Psychoacoustic impact assessment of smoothed AM/FM resonance signals

Antonio José Homsi Goulart
Computer Science Department
University of São Paulo
São Paulo - Brazil
antonio.goulart@usp.br

Joseph Timoney, Victor Lazzarini
Computer Science Deparment
Maynooth University
Maynooth, Co. Kildare - Ireland
joseph.timoney@nuim.ie
victor.lazzarini@nuim.ie

Marcelo Queiroz
Computer Science Department
University of São Paulo
São Paulo - Brazil
mqz@ime.usp.br

ABSTRACT

In this work we decompose analog musical resonant waveforms into their instantaneous frequency and amplitude envelope, and then smooth these estimations before resynthesis. Signals with different amounts of resonance were analysed, and different types and lengths were tested for the smoothers. Experiments were carried out with amplitude smoothing only, frequency smoothing only, and simultaneous smoothing of amplitude and frequency signals. The psychoacoustic impacts were evaluated from the point of view of dynamic brightness, tristimulus and spectrum irregularity. We draw conclusions relating the parameters explored and the results, which match with the sounds produced with the technique.

1. INTRODUCTION

Resonance is the tendency of a system to vibrate sympathetically at a particular frequency in response to energy induced at that frequency [1]. Resonances play an important role in computer music and a number of techniques have been proposed for their synthesis, including FOF [2], VOSIM [3], ModFM [4] and Phase Distortion [5]. Frequency modulation [6] methods are interesting because they can offer a flexible yet computationally inexpensive solution. However, care has to be taken in selecting appropriate modulation functions if the result is not to sound lifeless in comparison to their analog synthesizer counterparts. Instruments such as MiniMoog, Korg MS-20, TB-303 [7] present unique, readily identifiable resonant sounds. Transporting the compelling nature of the analog sound to the digital domain is an interesting problem, particularly if we want to preserve the feel of the sound, but not simply mimic it.

This work is a first investigation on the potential of using an AM/FM signal decomposition followed by smoothing and resynthesis for the modeling and synthesis of resonance signals. A link can be made with the modulation synthesis by decomposing a discretized analog signal into an AM/FM representation using an analytic signal approach [8]. This representation captures everything in two descriptors, the changing envelope (AM) over time and the frequency excursion around the fundamental (FM). These quantities display an 'average' behaviour for the collection of components in the signal at a particular time instant [9].

Albeit using different tools than ours, a similar work regarding assessment of analysis followed by modification of the parameters and resynthesis was performed in [10]. Acoustic musical instruments recordings were decomposed with Fourier analysis and different modifications were tested within an additive synthesis context for the comparison of the original and resynthesized sounds. Another work [11] explored AM/FM decomposition of musical instruments using energy separation, in order to analyse vibrato/tremolo and determine synthesis parameters for an excitation/filter model. In [12] the reverberation on voice recordings was analysed in terms of its impact on an AM/FM decomposition.

The extent of the modulations' variations is interesting regarding its relationship to the perceived sound, and it is a subject that has not received much attention in the literature. However, if we have a good perceptual intuition about these signals we should be able to design digital modulation-driven resonance generators that sound more exciting. A straightforward approach to achieve this is to form the decomposition of suitably chosen analog generated resonance signals and then manipulate their AM and FM quantities and observe the outcome in a series of controlled experiments.

Also, we highlight a different kind of approach for the processing of musically-interesting sounds, which is not as widely used in the computer music literature as the Fourier analysis and additive synthesis approach [13]. Analysis / resynthesis with AM/FM decomposition can also be seen as a different paradigm where we consider the sound signal in terms of a single harmonic oscillator model, with varying instantaneous amplitude and frequency. Whereas in the Fourier paradigm the signal is viewed as the resultant of superimposed oscillators, in the AM/FM model we might think of a mass-spring system with varying mass (e.g. where the body suspended from a spring is a liquid container with controllable inlet and outlet). We are then able to take advantage of the compactness of dealing with only a couple of signals to apply our desired modifications/adaptations to the sounds.

We start the paper by talking about the signals we used and their decomposition based on amplitude envelope and instantaneous frequency estimation, followed by a discussion regarding the process of smoothing and resynthesis. This is complemented by an explanation of the psychoacoustic metrics and the results we obtained, and our conclusions. All code was written in Octave and is available [1] alongside a comprehensive set of sound examples.

2. ANALOG SYNTHESIZER RESONANCES

Analog synthesizers are usually based on the subtractive synthesis model [7], where a raw rich excitation signal is the source for a modifier, typically a set of filters that will impose their characteristics and tailor the sound.

The TB-303 is an electronic bass synthesizer that was introduced in the 80s, being heavily explored by dance music producers since then. We chose the '303' based on our interest of working with waveforms containing strong resonances widely accepted as being musically relevant. The '303' filters are characterised by a sharp cutoff of 24 dB/octave [14] being able to produce very apparent and clean resonances. The '303' also features an "Env mod" knob, that adjusts the filter's envelope signal influence on the filter's cutoff frequency. By turning this control clockwise, the envelope will sweep the cutoff frequency over a greater range. When turned counter-clockwise, the filter's envelope will have very little affect on the filter's cutoff frequency [15].

Another control in the '303' is a "Decay" knob, which controls the filter's envelope decay time. Longer envelope decay times will allow high frequencies to pass through the filter for a longer amount of time. Turning this control counterclockwise will shorten the amount of time high frequencies can pass through the filter [15]. A "Cutoff" parameter sets the cutoff frequency of the low-pass filter [16] and the "Resonance" sets the Q factor, accentuating frequencies close to the cutoff [15].

We recorded lots of samples using the sawtooth waveform, exploring variations within the mentioned controls. Figure 1 illustrates as an example some periods of a very resonant waveform we used in the experiment. Notice the characteristic appearance of resonant waveforms, with a series of rapid large oscillations at the resonant frequency imposed on a periodic variation at the note pitch frequency.

Four waveforms were selected for the experiments, based on complementary settings of "Resonance", "Env mod" and "Decay". The "Cutoff" was always left fully open to avoid losing the contribution of the higher partials. The settings are summarized and labeled in Table 1.

Resonance	Env Mod	Decay	Label
60%	max	min	A
60%	max	60%	B
max	75%	25%	C
max	75%	75%	D

Table 1. Knob settings for the experiments

[1] www.ime.usp.br/~ag/dl/smc15files.zip

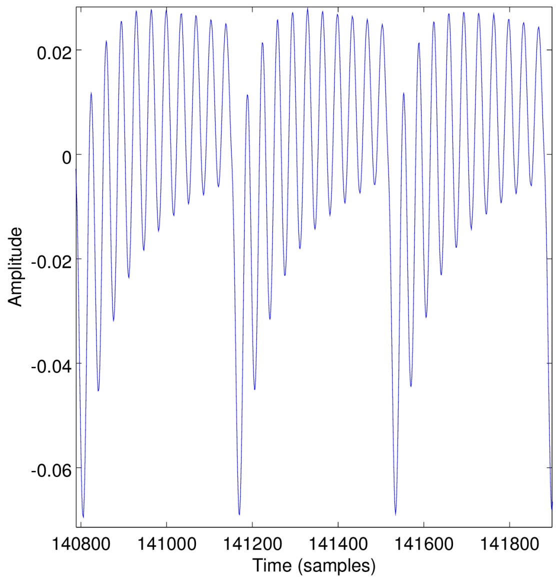

Figure 1. Some periods of a waveform generated with high resonance and decay time settings for the '303'

Waveforms A and B present the same values for a mild Resonance and a deep Env Mod, but the Decay value is minimal for A and medium for B. Waveform A is characterized by a fairly smooth sound, because although the values for Env Mod and Resonance are not small, the minimal Decay imposes a quick drop of the note, preventing the development of the modulation and resonance. The generous value for the Decay in Waveform B establishes a tailored resonance throughout the sound. Waveforms C and D are based on the maximum value for resonance and a strong Env Mod, differing by the small and large values for the Decay. The resonance generated in these cases is more aggressive, ringing for all the sound duration in Waveform C and predominately as a sweep in D.

3. AM/FM ANALYSIS

The AM-FM decomposition is a powerful method for the analysis of non-stationary signals [9]. Consider a signal

$$x(t) = a\cos(\omega t + \phi), \tag{1}$$

where a is the amplitude, ω is the frequency and ϕ an initial phase. The argument $(\omega t + \phi)$ is the instantaneous phase, and its derivative ω is the instantaneous frequency (IF).

We could also modulate both the amplitude and the frequency of the signal in (1) to give [8]

$$x(t) = a(t)\cos(\theta(t)), \tag{2}$$

where the phase derivative $\dot{\theta}(t) = f(t)$ is the IF of a signal, and $a(t)$ its instantaneous amplitude (IA). The IF can be described as the frequency of the sinusoid that locally fits the signal at instant t [9].

The AM/FM signal analysis is intended to decompose a signal into functions for the AM (related to the IA signal) and the FM (related to the IF signal). A number of techniques exists for that [17] [8]. In this work we assume a mono component signal and apply the analytic signal based approach based on the Hilbert Transform (HT) decomposition. The HT of a signal $x(t)$ is given by [9]

$$\hat{x}(t) = x(t) * \frac{1}{\pi t}. \tag{3}$$

This creates a 90° phase shifted version of the original, from which we build the analytic signal related to $x(t)$ as

$$z(t) = x(t) + j\hat{x}(t) = |z(t)|e^{j\theta(t)}. \tag{4}$$

For a signal of the form of (2), $|z(t)|$ and $\dot{\theta}(t)$ can be used as estimates for the AM and FM. Once we have these estimates for $a(t)$ and $f(t)$, we use them to resynthesize the original signal with the expression

$$y(t) = a(t)\cos\left(\int_{-\infty}^{t} f(\tau)d\tau\right) \tag{5}$$

4. SMOOTHING AND RESYNTHESIS

In this work we want to investigate the extent of the impact caused by modifications of the estimated signals for AM and FM, so before the resynthesis we apply smoothing on the AM and FM signals. The expression for the output signal is given by

$$y(t) = (a * w)(t)\cos\left(\int_{-\infty}^{t} (f * w)(\tau)d\tau\right), \tag{6}$$

being $w(t)$ the window and '$*$' the convolution operator. We also experimented the smoothing of only one signal at a time, either the AM, replacing $(f * w)(\tau)$ by $f(\tau)$ in (6) or the FM, replacing $(a * w)(t)$ by $a(t)$.

We tested two types of windows for the smoothing, namely the rectangular (Boxcar) and the Hanning windows. For each window, two lengths were experimented, 20 and 100 samples. We will refer to these configurations as B20, B100, H20 and H100, respectively. The types and lengths of the windows were chosen based on experiments comparing the sidelobes behaviour [18] and influence on sound for different lengths of widely used windows.

5. PSYCHOACOUSTIC EVALUATION

Besides the important subjective assessment of the results, some psychoacoustic metrics [19] were chosen to quantify objectively the results of smoothing the AM and FM signals. The dynamic brightness, tristimulus and spectrum irregularity of the sounds were observed. Now we will introduce the tool used to derive the harmonic estimation, and after that discuss the metrics.

5.1 Complex Signal Phase Evolution (CSPE)

For the derivation of the psychoacoustic metrics of a sound we need to know the contribution of its frequency components. The CSPE is a tool to decompose a signal into its sinusoidal components, working around the limitations of the Discrete Fourier Transform (or the Fast Fourier Transform). If sr and N respectively are the sample rate of the signal and N the size of the window for the analysis, the Fourier method presents a limited frequency resolution. Partials located exactly at multiples of $\frac{sr}{N}$ are clearly identified, but frequencies located far from these multiples are distorted in the analysis. Another problem related to the DFT/FFT is the time/frequency accuracy tradeoff, or uncertainty principle, where good resolution for one information comes with the detriment of the other (resulting in bias either in frequency or time).

In order to enhance the results of the FFT, the CSPE algorithm analyse the phase evolution of the components between N points frames of the signal and its time-delayed version [20]. An FFT analysis is performed on the signal, and another FFT is performed at a one-sample delayed version. The delayed version spectrum is multiplied with a complex-conjugate version of the signal, resulting in a frequency dependent function [21], from which we derive the spectral envelope as the set of values $a_k, k = 1..M$, where M is the highest relevant partial and a_k is the weight of partial k. More details about the CSPE and its mathematical development can be found in [20] and [21].

5.2 Some psychoacoustic metrics

Deriving metrics from audio excerpts, and correlating them to parameters that lead to these sounds, can enlighten comprehension, perception, and composition, as timbre attributes emerge [19]. Thus, here psychoacoustics evaluations help us quantify objectively the results of the modifications we introduced. Next we discuss the metrics used in this work.

5.2.1 *Brightness*

The brightness, or spectral center of gravity of a spectrum, is related to the spectral centroid, and "may be thought of as the harmonic number at which the area under the spectral envelope described by a_k is balanced" [22]. In such a way the brightness can be intuitively related to the most prominent portion of the spectrum of a sound. Specially when dealing with resonances, which impose a high selectivity on the spectrum, the brightness is an indication of the spectral localization of the resonance.

There are different definitions for the brightness of a sound. In this work we calculate it with the expression [22]

$$Br = \sum_{k=1}^{N} k a_k / \sum_{k=1}^{N} a_k \tag{7}$$

As noted by Beauchamp [22], brightness appears as a common feature in works that investigate timbral modification. It was shown [23] that brightness is usually the metric that presents larger variations within psychoacoustic experiments, suggesting that it is a strong measure to characterise sounds.

5.2.2 *Tristimulus*

According to Jensen [19], the tristimulus concept was introduced in [24] as a timbre equivalent to the color attributes in vision. It was used to analyse the transient behaviour of musical sounds. The tristimulus (T_{r1}, T_{r2} and T_{r3}) metrics are defined as [19]

$$T_{r1} = a_1 / \sum_{k=1}^{N} a_k \tag{8}$$

$$T_{r2} = (a_2 + a_3 + a_4) / \sum_{k=1}^{N} a_k \tag{9}$$

$$T_{r3} = \sum_{k=5}^{N} a_k / \sum_{k=1}^{N} a_k \tag{10}$$

Notice that $T_{r1}+T_{r2}+T_{r3} = 1$, so we can instantly see a lot of information about a specific instrument by plotting a T_{r3} x T_{r2} graph, like the example (from [19]) presented in Figure 2, which shows the case for some known acoustic instruments. If the point falls close to origin the instrument should present a strong fundamental, because T_{r2} and T_{r3} would be small, so T_{r1} would be close to 1. An analogous reasoning holds for the other corners, so close to the right corner of the triangle is the case of strong higher frequency partials, while the last corner would be one related to strong mid-range partials.

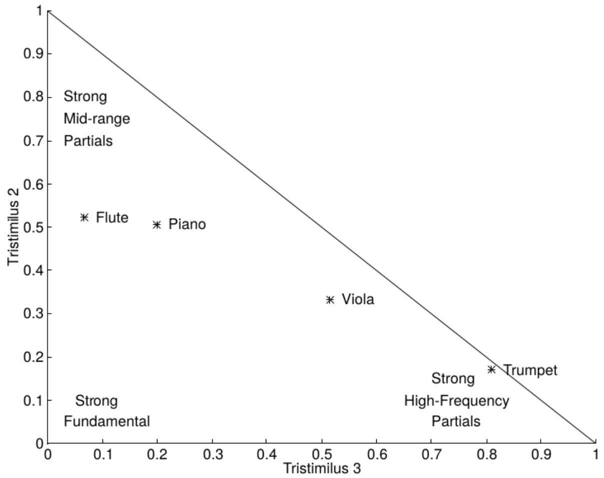

Figure 2. Visualising the tristimulus triangle. Source: [19]

A relation between the tristimulus metrics is also similar to another interesting psychoacoustic metric, the warmth, defined [25] as the energy contained in the partials up to 3.5 times the fundamental frequency of a sound, over the energy of the partials from 3.5 times the fundamental frequency to the uppermost harmonic. As highlighted in [25], it is likely that a processing which increases the warmth, or the tristimulus 2, will decrease the brightness, as the spectral centroid will become lower.

5.2.3 Irregularity

The irregularity of a spectrum is a measure of its “jaggedness”, as termed by McAdams *et al.* in [10], where they conclude that this metric is one of the most perceptually important parameter regarding timbre discrimination.

There are several formulas for the calculation of spectrum irregularity. We use the one given by [19]

$$Ir = \sum_{k=1}^{N} (a_k - a_{k+1})^2 / \sum_{k=1}^{N} a_k^2, \quad a_{N+1} = 0 \tag{11}$$

Since the expression for irregularity involves neighbouring partials, it reflects ripples in the spectrum, or the more extreme case of missing partials. An example of a sound with high spectral irregularity would be the square wave, or also the clarinet, which have components at odd harmonics only. These are known as having a hollow timbre, in contrast with the buzz timbre of an impulse train, which has zero irregularity.

5.3 Results

In this section we present and discuss some of the graphs we obtained in the study. All the graphs considering all the cases are available for download.

5.3.1 Brightness

Figures 3 and 4 show that the unprocessed waveforms start with a high brightness value and soon stabilise. Notice that the brightness values for A are smaller than those for C, and that was expected, due to the more modest settings used on the ‘303’ for its generation. A’s milder brightness can also be checked by listening to the samples.

When we consider the smoothed reconstructions for Waveform A (Figure 3) we see that the B100 is brighter than H100 during the early portion of the sound, but then they meet at a value smaller than the brightness for the pure waveform. B20 and H20 showed similar values, always close to the original. With Waveform C (Figure 4), the brightness values for all the 4 smoothers and the unprocessed case are practically the same.

Figure 5 shows the brightness values when only the AM is smoothed and Figure 6 when only the FM is smoothed. Notice that in this case the FM-only smoothing does not significantly impact the results, but the AM-only smoothing with the higher order windows augments the brightness. Similar behavior was observed for the other 3 waveforms.

Considering the processing of the highly resonant Waveforms (C and D), the brightness was not affected. Considering the milder resonance of A and B, the higher order smoothers imposed a smoothed brightness, but the small order smoothers augmented it.

It is interesting to notice that when we apply AM-only or FM-only smoothing in Waveform A with B100 and H100, the result is an increased brightness. However, when we apply both simultaneously the brightness is reduced, so a sort of cancellation happens, probably due to the creation of harmonics on the lower register.

5.3.2 Tristimulus

The tristimulus plots for the highly resonant C and D show a smaller variation comparing to the Waveforms A and B. Figure 8 show that tristimulus for C remains confined in the area where T_{r3} is high. Figure 7 show for A a larger excursion for T_{r2} and T_{r3} (and consequently, for T_{r1}).

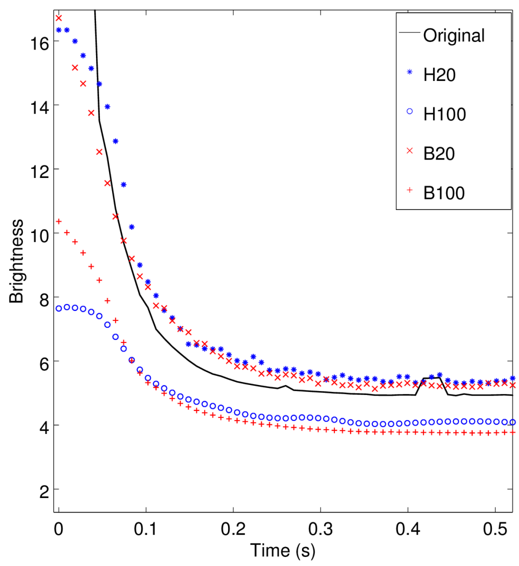

Figure 3. Comparison of brightness values over time for waveform A AM/FM smoothing using all the configurations considered. H20 smoother plotted with blue ‘$*$’, H100 with blue circles, B20 with red ‘x’, and B100 with red ‘$+$’. This legend code holds for all plots in this section.

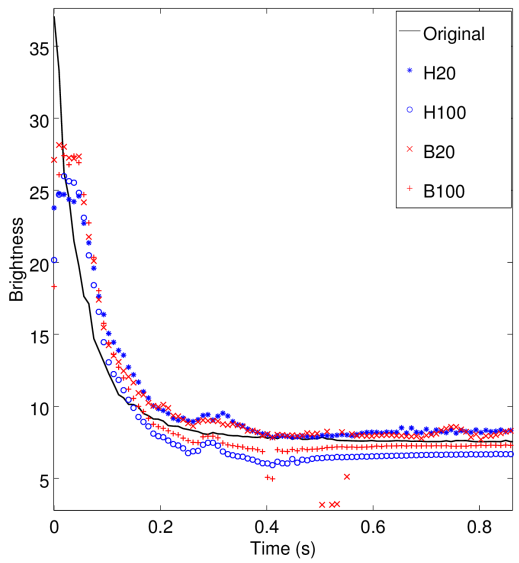

Figure 4. Comparison of brightness values over time for waveform C AM/FM smoothing using all the configurations considered

Figure 9 shows close values for tristimulus when considering only AM smoothing and the original case, but the FM-only smoothing (Figure 10) produces a larger variation for tristimulus. We can also check that the higher order smoothers were more effective for this variation than the low order ones.

Like what happened with brightness, we see that the tech-

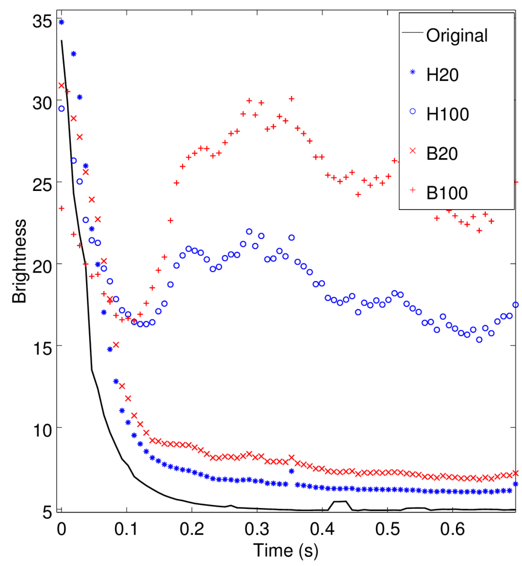

Figure 5. Comparison of brightness values over time for AM-only smoothing, using waveform A

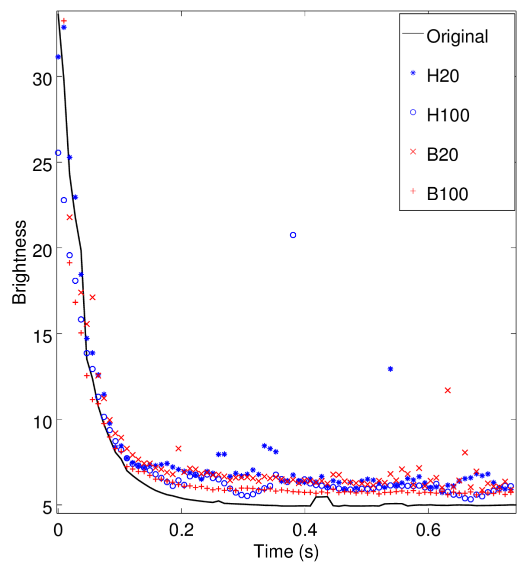

Figure 6. Comparison of brightness values over time for FM-only smoothing, using waveform A. The spikes observed are artefacts from the harmonics finding process

nique does not significantly affect the tristimulus for the waveforms with high resonance, but for mild resonance sounds the smoothers impose a variation proportional to their length. In contrast with the brightness values, here it is the FM smoothing that is the main source of variation in the processing.

5.3.3 Irregularity

The plots for the irregularity of Waveforms A (Figure 11) and B show a constant value throughout the sound. That is

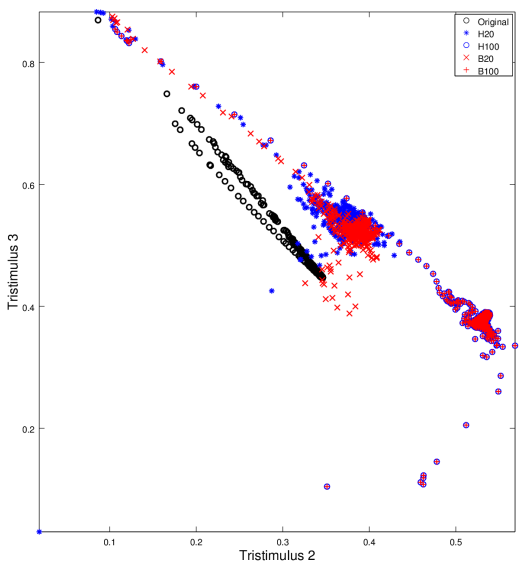

Figure 7. Tristimulus triangle for waveform A AM/FM smoothing with all the configurations considered

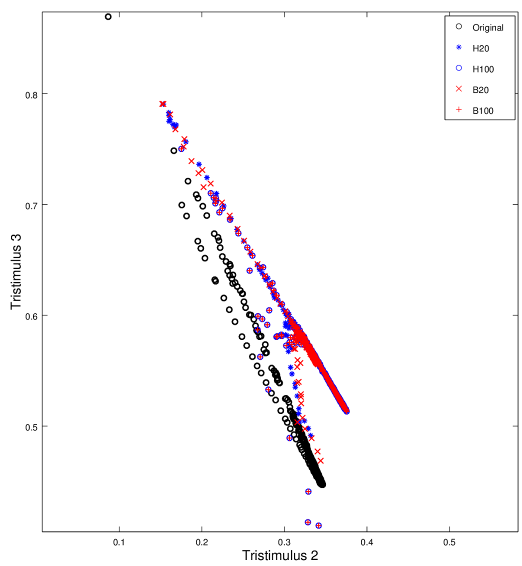

Figure 9. Tristimulus triangle for AM-only smoothing, using waveform A

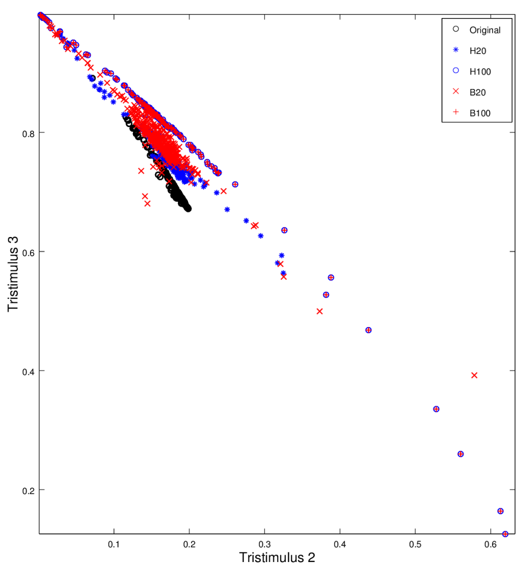

Figure 8. Tristimulus triangle for waveform C AM/FM smoothing with all the configurations considered

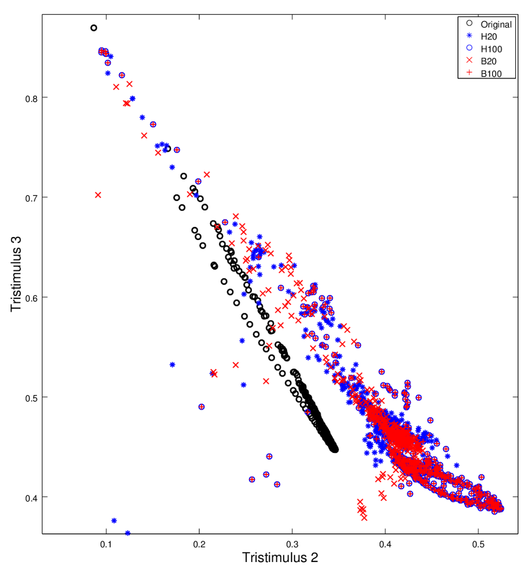

Figure 10. Tristimulus triangle for FM-only smoothing, using waveform A

not the case for C (Figure 12) and D, which show a periodically varying irregularity, with a large excursion.

Figure 11 shows that the H100 smoother matches the unprocessed case when processing A, while the small order smoothers decrease the irregularity and the B100 doubles it. The plots considering Waveforms C and D show similar values for all the smoothing configurations, with non-fluctuating values smaller than the originals.

Figures 13 and 14 show the cases for the AM-only and FM-only smoothing for the Waveform C. Notice that for the AM-only smoothing case the values are all similar, and smaller than the original, while for the FM-only smoothing only the higher order smoothers decreased the irregularity.

6. CONCLUSIONS

As a general trend, from the brightness and tristimulus points of view, it is not that effective to smooth signals with strong resonance, although it certainly changes the sound. Also, the tristimulus is more affected by smoothing of the instantaneous frequency component, while the irregularity is more affected by the AM smoothing.

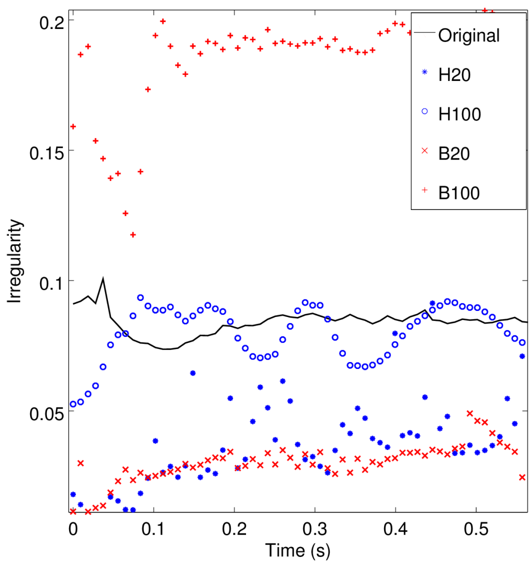

Figure 11. Comparison of irregularity value over time for waveform A AM/FM smoothing using all the configurations considered

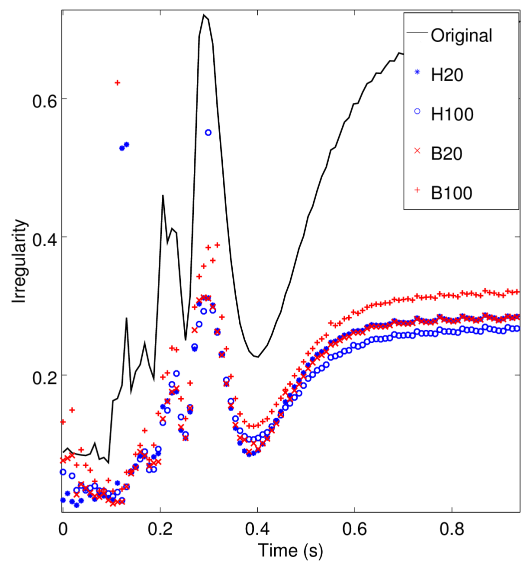

Figure 13. Comparison of irregularity value over time for AM-only smoothing, using waveform C. The spikes observed are artefacts from the harmonics finding process

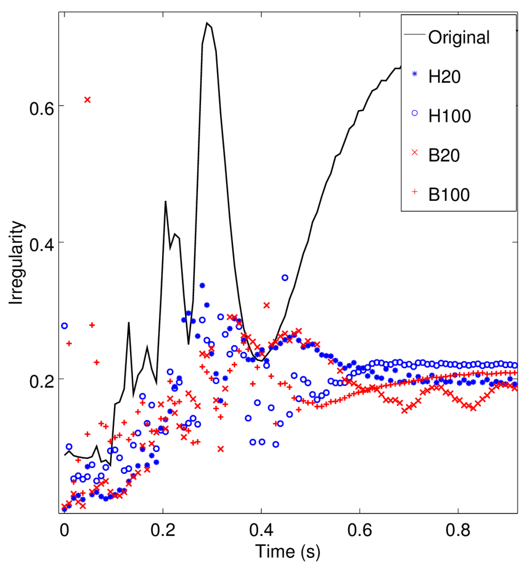

Figure 12. Comparison of irregularity value over time for waveform C AM/FM smoothing using all the configurations considered

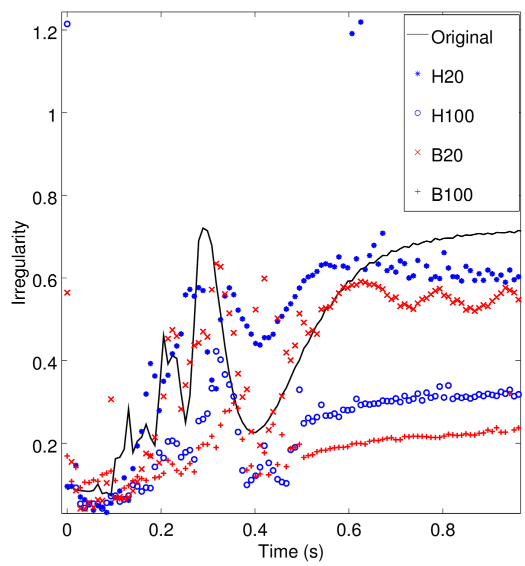

Figure 14. Comparison of irregularity value over time for FM-only smoothing, using waveform C. The spikes observed are artefacts from the harmonics finding process

According to the psychoacoustic metrics obtained in the study, it seems that the more irregular the input signal spectrum is, the more similar the perceptual outcome of smoothing will be in comparison to the original, regardless of the window used. Values will be typically lower, indicating a smearing or flattening of the spectrum, typical of modulations with high depth or index. There could be a relationship between this flattening and the spectrum whitening described in [26], but this needs further investigation.

The processing of modest and mild resonances, however, presents variations compared to the original case, and the new sounds obtained are musically interesting, indicating that AM/FM based techniques can be useful for the introduction of liveliness into flat resonance sounds. According to what was expected, the longer smoothers led to sounds with less artefacts in comparison to the sounds produced with the short smoothers. Also, the Hanning smoothers produced less artefacts than the Boxcar's.

Currently we are investigating possibilities to generalize the framework as a suite of AM/FM audio effects. It seems that all smoothers are an interesting possibility for the processing of sounds with a modest to medium resonance, as the overall original brightness shape is preserved, so the effect keeps a lot of the sound's original feel.

Acknowledgments

The research leading to this paper was performed during Antonio's internship period at Maynooth University, and was partially supported by CAPES (proc num 8868-14-0).

7. REFERENCES

[1] G. Loy, *Musimathics - Volume 1.* Cambridge, MA, USA: MIT Press, 2006.

[2] X. Rodet, "Time-domain formant-wave function synthesis," *Computer Music Journal*, vol. 8, no. 3, 1984.

[3] W. Kaegi and S. Tempelaars, "VOSIM - A new sound synthesis system," *Journal of the Audio Engineering Society*, vol. 26, no. 6, pp. 418–425, 1978.

[4] V. Lazzarini and J. Timoney, "Theory and practice of modified frequency modulation synthesis," *Journal of the Audio Engineering Society*, vol. 58, no. 6, 2010.

[5] M. Ishibashi, "Electronic musical instrument (patent US 4658691)," Patent 4 658 691, April, 1987.

[6] J. Chowning, "The synthesis of complex audio spectra by means of frequency modulation," *Journal of the Audio Engineering Society*, vol. 21, no. 7, pp. 526–534, 1973.

[7] M. Russ, *Sound synthesis and sampling*, 3rd ed. Burlington, MA, USA: Taylor & Francis, 2008.

[8] B. Picinbono, "On instantaneous amplitude and phase of signals," *IEEE Transactions on Signal Processing*, vol. 45, no. 3, 1997.

[9] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal-part 1: Fundamentals," *Proceedings of the IEEE*, vol. 80, no. 4, Apr 1992.

[10] S. McAdams, J. W. Beauchamp, and S. Meneguzzi, "Discrimination of musical instrument sounds resynthesized with simplified spectrotemporal parameters," *J. Acoust. Soc. Am.*, vol. 105, no. 2, pp. 882–897, 1999.

[11] R. Sussman and M. Kahrs, "Analysis and resynthesis of musical instrument sounds using energy separation," in *Proceedings of the 1996 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP-96)*, vol. 2, May 1996, pp. 997–1000 vol. 2.

[12] I. Arroabarren, X. Rodet, and A. Carlosena, "On the measurement of the instantaneous frequency and amplitude of partials in vocal vibrato," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 14, no. 4, pp. 1413–1421, 2006.

[13] J. L. Flanagan and R. M. Golden, "Phase vocoder," *Bell System Technical Journal*, vol. 45, no. 9, 1966.

[14] T. Stinchcombe, "Diode ladder filters," http://www.timstinchcombe.co.uk/index.php?pge=diode, online; accessed 27 April 2015.

[15] J. Flickinger, "Future retro - your guide for the revolution," http://www.future-retro.com/pdf/revolutionmanual.pdf, online; accessed 19 June 2015.

[16] R. Corporation, "TB-303 owner's manual."

[17] R. McAulay and T. Quatieri, "Speech analysis/synthesis based on a sinusoidal representation," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 34, no. 4, pp. 744–754, Aug 1986.

[18] A. H. Nutall, "Some windows with very good sidelobe behaviour," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 29, no. 1, pp. 84–91, Feb 1981.

[19] K. Jensen, "Timbre models of musical sound: From the model of one sound to the model of one instrument," Ph.D. dissertation, 1999.

[20] R. A. Garcia and K. M. Short, "Signal analysis using the complex spectral phase evolution method," in *Audio Engineering Society Convention 120*, May 2006.

[21] J. Wang, J. Timoney, and M. Hodgkinson, "Using apodization to improve the performance of the complex spectral phase estimation (CSPE) algorithm," in *Proceedings of the China-Ireland International Conference on Information and Communications Technologies*, 2009.

[22] J. Beauchamp, "Synthesis by spectral amplitude and 'brightness' matching of analyzed musical instrument tones," *J. Audio Eng. Soc*, vol. 30, no. 6, 1982.

[23] G. von Bismarck, "Sharpness as an attribute of the timbre of steady sounds," *Acta Acustica united with Acustica*, vol. 30, no. 3, pp. 159–172, 1974.

[24] H. F. Pollard and E. V. Jansson, "A tristimulus method for the specification of musical timbre," *Acta Acustica united with Acustica*, vol. 51, no. 3, pp. 162–171, 1982.

[25] T. Brookes and D. Williams, "Perceptually-motivated audio morphing: Warmth," in *Audio Engineering Society Convention 128*, May 2010.

[26] J. J. Wells, "Methods for separation of amplitude and frequency modulation in Fourier transformed signals," in *Proceedings of the International Conference on Digital Audio Effects (DAFx-10)*, Graz, Austria, 2010.