Dissonance Model Toolbox in Pure Data

Alexandre Torres Porres
MOBILE, ECA, USP
Av. Prof. Lúcio Martins Rodrigues 443
São Paulo, Brazil, 05508-020
porres@gmail.com

Abstract

This paper presents the development of a toolbox based on Dissonance Psychoacoustic Models departing from Terhardt's theory, which includes Sharpness, Roughness, Tonalness, Root Relationship & Pitch Commonality. Dissonance perception studies are found under many approaches in psychology and musicology. There is no final word on how each parameter contributes to its perception of Dissonance. A review of the theories has been discussed by the author. Here, the theory is only briefly discussed, the implementation section carries a bit more information on each attribute before the patches and creative applications are presented.

Keywords

Dissonance, Psychoacoustic Models, Live Electronics.

1 Introduction

This paper is part of a PhD program currently in progress. The research is on Psychoacoustic Dissonance Models to be applied over digital audio signals in realtime. The usage of this theory has been mostly applied in computer aided composition, but not much has been done for Live Electronics. Another contribution is to make such studies and concepts more accessible to composers and musicians.

The most familiar term is Roughness, which has usually been the only attribute present in Dissonance Models like Sethares' [1][1]. It's common for such attributes to be usually implemented independently, so part of this research relies on the investigation of many different implementations of each attribute, as to put them together in a Pure Data library dedicated to Dissonance Modeling.

1.1 Sensory Basis of Musical Dissonance

In Terhardt's theory [2], the sensory basis of Musical Dissonance depends on perceptual attributes divided into two groups. The "Sensory Dissonance" group refers to innate (objective) aspects of perception. The term was coined to distinct itself from subjective aspects of the "Musical Dissonance" concept (known to carry cultural factors) and made lots of research under a scientific approach possible [3]. But some authors have worked with an expanded notion of the sensory basis for Musical Dissonance [2, 4, 5] in which Sensory Dissonance is part of it. Complementary, "Harmony" is the second group. But there is still a clear separation from cultural aspects of perception [5 pp.48].

Nevertheless, inside a sensory approach, there are still further distinctions from lower to upper levels, such as: Psychoacoustics, Music Psychology and Cognitive Psychology. But an outline of Music Cognition research is not considered here. Many new outcomes are appearing, but a current review of this findings is not pertinent to the goal of research, as no applicable model has yet emerged from these researches.

Sensory Dissonance	Harmony
- Sharpness - Roughness - Tonalness	- Root Relationship - Affinity of Tones: Pitch Commonality

Table 1: Perceptual Attributes of Musical Dissonance according to Terhardt.

While the attributes of Sensory Dissonance can be applied to any kind of sound, the "Harmony" group deals with what is called "musical sounds", and that means basically sound spectra such as those of musical instruments – and not noise or speech.

Moreover, the Harmony group is related to sounds that evoke the perception of Pitch, and this does not discard sounds that have a weak and diffuse pitch image, like inharmonic sounds. But a very important detail is that it's related to chord structures with three or more notes (as in Root Relationship), and the relationship between successive tones or chords (as in Pitch Commonality). So it's very much related to the concept of "Harmony" in music – hence the given term.

2 Implemented Perceptual Attributes

2.1 Sharpness

The concept of Sharpness can also appear under the term of *Brightness* or *Density*, which are equivalent or closely related [6]. The

1 He was clearly aware of other attributes and describes them in his book while explaining his approach/choice.

sensation of sharpness depends on the quantity of a spectrum's energy in the high register.

Usually, Sharpness Models are based on data from spectral centroid. So a measure of sharpness increases with spectral centroid, and is like pitch and its relation to frequency. The model of Sharpness is actually defined as the perceptual equivalent to the spectral centroid, but computed using the specific loudness of the Critical Bands. The unit of sharpness is *acum*, which is latin for sharp. 1 *acum* is attributed to a 60dB narrow-band noise (less than 150Hz) at 1Khz.

$$S = 0.11 \frac{\sum_{z=1}^{nband} z \,.\, g(z) \,.\, N'(z)}{N} \qquad (1)$$

Where z is the band index, N is the total loudness, $N'(z)$ is the specific loudness and g(z) is a function = 1 if z < 15 and = 0.66 exp(0.171z) otherwise.

The model for Sharpness is then relatively simple, and regarded as a LLD (Low Level Descriptor). LLDs belong to an even lower and more objective level than other Psychoacoustic Attributes. They have been gaining accessibility in computer music, and there are a few Pd libraries around [7-9].

The chosen definition and model is from Zwicker and Fastl [4], an implementation in Pure Data was already available in Jamie Bullock 's LibXtract [9]. Because Sharpness is a rather simple and straightforward concept, there wasn't any change to be made, any enhancement to be proposed, or debate on it's accuracy, unlike the other attributes.

2.2 Roughness

Roughness is the physical correlate of Amplitude Fluctuations [10]. Slow fluctuations (at a rate lower than 20Hz) are known as "beats". Roughness or "fast beats" are for rates over 20Hz up to a Critical Bandwidth. So the value of Roughness always takes into account a frequency difference, or musical interval for that matter!

The results of Plomp & Levelt (see Figure 1) gives us a rule of thumb that maximum Roughness is perceived at a rate (or musical interval) that corresponds to one fourth of the Critical Bandwidth [11]. Since Helmholtz [12], Roughness is usually regarded as the main or only aspect of dissonance under a sensory approach. This is why Roughness in the graphic of Figure 1 is depicted as a consonance / dissonance measure.

Parncutt [13] offers an equation that fits the Plomp & Levelt's results from Figure 1 (see Equation 2). It gives us a vertically flipped graph of Figure 1, with the maximum Roughness being equal to 1.

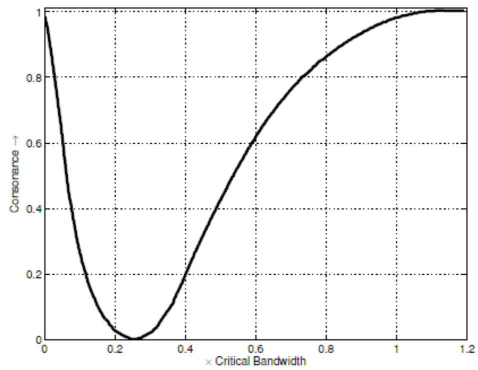


Figure 1: Roughness of simultaneous and equally loud sinusoids on vertical axis as dissonance. Frequency difference (in the Critical Band scale) on the Horizontal axis.

$$R = [e\,(b / 0.25)\exp(-b / 0.25)]^2 \qquad (2)$$

Where b is the frequency difference in the critical band scale (bark), and R = 0 if $b > 1.2$.

The Roughness of complex tones is measured by adding the result from each combination of pairs of partials. Dyads formed by complex tones are also measured by adding the values from each combination of partials from both of the complex tones in the dyad. See figure 2 below.

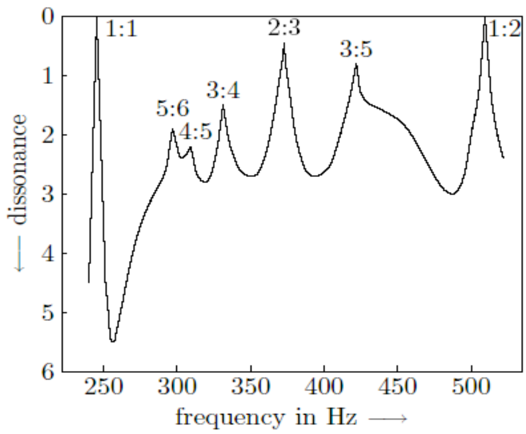


Figure 2: A Roughness Curve of complex tone dyads from Plomp & Levelt's model [11].

The above graph is much like Figure 1, but for complex tones and depicted over frequencies in Hz instead of the Critical Band scale. Maximum consonance points are marked with the frequency ratio (musical intervals) of the dyad. Such graphs are also known as "Dissonance Curves", being Roughness so closely related to Sensory Dissonance.

You can measure the Roughness of any sound without considering it to be a musical sound in a musical context. But, as Roughness is directly related to frequency intervals, it's a perfect feature to measure musical intervals in a musical context, and that's exactly what a graph such as Figure 2 is about. And because Roughness has been considered the most important attribute of Dissonance perception, the graph of Figure 2 has been considered a Dissonance Curve. More about Dissonance Curves in the Creative Applications section of this paper.

Many models have been developed over the years based on Plomp & Levelt [11], who only worked with equally loud pure tones. So the main difference and most controversial detail among all these models is how to compute Roughness for amplitudes that are not equal.

Regarding this issue, there are many proposals of how to change the Roughness value from Figure 1 according to an “amplitude weight function”. Clarence Barlow [14] has a nice revision of many Roughness models and developed a particularly careful one. What makes his model quite complete is that it accounts the masking effects and Robson and Dadson Equal Loudness Curves [15]. But as to compute Roughness for different amplitudes, Barlow simply extracts a quadratic mean of the amplitudes in Sones.

A different approach than from the ones based on Plomp & Levelt are based Amplitude Modulation [4, 16] and on a model of the peripheral auditory system, like the work of Pressnitzer [17], which is also implemented in Pd.

But then, Vassilakis [10] presented a revision of Plomp & Levelt models related to the unequal amplitudes issue, and proposed an amplitude weight equation based on the Amplitude Fluctuation Degree, which is also a revision of the usage of Amplitude Modulation Depth in Roughness modeling.

$$Amp(A_1.A_2) = (A_1.A_2)^{K_1}.K_2.\left(\frac{2.A_1}{(A_1 + A_2)}\right)^{K_3} \qquad (3)$$

Where $K_1 = 0.1$, $K_2 = 0.5$, $K_3 = 3.11$ and A_1 & A_2 are, respectively, the smallest and biggest of the amplitudes.

The Roughness model proposed here is based on Plomp & Levelt family and is greatly an implementation of Clarence Barlow's revision of the current models [14]. But we're proposing and investigating the inclusion of the Vassilakis’ equation.

This was actually first proposed in 2007 on a paper for the Pd Convention in Montreal [18], so more information about Roughness and a first description of the model can be found there. This was first presented as a Pd Patch, but now there is an object available, which has also been announced and published in previous publication by the author [19].

A new update version of this object is available now at the time of this publication, which is clearly an update and development of these previous work.

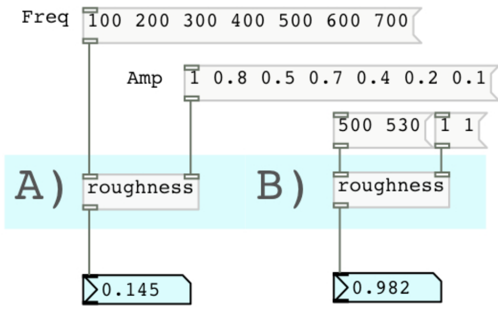


Figure 3: The [roughness] object works with lists of amplitudes and frequencies. A) is the Roughness of a harmonic complex tone, and B) the Roughness of two sinusoids (500 / 530 Hz).

2.3 Virtual Pitch, Tonalness (*Klanghaftigkeit*)

The term Tonalness refers to “the degree to which a sonority evokes the sensation of a single pitched tone” [20] – in a sense that sonorities with high Tonalness evoke a clear perception of pitch. As a component of Consonance, Tonalness is the “ease with which the ear/brain system can resolve the fundamental” [21], being the easier, the more consonant. Right next we'll see how this idea is strongly related to the Virtual Pitch theory. Virtual Pitch is a key feature of not only Tonalness, but also Root Relationship and Pitch Commonality. Parncutt's Pitch Commonality Model [20] gives us all of these attributes.

Besides Tonalness, Psychoacousticians have also used the term "tonality" and both have been defined and used in different ways, some difficulties have arisen because of the translation of the terms into and from German [22]. Huron [23] prefers Toneness. Terhardt adopts Tonalness as a translation of the German term *Klanghaftigkeit* and relates it to his Virtual Pitch theory. Parncutt is based on Terhardt’s theory and provides us with a model of Tonalness as part of his Pitch Commonality model, but translates Klanghaftigkeit as Sonorousness [20].

A Spectral Pitch is the perception of a sine tone component present in a complex tone. Sinusoids can only evoke Spectral Pitches. A Virtual Pitch comes into play to explain the pitch perception of complex tones (i.e. formed by a collection of Spectral Pitches).

For harmonic complex tones, the perception of Virtual Pitch is usually the same as the fundamental tone of that sonority. But if the fundamental is missing, it is still possible to have the same pitch perception – hence the term “Virtual” –, this is the most important implication of the theory. Another implication is that complex tones can evoke more than one Virtual Pitch sensation. And the more Virtual Pitches evoked, the less Tonalness we have.

Parncutt's Pitch Commonality Model [20] gives us primarily a Spectral and Virtual Pitch Weight. The first step is to apply loudness and masking functions on the spectrum. The amplitudes are then defined in *Audible Levels* (*AL*) in dB. Next, the model derives a spectrum with amplitudes in "*Pure Tone Audibilities*", which are actually Terhardt's Spectral Pitch Weight – a measure for the intensity of a Spectral Pitch – given as follows:

$$Sw(f) = 1 - \exp\{-AL(f) / AL_0\} \qquad (4)$$

Where $Sw(f)$ is the Spectral Pitch Weight of a sine tone component frequency in Hertz. $AL(f)$ is the Audible Level of this sine tone component in dB, the AL_0 was estimated experimentally at about 15dB.

The "*Audibility of a Complex Tone*" is the same as the Virtual Pitch Weight – a measure of the intensity of the perception of a Virtual Pitch. To find it, we need to look for harmonic patterns in the spectrum, so it may be regarded as the "measure of the degree to which the harmonic series, or part thereof, is embedded in the audible spectrum of a sonority at a given pitch" [20]. More than one harmonic series pattern can then be found, resulting in multiple Virtual Pitch sensations.

Parncutt uses a harmonic series template with ten components and sweeps it over the sonority to look for matches. Every time one or more harmonics from the template match a sine tone component in the spectrum, the pitch corresponding to the fundamental of the template gets a Virtual Pitch Weight. Check the mathematical formulation below:

$$Vw(f) = \Sigma_{1\text{-}n} [\text{sqrt}\ (Sw(f.n)/n)]^2 / Kt \qquad (5)$$

Where $Vw(f)$ is the Virtual Pitch Weight of the fundamental frequency on the template in Hertz, n is the number of the harmonic on the template, and Kt is typically about 3.[2]

Parncutt gives us both *Pure* and *Complex Tone Sonorousness*, which are, respectively, dependent on *Spectral Pitch Weight* and Virtual Pitch Weight. The latter is the equivalent to the Tonalness of a complex sonority. *Pure Sonorousness* or *Pure Tonalness* (*Tp*) is a quadratic sum of the Spectral Pitch Weights. *Complex Sonorousness* or *Complex Tonalness* (*Tc*) is given by the highest Virtual Pitch Weight. Both values can be normalized to one by multiplying, respectively, to 0.5 and 0.2.

A different Tonalness model is provided by Paul Elrich [21]. It is not based on a model of Virtual Pitch model, and is being considered as an alternate implementation.

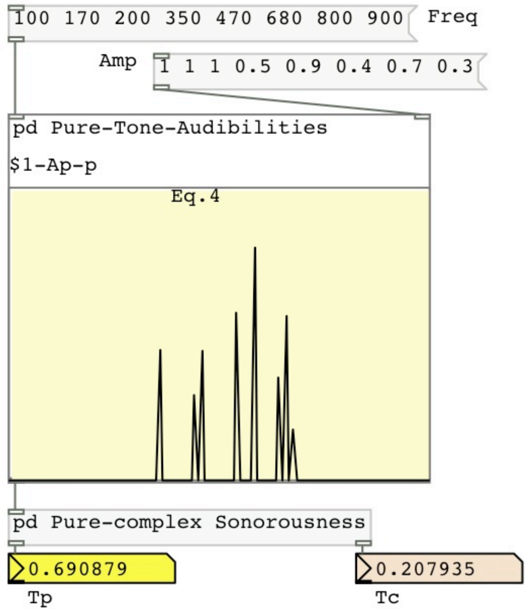


Figure 4: implementation of Tonalness/Sonorousness [20], Pure Tone Tonalness is Tp, and Complex Tone Tonalness is Tc. The implementation here is not as an object, but as an abstraction that shows the array with Spectral Pitch Weights (Pure-Tone-Audibilities) from Eq. 4, which is the spectrum after the loudness and masking function (note how the second partial of 170Hz has a smaller amplitude due to the masking effect)

2.4 Root Relationship (*Basse Fondamentale*)

The Virtual Pitch model also provides a way to find the probable root of a chord or spectrum. These implications of Virtual Pitch depend on the concept that complex tones and chords may evoke several Virtual Pitches (i.e. Pitch Multiplicity).

Further steps in the Pitch Commonality model [20] can provide us a measure of Root Relationship based on Pitch Salience. Both Spectral and Virtual Pitch Weights are mixed to provide a general Pitch Weight Profile – in the case of the same pitch having both a Spectral and Virtual Weight, the highest of them prevails.

$$M' = \Sigma\ Pw(f) / \max(Pw) \qquad (6)$$

$$M = M'^{\,Ks} \qquad (7)$$

$$Ps(f) = [Pw(f) / \max(Pw)]\ .\ [M / M'] \qquad (8)$$

Where $Pw(f)$ is the Pitch Weight of a frequency and $\max(Pw)$ the highest Pitch Weight. Ks *is* another free parameter which has a typical value of 0.5. M is the Pitch Multiplicity and $Ps(f)$ is the Pitch Salience of a frequency.

2 Kt is a free parameter that depends on the mode of listening, chosen such that resulting values of Virtual Pitch Weight are correctly scaled relative to Spectral Pitch Weight.

Pitch Salience is defined as the probability of consciously perceiving (or noticing) a given pitch. The most salient of the tones is considered to be the possible root. The Pitch Salience of a frequency depends on the Multiplicity, which is initially estimated as *M'*. Check equations 6 to 8.

Thus, a Model of Root Relationship can be provided as the maximum value of Pitch Salience, and it is considered more prominent if the maximum salience is much higher than the others. A Pitch Salience Profile is the set of Saliences for all frequencies, and is the fundamental data to calculate Pitch Commonality, described next.

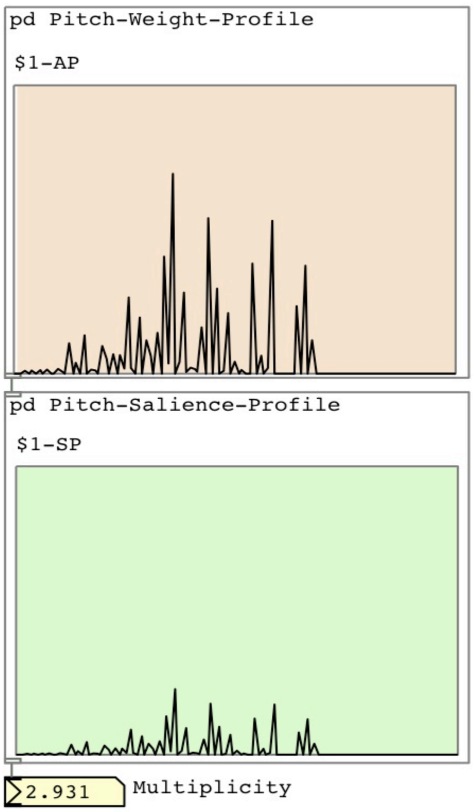


Figure 5: The Pitch Weight Profile fro,m the input of Figure 4 above, and The resulting Pitch Salience Profile below – as well as the Multiplicity measure.

2.5 Affinity of Tones: Pitch Commonality

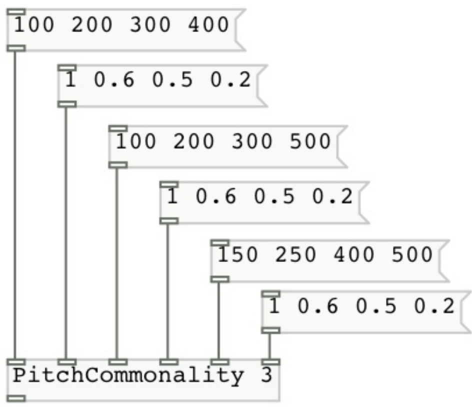
Affinity of Tones deals with the concept that a tone may be sensed as similar to another of a different pitch [2]. Terhardt raises an internal auditory sense as one of the aspects of Tone Affinity, mainly for the Octave and Fifth. But a second aspect is Pitch Commonality, which is based on the concept that different sonorities may evoke pitches in common as a result of their Multiplicity. So the more pitches are evoked in common, the bigger the Pitch Commonality we have.

Parncutt presents the Pitch Commonality Model as a Pearson correlation coefficient of the Pitch Salience Profiles of two different sonorities. It is equal to 1 in the case of equal spectra and hypothetically -1 for “perfect complementary sonorities”.

As for changes in the original model as described by Parncutt and available[3] in C code, the Pd implementation allows a finer division of tones, as it was originally a fixed array set of 120 pitch categories, in which each match a scale step in the 12-tone equal temperament over 10 octaves. This array of 120 Pitch Categories also applied for Root relationship and Tonalness.

The expansion considers a set of 720 pitches as it allows steps of 1/12 tone, which provides a very good approximation of Just Intonation intervals up to the 11th harmonic and other microtonal tunings. One example that approximately fits this division is Partch's tone system [24].

A much finer division is also possible. For example, steps of one cent (an array of 12000 elements) can account for a rather continuous frequency range. The harmonic template can also contain more partials than Parncutt's 10 harmonics set, and a typical chosen value is 16 harmonics.




```
1 0.839868 0.581165;
0.839868 1 0.664414;
0.581165 0.664414 1
```

Figure 6: The Pitch Commonality object, takes the number of spectra to compare as an argument and gives a list of all the Pitch Commonality combinations as a list.

The result from figure 6, which measures the Pitch Commonality of 3 pitches, gives as a result a list of 9 elements, which are shown in the form of a matrix instead. Lets label the Pitches from left to right as A, B & C. The first line is then; (A + A) , (A + B), (A + C). Second line is (B + A), (B + B), (B + C), and the third is

3 The code is available at: <http://www.uni-graz.at/richard.parncutt/parpitchcode.html>

(C + A), (C + B), (C + C).

2.6 Final Considerations

So Parncutt's Pitch Commonality model has goes over a few stages to reach the final step that's actually related to Pitch Commonality. And over the process we also have the Root Relationship as a byproduct to cover both attributes of the Harmony group.

The Tonalness measure is not actually a necessary step to calculate Pitch Commonality, but a complementary Sensory Consonance feature. It could then be completely ignored to open for a different Tonalness implementation such as Elrich's [21]. Sharpness and Roughness are not considered in Parncutt's model. And a complete dissonance model should include these other attributes.

One extra feature presented by Parncutt is a Pitch Distance Model based on the Pitch Salience Profile of two complex tones, but Pitch Distance/Proximity itself is not an attribute of Dissonance. It can be useful, though, for compositional purposes. Other similar features that relate to these perceptual concepts are equally welcome in this research, as the final goal is to have a set of patches and objects dedicated to composition and live electronics. Another such extra feature is Barlow's model of Harmonicity [14], which was also implemented in Pd.

3 CREATIVE APPLICATIONS

3.1 Dissonance Descriptors

Sharpness is a LLD and a Dissonance attribute. One creative possibility being applied with LLDs is to use them as timbre descriptors to match sounds with similar characteristics, in techniques such as concatenative synthesis and other similar processes – as the ones provided by William Brent [25]. The perceptual attributes of Dissonance Models can be applied in the same fashion in real time applications, and expand this process in conjunction with other LLDs.

By using any of the attributes as a descriptor, one can operate basic Live Electronic processes with such control data – like triggering events, DSP processes, and so on. So a more dissonant sound, such as described by a low Tonalness measurement can trigger any sample, or switch to a particular DSP transformation, etc.

The *sCrAmBlEd?HaCkZ!*[4] Performance by Sven König uses the matching of descriptors to reconstruct a live sound input from a sound bank of audio snippets. Miller Puckette's performance with Rogério Costa in the São Paulo Pure Data convention had a similar process, where both sax and guitar were fed to a buffer and re-used to reconstruct lines from live input.

Kind of similar processes are possible and expand these ideas. A high Roughness spectrum can trigger another spectrum alike or not. Or measure the Dissonance descriptors of a dyad/chord and form progressions by recalling a sound from the buffer or a sound bank that would have, for example, a high or low pitch commonality with a live input.

Such dissonance models have actually been used this way to assist the composer in defining the chord progression or organize the structure of a piece – such as the work of Clarence [14] and Sean Ferguson [26]. But besides the computer aided approach, new real time possibilities are available to be explored and discovered.

Nevertheless, the computer aided approach is still possible. It's not much what Pd was meant to be used for, as it is an offline process by concept (requiring the analysis and generation of tables and data to choose from and test). But even so, the Pd implementation can still help on that approach. Not only that, but it can provide a meeting point half way.

For example, several analysis can be done beforehand and stored in a data bank. That is actually how many processes in timbre matching or the *sCrAmBlEd?HaCkZ!* Performance works. As for the case with Dissonance descriptors, we can generate and store information like Dissonance Curves.

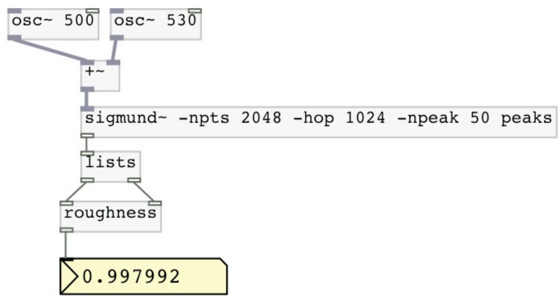


Figure 7: Roughness being measured in realtime. The input and result is about the same Figure 3 B). The analysis from [sigmund~] is used to generate lists of amplitudes and Frequencies. So in the same way it can be applied to provide input for the Tonalness/Pitch Commonality model in realtime.

3.2 Dissonance Curves

An expansion from a merely momentary description/measurement is the usage of Dissonance Curves, which are graphs of dissonance on the vertical axis over musical intervals on the horizontal axis. It was mentioned how Roughness ratings are fit for that, and Figure 2 is an example of it.

4 <www.popmodernism.org/scrambledhackz/>.

3.2.1 Roughness Curves

The Roughness Curve finds the most consonant intervals as the absence of Roughness, and this happens when the partials from a complex tone are aligned with the partials of another complex tone. The alignment of partials thus depends on the distribution of partials. On Figure 2 we have a harmonic complex tone, so the partials align in musical intervals that correspond to this harmonic relationship – i.e. just intervals.

So Roughness implies a strong relationship between the distribution of partials and matching musical intervals. In a sense that if you know which are the intervals of minimum Roughness, you can predict the structure of the spectrum.

The models based on Plomp & Levelt are more widely used for musical applications, which have been basically the measurement of musical intervals' "dissonance". This family of model was chosen in the Pd implementation because of that. These models don't need a digital signal over time as an analysis input, and can have input data in the form of frequency and amplitude lists. This allows the generation of curves such as Figure 2 without having to generate that signal and then analyze it.[5]

All you need is an algorithm like Sethares' [1] to generate such curves just by a snapshot of a spectrum, then duplicate it and shift the copy in different intervals over a specified range. Or have two different spectra and have one fixed while the shifts.

Figure 8: Roughness Curve of a [triangle~] oscillator, more on Figure 9.

5 It's also possible to perform analysis of temporal sounds with FFT anyway.

A similar algorithm can generate a curve from 2 different spectra, keeping one of them still while the other shifts.

3.2.2 Autotuner (Adaptive Tuning)

One main usual application of Dissonance Curves is on tuning theory, as this is a perfect tool for finding the most consonant or dissonant interval according to a specific spectrum. More about this can also be found on the previous PdCon paper [18]. By that time, a Roughness model was implemented as a patch, and was applied in an "Adaptive Tuning" module, which was now updated to a newer version, but you can still check that previous publication for more info.

The "Adaptive Tuning" concept given by Sethares [1] is basically an Autotuner, which is a more common term that I prefer now. It is based on the generation of a Dissonance Curve of a spectrum, then dissonant and consonant intervals are found. So, for any note input in a scale, you can automatically re-tune it to a scale step from the curve.

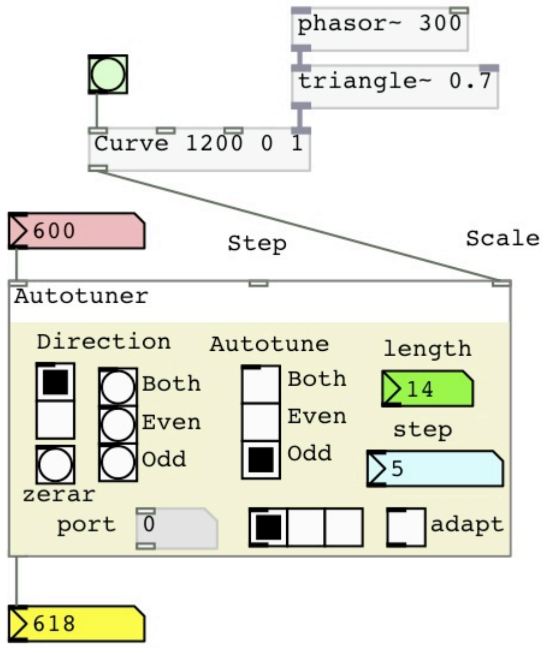


Figure 9: Autotuner module being fed with the scale from Figure 8.

The Autotuner can receive any scale list with intervals in cents. The [Curve] object is doing that from maximum and minimum points of the Roughness Curve from Figure 8. So it always alternates between maximum (odd) and minimum (even) intervals. You can chose to adapt to the closest even/odd step or both. We see that the equally tempered tritone of 600 cents was retuned to 618, which is actually a bit rougher.

3.2.3 Finer Dissonance Curves

Roughness has been solely used to account for a Dissonance Curve, but it isn't enough to describe this complex and multiple perceptual phenomenon. Roughness is not the only attribute of Sensory Dissonance, and besides Sensory Dissonance there's also Harmony attributes that account for Musical Dissonance. So how much each attribute contributes to the overall measurement of Dissonance?

This is still in debate, and recent researches show that Harmonicity plays a more important role than previously considered [27], being Harmonicity pretty close and related to Tonalness. But further research is needed to outline and model the perception of Dissonance in a finer way. Roughness is an important cue, but it can't be solely responsible for a so called Dissonance Curve.

With other attributes than Roughness, we can refine or expand the concept of Dissonance Curve to other features, such as points of maximum and minimal Tonalness, and a measure of the Pitch Commonality between two spectra over a specified range.

And as it gets more complex, it relates a lot to the computer aided process as generating all these tables aren't that fast yet to be constantly computed in realtime. On the other hand, previous analysis and the generation of a data bank is useful for further matching in realtime. This expands the idea of comparing just sounds in a bank by adding another dimension of the transposition of such sounds, and the generation of curves for more than one attribute.

Now other tools can come into play in this process. One is related to spectral transformation os sounds, and another is a pitch shifter that can provide transposition of sounds. The author uses a Phase Vocoder to transpose a pitch but it can also mix it with the untransposed original sound. More about it on the next section.

3.3 Spectral Transformations

3.3.1 Complex Modulation and Pitch Shift (compress and expand)

The technique of complex modulation, also known as Single Side Band Modulation, readily comes in Pd's audio examples patches (H09.ssb.modulation), which are part of Miller Puckette's book[6].

It performs a linear shift in the spectrum up or down. If shifted upwards, the relationship between partials become proportionally narrower. So if the spectrum is transposed back to match the original fundamental, we can say we have compressed the spectrum. Conversely, the spectrum can be expanded (partials become proportionally wider) if the spectrum

For the pitch shift you need first a pitch tracker like [sigmund~], and then a tool such as the phase vocoder to promote the Pitch Shift. The author has developed a phase vocoder patch with many capabilities in two versions[7]: one for realtime live input, and another that loads previously recorded samples. See the live version on Figure 12.

3.3.2 Arbitrary control of partials

Another patch provided by the author relies on re-synthesis based on [sigmund~] to perform arbitrary control of individual partials. The data from [sigmund~] feeds a bank of oscillators that total up to one hundred, which is a reasonable number of oscillators for this purpose.

The detuning of partials is possible via the manipulation of a detuning table ($0-Detune in Figure 10), which can also be controlled by sliders and MIDI data. Each point in the array correspond to a partial number in ascending order. A detuning generator (below the $0-Detune table in Figure 10) also performs a compression or expansion of the partials, in a similar fashion than the one possible via complex modulation.

As you can perform any arbitrary manipulation, any kind of deviation function can be applied. One easy, for example, is to send 'sinesum' or 'cosinesum' commands to the detuning table in Pd. But the most interesting theoretical application is what Sethares calls Spectral Mapping [1].

3.3.3 Spectral Mapping

This technique allows us to change the relationship between partials to match a particular tuning. Like a Roughness Curve gives us musical intervals (or a 'scale' per se) that matches a spectrum. You can say that Spectral Mapping aims for the opposite, and that is to get a spectrum that matches a given tuning/scale.

For example, if you have a harmonic spectrum and detune the second partial (which is an octave above the fundamental) 50 cents

6 <http://crca.ucsd.edu/~msp/techniques.htm>.

7 First available as examples of a Computer music course by the author based on Pd examples, as presented in the last Pd Convention [28].

upwards, the alignment of partials will match in that same interval 50 cents over an octave instead of the octave, which will actually sound much rougher.

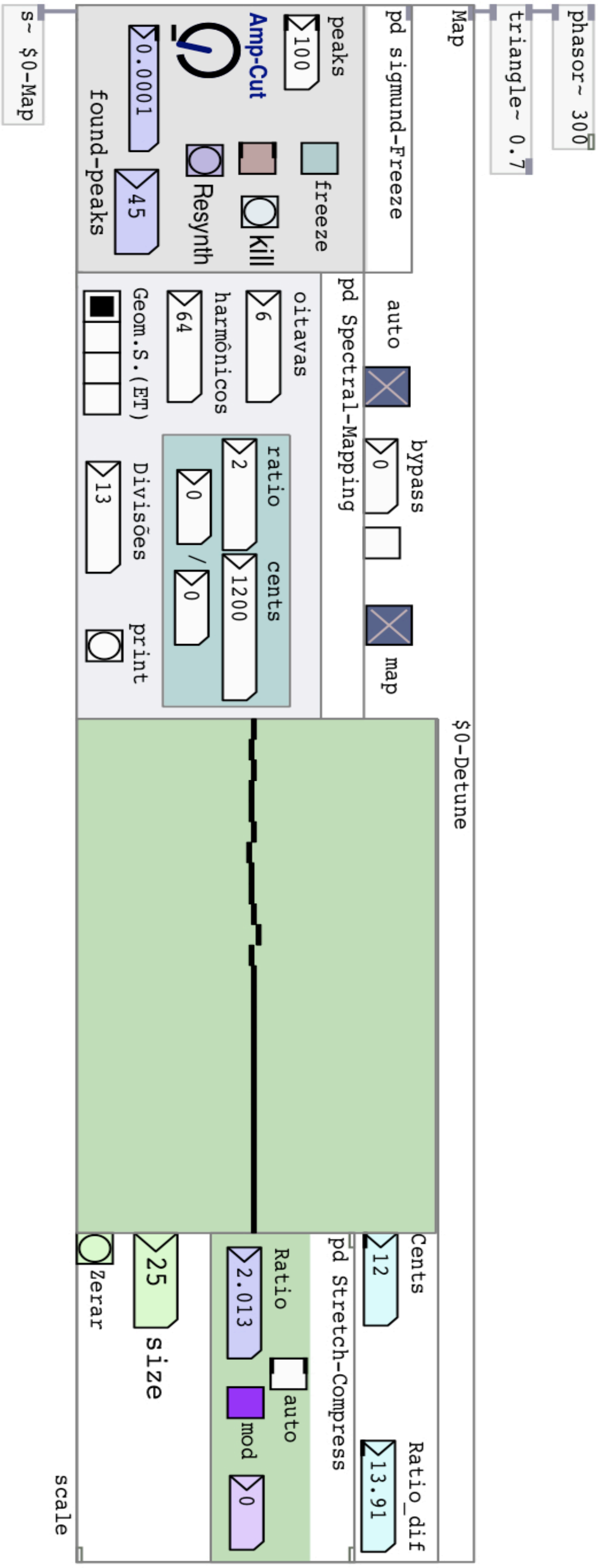


Figure 10: Arbitrary Control of partials via re-synthesis and the Spectral Mapping Module that generates a new series of partials according to a scale.

Tuning systems based on Equal Temperament can never perfectly match harmonic spectra. Our equal temperament of twelve tones enable a “tolerable” mismatch, but if you divide by eleven or thirteen, you'll find trouble. Just Intonation provides the best fit as they're exactly harmonic intervals, and not approximations like Equal Temperaments. A spectrum that matches a “weird” scale is then needed.

For that, the author provides a Spectral Mapping module as an abstraction in Pd, and it generates a new detuned series of partials according to a scale, which is sent to the a table ($0-Detune) that retunes the partials from the original input into the new series.

Among the possibilities, you can divide the octave in any equal number of steps, but also any other interval given in ratios or cents. This allows non octave tunings such as a twelfth [3:1] divided into 13 equal steps, which is a famous Bohlen-Pierce tempered scale[8].

Other unequal divisions are possible such as harmonic and arithmetic divisions or both. Although conceptually they generate Just Intonation intervals, these scales can also be used to mistune the partials in a harmonic series. The right outlet sends the generated scale, which can go then into the Autotuner module, or into the Phase Vocoder, which also has a built-in Autotuner (see Figure 12).

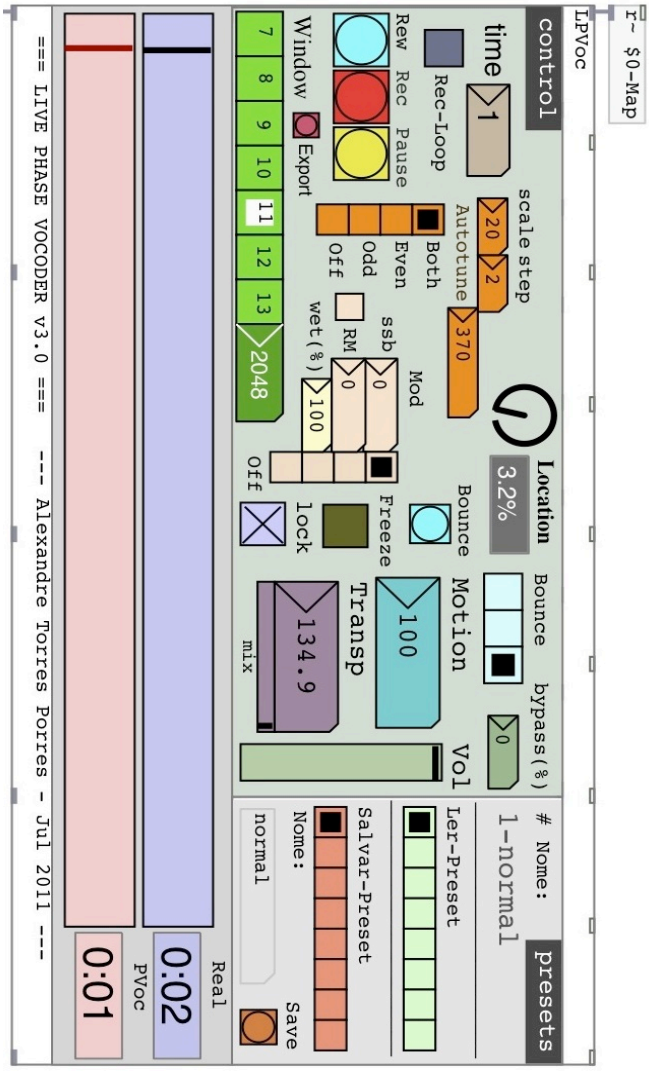
8 <http://www.huygens-fokker.org/bpsite/>

Figure 11: Roughness Curve of the output from the Spectral Mapping module of Figure 10.

On Figure 10, the live input from a [triangle~] oscillator (which is the same used to generate the curve from figure 8) is analyzed by [sigmund~] and re-synthesized via an oscillator bank. The Spectral Mapping module detunes this harmonic spectrum according to a division of the octave in thirteen equal steps. Figure 11 shows the Roughness Curve of the resulting audio signal. The derived scale can be also fed into the Autotuner or the Phase Vocoder.

The Phase Vocoder has several performance features, being the main one to generate “canons” of the recorded live session at different speed and/or transpositions. It can also loops and bounce backwards, as well as other convenient features. But it's pertinent to mention that we can use it to generate dyads as a mix of the shifted and the original signal.

As we have the original untransposed signal, we can Pitch track it. This permits a built-in module of complex/ssb modulation that can compress and expand the spectrum as previously discussed. But it also allows a proper autotuner, that keeps a track of the untransposed signal and shifts it to intervals of a given scale.



The Phase Vocoder is very useful to replay a recorded buffer in different tempos and transpositions. But now, with the Spectral Mapping technique, it can also replay in a different tuning system a transformed spectra that fits to that scale. That is what we see on figure 12. It got the scale from the Roughness Curve, and is retuning the pitch to the closest scale step. The relative fundamental of the scale needs to be specified and it's C if not.

The input is transformed version of the [triangle~] oscillator at 300Hz. So it retunes it to the frequency corresponding to the nearest interval from the scale given by the Roughness Curve. 300Hz is approximately D one sixth of a tone sharper than C (about 237 cents).

The Roughness Curve, as you can see in Figure 11, found an interval at 370 cents, which is somewhat like a “major third” in this system, and the exact interval is in fact about 369.2 cents – so we can see that it's actually working as expected by the theory. The built-in autotune module is then shifting this input (of around 270) almost 135 cents up to match the 370 cents step.

Another way to use the Phase Vocoder is as a dyad generator, by setting the “mix slider” below the “Transp number box” in halfway. Or just a plain pitch shifter. So instead of retuning the recorded data to the closest scale step, you can generate dyads with the original untransposed signal or shift it any way you want.

As a different similar possibility, you can have data coming from your MIDI keyboard into the autotuner module from figure 9. For the examples here exposed, the autotuner got a MIDI input of 700 cents (perfect fifth and shifted it to 740 cents, which can also be seen as a consonance in Figure 11, and is also very close to the actual interval in theory, which is about 738.5 cents.

3.3.3 Other Transformations and synthesis

One other idea to have more of an arbitrary control without the usage of an oscillator bank is to apply the vocoder/convolution technique, which is often used to filter an input to match the spectral imprint of another source.

This is what is behind the many autotune videos on youtube nowadays, where they usually force a melody with a harmonic spectrum over the voice. But you can have other targeted spectra, which may allow a sort of Spectral Mapping.

But this is a link to mention that just any kind of spectral manipulation that you can perform, including ones that you don't even quite understand, might be applied as the curves will tell you what you can do with whatever you got. For example, any random Ring or Amplitude Modulation can generate something interesting and applicable for the creative applications here exposed. And by the way, a Ring and Amplitude Modulation module are also possible in the Phase Vocoder abstraction, and it also corrects the Pitch up or down to sustain the same fundamental.

And lets not forget that all of this applies to synthesis techniques, so again there's Amplitude Modulation, and also the more complex results of Frequency Modulation, Waveshaping, whatever.

4 Final Discussion

Most of this research so far has been mainly concerned with the implementation of the models and the psychoacoustic theory. This is a major problem on itself as there's still a good debate on how each parameter affects the perception of Dissonance.

Not to mention that there isn't a clear straightforward idea of what a complete Dissonance Model is yet. More than that, even regarding singular attributes, there are still ongoing debates on how each one can be improved or more accurate. This paper does not properly address this issue. But the final PhD dissertation will.

Regarding the issues on each attributes. An investigation and further validation of Vassilakis' work is on progress. By applying his formula, the Roughness Curves have a much smaller result for intervals such as the major seventh (the Roughness Curve graphs on this paper are from Barlow's model).

A graph such as proposed by Barlow looks much more like what one would expect a Dissonance Curve to look like. But then, Roughness is not the only attribute of Dissonance. So if Vassilakis is in fact more accurate, it needs to be combined with another attribute such as Tonalness to derive a more intuitive Dissonance Curve.

The Tonalness model by Paul Elrich [21] provides curves that are also much more intuitively like the idea of a Dissonance Curve. It needs to be confronted with the alternate process behind the model provided by Parncutt.

The goal of research is not to put a final rest on the debate of Dissonance Modeling, but generate a state of the art review of this theory. Raising some questions, and some perceptual tests are being considered for that matter. Some discrepancies in the models are still under investigation, and final conclusions will be expressed also in the thesis.

The application of this theory in compositional practice is also incipient. We can see Hindemith's system of Dissonance [29] as one of the first important references, but without any psychoacoustic modeling theory. From the few examples based on psychoacoustic theory and models, the work of Clarence Barlow [14] and Sean Ferguson [26] must be highlighted, but they have worked in a computer aided process and not in real time yet.

One contribution of this study is to make this theory more available to musicians and composers, and also provide it in the form of open source tools implemented in Pure Data.

Even though it is still a research in development, the implementation and some creative applications could already be here exposed. And other creative possibilities shall arise by the pace this research becomes more available to creative musicians. It's certainly a tool with great creative potential.

Sethares et al [30] also provides a spectral toolbox for MAX/MSP. Sethares' theoretical work is of great importance on this research, but his implementations weren't actually taken into account, and the final products differ for that matter. Anyway, although MAX isn't free, the spectral toolbox code is available under the GNU General Public License v2.0.

One main difference to the work of Sethares is that his tools are solely based on his Roughness Model, and the fact that the patches can't be edited as they are this ready made and previously coded interfaces, and the result is more user friendly too, of course.

This paper focused on creative applications for real time live input manipulation, such as from musical instruments. Regarding Spectral Mapping, synthesis techniques can be more stable and easier to tame and apply in practice.

The results for input such as the [triangle~] oscillator are accurate as exposed. The challenge now is to make it more satisfactory for real instruments. As one might expect, a sound source that has a very rich spectrum with may transients can result in a chaotic mess. An idea currently in progress is to detect and segment attacks out of the Spectral Mapping transformation.

5 Acknowledgements

This work is part of a research group funded by the agency Fapesp, proc. n.o 2008/08632-8. Thanks to Sean Ferguson for the supervision of this research during an internship at McGill in 2010, Bruno Giordano for the help designing perceptual tests, and Mathieu Bouchard for the help with the object coding during that time.

References

[[1] W. Sethares, Tuning, Timbre, Spectrum, Scale (2nd ed.), Springer-Verlag, 2004.

[2] E. Terhardt, “The concept of musical consonance: A link between music and psychoacoustics”, Music Perception, 1984, Vol.1, pp. 276-295.

[3] L.A. DeWitt, R.G. Crowder, “Tonal fusion of consonant musical intervals: The oomph in Stumpf”, Perception & Psychophysics, 1987, Vol. 41, No. 1, pp. 73-84.

[4] E. Zwicker and H. Fastl, Psychoacoustics. Springer-Verlag, 1990.

[5] R. Parncutt, Harmony: A Psychoacoustical Approach. Springer-Verlag, 1989.

[6] E. Boring, S. Stevens, “The nature of tonal brightness”, In Proc. National Academy of Science, 22, 1936, pp. 514-521.

[7] A. Monteiro, A Pure Data Library for Audio Feature Extraction <https://sites.google.com/site/pdescriptors/>.

[8] W. Brent, “A Timbre Analysis And Classification Toolkit for Pure Data” <http://williambrent.conflations.com/papers/timbreID.pdf>

[9] J. Bullock, “LibXtract: a lightweight feature extraction library” <http://www.postlude.co.uk/postlude/downloads/LibXtract:_a_lightweight_feature_extraction_library.pdf>

[10] P.N. Vassilakis. “Perceptual and Physical Properties of Amplitude Fluctuation and their Musical Significance.” Ph.D. Thesis, 2001, UCLA.

[11] R. Plomp, W.J.M. Levelt, “Tonal Consonance and Critical Bandwidth”. Journal of the Acoustical Society of América, 1965, N. 38, pp. 548-568.

[12] H. Helmholtz, On the Sensations of Tone as a Psychological basis for the Theory of Music (2nd ed.), Dover Publications, 1954.

[13] R. Parncutt “Parncutt's implementation of Hutchinson & Knopoff (1978)” <http://www.uni-graz.at/richard.parncutt/rough1doc.html>

[14] C. Barlow, “Von der Musiquantenlehre” Feedback Paper 34, 2008. Feedback Studio Verlag.

[15] D.W Robinson, R.S. Dadson, “A re-determination of the equal-loudness relations for pure tones”, British Journal of Applied Physics, 1956 Volume 7, Issue 5, pp. 166-181.

[16] E. Terhardt, “On the Perception of Periodic Sound Fluctuations (Roughness)” Acustica, 1974, 30, pp. 201-213.

[17] D. Pressnitzer, “Perception de rugosité psychoacoustique: d’un attribut élémentaire de l’audition a l’écoute musicale” Ph.D. thesis, 1999, Université Paris 6.

[18] Porres, A.T., Manzolli, J. (2007). “A roughness model in Pd for an adaptive tuning patch controlled by antennas,” *Pure Data Convention 2007*; 6 pages. Montréal, Québec, Canada.

[19] Porres, A.T. and Pires, A.S. (2009). “Um External de Aspereza para Puredata & MAX/MSP,” in *Proceedings of the 12th Brazilian Symposium on Computer Music 2009* (SBCM 2009). Recife, PE: Brazil.

[20] R. Parncutt, H. Strasburger, “Applying psychoacoustics in composition: “Harmonic" progressions of "non-harmonic" sonorities”, *Perspectives of New Music*, 1994, 32 (2), pp. 1-42.

[21] P. Erlich, On Harmonic Entropy, <http://sonic-arts.org/td/erlich/entropy-erlich.htm>.

[22] S. Fingerhuth, Tonalness and consonance of technical sounds, Logos Verlag Berlin GmbH, 2009.

[23] D. Huron. Tone and Voice: A Derivation of the Rules of Voice-Leading from Perceptual Principles. *Music Perception* Fall, 2001, Vol. 19, No. 1, pp. 1–64.

[24] H. Partch, Genesis of a Music. Da Capo, 1974.

[25] W. Brent, “Perceptually Based Pitch Scales in Cepstral Techniques for Percussive Timbre Identification : a comparison of raw cepstrum, mel-frequency cepstrum, and Bark-frequency cepstrum”, In Proc. ICMC 2009.

[26] S. Ferguson, “Concerto for Piano and Orchestra”, Ph.D. Thesis. 2000. McGill.

[27] J.H. McDermott, J.H., Lehr, A.J., Oxenham, A.J. (2010) Individual Differences reveal the basis of consonance. Current Biology, 20, 1035-1041.

[28] Porres, A.T. Teaching Pd & Using it to teach: Yet Another Pd Didactic Material for Computer Music. In Proc. PdCon09.

[29] P. Hindemith, The Craft Of Musical Composition, *Book 1: Theoretical Part*. Associated Music Publishers, Inc., 1945.

[30] W. A. Sethares, A. J. Milne, S. Tiedje, “Spectral Tools for Dynamic Tonality and Audio Morphing” Computer Music Journal, 2009, Volume 33, Number 2, Summer 2009, pp. 71-84.