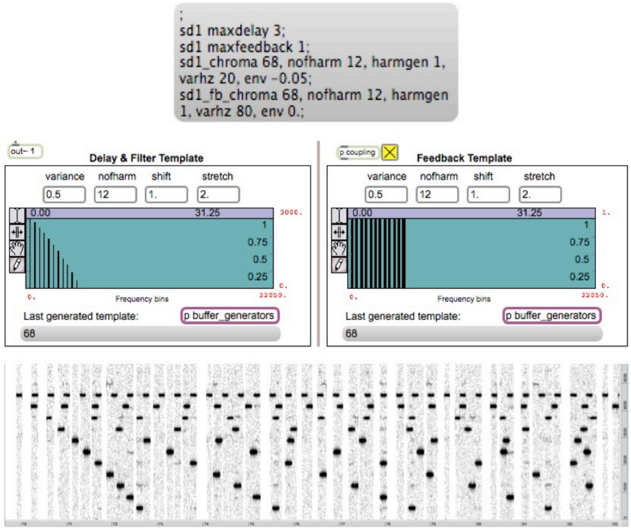

Figure 10: *Chromax message box, template and sonogram in response to a noise burst realizing a decelerating "zero-phasing" rhythmical pattern.*

Dynamic controls either perform a cross-fade between a given number of templates over a specified duration (spectral interpolation between two templates), or dynamically regenerate templates between two patterns of the same base type to simulate *glissandi*. To avoid audio glitches, the time-step used for this kind of dynamic generation is automatically synched to the user-set analysis/re-synthesis parameters in the DSP. Regeneration schemes have been extensively optimized to make the technology embeddable in large real-time systems.

2.6. Interface

The user interface consists in a collection of Max messages that set all the parameters needed by *Chromax* through simple message-passing. For a complete list and further details, the user can refer to the documentation available within the Max object.

3. IMPLEMENTATION

Chromax is an external Max object written by Arshia Cont. It dynamically generates FFT templates into buffers that can be used in later spectral processing. The object has the ability to further synchronize its dynamic generation to the spectral processing in the Max world. In the case of a Spectral Delay, it provides temporally controlled buffers for delays and feedback values, and uses the analysis parameters of the Spectral Delay for rhythmic quantization (described in Section 2.3) and synchronous dynamic generation (Section 2.5).

The spectral delay patch by itself is an enhancement by Carlo Laurenzi of earlier implementations using *Gen~* [1] to comply with the needed bin-synchronous real-time computation. Upgrading the earlier buffer-based to a sample-synchronous implementation using Gen~ has greatly improved efficiency and the fine compositional control of the module, especially for long-term rhythmical composition, which was otherwise imprecise or musically impossible.

3.1. Further developments

The compositional experience with *Chromax* and its usage in concerts by various composers have emphasized the importance of high-level controls for real-time DSP, as well as important future work for its enhancements.

A major drawback of such analysis/re-synthesis approach is in the hard-to-predict detailed spectral content of the input sound in live performances. Sometimes, the template does not match the spectrum of the incoming sound, and therefore generates dynamically irregular patterns. For monophonic, pitched sounds, a preliminary f0 analysis might yield useful information to construct a satisfactory template.

However, an automated spectral exciter, tuned to the structure of a given *Chromax* template would be highly appreciated.

More vocabularies should also be incorporated into the spectral generator, in order to have the much needed *Formant atoms* produce rich and dynamic structures.

4. COMPOSITIONAL ISSUES: BETWEEN SOUND PROCESSING AND SOUND SYNTHESIS

The main reason that motivated us to develop this approach to a spectral delay was a request by Marco Stroppa to have a processing tool matching the compositional refinement that sound synthesis allowed him to achieve [8]. Since this is not a purely scientific principle, a collaboration at the intersection of signal processing and composition is mandatory. By fine tuning dynamic control parameters and using multiple f0's these sonic processes could be successfully generated.

As each frequency band tends to be resynthesized by a small amount of bins, especially in the middle-low range, which is the most musically significant, the result is comparable to a narrow mixture of sine tones.

However, the original sound can still be recognized because it determines the overall spectrum. *Chromax* lies therefore at the crossroad between sound processing and sound synthesis. The mutual exchange between all these realms was extremely beneficial to the specification of the final outcome.

5. REFERENCES

[1] *Extend the power of Max.* http://cycling74.com/products/gen/. Accessed April 11, 2013.

[2] Jussi Pekonen and Vesa Välimäki, Jonathan S. Abel and Julius O. Smith, "Spectral delay filters with feedback and time-varying coefficients", *Proc. Digital Audio Effects (DAFx'09),* Como, Italy, Sep. 2009, pp. 157-164.

[3] *Sound Design with Spectron*, iZotope, 2005. Available at http://www.izotope.com/products/audio/spectron/SpectronSoundDesignGuide.pdf. Accessed April 11, 2013.

[4] *NI-Spektral Delay*, Native Instruments, 2007. Available at http://www.native-instruments.com/index.php?id=k3spektraldelay_us. Accessed April 11, 2013. Discontinued since 2007.

[5] Dr. David Kim-Boyle, "Spectral delays with frequency domain processing", *Proc. Digital Audio Effects (DAFx'04),* Naples, Italy, Oct. 2004, pp. 42-44.

[6] J. Gibson, *Spectral Delay as a Compositional Resource*, Toronto Electroacoustic Symposim, 2009, Toronto. Available at http://cec.sonus.ca/econtact/11_4/gibson_spectraldelay.html. Accessed April 11, 2013.

[7] S. Mc Adams, "Spectral fusion and the creation of auditory images", in Music, Mind and Brain, M. Clynes Ed. Plenun Press, 1982, pp. 279-298.

[8] M. Stroppa (2000) "High-level musical control paradigms for digital signal processing", *Proc Digital Audio Effects (DAFx'00)*, Verona, Italy. Available at http://profs.sci.univr.it/~dafx/Final-Papers/pdf/Stroppa.pdf. Accessed April 11, 2013.

[9] C. Agon, J. Bresson, M. Stroppa, "OMChroma: Compositional control of sound synthesis", *Computer Music Journal*, 35(2), 2010, pp. 67-83.

[10] Marco Stroppa, *...of Silence* (2007), for saxophone and chamber electronics. Ricordi Edition, Milan.

[11] Marco Stroppa, *Re Orso* (2012), musical legend for 4 singers, 4 actors, 11 instruments, invisible voices and sounds, spatialisation and acoustic totem. Ricordi Edition, Milan.

[12] Xavier Rodet, Yves Potard, Jean Baptiste Barrière, "The Chant Project: From the Synthesis of the Singing Voice to Synthesis in General", *Computer Music Journal*, 8(3), 1984, pp. 15-31.