University of Huddersfield Repository

MacCallum, J. and Einbond, A.

Real-time Analysis of Sensory Dissonance

Original Citation

Real-time Analysis of Sensory Dissonance

John MacCallum and Aaron Einbond

Center for New Music and Audio Technologies (CNMAT)
Department of Music
University of California, Berkeley
{johnmac, einbond}@berkeley.edu

Abstract. We describe a tool for real-time musical analysis based on a measure of roughness, the principal element of sensory dissonance. While most historical musical analysis is based on the notated score, our tool permits analysis of a recorded or live audio signal in its full complexity. We proceed from the work of Richard Parncutt and Ernst Terhardt, extending their algorithms for the psychoacoustic analysis of harmony to be used for the live analysis of spectral data. This allows for the study of a wider variety of timbrally-rich acoustic or electronic sounds which was not possible with previous algorithms. Further, the direct treatment of audio signal facilitates a wide range of analytical applications, from the comparison of multiple recordings of the same musical work to the real-time analysis of a live performance. Our algorithm is programmed in C as an external object for the program Max/MSP.

Taking musical examples by Arnold Schoenberg, Gérard Grisey and Iannis Xenakis, our algorithm yields varying roughness estimates depending on instrumental orchestration or electronic texture, confirming our intuitive understanding that timbre affects sensory dissonance. This is one of the many possibilities this tool presents for analysis and composition of music that is timbrally-dynamic and microtonally-complex.

1 Introduction

1.1 Dissonance and perception

Since the time of Rameau and Helmholtz several psychoacoustic models have been proposed for the perception of dissonance. Ernst Terhardt observed that musical consonance is the product of sensory consonance (absence of sensory dissonance) and harmonicity (the similarity of a sound to the harmonic series) [11, 12]. Sensory dissonance comprises a number of psychoacoustic factors including *roughness*, the beating sensation produced when two frequencies are within a critical bandwidth, which is approximately one third of an octave in the middle range of human hearing [6]. The partials of complex tones, in which several components are fused into a single percept, can also produce roughness when they fall within a critical bandwidth. As a result, the timbre of complex tones can effect our experience of roughness. Richard Parncutt has further extended and

developed Terhardt's theory by proposing a cognition-based measure of roughness of complex tones [7, 8]. Despite the availability of these tools, few current musical theoretical techniques take advantage of them to analyze larger musical structures as they unfold in time.

1.2 Taking timbre into account

Previous models such as those by Richard Parncutt [7, 8] and Kameoka and Kuriyagawa [3], include only a rudimentary treatment of the timbre of complex tones. They model each pitch as a generalized instrumental timbre with the first several partials of the harmonic series in decreasing amplitudes. In the case of Parncutt, these partials are further rounded to the nearest equally-tempered frequency. We can improve on this model, including timbre in a more flexible and faithful way.

Rather than using a prescribed harmonic series to model each pitch of an instrumental chord, we directly analyze audio recordings of the sound in question. We then use Fourier-transform-based analysis tools fiddle$\sim$ [10] and iana$\sim$ [13], running in the computer program Max/MSP, to retrieve frequencies and amplitudes of the partials making up the recording. We further revise Parncutt's algorithm to treat the precise frequencies of the partials available from this data, rather than rounding them to equally-tempered pitches. We do not limit our calculation to a small number of partials; instead we include all relevant partial data available from spectral analysis. By using precise frequencies rather than idealized harmonics, we make it possible to analyze sounds that contain inharmonic spectra, for example bells or electronically-generated sounds.

1.3 Benefits of real-time audio

Previous dissonance measures assume that roughness is additive. However perceptual effects such as masking are not accommodated by an additive model. Dense sonorities, such as clusters, may be perceived in a non-additive way, as they do not present a single clear amplitude modulation, or "beating," frequency. While there are many cultural and contextual factors that may contribute to this "smoothing" of dense sonorities, it would be attractive to include the effect in the roughness estimate itself. By analyzing audio recordings, we make it possible to model this effect. Fiddle$\sim$ automatically mimics masking: when a loud partial is present near a quiet partial, or a large cluster of partials are present in close proximity, fiddle$\sim$ is unable to resolve the separate components, like the human ear. This therefore reduces the contributions of masked frequencies to the roughness estimate. While the precise parameters of the analyses must be fine-tuned to better match perceptual data, the results are already promising.

Another potential benefit of analyzing the audio signal is that different recordings of the same piece may be analyzed separately. This is impossible with previous measures that proceed from the printed page. However it also raises questions of experimental control, for example how to compare two performances recorded with different equipment or in different acoustic spaces.

Audio data allow for analysis to be carried out in real time, as a recording or live electronic sound is played. This suggests application to live electronic music performance, especially improvisation.

2 Implementation

To leverage real-time audio we implement our algorithm as an external object for Max/MSP.

2.1 Roughness Computation

As described in MacCallum *et al.* [4], the roughness algorithm we implement is a modified version of Parncutt [7, 8]. The principal modifications to Parncutt's method are to work with continuous frequency rather than rounding to equal temperament and to analyze sounds directly rather than to model synthetic timbres above a given fundamental. Our object accepts a list of frequency amplitude pairs and returns a single roughness value which is the sum of the roughness of all frequency components

$$\rho = \frac{\sum_{j,k}^{n} a_j \cdot a_k \cdot g(f_{cb})}{\sum_{j}^{n} a_j^2} \tag{1}$$

where a_j and a_k are the amplitudes of the components, and $g(f_{cb})$ is a 'standard curve' developed by Parncutt that models the experimental data of Plomp and Levelt [9]

$$g(f_{cb}) = \left(e(f_{cb}/0.25) \cdot e^{(-f_{cb}/0.25)} \right)^2 . \tag{2}$$

and f_{cb} is the critical bandwidth around the mean frequency of the two components. The literature abounds with formulæ that describe the critical bandwidth; our software implements those of Hutchinson and Knopoff, and Moore and Glasberg [6] (equations 3 and 4 respectively).

$$f_{cb} = 1.72 f_m^{0.65} \tag{3}$$

$$f_{cb} = 0.108 f_m + 24.7 \tag{4}$$

2.2 Peak Extraction

Raw spectral data must be passed through a peak-extraction algorithm to remove noise. For this we use fiddle∼ [10] which has consistently produced results that correspond well to our intuition, although future research plans include the implementation of a peak-extraction algorithm specifically designed for our purposes. Fiddle∼ is particularly useful when analyzing noisy signals such as parts of *Mycenae-Alpha* (see below) and György Ligeti's *Atmosphère*. In the latter example, although the composition begins with a sonority made up entirely of

minor seconds, the sensory experience is far from the extreme roughness that we might expect. Because there are few peaks that fiddle∼ can differentiate from the rest of the spectrum, the resulting roughness calculation is low and corresponds well with our experience of a smooth sound mass.

2.3   User Interface

Critical Bandwidth Formulæ We allow the user to choose between two of the more commonly used critical bandwidth formulæ. The default formula (used to produce the analyses in section 3) is that of Moore and Glasberg which is more recent than that of Hutchinson and Knopoff, used in Parncutt's original work.

Modeling Experimental Data In equation 1 we see that the roughness of each pair of components is the product of their amplitudes weighted by a curve (equation 2) that models Plomp and Levelt's experimental data. Although these data correspond well to our experience and the curve fits the data well, the user can instead define his or her own roughness curve by inputting a list of x-y pairs over which the object will interpolate. The user also has access to the outer-most exponent in equation 2 to make adjustments to the steepness of the curve.

In addition to the modified version of Richard Parncutt's algorithm, our object also implements the algorithm of Kameoka and Kuriyagawa [3]. Although we analyze the following examples using the former method, the user has the option of switching between the two. For a comparison of the two models see Mashinter [5].

Non-Real-Time Analysis Using SDIF data The roughness object can be linked to an SDIF-buffer in Max/MSP for processing an SDIF (Sound Description Interchange Format) file containing sinusoidal tracks (1TRC), harmonic partials (1HRM) or resonance models (1RES) [14]. This allows the spectral analysis and peak extraction to be done with software outside of Max/MSP such as AudioSculpt, AddAn, ResAn, or SPEAR.

3   Practical Examples

Three analytical examples highlight the diverse musical applications of our algorithm: Arnold Schoenberg's *Fünf Orchesterstücke* Op. 16 No. 3, "Farben," Gérard Grisey's *Partiels* for 16 musicians (1975) and Iannis Xenakis' *Mycenae-Alpha* (1978). Each work highlights an analytical problem difficult to resolve using traditional musical-theoretical tools. In "Farben," two different orchestrations of the same chord look similar on paper but produce different spectra when excerpts of the audio recording are compared. While *Partiels* is scored for an ensemble of orchestral instruments, the instruments are called on to play
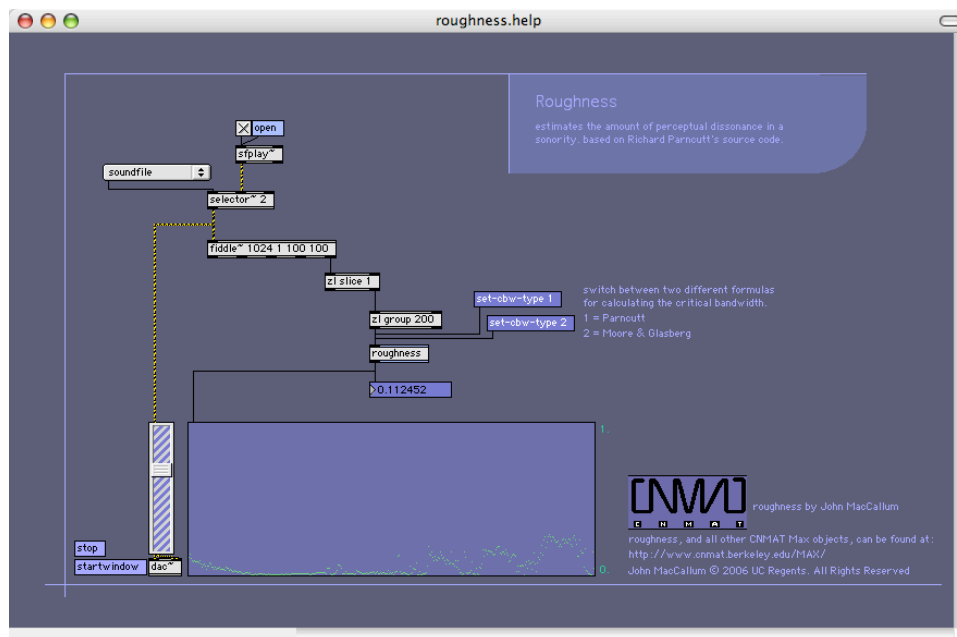
Fig. 1: Max/MSP help file for the roughness object.

microtones and extended techniques which are unstable and may vary from performance to performance. We analyze two recordings to show these differences. *Mycenae-Alpha* is even more challenging analytically: as a purely electronic work it does not exist as a traditional score. Furthermore, most of its sonic material is dense and noisy and it cannot be easily abstracted to discrete notes and rhythms. In this case the audio recording is the necessary starting point for analysis.

3.1 *Farben*

The first example of a musical use of the algorithms is taken from Arnold Schoenbergs *Fünf Orchesterstücke*, Op. 16 No. 3, "Farben." The opening alternates between two orchestrations of the same chord, creating a subtly-shifting klangfarbenmelodie (tone-color melody). The first orchestration is for Flutes, Clarinet, Bassoon, and Viola, and the second is for English Horn, Trumpet, Bassoon, Horn, and Contrabass. The first orchestration yeilds a roughness value of 0.165365, while the second yeilds 0.207284, on a scale of 0-1 (2). This is in contrast to pure sine tones, which give 0.022774. The low value for sine tones reflects the fact that the chord contains intervals mostly greater than a critical bandwidth apart, so roughness comes mostly from the overtones of the written pitches. The fact that the second orchestration is more rough than the first corresponds to our experience of the second chord, containing bright brass instruments, as more dissonant than the first.
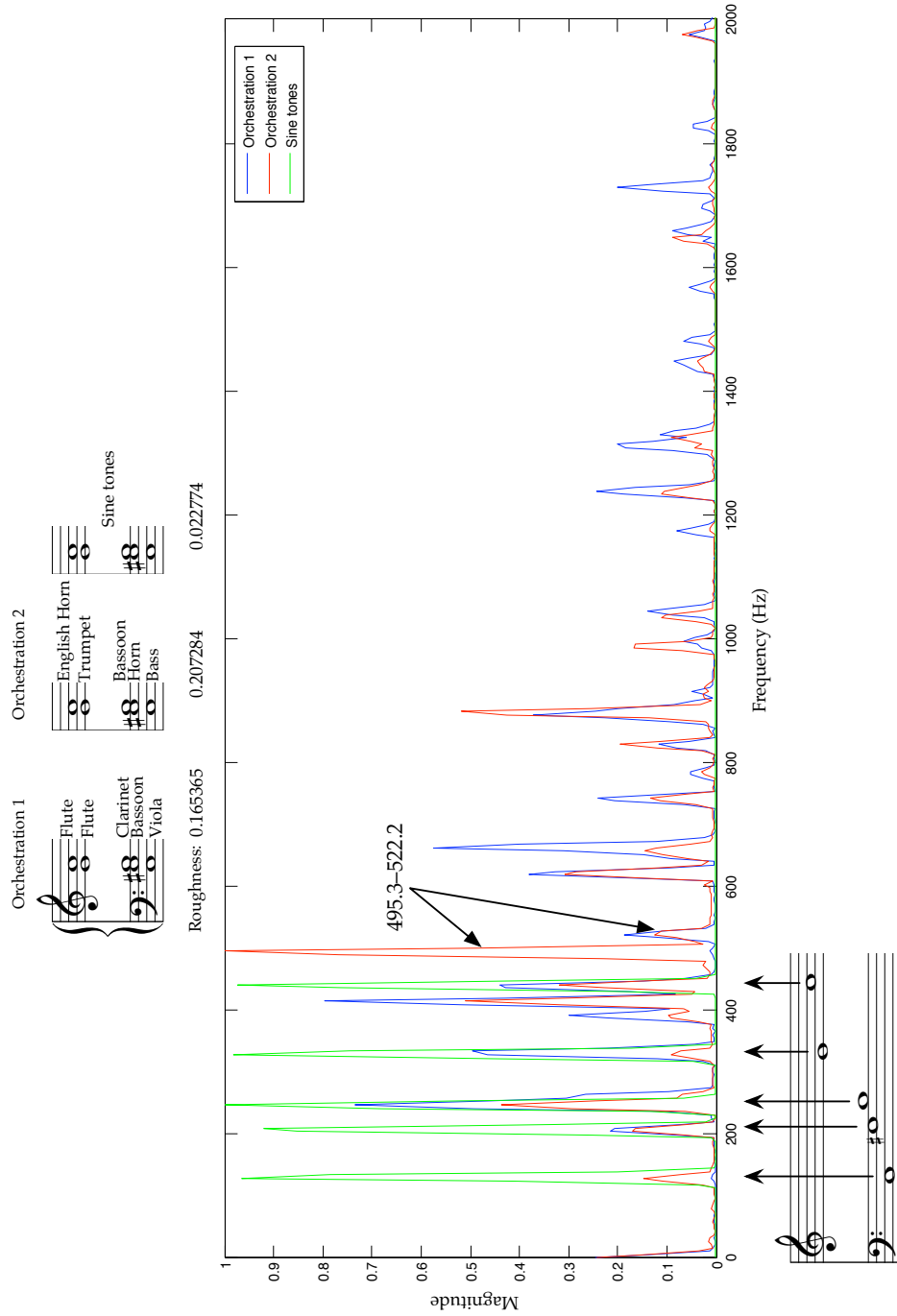
Fig. 2: The "Farben" chord from Schoenberg's Fünf Orchesterstüke, Op. 16 in its original orchestrations (blue and red) and sinusoids (green) showing differences in roughness estimates.

3.2 *Partiels*

Partiels is one of the seminal compostions of the "Spectral" movement and is characterized by its microtonal detail and timbral complexity. Figures 3 and 4 are plots of the roughness value of two different recordings of *Partiels*, as output by our computer patch in real time while the recordings of these pieces are played. Each roughness estimate is averaged over the previous 2.5 seconds of music, which corresponds to approximately 200 individual roughness calculations (one calculation per frame of spectral data from fiddle~).

Figures 3 and 4 plot the first three and a half minutes and four minutes respectively. The timings of the same section of the score differ due to different performance tempi. The periodic attacks of contrabass and trombone are visible as troughs in both of the roughness plots. The peaks in between correspond to tutti sustained chords which gradually evolve from harmonic at the opening (less rough) to inharmonic (more rough) to flickering string sonorities in the middle register (less rough). A second excerpt from the third to the sixth minute (figure 5) follows the opposite path, as dense brass and wind clusters and multiphonics in the low register (more rough) expand out into widely-spaced wind and string chords in the upper register (less rough). The motif of inhalation and exhalation, discussed by Grisey in his own analysis of the piece, is clearly visible in the roughness plots, as is the larger trend of increase and decrease in tension governing this section.

A comparison of figures 3 and 4 reveals the value of working directly with recordings. Given the extended techniques that Grisey requires of his performers and the balance issues inherent in the composition, it is not surprising that the two plots differ substantially while retaining a similar overall shape. The gradual rise in roughness in the Valade performance as compared to Knox could reflect different interpretive decisions in performance as well as different nuances in playing technique. However the roughness plots of the two performances and recordings must be compared with caution. Although we can compare the general shapes of figures 3 and 4 we cannot infer that the peak at 120" in figure 3, for example, is more rough than the peak at 160" in figure 4. Different recording noise levels, instrumental volume, and room acoustics are among the many factors that can account for these differences of scale.

3.3 *Mycenae-Alpha*

Mycenae-Alpha is the first work Xenakis realized with his UPIC graphical synthesis system. Therefore it is conceived outside the bounds of traditional music notation. The dense textures and lack of tempered or clearly-defined pitch, make traditional pitch-based analysis of this piece impossible. In our analysis of the work's first three minutes (see Figure 6), we plot roughness estimates averaged over 2.5 seconds (corresponding to 200 roughness calculations). The roughness plot recovers several prominent features of the work's profile: the sudden drop to a quiet low sound at 60 seconds and the entrance of heterophonic glissandi at 110 seconds.
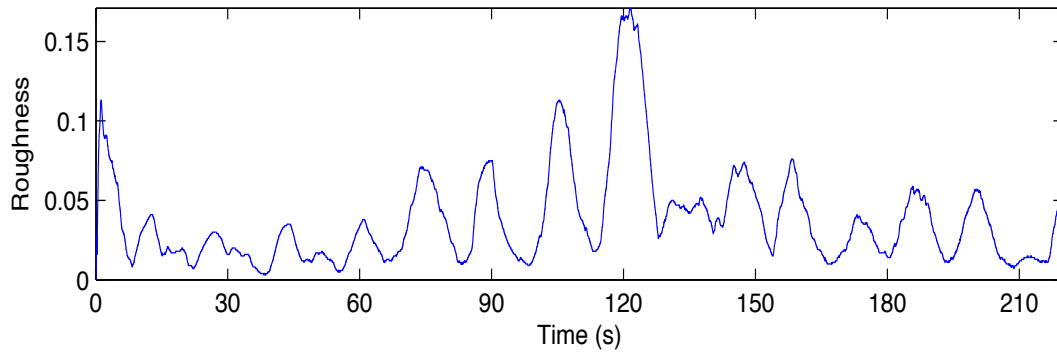
8

Fig. 3: Roughness in Grisey's *Partiels* (0" − 3'40") performed by Pierre-André Valade and Ensemble Court-Circuit.

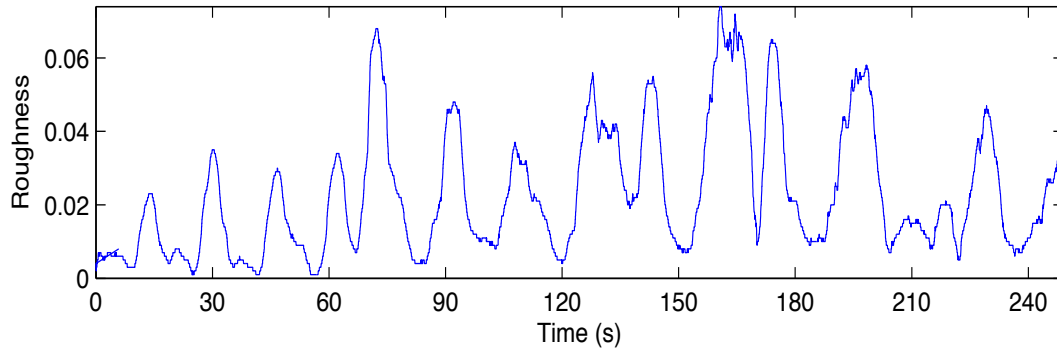

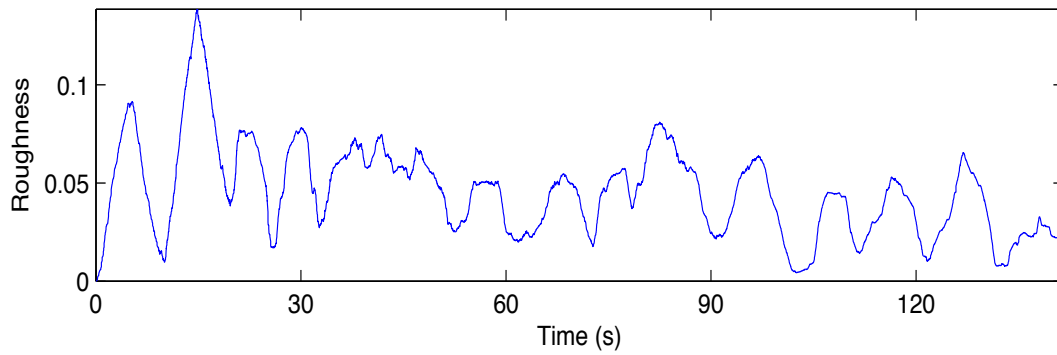Fig. 4: Roughness in Grisey's *Partiels* (0" − 4'10") performed by Garth Knox and the Asko Ensemble.

Fig. 5: Roughness in Grisey's *Partiels* (3'40" − 6'02") performed by Pierre-André Valade and Ensemble Court-Circuit.
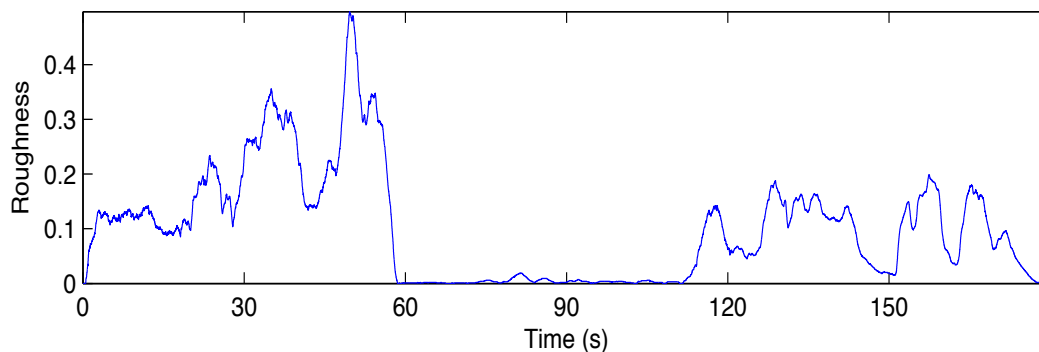
Fig. 6: Roughness plot for Iannis Xenakis' *Mycenae-Alpha* (0" − 3').

4   Conclusions and Future Work

The roughness object described provides a real-time method for the analysis of sensory dissonance. This allows for the analysis of the sonic trace of an audio recording rather than the printed representation and can illuminate the extreme timbral variations found between different performances of many contemporary works. Additionally, it can be used to analyze electronic and improvised music, two genres that often defy traditional analytical tools.

While the algorithms described work well under certain conditions, we plan to improve upon their shortcomings (see Mashinter [5]) through future research to analyze amplitude modulation directly.

References

1. Bregman, A. S. *Auditory Scene Analysis: The Perceptual Organization of sound.* Cambridge, Massachusetts: The MIT Press, 1990.
2. Greenwood, D. D. Critical bandwidth and the frequency coordinates of the basilar membrane. *The Journal of the Acoustical Society of America* 33, 13441356, 1961.
3. Kameoka, A. and M. Kuriyagawa. "Consonance theory, parts I and II," *Journal of the Acoustical Society of America*. 45: 6, 1969.
4. MacCallum, J., J. Hunt, and A. Einbond. "Timbre as a Psychoacoustic Parameter for Harmonic Analysis and Composition," *Proceedings of the International Computer Music Conference*, 2005.
5. Mashinter, K. "Calculating Sensory Dissonance: Some Discrepancies Arising from the Models of Kameoka & Kuriyagawa, and Hutchinson & Knopoff," *Empirical Musicology Review*, 1:2, 2006.
6. Moore, B.C.J. and B.R. Glasberg. "A revision of Zwicker's loudness model," *Acustica* 82, 1996.
7. Parncutt, R. "Harmony: A Psychoacoustical Approach," Berlin: Springer-Verlag, 1989.

10

8. Parncutt, R. and Strasburger, H. "Applying Psychoacoustics in Composition: "Harmonic" Progressions of "Nonharmonic" Sonorities," *Perspectives of New Music* 32:2, 1994.
9. Plomp, R. and W. J. M. Levelt. "Tonal consonance and critical bandwidth," *Journal of the Acoustical Society of America* 38, 1965.
10. Puckette, M. and T. Apel. "Real-time audio analysis tools for Pd and MSP," *Proceedings of the International Computer Music Conference*, 1998.
11. Terhardt, E. On the perception of periodic sound fluctuations (roughness). *Acustica* 30, 20113, 1974.
12. Terhardt, E. "Ein psychoakustisch begründetes Konzept der musikalischen Konsonanz," *Acustica* 36, 1976.
13. Todoroff, T., E. Daubresse, and J. Fineberg "Iana∼ (a real-time environment for analysis and extraction of frequency components of complex orchestral sounds and its application with a musical realization)," *Proceedings of the International Computer Music Conference* 1995.
14. Wright, M., A. Chaudhary, A. Freed, D. Wessel, X. Rodet, D.Virolle, R. Woehrmann, and X. Serra., "New applications of the sound description interchange format," *Proceedings of the International Computer Music Conference*, 1998.