AUTOMATED REPRESENTATIONS OF TEMPORAL ASPECTS OF ELECTROACOUSTIC MUSIC : RECENT EXPERIMENTS USING PERCEPTUAL MODELS

Dr David Hirst
School of Contemporary Music
Faculty of VCA and MCM, University of Melbourne
dhirst@unimelb.edu.au

ABSTRACT

Within this paper we firstly examine the determination of a number of temporal aspects of Electroacoustic Music, and their representations. Then various automated segmentation methods, for Harrison's *Unsound Objects*, are investigated. We find the multi-granular approach outlined by Lartillot *et al.*, combined with the use of MFCCs, is a very efficient and salient segmentation strategy for music structured predominantly according to timbre. Further, the 'Contrast' parameter is both versatile and effective in determining the granularity of segmentation.

INTRODUCTION

Traditional Electroacoustic Music is a studio-based artform involving the mixing of field recordings, processed field recordings, and synthesized sounds. Electroacoustic Music can also include performance of live electronic instruments in the form of laptops or other electronic devices and/or sensors.

This paper concentrates on studio-based Electroacoustic Music. Being largely an aural tradition, there is no widely accepted standard of notation or representation for this kind of music, either in the creation of the music, or the analysis of this kind of music. Our work seeks to explore ways in which signal analysis and/or perceptual models can assist in automating some aspects of the analysis of Electroacoustic Music in order to augment the aural analysis that is the predominant analytical method for this style of music.

Here we set out three recent attempts to automate analytical aspects of Electroacoustic Music associated with the temporal dimension of the music:

1. The representation of a measure of the activity within a section of an Electroacoustic musical piece, and the associated density of musical events.
2. The use of auditory models to derive a 'Rhythmogram Representation' of both short and long sections of music within a work.
3. Segmentation of Electroacoustic Music works, over a longer time-span, using the Music Information Retrieval Toolbox (MIRToolbox).

MEASURING SONIC ACTIVITY

The Problem Defined

While undertaking a recent analysis of Jonty Harrison's electroacoustic musical work, Unsound Objects [1] the initial phase involved analysing the acoustic surface to identify sound objects. The next phase required an examination of relationships between sound objects, giving rise to the following question: What propels the work along from moment to moment, section to section, scene to scene ? To help answer this question, I observed that an increase in sonic activity seems to elicit expectation in the listener that an important event is about to occur. There is a tension build up that seems to require a release the longer the build up goes on. But how can we measure something I have called "sonic activity" and, even better, how can we display sonic activity easily within a work ? Can some form of signal processing be used and be represented to assist in the interpretation of electroacoustic musical works ?

The Analytical Process

With Electroacoustic Music, the first part of an analysis can be described as analysing the acoustic surface. This involves "segmentation". Large scale segmentation into sections, and then small-scale segmentation of sound events from each other. In the analysis of Unsound Objects, the spectrogram and audio waveform displays were useful for the process. Sound events were annotated on the spectrogram and it was possible to get a time-stamped listing of the annotation layer, using the program Sonic Visualiser [2], which was then imported into a spreadsheet program (Microsoft Excel) and printed as a listing of all the annotations. The visual screens and printed time-stamped sound object listings became the data that facilitated detailed identification and specification of sound events within the aurally identified sections of the work.

The next phase of the analysis involved moving beyond the acoustic surface to examine structures, functions and motions between sound events. By “zooming out” to look at longer sections of the work, or carrying out “time-span reduction”, we can observe changing sonic patterns over the course of the work. We can look at the different sections and ask questions like: What propels the work along from moment to moment, section to section, or scene to scene ? To help answer this question, we can observe that an increase in sonic activity seems to elicit expectation in the listener that an important event is about to occur. But how can we measure and, even better, display activity within a work ? Well the Sonic Visualiser program provides access to a suite of plugins of signal analysis. In the Unsound Objects article, I postulated that the type of analysis that seems to correlate best with sound object activity is a plot of “spectral irregularity” versus time.

There are several different methods for calculating the irregularity present within a spectrum, but essentially they both give a measure of the degree of variation of the successive peaks of the spectrum. Jensen, for example, calculates the sum of the square of the difference in amplitude between adjoining partials [3]. What I am postulating here is that where there is a large variation across the spectrum, partial to partial, then this can provide us with a depiction of a high degree of activity. Figure 1 depicts a spectral irregularity plot for the whole of *Unsound Objects*.

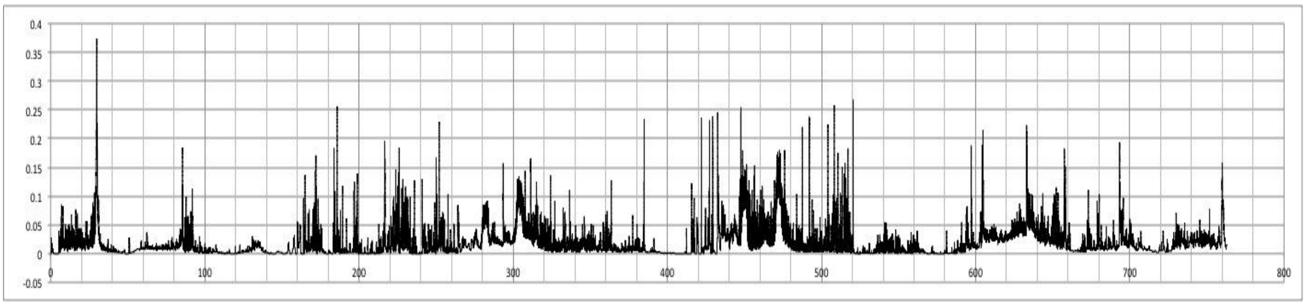

Figure 1 : A spectral irregularity plot for the whole of *Unsound Objects*.

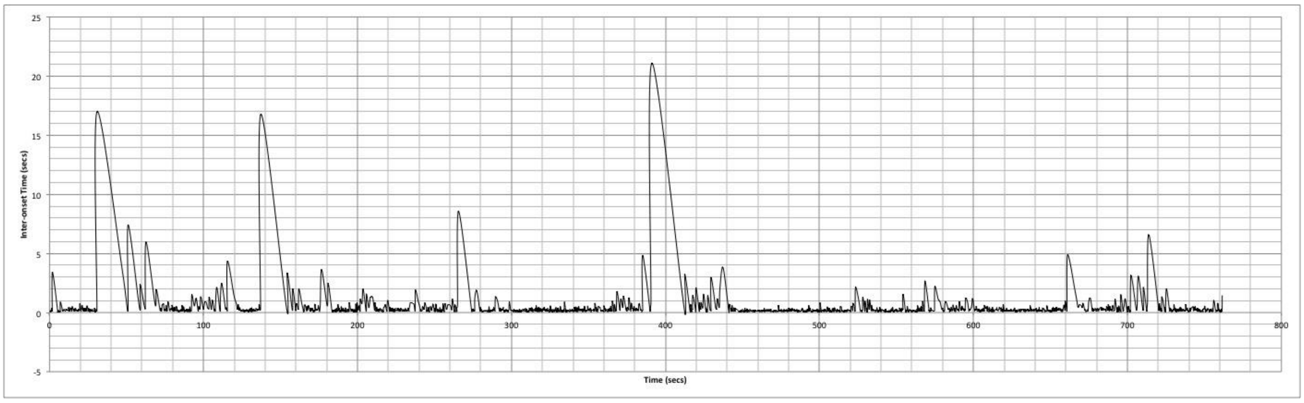

Figure 2 : Plot of Inter-onset Time *vs* Time (secs) for the whole of *Unsound Objects*.

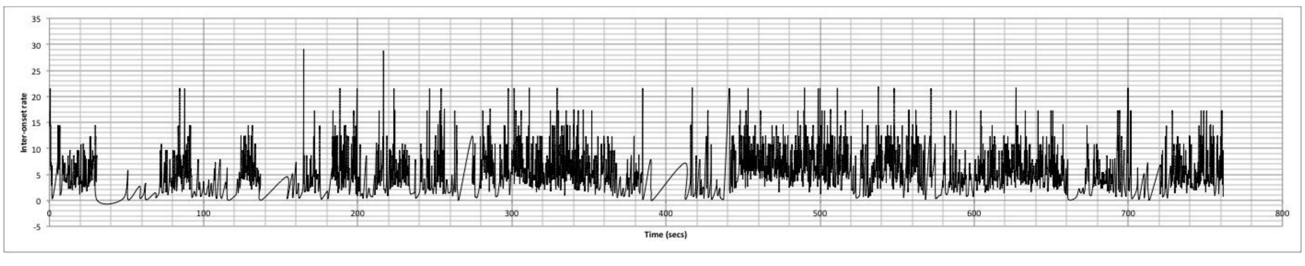

Figure 3. Plot of Inter-onset Rate vs Time (secs) for the whole of *Unsound Objects*.

The analysis of *Unsound Objects* then combined the use of spectral irregularity plots with aurally identified sections, within the work, to provide a detailed analysis of "activity" and to tabulate "sound types" for each section. This table showed "activity amount and type" and "selected sound object types". The work actually divides into two main halves and after the two halves were compared, a summary of sonic archetypes (in the form of mimetic archetypes and structural archetypes), sound transformations, functional relations, and sonic activity were discussed.

Determining Activity

The aim of the next study [4] was to seek an alternative method to the use of "spectral irregularity" for measuring activity in electroacoustic music.

In essence, activity could be defined as the number of sound events in a given time period. Therefore we are interested in the onset time of each sound event, and its duration. Let's start with onset time. What signal analysis tools exist for determining sound event onset time within a musical work ?

The program Sonic Visualiser, has a number of tools within it to perform such an analysis. Aubio onset detection (`aubio.org`) has eight different types which all produce a single list of time "instants" (vertical lines when plotted) of individual start times. This output can be exported to a spreadsheet. Their algorithm can be varied to suit the source material. The Queen Mary, University of London, in-built Sonic Visualiser onset detection algorithm lists three types of onset detector, but these are just the one detector with lots of variables: Program; Onset detection function type; Onset detection sensitivity; Adaptive whitening; Channel options for stereo files; Window size; Window increment; and Window shape. Output is an "onset detection function" which is a probability function of a "note" onset likelihood.

In developing a method for the detection of onsets in *Unsound Objects*, combining several forms of representation was found to provide a more reliable guide to data gathering rather than using any single plot. After some experimentation, the following combination was employed, using the Queen Mary algorithms:

1. RMS Amplitude.
2. Smoothed detection function: Time Values (displays probability function of onsets).
3. Note onsets: Time Instants. Program: Soft Onsets; Onset detection function: Complex Domain; Onset detection sensitivity: 60%; Adaptive whitening: Yes.

This resulted in the onsets (#3 above) aligning pretty well with the smoothed detection probability (#2 above), but with some low level noise swells failing to trigger the onset detector (#3 above).

The "time instants" data (#3 above) was exported, then imported into an Excel spreadsheet in order to be able to make further calculations such as "inter-onset times" (the time between onsets). Figure 2 shows a plot of Inter-onset Time versus Time for the whole of *Unsound Objects.* Its peaks show us where there are long breaks in the work, and give a pointer to how the work may be divided up in analysis.

Displaying time instants, however, only progresses us part of the way to obtaining a measure of event "activity". Inter-onset "rate" was then calculated and plotted, as shown in Figure 3. This provides us with a measure of the number of onsets per second, which, in turn, provides a guide to the amount of event initiation activity at a particular time within the work.

Implications of Activity Plots

Determining inter-onset time can give us a plot (Figure 2) that is useful in showing the main sections within a work. Calculating its reciprocal, inter-onset rate can generate a graph that provides some measure of the varying activity within an electroacoustic work (Figure 3). If we had graphed Figure 3 at the beginning of the analysis, we would have observed that the piece does divide into two, with little activity between about 390 and 410 seconds. The first half begins with three bursts of activity, followed by a longer, more active phase of increasing activity until the "mid-break". The second half is more continuously active until around 660 seconds, where the work has several less active periods, perhaps in preparation for the end of the piece.

In the previous analysis of *Unsound Objects*, sections were first determined aurally, then superimposed over the irregularity plot. Comparing the plot of inter-onset rate (Figure 3) with the irregularity plot (Figure 1) we can see that the piece appears to be much more active in Figure 3 than Figure 1, especially in the second half. The question remains as to which is a better measure of "activity" ? The inter-onset rate is probably a more accurate method, but it seems exaggerated. This is possibly because it doesn't take into account the loudness of the events. Perhaps if this plot (Figure 3) was modified by the RMS amplitude, then a more useful picture of "effective activity" may emerge. There are also inherent definition problems for "iterative" sound events, such as drum rolls or machine sounds. Is such a sound type one long event or many short events ? This phenomenon may skew the events per second data.

In terms of automating analysis, the inter-onset time plot (Figure 2) is very effective in identifying sections in a long musical piece, while the inter-onset rate (Figure 3) does provide a measure of active versus inactive depiction for various passages in a long piece.

The next step in this work was to examine activity and other temporal measures in other works, including more rhythmical pieces.

RHYTHMOGRAM REPRESENTATIONS

This section of our paper introduces work that is well documented in a paper from the ICMC in 2014 [5], but it will be very briefly summarized here to place our subsequent work on automated segmentation into a context of our ongoing work, and to demonstrate some contrasting and varied representations.

Having investigated activity plots, the aim of the next stage of our work was to continue our Segregation, Integration, Assimilation, and Meaning (SIAM) approach of employing a cognitive model [6], in combination with signal processing techniques, to analyse the "raw" audio signal, and more specifically, to depict time-related phenomena (beat, rhythm, accent, meter, phrase, section, motion, stasis, activity, tension, release, etc.). Such depictions should assist or enhance aural analysis of, what is essentially, an aural art-form.

After an extensive literature search, the use of the "rhythmogram" in the analysis of speech rhythm, and the analysis of some tonal music, seemed to fulfill the requirement of a cognition-based method that uses an audio recording as its input signal to produce a plot of the strength of events at certain time points.

The Rhythmogram

In my ICMC 2014 paper [5], I provided a thorough explanation of the rhythmogram, so I will only briefly summarise it here. The framework is documented in Todd [7], Todd & Brown [8] and Marr [9]. It makes use of quite a traditional auditory model where outer and middle ear responses are modelled by filtering, then gammatone filters model the basilar membrane. This is followed by the Meddis [10] inner hair cell model, which outputs the auditory nerve firing probability. It is then summed and processed by a multi-scale Gaussian low-pass filter system. Peaks are detected, summed and plotted on a time constant versus time graph, resulting in a plot known as a rhythmogram.[1]

Figure 4 shows an example rhythmogram for a repeating pattern of three short 50ms tones, followed by a 550ms period of silence, lasting 7 seconds.

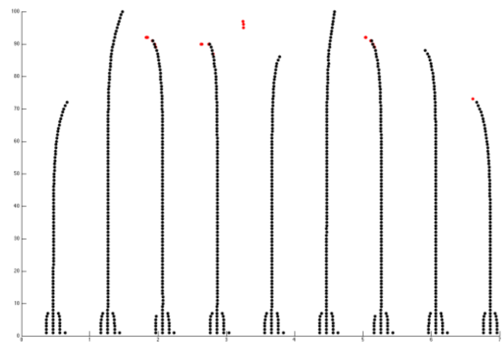

Figure 4. Rhythmogram for a repeating pattern of three short 50ms tones, followed by a 550ms period of silence.

Notable features of the rhythmogram model are:

- Consideration of sensory memory consisting of a short echoic store lasting up to about 200 to 300 ms and a long echoic store lasting for several seconds or more2.
- Each filter channel detects peaks in the response of the short-term memory units.
- The sum of the peaks is accumulated in a simplified model of the long echoic store.
- An "event" activation is associated with the number of memory units that have triggered the peak detector and the height of the memory unit responses.
- The hierarchical tree diagrams of Lerdahl and Jackendoff [12] have visual similarities to rhythmogram plots and so rhythmograms may help the researcher with gaining insights into the hierarchical structure of a musical work under investigation.
- Not only does the rhythmogram model detect the onsets of events, but it can represent other rhythmic grouping structures based on inter-onset times, changes in rhythm, and meter.
- Changing the analysis parameters allows the researcher to "zoom in" or "zoom out", to focus on short-term rhythmic details, or provide a representation of an entire section, or even a complete work.

In the case of the final point above, both of these levels of focus have been explored, and a summarised illustration of both short-term and long-term structures will be recapitulated briefly here.

Analysis of Normandeau's Electroacoustic works

This study utilised the MATLAB code written by Guy Brown, and adapted by Vincent Aubanel for the LISTA

[1] A version of Silcock's schematic [11] for Todd and Brown's model is shown in the Hirst (2014) ICMC paper [5].

[2] Todd (1994), pp. 34-35.

project [13]. The code makes use of the fact that it is possible to increase the efficiency of the computation and still obtain a useful, meaningful rhythmogram plot by using a rectified version of the input signal directly, i.e. bypassing the Gammatone filterbank and inner hair cell stages[3].

The electroacoustic works which were chosen for analysis in this study are collectively known as Robert Normandeau's *Onomatopoeias* Cycle, a cycle of four electroacoustic works dedicated to the voice. The *Onomatopoeias* Cycle consists of four works composed between 1991 and 2009, which share a similar structure of five sections and are of a similar duration of around 15 minutes. The works have been documented by Alexa Woloshyn [14], and by Normandeau himself, in an interview with David Ogborn [15].

Two types of analysis were performed. The first is a detailed rhythmic analysis of a short segment of one of the works. The second analysis zooms out to examine the formal structure of three pieces in the cycle and make comparisons.

Detailed analysis of a short segment of *Spleen*

The work chosen for detailed rhythmic analysis was the second work in the cycle called *Spleen* [16]. This work[4] was chosen as it has a very distinctive beat in various sections and it is slightly unusual for an electroacoustic work in that respect. Figure 5 shows a rhythmogram for the 13.5 second segment of *musique et rythme* from Normandeau's *Spleen*. The X-axis is time (in secs) and the Y-axis is filter number (from 1 to 100). For the full test parameters see [5]. For now we note that the minimum time constant was 10 msec, and the maximum time constant was 500 msec for this test.

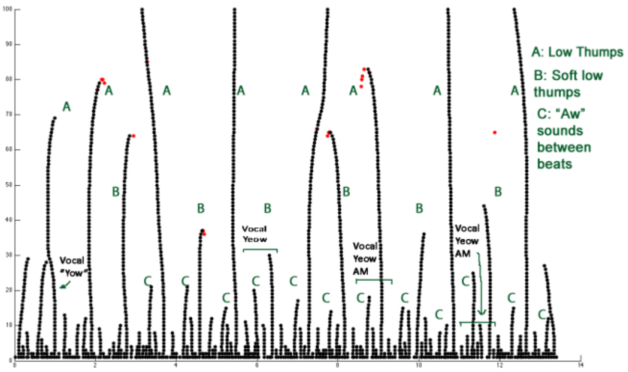

Figure 5. Rhythmogram for 13.5" of *musique et rythme* from *Spleen*.

[3] See Todd (1994) "Appendix A.3.3 Input" p. 65.

[4] The first two mins of *musique et rythme* can be heard via the link on the electrocd site: `http://www.electrocd.com/en/cat/imed_9920/`

Labelled as 'A' in Figure 5, the tallest spikes correspond with a "low thump", somewhat like a bass drum. Using these spikes we could even infer a tempo from their regularity. Labelled as 'B' and "soft low thumps" in figure 5, these softer peaks (B) are interspersed between the louder peaks (A) and are equidistant.

To summarise our observations further we can note that there is a rhythmic background of regular beats, consisting of low thumps, arranged in a hierarchy with softer low thumps interspersed. The "tempo" is around 66 bpm. An implied duple meter results from the loud-soft thump beats alternating.

Against this regular background is a foreground of vocal "yow" shouts. Less regular in their placement, the shouts become elongated to "yeow", and then amplitude modulated to add colour and variety. Although less regular in their placement, the "shouts" always terminate on a "thump" beat and thereby reinforce the regular pulse.

There are finer embellishments too, labelled 'C' in figure 5. This third level of spikes in the rhythmogram depicts events that are placed between thump beats and have a timbre that is somewhere between a saw and a squeaky gate. I'll describe these events as "aw" sounds, and they function as an upbeat to the main thump beat. This "one and two and three and four" pattern has a motoric effect on the passage. The presence of further, shorter, and regular spikes is an indication of more sound events which function to embellish the basic pattern.

Looking at the rhythmogram as a whole, for this passage, we can observe that it tells us there are regular time points in the sound, there is a hierarchy of emphasis in the time points (implying some meter), and a further hierarchy in the sense that there is a background of a regular part (the thumps) and a foreground of less regular vocal shouts. Both the background and the foreground have their own embellishments - anticipation of the events in the case of the former, and an increase in length and use of amplitude modulation, in the case of the latter.

Comparison of whole works from the cycle

The second part of this study involved the use of the rhythmogram in the representation and analysis of whole works. It turns out that the works of Robert Normandeau are ideally suited to this application as well. The *Onomatopoeias Cycle* comprises four works, which consist of the same basic form. Normandeau used the same timeline, but different samples, to create a cycle of works. In 1991 he composed the piece *Éclats de Voix* using samples of children's voices [15]. In 1993 came *Spleen* using the voices of four teenage boys, and in 1995 *Le renard et la rose* used the same timeline with adult voices. The final piece in the cycle is *Palimpseste*, from 2005, and it is dedicated to old age. The first three works were analysed, and rhythmograms were created for them.

As these works are each about 15 minutes long, a different set of analysis parameters was required from the analysis of just a 13.5 second excerpt. After a lot of experimentation, a suitable set of parameters was found. The reader can see [5] for further details, but significantly, the minimum time constant was 0.6 seconds, and the maximum time constant was 30 seconds. These parameters represent a “zoomed out” temporal view of the three pieces.

Figure 6 depicts the rhythmogram (Time *vs* Filter No.) for *Éclats de Voix* for its full duration of around 15 minutes. The alternating grey and white areas mark out the five sections that each piece is divided into - as tabulated by Woloshyn in her paper [14].

There is not the space within the confines of this paper to show the Rhythmograms for all three Normandeau works in the cycle. Neither is there the space to go into our detailed findings, however we can make some indicative comparisons in summary here.

Comparing *Spleen* with *Le renard* we observed similarities between the rhythmic profiles of sections 1, 3, 4 and 5. Comparing the rhythmograms from *Éclats de voix* and *Spleen*, there are some similarities of shape, especially in sections 3, 4 and 5. *Éclats* is more busy than *Spleen*, which is busier than *Le renard et la rose*. Finally, the contrasts become more exaggerated with each piece.

Remarks About Rhythmograms

This initial use of the rhythmogram in the analysis of electroacoustic music has demonstrated that the algorithm is capable of displaying the temporal organization of a short segment with details that may enhance analysis through listening. The algorithm is also flexible, given the careful selection of analysis parameters, in the sense that it can also be used on entire pieces to help elicit information regarding more formal temporal organisational aspects, and to make comparisons with other works.

Some of its short-comings are that it can’t solve the separation problems of polyphonic music, rhythmograms can be awkward to interpret, and they also rely on aural analysis. Careful selection of analysis parameters is crucial in obtaining meaningful plots.

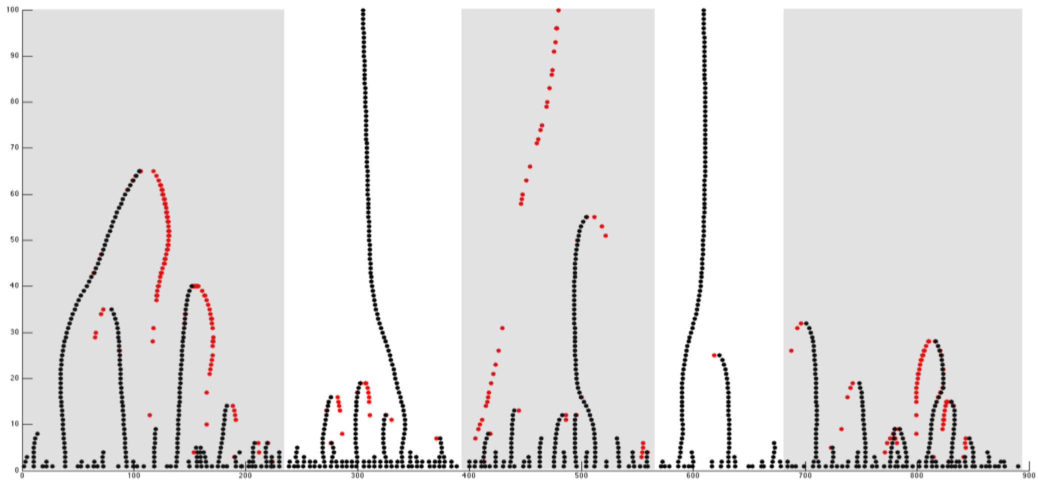

Figure 6. Rhythmogram of the whole of *Éclats de voix* from Normandeau’s *Onomatopoeias* cycle.

AUTOMATED SEGMENTATION OF ELECTROACOUSTIC MUSIC

Following on from the investigation of the rhythmogram, the work on the entire Normandeau pieces brought up the research question of whether the segmentation of entire pieces into their sectional constructs could be automated somehow.

Recalling from section 2.2 above, the analysis of *Unsound Objects* began with **analysing the acoustic surface**. This process involves large-scale segmentation into sections, and then small-scale segmentation of sound events from each other.

To explore such segmentation, signal analysis routines from the MIRToolbox [17] were investigated as they represent a collection of auditory perceptual models on the one hand, and a modular approach in their selection and combination, on the other hand.

Automated segmentation Model

For large-scale segmentation, a method for media segmentation, proposed by Foote and Cooper [18], was used as a model. Their method focuses on the notion of self-similarity. Essentially, the spectrum of every time-segment of an audio work is compared with every other time-segment spectrum, and a “similarity matrix” is created for the whole work. Foote and Cooper [18] describe how the work can be divided into sections from the similarity matrix through the construction of a “novelty curve”: ‘To detect segment boundaries in the audio, a Gaussian-tapered “checkerboard” kernel is correlated along the main diagonal of the similarity matrix. Peaks in the correlation indicate locally novel audio, thus we refer to the correlation as a novelty score’.

Large peaks detected in the resulting time-indexed correlation are then labeled as segment boundaries.

Foote and Cooper go on to describe how they calculate similarity-based clustering to derive the signature of a musical piece, but our work has only proceeded as far as testing the segmentation technique within the electroacoustic musical realm.

Automated Segmentation in Practice Method I

Figures 7 and 8 demonstrate an example of a “novelty curve” and its accompanying segmented audio for the first 3 minutes of Harrison’s *Unsound Objects* [19]. Figure 9 shows the sections derived by a human listener superimposed over the spectral irregularity plot for the same extract of *Unsound Objects*. Figure 9 is included for the sake of comparison between automated methods and a human analyst.

Using this segmentation method, the “kernel size” was manipulated to produce section lengths approximating the manual analysis. With a kernel size of 1250 samples, 7 segments were created in the first 3 minutes.

Comparing figures 8 and 9 we can observe that automated segments 1 and 2 (Figure 8) match Section 1 of the manual analysis pretty well (Figure 9). Similarly automated segments 3 and 4 seem to match Section 4, automated 5 and 6 line up with Section 3, and automated segment 7 matches the manual Section 4. At first glance then, this seems quite a useful method of segmentation. However, in deriving this representation, a convolution computation time of nearly 16 minutes is required for a “kernel size” of 1250 samples in the similarity matrix (quite a large kernel size). Clearly a more efficient method was needed.

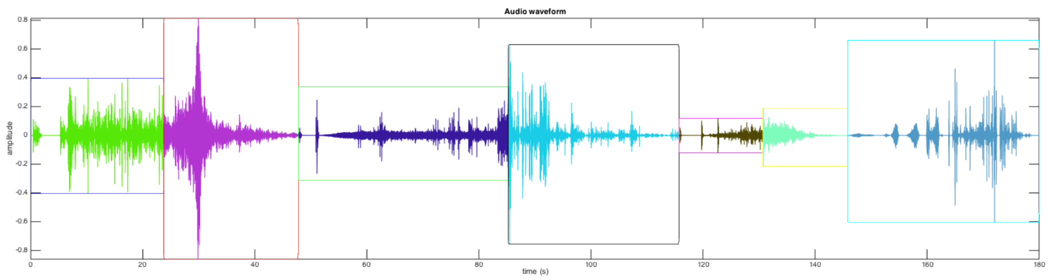

Figure 7. Novelty curve for the first 3 minutes of *Unsound Objects* – Method I.

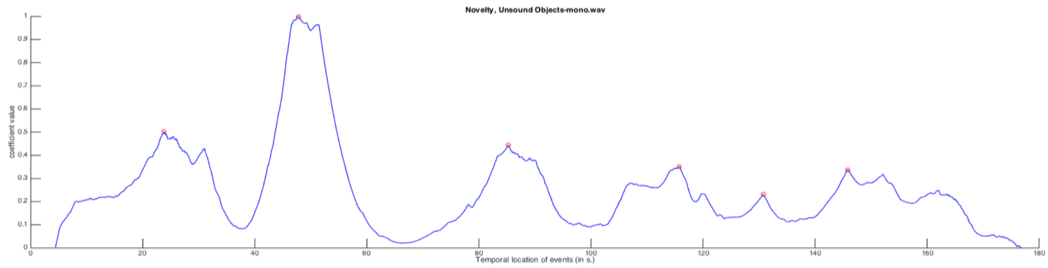

Figure 8. Audio waveform segmented using the novelty curve for the first 3 minutes of *Unsound Objects* – Method I.

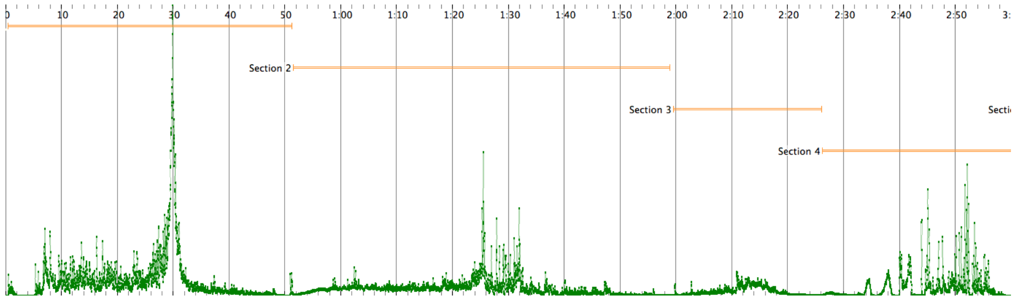

Figure 9 : Irregularity plot with section specification notated by a human listener for the first 3 minutes of *Unsound Objects*.

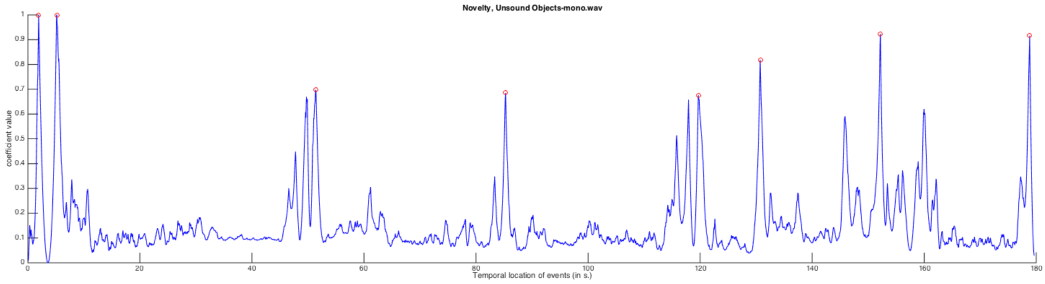

Figure 10 : Novelty curve for the first 3 minutes of *Unsound Objects* – Method II.

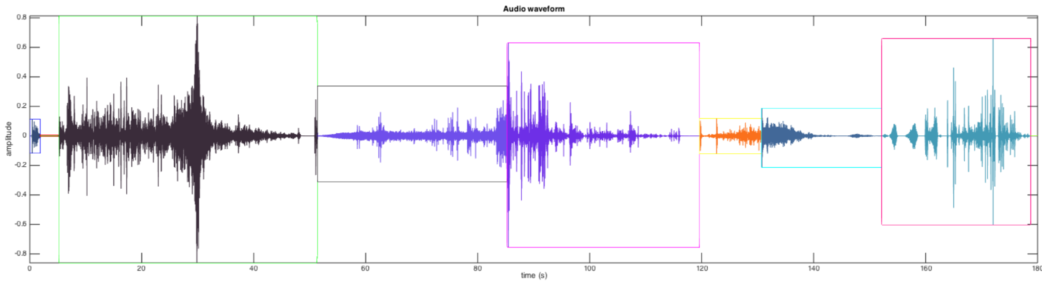

Figure 11 : Audio waveform segmented using the novelty curve for the first 3 minutes of *Unsound Objects* – Method II.

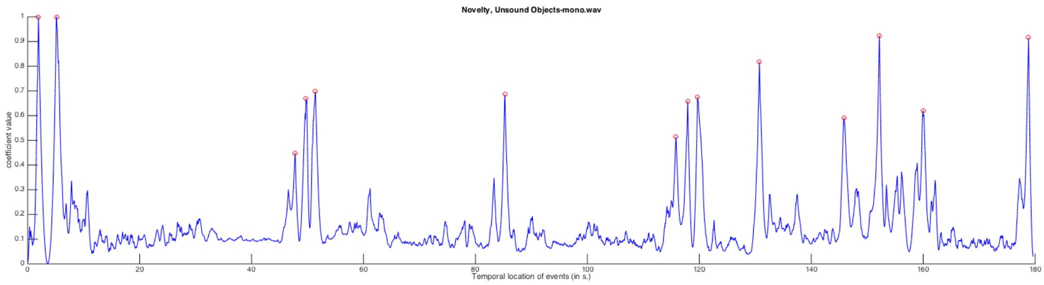

Figure 12 : Novelty curve for the first 3 minutes of *Unsound Objects* – Method II, lower Contrast value.

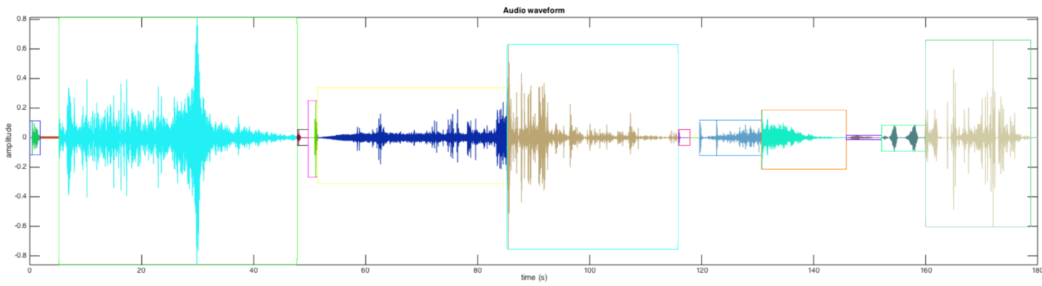

Figure 13. Audio waveform segmented using the Figure 12 novelty curve – Method II, lower Contrast value.

Automated Segmentation in Practice Method II

In Method I, segments are determined from peaks in the *novelty curve*. The *novelty curve* represents the probability along time of the presence of transitions between successive states, indicated by peaks, as well as their relative importance, indicated by the peak heights. For electroacoustic music, we use the spectrum as input to the similarity matrix specification routine. The Kernel based approach is described by Foote and Cooper [18] as follows: 'Novelty is traditionally computed by comparing – through cross-correlation – local configurations along the diagonal of the similarity matrix with an ideal Gaussian checkerboard kernel.' That is, every segment of the piece is compared with every other segment to look for similarities and differences. The sequence of operations is: audio in - spectrum - similarity matrix - novelty - convolution - peaks - segmented audio display - novelty score display.

Method II makes use of the simpler, multi-granular approach outlined by Lartillot, Cereghetti, Eliard & Grandjean [20]: 'For each instant in the piece, novelty is assessed by first determining the temporal scale of the preceding homogeneous part as well as the degree of contrast between that previous part and what just comes next. The idea is to estimate the temporal scale of the previous ending segment as well as the contrastive change before and after the ending of the segment. The novelty value is then represented as a combination of the temporal scale and the amount of contrast'.

Using this multi-granular approach, the following MIRToolbox command yields the novelty curve shown in figure 10 and the segmented audio given in figure 11:

```
mirsegment(a,'Novelty','MFCC','Rank',1:10,'Contrast', 0.6)
```

Note that this method also uses the first ten Mel-Frequency Cepstral Coefficients (MFCCs) in order to decrease computation time, and the 'Contrast' level is set at 0.6. With this 'Contrast' value there are 8 segments identified in figure 11. These segments correlate quite well with the 4 sections shown in Figure 9 in the following way : Section 1 (segments 1-3); Section 2 (segments 4-5); Section 3 (segments 6-7); and Section 4 (segment 8).

It is also possible to vary the 'Contrast' parameter to segment on a shorter-term or longer-term event basis – using the same novelty curve. 'Contrast' is defined as: 'A given local maximum will be considered as a peak if the difference of amplitude with respect to both the previous and successive local minima (when they exist) is higher than the threshold value specified'.

For example, by halving the 'Contrast' value to 0.3 (Fig. 12), six additional peaks in the novelty curve are included, and the audio is segmented into 14 segments (Fig. 13). This provides an effective means to vary segmentation from large sections to individual events, depending on the 'Contrast' value. In our examples, segmentation is on the basis of timbre, however pitch, rhythm and meter could also be used.

In contrast to the 16 minutes required to calculate segmentation using Method I, Method II is at least four times faster and more efficient.

CONCLUSIONS

Within this paper we have examined the determination of a number of temporal-related analytical aspects of Electroacoustic Music, and their representations. We calculated onset times, inter-onset times, and inter-onset rate for Harrison's *Unsound Objects.* We explored the use of the "rhythmogram" as a means of hierarchical representation in the works of Normandeau's *Onomatopoeias* cycle.

Finally we investigated various automated segmentation methods for *Unsound Objects.* We found the multi-granular approach outlined by Lartillot et al, using MFCCs, was a very efficient and salient segmentation strategy for music structured predominantly according to **timbre** (as opposed to pitch or rhythm). Further, the 'Contrast' parameter is effective in determining the granularity of segmentation – short events to long sections.

Acknowledgments

Many thanks to Vincent Aubanel, who generously shared his rhythmogram code, and to Olivier Lartillot and Petri Toiviainen for making their MIRToolbox publically available.

REFERENCES

[1] D. Hirst, "Connecting the Objects in Jonty Harrison's Unsound Objects*." eOREMA Journal, Vol 1, April, 2013*.

http://www.orema.dmu.ac.uk/?q=eorema_journal

[2] C. Cannam, C. Landone, and M. Sandler, "Sonic Visualiser: An Open Source Application for Viewing, Analysing, and Annotating Music Audio Files." In *Proceedings of the ACM Multimedia 2010 International Conference*, pp. 1467-1468.

[3] K. Jensen, *Timbre Models of Musical Sounds*, Ph.D. dissertation, University of Copenhagen, Rapport Nr. 99/7, 1999.

[4] D. Hirst, "Determining Sonic Activity In Electroacoustic Music." In *Harmony Proceedings of the Australasian Computer Music Conference 2014.*

Hosted by The Faculty of the Victorian College of the Arts (VCA) and the Melbourne Conservatorium of Music (MCM). 9 – 13 July 2014. pp 57-60.

[5] D. Hirst, “The Use of Rhythmograms in the Analysis of Electro-acoustic Music, with Application to Normandeau’s Onomatopoeias Cycle.” *Proceedings of the International Computer Music Conference 2014*. Athens, Greece, 14-20 Sept, 2014. pp 248-253.

[6] D. Hirst, *A Cognitive Framework for the Analysis of Acousmatic Music: Analysing Wind Chimes by Denis Smalley* VDM Verlag Dr. Muller Aktiengesellschaft & Co. KG. Saarbrücken, 2008.

[7] N. Todd, The auditory “Primal Sketch”: A multiscale model of rhythmic grouping, *Journal of New Music Research*, 23: 1, 25-70. 1994.

[8] N. Todd, and G. Brown, Visualization of Rhythm, Time and Metre. *Artificial Intelligence Review* 10: 253-273. 1996.

[9] D. Marr, *Vision.* Freeman. New York. 1982.

[10] R. Meddis, Simulation of Auditory-Neural Transduction: Further Studies. *J. Acoust. Soc. Am*. 83(3): 1056-1063. 1988.

[11] A. Silcock, *Real-Time ‘Rhythmogram’ Display*. Report submitted in partial fulfillment of the requirement for the degree of Master of Computing with Honours in Computer Science, Dept. of Computer Science, University of Sheffield. 2012.

[12] F. Lerdahl, and R. Jackendoff, *A Generative Theory of Tonal Music*. Cambridge, Mass. MIT Press. 1983.

[13] G. Brown, and V. Aubanel, Rhythmogram MATLAB code written and adapted for the Listening Talker, LISTA, Project (See `http://listening-talker.org/`)

[14] A. Woloshyn, Wallace Berry’s Structural Processes and Electroacoustic Music: A Case study analysis of Robert Normandeau’s “*Onomatopoeias*” cycle. *eContact!* 13(3) 2010.

`http://cec.sonus.ca/econtact/13_3/woloshyn_onomatopoeias.html`

[15] D. Ogborn, Interview with Robert Normandeau. *eContact!* 11(2) 2009.

`http://cec.sonus.ca/econtact/11_2/normandeauro_ogborn.html`

[16] R. Normandeau, *Spleen*. On music CD *Tangram*. Empreintes DIGITALes, Montréal (Québec), 1994, IMED-9419/20-CD.

[17] O. Lartillot and P. Toiviainen, “A MATLAB Toolbox For Musical Feature Extraction From Audio” in *Proc. of the 10th Int. Conference on Digital Audio Effects (DAFx-07)*, Bordeaux, France, September 10-15, 2007. Pp DAFX 1-8.

[18] J. Foote and M. Cooper, "Media Segmentation using Self-Similarity Decomposition," in *Proc. SPIE Storage and Retrieval for Multimedia Databases*, Vol. 5021, pp. 167–75, January 2003, San Jose, California.

[19] J. Harrison, *Unsound Objects*. On *Articles indéfinis.* IMED 9627. emprintes DIGITALes, Audio CD, 1996.

[20] O. Lartillot, D. Cereghetti, K. Eliard, and D. Grandjean, “A simple, high-yield method for assess- ing structural novelty”, *International Conference on Music and Emotion*, Jyväskylä, 2013.