“VIRTUAL TETTIX” : CICADAS’ SOUND ANALYSIS AND MODELING AT PLATO’S ACADEMY

Anastasia Georgaki
Music Department
University of Athens, Greece
`georgaki@music.uoa.gr`

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

ABSTRACT

This paper deals with the acoustic analysis of timbral and rhythmic patterns of the Cicada Orni sound activity, collected at the Plato Academy archaeological site during the summer period of 2014, comprising the Tettix soundscape database.

The main purpose here is to use sound analysis for understanding the basic patterns of cicada calls and shrilling sounds, and subsequently use the raw material provided by the Tettix database in a statistical modeling framework for creating virtual sounds of cicadas, allowing the control of synthesis parameters spanning micro, meso and macro temporal levels.

1. INTRODUCTION

The Plato Academy soundscape database has been collected in the context of the TETTIX project and consists of several sound recordings of cicadas, singing both individually and collectively, which were taken at the Gymnasium portion of the archaeological site of Plato’s Academy (Athens) during the period of July through September 2014. Important philosophical and poetic references deal with the diachronic aesthetic value of this particular insect in Ancient Greek mythology, as also in Classical, Byzantine and Modern Greek literature (including Hesiod, Anacreon [1], Aesop, Homer, Plato, Aristotle, Thucydides, Ritsos, Elytis, Seferis); this discussion is part of an interdisciplinary research domain that goes under the label of tettigology [1], where entomologists, sociologists, cultural anthropologists and composers try to investigate the strange sonic events and chorusing of cicadas in different geographical areas during summer [2, 3].

This research, which deals with the rhythmic and timbral analysis of the Cicada Orni, has been inspired by the stochastic model of cicadas proposed by Iannis Xenakis. More precisely, Iannis Xenakis has continuously referred to the sonic texture and stochastic effect of the cicadas in order to describe the granular clouds of sounds he experienced during the war [4] [2]. Within this context, Xenakis underlines the fact that the statistical characterization of certain sonic events (e.g., a demonstration crowd or shouting guns) can be very similar when separated from their political or moral context, and when applied to the cicada chorus, this fact indicates the passage from total order to total disorder [6].

Many other composers have also been inspired by the cicada soundscape in different ways: Bartók evokes the cicada in his piano suite of 1926 (*Szabadban* or Out of Doors), Ligeti evokes the stochastic behavior of cicadas in his *Poème Symphonique* (1962) for 100 metronomes, Luc Ferrari and J. C. Risset use elaborated cicada sounds in their electroacoustic music works, and David Rothenberg tries to interact with the cicada choruses by playing music in nature [5].

Previous work is related typically to either isolated analysis or synthesis of cicada singing, and include chaos-theoretical models for Asian Cicadas [6], physical modeling of the sound-producing mechanism of the cicadas [7], assessing the correlation between sound level activity and size of cicada populations [8] and measuring the mechanical response in the female cicada *tympani* [9]. In counterpoint to these works, we propose here a three-layered hierarchical analysis-synthesis model for cicada sounds based on Gaussian and Markov statistical models.

The TETTIX project [3] is primarily focused on acoustic and musical issues associated to the collected signals, such as spatial distribution, diversity of timbres and rhythms, antiphonal and choral organizations, among others. These issues require musical analysis, statistical analysis and both analytic and synthetic methods of signal processing, which are used to create representation models and manipulation tools that will allow a thorough exploration of the artistic potential of timbres and rhythms associated to this soundscape.

---

[1] Tettigology (from the Ancient Greek word Τέττιξ, or Tettix, for cicada) has its roots in several ancient civilizations, from China to Greece; more specifically in Ancient Greek culture, the cicada has appeared on diverse fields, such as literature, visual arts, folklore, scientific writing, philosophy and religion. The cicada life cycle and its characteristic song has served as inspiration for many philosophers, poets and musicians of Ancient Greece.

[2] Xenakis was the first composer to use stochastic distributions as a tool for creating a music (as he writes in his book Formalized Music) which was similar to the sound of cicadas, or to that of rain falling on a tin roof.

[3] The TETTIX project deals with the cartography of cicadas singing in Greece and with the acoustic analysis of the soundscape in archaeological areas; evolutionary models of the cicada singing in different areas of Athens have been designed; this material has been used in artistic explorations of the cicada choruses and in sound installations.

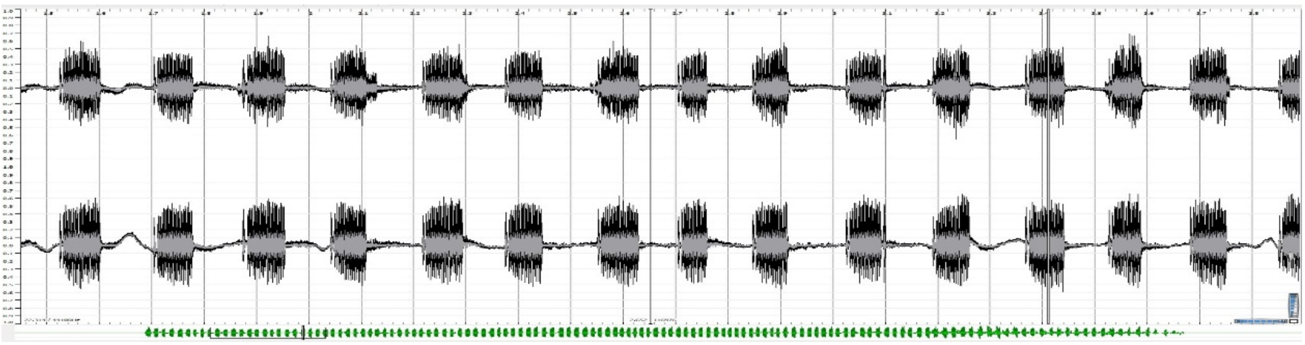

Figure 1. Echemes and interechemes of the Platonic Cicada Orni.

1.1 From the echeme to the Cicada choir

The sound-producing mechanism of the cicadas[4] has been studied since the 18th century, but only after sound recordings and electronic means in the 1950's the exact function of the morphological structures during cicada singing has been established [10–12]. This mechanism consists in a pair of membranes called *Tymbales* which are excited by the action of muscles with the same name, often in alternated motion; the sound thus produced is amplified by air bags positioned directly under the Tymbales. When this mechanism is set into motion the cicada emits its typical sound, which is referred to here as an echeme[5]. Figure 1 shows an example of a series of short echemes.

The sound produced by the cicada has several functions, such as

- a danger alarm for other cicadas,
- a defense mechanism from predators,
- calling and mating with other cicadas, and
- establishing territories,

which explains the emergence of a large number of rhythmic and timbral patterns easily discernible by the human ear [13–16].

The TETTIX project collected several examples of such patterns through sound recordings made in the archaeological site of Plato's Academy, in Athens, during the summer of 2014 (July, August and September). All these recordings were labeled according to diverse conditions that affected the sound material, such as the date and time, temperature, whether it was an individual or collective emission, and also the type of tree where the cicadas were. This material was analyzed according to their temporal macro-structure (density and duration of emissions, and intervals between *echemes*), and to their spectral distribution (spectral centroid, bandwidth and quartiles), to be used as sound material in the context of granular synthesis and swarm models [2].

Future plans for artistic use of this sound material in compositions and installations are coupled with the timbral and rhythmic modeling described in this paper. Timbral modeling is accomplished with parametric sound synthesis models, whereas rhythmic modeling is approached with statistical models, as discussed below.

The proposed modeling of timbre in Plato's Academy soundscape extends on previous studies of this type of signal [7, 11, 14] and consists of three levels of representation. At the lowest level, it provides statistical modeling of spectra in quasi-periodic portions of the signal, aiming at representing observed fragments at a micro-temporal scale (where quasi-periodicity holds), and providing a Gaussian spectral model and an amplitude-modulation model as functions of the signal type (according to date, time, temperature, number of cicadas and tree type).

Rhythmic modeling extends the formerly conducted statistical analysis, establishing a second-order Markov model. Taken together, these models comprise a three-level hierarchical statistical synthesis model. The upper level deals with macro-temporal transitions between different types of emissions, the intermediate level deals with meso-temporal subtler variations within a given emission regime, and the lower level represents the dynamically evolving spectra of the cicada song.

Timbral and rhythmic statistical modeling are viewed as low-level representations for the creation of synthetic instruments here referred to as virtual cicadas. The use of virtual cicadas in a musical context will involve also the consideration of musical issues dealing with rhythm and polyphony, and specifically to phrasing and choral structures that can be observed in the recordings.

Two different categories of sounds can be identified in the emissions of the Cicada Orni species: a continuous component (i.e. a long echeme or emission) and a broken component (separate echemes or emissions). These emissions may be viewed as phrases, which can be either simple and repetitive (monotonous or alternating short and long emissions) or syncopated (i.e. having a more complex rhythmic structure).

Another very important aspect of the songs of cicadas are choral relationships [16]. Concerning the choral organization in most cicadas species two main categories of interaction are to be found: synchrony and alteration. Synchronicity, which was formerly viewed as a cooperative mechanism among males, never seems to be perfect, and

---

[4] Technically only the male cicada has this mechanism, while female cicadas have corresponding sound receivers (tympani).

[5] An echeme is essentially an uninterrupted burst of sound, which can be as short as a tick or a click or it may continue for a longer period. When an echeme is viewed at an expanded time scale the separate pulses created by the buckling of the tymbale ribs can be seen.

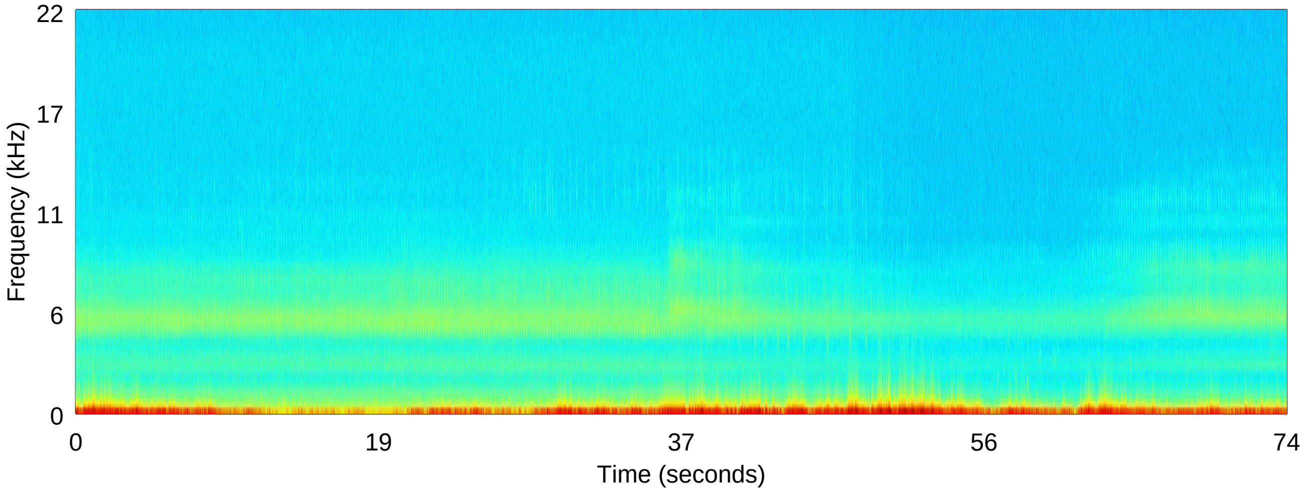

Figure 2. Spectrogram of a sample recording from the TETTIX database made on 24/7/2014. Colors (from blue to red) represent energy on a log scale.

this is related to the psychoacoustic precedence effect [6] . In cicadas, female phonotaxis is influenced by the precedence effect in the sense that the first of two or more closely synchronized calls is preferred. Thus, males are selected to adopt a timing mechanism of signal jamming activities, averting following calls in a synchronizing or alternating fashion [16]. In choral emissions we can also discriminate the domino effect and the effect of the last word [14].

2. TIMBRE AND RHYTHMIC MODELS

Timbre and rhythm in cicada singing are here represented in a three-layer hierarchical model, which considers the dynamic evolution of spectra on a micro-temporal scale (few milliseconds), the perceived fast-evolving fluctuations in amplitude (akin to tremolos) on a meso-temporal scale (on the order of a second), and longer-term evolutions, such as beginnings and ends of echemes and also variations in intensity and number of cicadas, on a macro-temporal scale (many seconds or minutes). As a running example we will use one particular recording, made on July 27th 2014, whose spectrogram is presented in Figure 2. It is sampled at 44100 Hz, and is windowed in frames of 2048 samples or approximately 46 milliseconds.

2.1 First layer: micro-temporal model

One distinguishing aspect of the spectrogram of Figure 2, which is not uncommon in several other recordings, is the presence of strong low-frequency components, corresponding to the red and yellow fringe at the bottom of the figure. These are possibly due to effects of the wind and other environmental sounds during the recording, and hide the real cicada call that requires analysis and modeling, which corresponds more closely to the greenish band around 6 kHz. The presence of these effects has an evident impact on the time-domain representation of the signal, as illustrated in the upper-left graph in Figure 3.

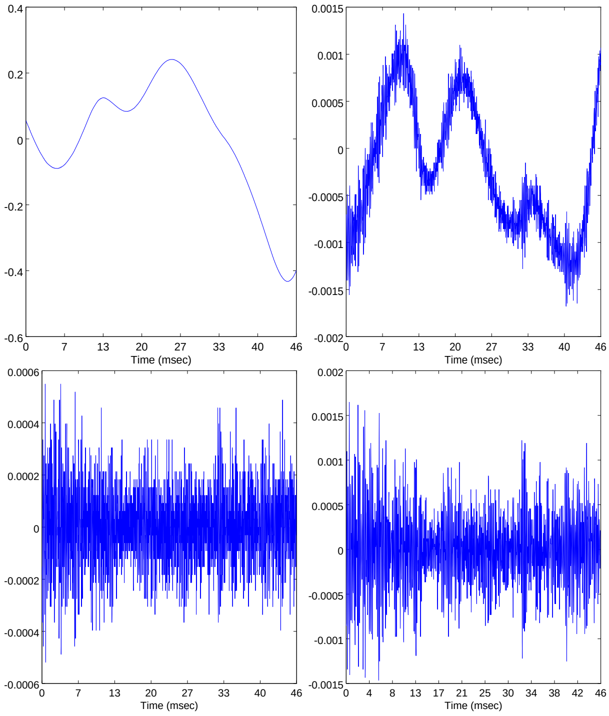

Figure 3. A windowed portion of the signal from Figure 2, 46 ms long (upper-left), a first high-passed filtered version (upper-right), the twice-filtered version (lower-left) and the corrected low-pass filtered version (lower-right) containing predominantly the cicada song.

We see a slowly-varying, apparently smooth profile, which correspond to very low frequencies (up to 60 Hz). The signal of interest lies hidden in much subtler fluctuations around this apparently smooth profile. By filtering once (Figure 3, upper-right) and twice with a high-pass (difference) filter (Figure 3, lower-left) the actual content that requires modeling starts to emerge. As it turns out, a further low-pass filtering step is required to deemphasize higher-

---

[6] The precedence effect describes a phenomenon by which two sounds arriving from different directions are perceived as a single auditory event, whose spatial location is determined mainly by the location of the leading sound.

frequency components that were distorted by the two-zero high-pass filter, producing the last graph in Figure 3 (lower-right).

On the spectral domain (Figure 4), we can see the discrepancy between the lower frequencies in the original signal (Figure 4 upper-left, spectrum starting at 0 kHz) and the higher frequencies (Figure 4 upper-right, same spectrum starting at 1 kHz). The twice low-pass filtered signal has the spectrum shown in Figure 4 lower-left, which illustrates the difference filter effect of distorting the higher-frequency components in comparison with the frequency range of interest (around 6 kHz). This is further compensated by a two-zero low-pass (averaging) filter, which produces the spectrum of Figure 4 (right), closer to the original spectrum in that range (Figure 4 lower-right), and that corresponds on the time-domain to the signal displayed in the lower-right graph of Figure 3.

This result, obtained by the application of simple filters (or equivalently a single four-zero pass-band filter), is preferred over the application of an ideal pass-band filter for several reasons, among them the lack of knowledge of the precise boundaries of the frequency range of interest, and the rippling (Gibbs) effects and spectral leakage that would be introduced by the use of a much sharper filter.

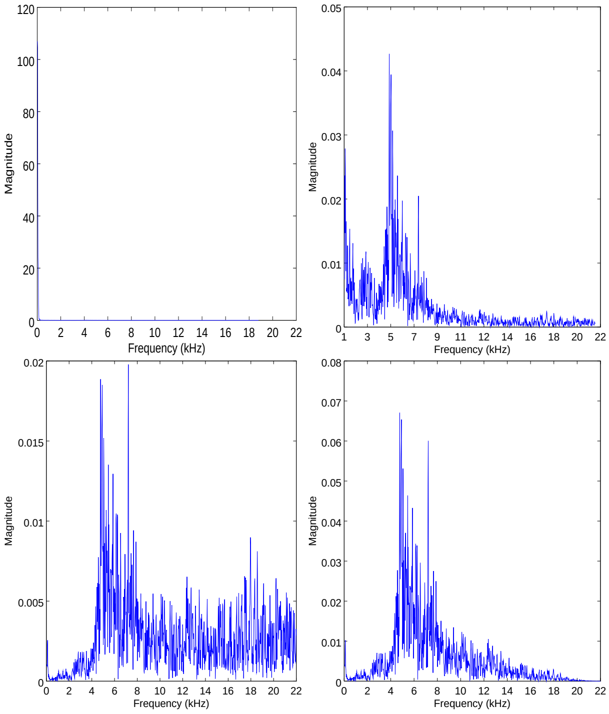

Figure 4. Spectra of the original window (upper-left), same spectrum starting at 1 kHz (upper-right), twice high-pass filtered signal (lower-left) and final low-pass filtered version (lower-right).

The next step in micro-temporal analysis is considering a sliding window along a relatively stable portion of the signal, to capture the dynamic aspects of the spectra observed. In Figure 5 (left) we see the same window which has been treated by the four-zero filter, and the average of the next 1000 consecutive windows, representing the average spectral pattern for 46 seconds of a steady cicada call (Figure 5, center). For each frequency, the variance of the energy around these mean values, as the window slides through these 46 seconds, is represented in the right-most graph of Figure 5. It can be seen that a somewhat similar profile is observed in the variance spectrum, meaning the higher the amplitudes of the spectral components the higher also the variance observed in consecutive windows, although variances are significantly smaller outside the range between 4 kHz and 10 kHz.

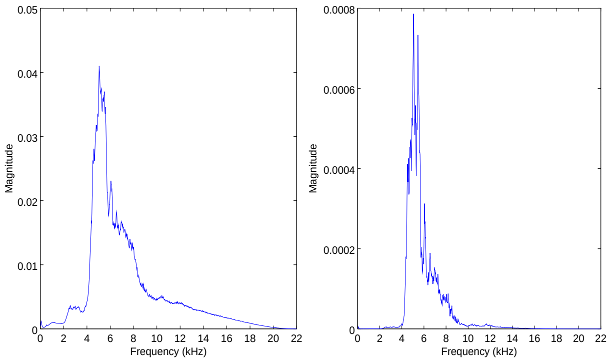

Figure 5. Average spectrum of 1000 consecutive windows (46 seconds) of a steady emission (center) and variance spectrum for the same windows (right).

These average and variance spectra are the input to a synthesis engine that produces virtual cicada timbres that serve as input for the subsequent meso and macro-temporal refining synthesis stages.

2.2 Second layer: meso-temporal model

Moving on to the meso-temporal scale, we want to model the amplitude fluctuations that are perceived as shrilling, an effect somewhat similar to tremolo in the sense that it can be viewed as a form of amplitude envelope applied to a steadier signal, although in the cicada call these are by no means periodic or quasi-periodic fluctuations. As opposed to the micro-temporal modeling, here these relatively smooth variations are the relevant part to be modeled, as an amplitude envelope that is going to be applied to the micro-temporal synthesis engine.

At this scale the superposition of undesired low-frequency effects (such as wind) with the shrilling is an important issue to be addressed. The two-zero high-pass filter applied in the first layer of the model destroys most of the low-frequency amplitude fluctuations we want to model, whereas not applying any high-pass filters will contaminate the meso-temporal model with external factors. A compromise solution is using a softer one-zero high-pass filter before further low-pass filtering. Figure 6 (left) shows an example of this soft high-pass followed by a 48-order low-pass filter applied to a window of size 2^{15} (about 743 milliseconds). Its spectrum, shown in Figure 6 (right), allows the identification of the range of frequencies (up to 50 Hz) predominantly involved in the shrilling effect.

Since this shrilling is typically not steady it wouldn't make much sense to repeat the model of the micro-temporal scale by considering a sliding window of size 2^{15}. Instead, here

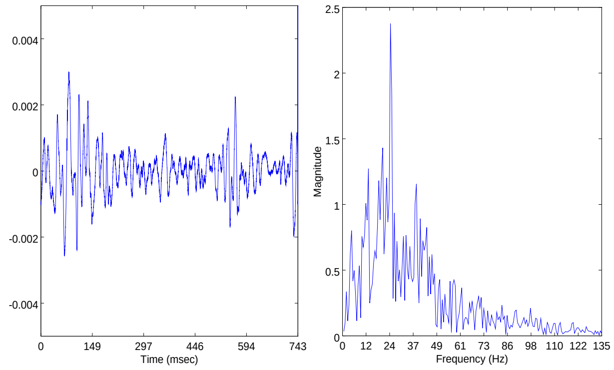

Figure 6. Low-passed amplitude values (left) for the meso-temporal scale representation (743 ms) and corresponding spectrum (right).

we consider a basic sinusoidal model with Gaussian random variables modeling amplitude and frequency values, which are estimated from the TETTIX database for each steady emission type. Essentially this Gaussian model represents frequency and amplitude of the main component[7] of the spectrum of the above processed signal, observed in consecutive windows of 2048 samples or 46 ms. During synthesis these random values are generated once for each 46 ms window, and are linearly interpolated to recreate the smoothness of the amplitude envelope that is applied to the micro-temporal synthesis engine.

2.3 Third layer: macro-temporal model

Moving now to the macro-temporal level, here considered as the large-scale fluctuations in amplitude observed over many seconds, including the occurrence of rhythmic patterns such as stops and resumes in singing (echemes), we have in Figure 7 the RMS envelope (blue line) computed on windows of size 8192 (186 ms) and spanning 2^{21} samples or 47.554 seconds from the initial portion of the recording displayed in Figure 2. This amplitude profile allows us to recognize moments where the emission stops (between 10 and 20 seconds) and also the occurrence of interrupted echemes (especially visible from 30s to 45s).

Considering the kind of profile exhibited by those RMS values over large-scale temporal ranges, of which Figure 7 is a reasonable example, we propose to use a second-order Markov model for updating amplitude information on the synthesis macro-temporal engine.

This model begins with the quantization of amplitude profiles using N linearly spaced amplitude values, which become nodes in the Markov chain, and then generating transition matrices from the observed RMS data in the TETTIX database. For instance, in the above example we might have N=5 different quantized amplitude levels (green dots in Figure 7), each associated with a node in the Markov chain, and second-order transition probabilities reflecting the observed temporal sequence of quantized amplitude levels. Specifically, each triple of adjacent quantized values (a_1, a_2, a_3) observed in the sequence increases a counter $P_{a_1 a_2 a_3}$ in the transition matrix, and finally each line is normalized to satisfy the condition $\sum_{a_3} P_{a_1 a_2 a_3} = 1$.

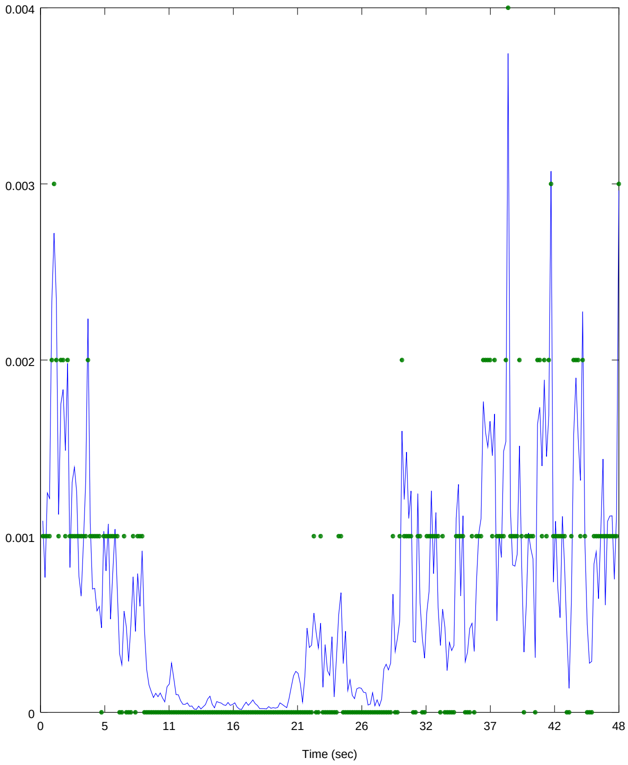

Figure 7. RMS values (blue line) on a larger temporal scale (47.554 s) and quantized values (green dots) used in the Markov model.

In order not to produce over-fitted overly-sparse transition matrices (N too large), and also not overly-dense matrices with poor fitness with respect to the data (N too small), it is important to choose intermediate values for N, which of course depend on the length of the observed RMS data available. Considering that the transition matrix is of size N^3 and that the RMS sequence is of size M, having a 50% occupation of the matrix would require a maximum value of N of the order of $\sqrt[3]{2M}$ (assuming all transitions are different); using half of this estimate, i.e. $N = \frac{\sqrt[3]{2M}}{2}$, makes the observed transitions relatively 8 times higher, with a more evenly-distributed transition matrix for the Markov model.

This second-order Markov model is used in the third-level of the synthesis hierarchy in order to recreate both rhythmic patterns (when the Markov chain jumps between low-level and high-level values of amplitude) and choral aspects such as the variation in the number of insects leading to sudden amplitude changes (assuming there is not also a sudden variation in timbre, in which case the model assumes this is a different emission type and redefines the synthesis engine on all three levels of the hierarchy).

---

[7] In Figure 6 (right) we can identify this main component around 25 Hz.

3. FROM MICROSTRUCTURE TO MACROSTRUCTURE: CREATING HETEROTOPIAN VIRTUAL CICADA SPACES

The rhythmic and timbral analysis and synthesis model of the cicada orni call song presented in the previous section serves as a starting point for discussion on the use of this sound synthesis model in artistic applications. For the moment, we will focus on the use of this cicada call through signal processing techniques in order to reconstruct the Plato's Academy soundscape, in an allegorical way, based on the notion of heterotopia[8] in time or in space.

One interpretation of the term heterotopia corresponds to a real place where several other spaces overlap. In this way, Plato's Academy may be transformed into a hybrid real/virtual environment with real Greek cicadas and virtual cicadas from around the world. Another interpretation considers heterotopia in time, where space is treated like a sound museum and sonic events from different moments are used to create heterochronistic metaphors (for instance the suggestion of summer warmth brought by the virtual cicada call during winter). This kind of illusory space is easily recreated by mixing virtual cicada sounds of different types and virtual cicadas from different places using the Pd environment. The three-level hierarchical virtual cicada model previously presented serves as basic building blocks for the construction of a heterotopical soundscape, both in the temporal and spatial senses discussed.

Time and space also serve as metaphors for compositional aspects of heterotopical soundscapes. In the temporal domain different rhythmic patterns can be elaborated by using deterministic and stochastic models to control the parameters of the virtual cicadas, particularly aiming at creating choral effects. In order to make the perception of cicada choirs easier the domino effect and the precedence effect can be explored, as also synchronization and alternation of virtual cicada voices in a polyphonic texture.

In the spatial domain some relevant aspects are the use of different phases, counterpoints and antiphonies between different sound source positions corresponding to a spatially distributed choir of cicadas. Extrapolating this time and space metaphor and extending the spatial aspect to the frequency domain, virtual cicada choirs can be made to produce vowel-like alterations of the original pattern by coloring chosen central frequencies with chosen bandwidths (FOF synthesis).

This virtual sonic environment based on the heterotopian cicada song call can also be treated as a metaphor for the duality and contradictions of the cicada's autochtony[9].

4. CONCLUSIONS

In this paper we have discussed the process of sound analysis and statistical modeling of the Cicada Orni sounds collected by the TETTIX project at Plato's Academy in Athens during the summer of 2014. This sonic database has been analyzed on micro, meso and macro-temporal levels, and on each level a specific synthesis model has been proposed, which combined produce a three-level hierarchical synthesis model for virtual cicadas.

For the micro-temporal level a Gaussian model based on average and variance spectra has been developed, which allows the representation of several different timbres appearing on the recordings, due to changes in temperature, time of the day, emission type and tree type, among others. For the meso-temporal level a Gaussian sinusoidal model has been used to represent the shrilling that occurs and adds complex patterns of amplitude variation on the lowest frequency range (0-50 Hz). On the macro-temporal level a second-order Markov model was proposed for handling both the rhythms appearing due to interrupting/resuming echemes, as well as variations in amplitude due to choral structures observed in the data.

As further research the following topics are planned to be tackled in the near future:

Augmented Aurality: we intend to compose a collaborative soundscape for mobile phones, that will augment the sensorial dimensions of the experience of hearing the cicadas. The participants will be introduced to the processes of soundscape composition, sound design and sound mapping, within the framework of site-specific artistic practice with the use of innovative locative media applications. In this framework mobile phones can be used by people near a place where cicadas sing to acquire and also mix in real time both real and virtual cicadas in an utopian music environment, instrumentalizing a particular form of human-nature interaction.

Evolutionary heterochronistic model: we are planning to develop a computational model for representing cicada sonic behavior evolving over longer periods of time (several months or years). Making such an extended database and the corresponding computational model available online would allow the exploration of real and processed sound of cicadas in the context of an evolutionary model, and also to extract different patterns from this evolution.

Multi-ethnic Heterotopical Soundscape: through modeling and using different species of cicadas that exist in specific geographical sites (for instance mixing the Brazilian Fidicnoides Picea or Quesada Gigas with the Greek Cicada Orni), it would be possible to create multi-ethnic heterotopical soundscapes that would not be observed anywhere in the real world due to ecological and biological constraints of these species.

Acknowledgments

The second author acknowledges the partial support of FAPESP grant 2014/25686-5.

---

[8] The term heterotopia, introduced by philosopher Michel Foucault in a lecture delivered in Paris on March 14th, 1967 [17], has been used to describe spaces that have added layers of meaning and refer to other spaces, or spaces of otherness, which are neither here nor elsewhere.

[9] Cicadas used to be the emblem of Athenian autochtony.

5. REFERENCES

[1] Anacreon and M. L. W. (Ed.), *Carmina Anacreonta*. Leipzig: B. G. Teubner, 1984.

[2] A. Georgaki, “Listening to the cicada chorus in the plato academy: soundscape research,” in *Filigrane. Musique, esthétique, sciences, société*, Musique et écologies du son, Université Paris VIII, 2014.

[3] ——, “Ο χορός των τζιτζικιών της σύγχρονης Αθήνας (the dance of the cicadas in contemporary athens, in greek).” in *Journal Highlights*, Athens, 2003.

[4] I. Xenakis, *Formalized music: thought and mathematics in composition*. Pendragon Press, 1992, no. 6.

[5] D. Rothenberg, *Bug music: how insects gave us rhythm and noise*. Macmillan, 2013.

[6] T. P. Benko and M. Perc, “Deterministic chaos in sounds of asian cicadas,” *Journal of Biological Systems*, vol. 14, no. 04, pp. 555–566, 2006.

[7] T. Smyth and J. O. Smith, “A musical instrument based on a bioacoustic model of a cicada,” *Proceeding of ICMC 2001*, 2001.

[8] I. J. Patterson, G. Massei, and P. Genov, “The density of cicadas cicada orni in mediterranean coastal habitats,” *Italian Journal of Zoology*, vol. 64, no. 2, pp. 141–146, 1997.

[9] J. F. C. Windmill, J. Sueur, and D. Robert, “The next step in cicada audition: measuring pico-mechanics in the cicada’s ear,” *The Journal of experimental biology*, vol. 212, no. 24, pp. 4079–4083, 2009.

[10] M. F. Claridge, “Acoustic signals in the homoptera: behavior, taxonomy, and evolution,” *Annual review of entomology*, vol. 30, no. 1, pp. 297–317, 1985.

[11] D. Young and H. Bennet-Clark, “The role of the tymbal in cicada sound production,” *The Journal of experimental biology*, vol. 198, no. 4, pp. 1001–1020, 1995.

[12] H. C. Bennet-Clark, “How cicadas make their noise,” *Scientific American (USA)*, 1998.

[13] P. C. Simões, M. Boulard, M. T. Rebelo, S. Drosopoulos, and M. F. Claridge, “Differences in the male calling songs of two sibling species of cicada (hemiptera: Cicadoidea) in greece,” *Eur. J. Entomol*, vol. 97, pp. 437–440, 2000.

[14] J. Sueur and T. Aubin, “Acoustic communication in the palaearctic red cicada, tibicina haematodes: chorus organisation, calling-song structure, and signal recognition,” *Canadian journal of zoology*, vol. 80, no. 1, pp. 126–136, 2002.

[15] G. Pinto-Juma, P. C. Simões, S. G. Seabra, and J. A. Quartau, “Calling song structure and geographic variation in cicada orni linnaeus (hemiptera: Cicadidae),” *Zoological Studies*, vol. 44, no. 1, pp. 81–94, 2005.

[16] M. D. Greenfield, “Synchronous and alternating choruses in insects and anurans: common mechanisms and diverse functions,” *American Zoologist*, vol. 34, no. 6, pp. 605–615, 1994.

[17] M. Foucault, “Des espaces autres,” *Empan*, vol. 54, no. 2, pp. 12–19, 2004.