Emergent Timbre and Extended Techniques in Live-Electronic Music: An Analysis of *Desdobramentos do Contínuo* Performed by Audio Descriptors

Danilo Rossetti (Universidade Estadual de Campinas, Campinas, São Paulo, Brasil)
danilo.rossetti@nics.unicamp.br

William Teixeira (Universidade Federal do Mato Grosso do Sul, Campo Grande, Mato Grosso do Sul, Brasil)
william.teixeira@ufms.br

Jônatas Manzolli (Universidade Estadual de Campinas, Campinas, São Paulo, Brasil)
jonatas@nics.unicamp.br

Abstract: In this article, an analysis of the piece *Desdobramentos do contínuo* for violoncello and live-electronics is addressed concerning instrumental extended techniques, electroacoustic tape sounds, real-time processing, and their interaction. This is part of a broad research about the computer-aided musical analysis of electroacoustic music. The objective of the analysis of this piece is to understand the spectral activity of the emergent sound structures, in terms of which events produce huge timbre variations, and to identify timbre subtle nuances that are not perceptible on a first listen of the work. We conclude comparing the analyses results to the compositional hypotheses presented in the initial sections.

Keywords: Computer-aided musical analysis, Live-electronic music, Extended-techniques, Audio-descriptors, Emergent timbre.

Timbre emergente e técnicas estendidas na música eletroacústica mista em tempo real: uma análise de *Desdobramentos do contínuo* realizada a partir de descritores de áudio.

Resumo: Neste artigo, uma análise da peça *Desdobramentos do contínuo* para violoncelo e sons eletroacústicos é realizada tendo em vista as técnicas estendidas instrumentais, sons eletroacústicos em suporte fixo, processamentos em tempo real e a interação entre eles. Esta é parte de uma pesquisa mais abrangente sobre análise musical com suporte composicional da música eletroacústica. O objetivo da análise dessa peça é entender a atividade espectral das estruturas sonoras emergentes, em termos de quais eventos produzem grandes variações no timbre, e identificar variações sutis que não seriam perceptíveis numa primeira audição da composição. Concluímos comparando os resultados das análises com as hipóteses composicionais apresentadas nas seções iniciais.

Palavras-chave: Análise musical com suporte computacional, Música mista em tempo real, Técnicas estendidas, Descritores de áudio, Timbre emergente.

1. Introduction

Audio Parametric Descriptors are tools that extract different information from audio recordings. The objective of this procedure is to analyze these data aiming to understand some features related to human auditory perception and to perform a classification of the evaluated pieces and musical styles. This research area is known as MIR (Music Information Retrieval) and the obtained analyses results until now are available in the MIREX web page[1] (Music Information Retrieval Evaluation eXchange).

The use of audio descriptors for musical classification was already employed in previous researches, such as in Peeters (2004), Pereira (2009), and Peeters *et al.* (2011). The Interdisciplinary Nucleus for Sound Communication of UNICAMP (NICS) developed similar research in the past few years, resulting in the works of Monteiro (2012), and Simurra & Manzolli (2016a, 2016b). In relation to the use of audio descriptors specifically for the analysis of contemporary music, we mention the work of Malt & Jourdan (2008, 2009) and Rossetti & Manzolli, 2017.

The general objective of this article is to contribute for the area of computer-aided analysis of live electroacoustic music. Specifically, an analysis of Rossetti's work *Desdobra-*

ROSSETTI, D.; TEIXEIRA, M., WILLIAM J. Emergent Timbre and Extended Techniques in Live-Electronic Music: An Analysis of
Desdobramentos do Contínuo Performed by Audio Descriptors
Revista Música Hodie, Goiânia, V.18 - n.1, 2018, p. 16-30

-mentos do contínuo (2016), for violoncello and live-electronics is performed by Audio Descriptors. We first present the work contextualization and a discussion about the instrumental extended techniques employed in the instrumental writing. For the analysis, parametric Audio Descriptors are adopted to get data from the audio recording. These Audio Descriptors constitute an analysis model that can possibly be used to analyze other electroacoustic works. Our analysis will be centered on the behavior of the Spectral Flux, Energy Mean, Spectral Centroid, Loudness and Spectral Flatness Descriptors whose definitions will be detailed further up.

We understand that Audio Descriptors are tools very useful to analyze different aspects of sonority in any kind of music, including classical solo, chamber or orchestral pieces. In contemporary music, these tools are important to reveal timbre qualities related to extended instrumental techniques among other applications. These techniques are associated with the production of sounds with transient and noisy characteristics, which exceed the common sound models applied to the instruments (TEIXEIRA, 2017, p. 210).

In relation to live-electronic music, we did observe that musical analyses based only on the score do not encapsulate the entire range of possibilities the pieces have. Furthermore, the interaction between the fixed tape sounds, the dynamic part of the sound processing and the instrumental parts produce dynamic elements that are beyond the score notation (MANZOLLI, 2013; MCADAMS, 2013). We name this phenomenon *emergent timbre*, which is related to the different sound layers that are in constant interaction with the performance of live-electronic pieces. This complex interaction of sound structures and processes brought by the operations and procedures of the electroacoustic music generate sonorities that emerge during the moment of the performance. These interactions can be indicated in the musical score, but they are also dependent on innumerous technical and interpretative variants. The emergent timbre is a structure formed by the amalgam of the different performed sound layers and is perceived as a *Gestalt* unity. The proposed analytical methodology, which will be detailed in the next sections, has the objective to contemplate these aspects related to live-electronic pieces.

2. Work Contextualization

Desdobramentos do contínuo is a work for violoncello and live electronics composed in 2016 by Danilo Rossetti. It is the last work included in his doctoral thesis (ROSSETTI, 2016), which investigates interaction and convergence possibilities between acoustical instruments and electroacoustic treatments (ROSSETTI, 2017). This work is dedicated to William Teixeira, who participated in its development, which involved rehearsals, cello recordings, and audio analyses.

The general form of the work contains two parts that differ from each other mainly concerning the employed electroacoustic treatments. These treatments can be implemented in real-time (morphological transformations of the violoncello sound captured live along the performance) or in fixed tape sounds (MENEZES, 1999, p. 17-18), which are audio manipulations involving phase-vocoder and convolution processes from pre-recorded cello phrases.

In relation to the real-time treatments, processes like granulation, microtemporal decorrelation, and dephasing were employed. It is conceived an ambisonic spatialization of the electroacoustic sounds, creating a diffused sound field that surrounds the listener in the moment of performance. This spatialization is planned for an eight-speakers model, however quadraphonic and stereo versions of the piece also can be performed. The integra-

ROSSETTI, D.; TEIXEIRA, M., WILLIAM J. Emergent Timbre and Extended Techniques in Live-Electronic Music: An Analysis of
Desdobramentos do Contínuo Performed by Audio Descriptors
Revista Música Hodie, Goiânia, V.18 - n.1, 2018, p. 16-30

tion of real-time electroacoustic treatments with an ambisonics spatialization is achieved by the utilization of the *process~* object, belonging to the High Order Ambisonics Library (HOA). This library was developed by the CICM of *Université* Paris 8 (GILLOT, 2012-13). In *Desdobramentos do contínuo*, the architecture of the patch was implemented in Max MSP.

The objective of overlapping both fixed tape sounds and real-time treatments of violoncello sound was to explore different possibilities of the electroacoustic universe. The adopted compositional hypothesis was that this combination would be complementary in terms of sound morphology (ROSSETTI, 2017, p. 274-278). So, the overlapped sounds would merge together into a single timbre. In this process, tape sounds have a continuous and similar development; on the other hand, real-time treatments generate sounds with discontinuous granular characteristics. In the analysis performed further on, these questions will be verified.

Next, the instrumental part of *Desdobramentos do continuo* will be discussed, focusing on extended techniques and the resultant sound morphology.

3. Instrumental extended techniques

The role of instrumental writing within this piece's discourse[2] is immense, but perhaps not in the way expected from a piece of music regularly written for an acoustic instrument with live-electronics. The concept of the piece started with an attempt to escape from two extremes usually noticed in pieces written within the genre.

On the one hand, a compositional trend is identified where musical instrument functions simply as a signal generator, with the electronic synthesis being the most prominent character of the development of musical discourse. The instrumental gesture functions almost in subordination to the electronic gesture and the function of the last is constituting continuity to the insertions, almost disparate, of the former.

On the other hand, it is possible to notice another extreme, where the instrumental gesture alone assumes the role of generator of musical material, and the electronic support works just like that, a support, a kind of effects box that only ornaments the music, almost autonomous, executed by the acoustic instrument. In this case, electronics create only small inserts of effects, and sometimes act like a tape even in live-electronics, executing another set of materials without any interaction with the instrumental gesture.

Desdobramentos do contínuo comes from this attempt to overcome such extremes, starting from the interaction between the two sound sources as basic writing material. It is important to state that, because, in this sense, instrumental writing works in a way not so much "soloistic", but much more like chamber music, since musical materials are generated by both and often, from the interaction of both. During the piece, there is a feeding of new gestures, where it is up to the instrumentalist to be able to respond instantaneously to the stimuli produced by the electronic source, including the sort of sound produced by extended techniques; like chamber music, these stimuli never repeat themselves, because they are also responses, in turn, of events previously produced by the musical instrument and which are never identical. This is the great beauty and difficulty presented by the proposal and that ends up giving a great dynamism to musical discourse. Understood in this totality, the discourse assumes more fully its vocation of interacting with, for and through its agents.

Even so, instrumental writing brings difficulties of a very advanced technical level that needs to be solved for the effectiveness of the mentioned interactions. One of the first questions that arise when musical score is read is the presence of three levels of bow pres-

ROSSETTI, D.; TEIXEIRA, M., WILLIAM J. Emergent Timbre and Extended Techniques in Live-Electronic Music: An Analysis of
Desdobramentos do Contínuo Performed by Audio Descriptors
Revista Música Hodie, Goiânia, V.18 - n.1, 2018, p. 16-30

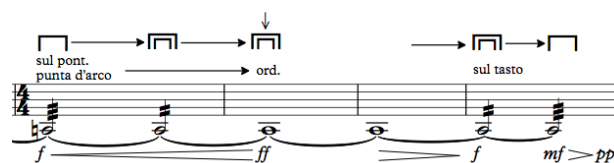sure, as in the following passage, Figure 1:

Fig. 1: Measures 31 to 34 of the score.

Although the performance of this kind of sonority is already very settled within the writing for strings in the contemporary repertoire, the piece brings a new issue that is the execution of long passages in *legato* with these different levels. This requires more than just changing the 90° angle of the bow in relation to the string, which usually generates a distorted sound, but it requires the extra weight of the interpreter. Making the three levels sound distinctive and at the same time homogeneous throughout the piece results in the fact that the piece requires a lot of the musician's physique when played in its entirety.

Another very demanding gesture it is in the "rebounds section", so to speak, whe-re different kinds of *ricochet* bow-strokes are prescribed in different rhythmical structures and with different numbers of notes per stroke, as in Figure 2:
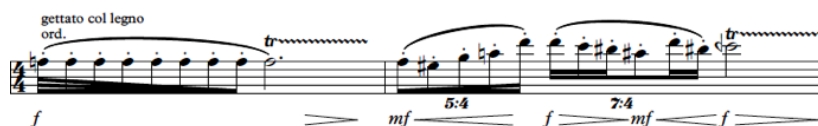

Fig. 2: Measures 43 and 44 of the score.

The aim here is to make the sounds always respond to the granular sound in the electronic synthesis, so the duration of each note must be at the same time proportional to the duration written, fluent in the gestural flow and as short as a granular sound must be to put the sound inside the bigger sonority.

The last passage that worth to be mentioned due its odd instrumental technique is this in Figure 3, but that occurs in other sections at the end of the piece:

Fig. 3: Measures 145 of the score.

This is a good example where traditional technique must be expanded not because a different timbre is required, but due to a new musical context, defining the difference be-tween an *extended technique* and an *extended sonority*; here a regular rush passage is full of notes written in hard string skipping and in the same direction of the bow, everything in a *crescendo* gesture. The result of such requirements among the microtonal pitches is an only and single sonority, almost like the Mannheim rockets in the Classical period, but re-viewed in order to get prominence to sound instead of only notes.

ROSSETTI, D.; TEIXEIRA, M., WILLIAM J. Emergent Timbre and Extended Techniques in Live-Electronic Music: An Analysis of
Desdobramentos do Contínuo Performed by Audio Descriptors
Revista Música Hodie, Goiânia, V.18 - n.1, 2018, p. 16-30

4. Audio Descriptors Model

To analyze *Desdobramentos do contínuo*, a model formed by different types of audio descriptors was determined. This model included descriptors that provide temporal, spectral, energy, and psychoacoustic features. The selected descriptors (detailed below) were Spectral Flux, Energy Mean (RMS), Spectral Centroid, Loudness and Spectral Flatness. The computational environment used for the descriptors calculation was the *Pdescriptors Library,* designed by Adriano Monteiro in PureData software (MONTEIRO, 2012), and revised by Gabriel Rimoldi.

According to Pereira (2009, p. 17) and Monteiro (2012, p. 27-29) the Spectral Flux $F_{(i)}$ is a measure of how quickly the power spectrum is changing. It is described by the magnitude difference between two successive analysis windows (X_i and X_{i-1}). This Descriptor provides lower values when the spectrum remains relatively invariable; on the other hand, it provides higher values when huge variations between successive frames are found. The Spectral Flux does not depend upon overall power (since the spectra are normalized), or phase considerations (since only magnitudes are compared). The $F_{(i)}$ is calculated from the expression below:

$$F_{(i)} = \sum_{k=1}^{k} \{log_{10}[X_i(k)] - log_{10}[X_{i-1}(k)]\}^2 \qquad \textit{(Eq. 1)}$$

where $X_i(k)$ and $X_{i-1}(k)$ are the frequency amplitude of two successive analysis windows.

According to Monteiro (*Op. Cit.*, p. 31) the Energy Mean $M_{(i)}$ or RMS (root mean square) is the root mean square (the arithmetic mean of the squares) of amplitude values in a window analysis. The RMS is also known as the quadratic mean and its values describe the energy envelope profile of a sound. The $M_{(i)}$ is defined by the following equation:

$$M_{(i)} = \sqrt{\frac{\sum_{k=1}^{K} x^2_i(k)}{K}} \qquad \textit{(Eq. 2)}$$

where $x_i(k)$ for $k = 1$ to K are the amplitude values of the $_i th$ window digitalized signal.

According to Agostini, Longari and Pollastri (2003, p. 7) the Spectral Centroid $C_{(i)}$ is the barycenter of the energy distribution belonging to the spectral envelope of a sound. It is calculated as the weighted mean of the frequencies present in the signal, where $X_i(k)$ are the magnitudes extracted from the Discrete Fourier Transform of the *i* window, and *k* is the half of the adopted spectral components of the Transform. Perceptively it is related to the sound brightness perception: higher values characterize the predominance of high frequencies in the signal (in Hertz) and lower values characterize the predominance of lower frequencies, in terms of energy. The Spectral Centroid $C_{(i)}$ is calculated from the following expression:

$$C_{(i)} = \frac{\sum_{k=1}^{k} k * |X_i(k)|^2}{\sum_{k=1}^{k} |X_i(k)|^2} \qquad \textit{(Eq. 3)}$$

where $X_i(k)$ for $k = 1$ to K are the frequency amplitude of the analysis windows.

According to Pereira (*Op. Cit.*, p. 19-20) and Monteiro (*Op. Cit.*, p. 138-139) the Loudness $L_{(i)}$ is a psychoacoustic measure related to the perception of sound amplitude. It is variable according to different frequency bands (as demonstrated by the Fletcher and Mun-

ROSSETTI, D.; TEIXEIRA, M., WILLIAM J. Emergent Timbre and Extended Techniques in Live-Electronic Music: An Analysis of
Desdobramentos do Contínuo Performed by Audio Descriptors
Revista Música Hodie, Goiânia, V.18 - n.1, 2018, p. 16-30

son curves, in 1933) and describes the auditory sensation of amplitude variation of a given sound. The Loudness $L(i)$ of a spectral analysis window is determined by Eq. 4, according to Pereira's model (2009, p. 19), and the Fletcher and Munson curves are included in Eq. 4 from the W[k] factor, which modulates the $X_i(k)$ values. This formula is presented in Eq. 5:

$$L_{(i)} = \sum_{k=1}^{K} |X_i[k]|^2 10^{w[k]/20}$$

<div align="right">(Eq. 4)</div>

where $X_i(k)$ for $k = 1$ to K are the frequency amplitude of the analysis windows.

$$W(k) = -0{,}6 * 3{,}64 * f(k)^{-0{,}8} - 6{,}5 * e^{-0{,}6*(f(k)-3{,}3)^2} + 10^{-3}*f(k)^{3{,}6}$$

<div align="right">(Eq. 5)</div>

where the frequency $f(k)$ is measured in kHz, is defined as $f(k) = k.d$, and it is the difference between two consecutive spectral bins in kHz.

According to Peeters (*Op. Cit.*, p. 20) and Monteiro (*Op. Cit.*, p. 126-127) the Spectral Flatness Descriptor quantifies the amount of noise found in a sound signal (noisiness), in opposition to the measure of "tonal quality". An extremely high Spectral Flatness is found in the white noise (1,0 value), on the other hand, the lower level of Spectral Flatness is found in a pure harmonic tone (e.g. an additive synthesis timbre formed by sine waves). This Descriptor is calculated from the ratio of the geometric mean to the arithmetic mean of the energy spectrum value and is computed for several frequency bands[3]. It is important to highlight that the Spectral Flatness Descriptor is not dependent on the intensity of a sound signal. This means that a sound with extremely low intensity have can high values of Spectral Flatness, and a sound with high intensity can have low values of this Descriptor. The equation that is used to compute the values of the Spectral Flatness is presented below (Eq. 6):

$$S_{(n)} \frac{(\prod_{k \in N} X(k))^{1/k}}{\frac{1}{K} \sum_{k \in N} X(k)}$$

<div align="right">(Eq. 6)</div>

where $X(k)$ is the frequency amplitude for $k \in N$ band.

These chosen audio descriptors will be applied to the audio recording of the piece whose analysis will be presented next. For the tape analyses, Spectral Flux, RMS, Spectral Centroid and Loudness Descriptors will be applied. For the analysis of the real-time processing and the emergent timbre, we developed an approach based on Spectral Flux and Spectral Flatness results, in order to discuss the interaction between instrumental and electroacoustic sounds that constitute the whole generated timbre.

5. Analysis by Audio Descriptors

In the audio of the performance used for this analysis[4], the entire piece lasts 12". The first part goes from the beginning to 5'30", and the second part from 5'30" to its ending. In the first part, the fixed tape sound corresponds to a phase vocoder that stretches the spectrum of a given sound and repeats it continuously as a loop. During this part, the sound is sent to a granulator that has six different presets containing a sort of parameter values (such as grain size and rarefaction). These presets determine the direction of the sound mass evolution, whose perception changes gradually from a continuous timbre to a grainy sound cloud considerably rarefied.

In the second part, the tape sounds were originated from the convolution between different pre-recorded cello sounds. In total five sounds of different durations were generated by this process (which have respectively 35", 26", 50", 62" and 78" of duration). As

ROSSETTI, D.; TEIXEIRA, M., WILLIAM J. Emergent Timbre and Extended Techniques in Live-Electronic Music: An Analysis of
Desdobramentos do Contínuo Performed by Audio Descriptors
Revista Música Hodie, Goiânia, V.18 - n.1, 2018, p. 16-30

common perceptual features among them, all these sounds have continuous spectral evolutions in time. It is important to remark that during the entire piece besides the tape sounds, the cello sound is granulated in real-time (its parameters are constantly modified), and the electroacoustic timbre (formed by these layers) is spatialized through high-order ambisonics models.

5.1 Analysis of Fixed Tape Sounds

In this section, the looped phase-vocoder sound of the first part (which changes gradually in time) and the five tape sounds of the second part, generated by convolution processes, will be analyzed and discussed.

In the first part, the phase vocoder generated sound evolves directionally from a continuous texture to a grainy sound mass that gradually becomes more discontinuous in perception. Our audio descriptors model was applied to the audio and the resultant graphics with normalized values are presented in Figure 4.
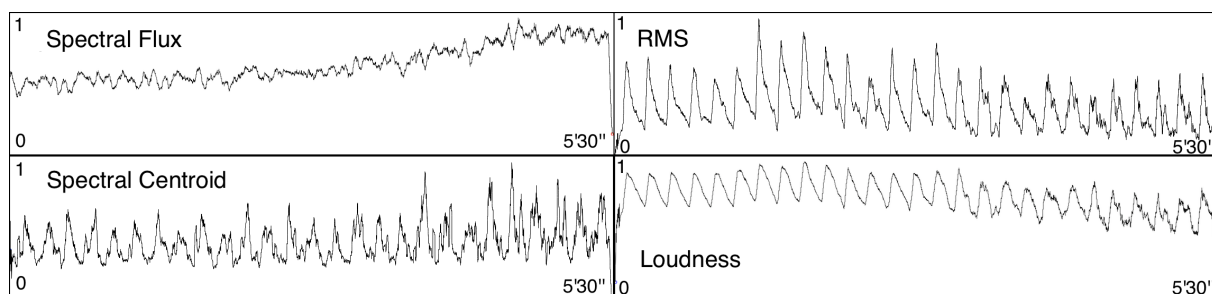

Fig. 4: Descriptor model applied to the phase vocoder sound.

As shown in the Figure above, the Spectral Flux and Spectral Centroid curves have an overall increasing perception. At the same point (around 3'10"), both curves start to have higher values. Here, the growth of the Spectral Flux curve is more consistent and constant, meaning that there are more intensity variations and spectral activity between successive frames. The centroid curve has also fewer and weaker peaks in the beginning, with an overall increasing frequency brightness perception during its evolution. We observe that both curves have stronger peaks at the end of their sound evolution.

The Energy Mean evolution (RMS) presents few peaks of energy that appear periodically. The higher peak arrives at 1'02", and then there is an overall decreasing perception. The Loudness curve has a similar behavior with an increasing pattern from 0 to 1'. At this point, it also decreases gradually. From these observations, we assume that the Spectral Flux and Spectral Centroid Descriptors have a convergent behavior, the same happening with the RMS and Loudness Descriptors. The formers give us information about the spectral movement of the sound timbre, and the last inform us about the sound intensity perception.

We assume that the variations found in Spectral Flux and Spectral Centroid curves are related to the granulation parameters applied to the phase-vocoder sound. In the first part of the work, six different granulation preset values were applied. On them, while the feedback rate and grain delay remain constant, the grain size decreases from 400 to 75ms, while rarefaction rate increases from 0 (a totally continuous sound in perception) to 0,8 (indicating an amount of 80% of silence in the totality the diffused sound mass).

In relation to the grainy cloud perception, it is important to emphasize that bigger grains generate sonorities that privilege the sustained parts of the sounds (normally characterized by the presence of a fundamental frequency and upper partials). Smaller grains

ROSSETTI, D.; TEIXEIRA, M., WILLIAM J. Emergent Timbre and Extended Techniques in Live-Electronic Music: An Analysis of
Desdobramentos do Contínuo Performed by Audio Descriptors
Revista Música Hodie, Goiânia, V.18 - n.1, 2018, p. 16-30

have a prominent presence of attack transients. For this reason, from a sound morphology standpoint, grainy clouds formed by smaller grains sizes (of less than 100ms) have a noisier auditory perception (ROSSETTI, 2016, p. 284-285).

On the second part of *Desdobramentos* five tape sounds were addressed (Seq. 1 to 5)[5]. As a common feature of all these tape sounds they all have a continuous evolution. However, it is desired to investigate if they have different evolution characteristics. In this sense, Audio Descriptors can support the evaluation of timbre qualities of these sounds, in order to describe their behavior. We applied the presented Descriptors model to each sound and extracted the normalized (from 0 to 1) arithmetic average of each descriptor value (Tab. 1). This strategy was adopted to obtain significant data, in order to compare the evolution of the Descriptors applied to the sounds.

Sound/Descriptor	Seq. 1	Seq. 2	Seq. 3	Seq. 4	Seq. 5
Flux	0,32	0,4074	0,3111	0,2733	0,3051
RMS	0,183	0,3612	0,2962	0,4797	0,4094
Centroid	0,084	0,3515	0,2388	0,2581	0,3816
Loudness	0,63	0,6051	0,7636	0,7291	0,7113

Tab. 1: Normalized averages of the Audio Descriptors of the Second Section.

From Tab. 1, it is possible to verify that the five sequences of the tape sounds show a gradual increase of the Energy Mean. This behavior is more prominent in Seq. 4 and 5. Therefore there is more spectral energy at the end of the piece. These spectral changes act in the perception as an increase in intensity and sound density. Regarding Spectral Centroid values, we observe that the five tape sounds are organized in three brightness levels: low, middle and high. The low brightness level is assigned to the Seq. 1, the middle brightness level is assigned to Seq. 3 and Seq. 4, and the high brightness level is related to Seq. 2 and 5. Finally, the normalized Loudness average values are concentered in a middle-high level. Seq. 1 is nearer to the middle level, while Seq. 3, Seq. 4 and Seq. 5 have higher intensity perception. Next, in Figure 5, a histogram graphic is presented, showing the descriptors average values related to each tape sound, in complementation to Tab. 1.
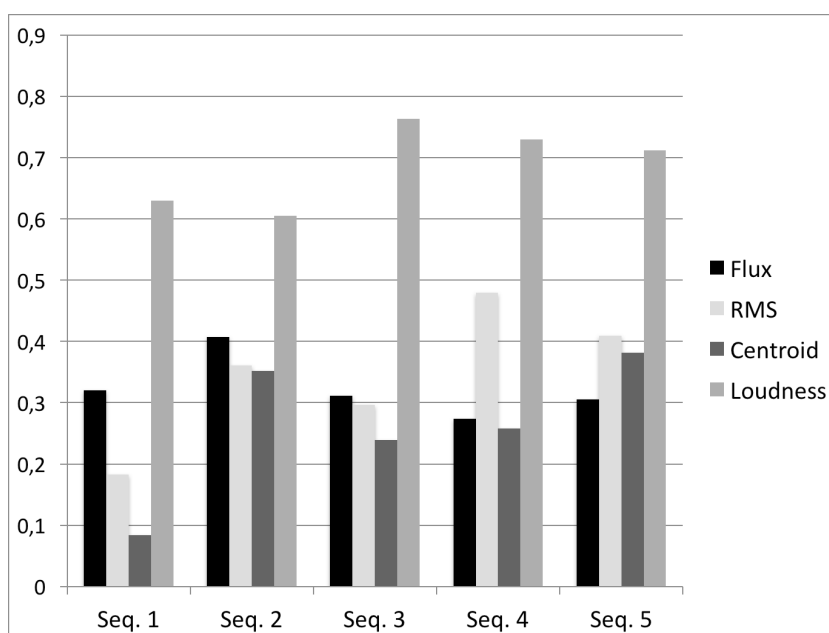

Fig. 5: Descriptor values for each sound.

ROSSETTI, D.; TEIXEIRA, M., WILLIAM J. Emergent Timbre and Extended Techniques in Live-Electronic Music: An Analysis of
Desdobramentos do Contínuo Performed by Audio Descriptors
Revista Música Hodie, Goiânia, V.18 - n.1, 2018, p. 16-30

Taking into account the Figure above, some observations can be made, in consideration of a global view of the Descriptor values of each sound. Seq. 1 has the lower RMS, lower Centroid and one of the lowest Loudness values. Seq. 2 has the higher Flux, high Centroid, and the lower Loudness. Seq. 3 has the higher Loudness, average Centroid, and average-low RMS. Seq. 4 has the higher RMS, high Loudness, average Centroid and the lower Flux. Seq. 5 has the higher Centroid value, high RMS, and low Flux.

5.2 Analysis of the emergent timbre

In this analysis, we focus on the real-time granulation of the violoncello sound, its acoustical sonority, and extended techniques, merged together with the electroacoustic tape sounds. For this purpose, two different excerpts[6] of the piece will be addressed. These excerpts were chosen from the Spectral Flux and Spectral Flatness analyses of the entire piece. The first excerpt corresponds to a moment where we find a growth of both mentioned curves, and, in the second excerpt, a huge distance between them is found (high values for the Spectral Flatness and low values for the Spectral Flux Descriptor). We will search to explain why this kind of behavior is found in those Descriptors from both excerpts.

As we can see in Figure 6[7] (Sonogram of the entire piece, Spectral Flux – in red – and Spectral Flatness – in orange – curves), there is only one moment where both curves grow together: from 2'03" from 2'35" of the recording. Otherwise, we find more than one moment where there is a huge distance between Spectral Flux and Spectral Flatness curves. Then, we decide to address the excerpt between 6'46" and 7'37" due to the generated timbre of this part that differs from the first example in terms of musical intention. The two analyzed excerpts are circulated in yellow in Figure 6.
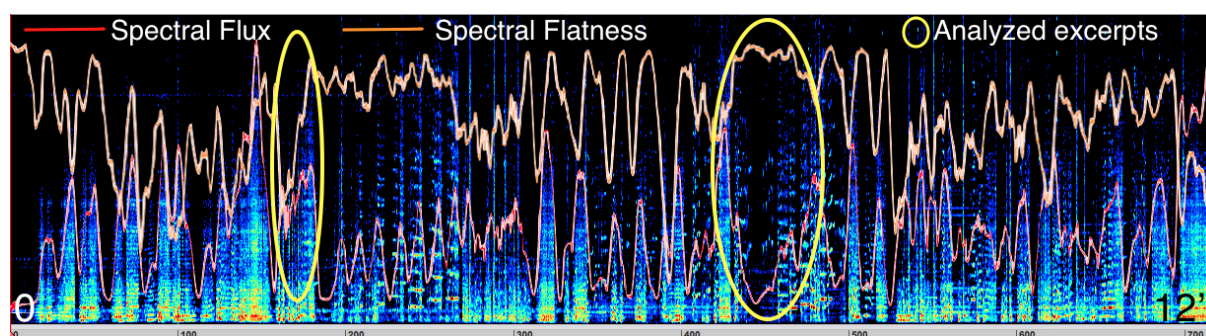

Fig. 6: Desdobramentos do contínuo sonogram and Spectral Flux and Spectral Flatness curves: two circulated excerpts to be analyzed.

In the first excerpt, from 2'03" to 2'35" (measures 26-34), we have in the electroacoustic sounds a loop in the phase-vocoder with a continuous texture, besides the real-time violoncello's granulation. In the instrumental writing (Fig. 7), relatively fast musical phrases are combined with sustained notes, which are modulated from effects such as *molto vibrato* and *tremolo* with *sul tasto* to *sul ponticello* bow positions, and from normal to exaggerate bow overpressure above the string. The intention of these effects is to produce timbre variations in the real-time generated electroacoustic sounds and, consequently, in the timbre amalgam. In Figure 7, the score of this excerpt is shown.

ROSSETTI, D.; TEIXEIRA, M., WILLIAM J. Emergent Timbre and Extended Techniques in Live-Electronic Music: An Analysis of
Desdobramentos do Contínuo Performed by Audio Descriptors
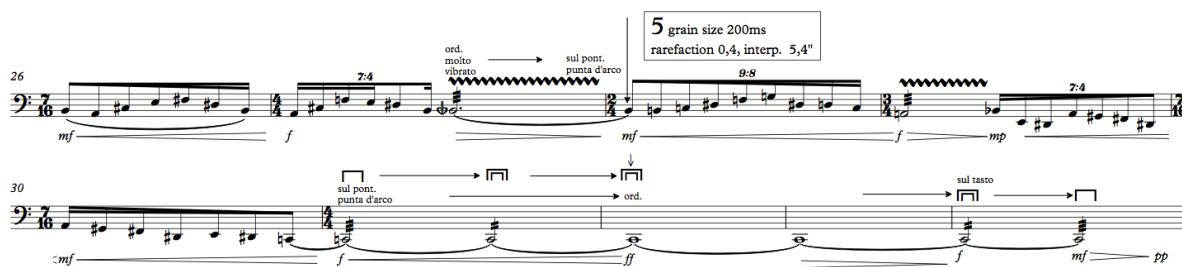Revista Música Hodie, Goiânia, V.18 - n.1, 2018, p. 16-30

Fig. 7: Measures 26 to 34 of the score.

In Figure 8 we find the sonogram, Spectral Flux (red) and Spectral Flatness (orange) curves. Since the electroacoustic part remains relatively constant (as a continuous sound layer), we admit that the perceptive timbre variations come from the violoncello sounds and these variations are related to the perceived different timbre densities.
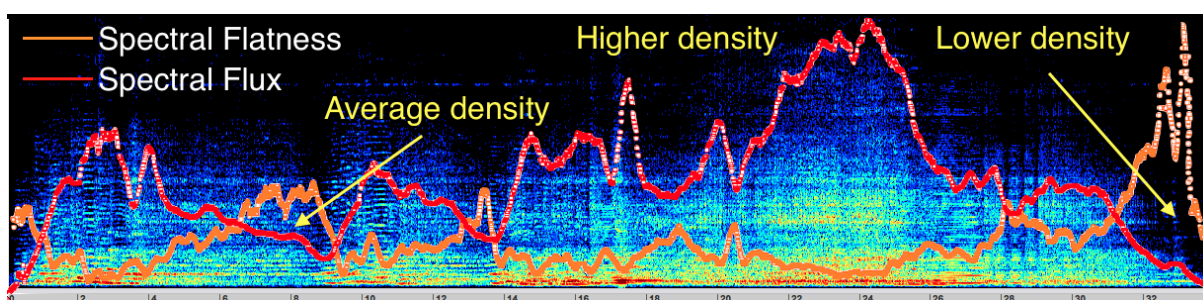

Fig. 8: Sonogram, Spectral Flux and Spectral Flatness curves of excerpt 1.

In the beginning of this excerpt, the violoncello phrases produce an average density timbre, with dynamics varying from *mp* to *f*. Here, Spectral Flux and Spectral Flatness are relatively close, alternating between which Descriptor has a higher value at each point. When we look at the higher intensity part, it corresponds to measures 31-32, where there is a C2 *tremolo sul ponticello* with exaggerated overpressure above the C string, varying from *f* to *ff*. These effects produce a noisy dense timbre, characterized by a huge spectral movement, indicated by high Spectral Flux values. At the end of this section (measure 34), the same sustained pitch is still in *tremolo*, but the bow position is now on *sul tasto* with normal pressure, and the dynamics are decreasing to *pp*. In this situation, the spectral movement is very low; on the other hand, the Spectral Flatness reaches the higher point. Here, the violoncello does not perform fast alternating pitches, but only a *tremolo* in C2 pitch, which decreases in intensity from *mf* to *pp*, at the same time that the bow pressure establishes in the normal level. For the high Spectral Flatness values found, we assume that the *tremolo* effect and the overpressure produce a noisy inharmonic timbre, which is complemented by the electroacoustic tape sound that comes from the phase-vocoder.

The second chosen excerpt corresponds to measures 91 to 103 of the score (6'46'' to 7'37'' of the recording), which is shown in Figure 9, where high values of Spectral Flatness and low values of Spectral Flux were computed. As we see in this Figure, several different violoncello extended techniques are performed, such as trills, *gettato col legno*, *tremolo*, artificial harmonics, and bow overpressures in different levels with positions *sul tasto*, *ordinario* and *sul ponticello*. From a simple description of the employed instrumental techniques, it becomes not clear why this kind of timbre perception indicated by the Descriptor's curves is originated.

ROSSETTI, D.; TEIXEIRA, M., WILLIAM J. Emergent Timbre and Extended Techniques in Live-Electronic Music: An Analysis of
Desdobramentos do Contínuo Performed by Audio Descriptors
Revista Música Hodie, Goiânia, V.18 - n.1, 2018, p. 16-30

For a more detailed investigation of the produced timbre amalgam, in Figure 10 we show the sonogram, Spectral Flux (in red) and Spectral Flatness (in orange) curves of this excerpt. From this analysis, we highlight three moments in this excerpt. The first moment presents average values for the Descriptors and goes from measures 91 to 94 of the score, where trills, *gettato col legno* and *accelerandi* figurations in the violoncello merge together with its real-time granulation. The second moment is found in the measures 95 to 97, where the cellist performs an artificial harmonic *glissando* with *tremolo* and overpressure bow. Here, due to the huge spectral activity and intensity produced, high values of the Spectral Flux Descriptor are detected. On the other hand, this high instrumental activity produces low Spectral Flatness values, since the generated spectrum has prominent harmonic features, although the "noisy" and dense perceived instrumental sound.

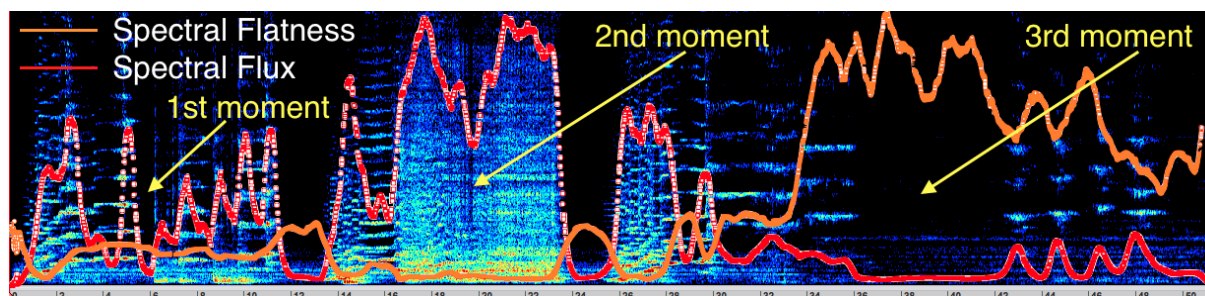Fig. 9: Measures 91 to 103 of the score.

Fig. 10: Sonogram, Spectral Flux and Spectral Flatness curves of excerpt 2.

The third moment is related to measures 101 to 103, where high values of Spectral Flatness and low values of Spectral Flux are detected. The higher difference between both Descriptors is found in measure 101, where the violoncello does not play and we listen only to the electroacoustic tape sound with low intensity. From these results, we assume that the tape has an inharmonic spectral configuration that is approximated by the white noise. The difference between these Descriptors decreases in measures 102-103 because there are artificial harmonics with *tremolo* played by the violoncellist in low intensity. Since both structures are permeable[8] and clearly perceived, average-high values for Spectral Flatness and average-low values for the Spectral Flux are detected.

After the segmentation of these two excerpts of the piece based on Spectral Flux and Spectral Flatness Descriptors information, we applied the five Descriptors of our analytical model to both excerpts and extracted the normalized average values of them. With these data, it became possible to compare the behavior of these Descriptors in the analyzed parts and to extract information about the emergent timbre. These values are shown in Figure 11.

ROSSETTI, D.; TEIXEIRA, M., WILLIAM J. Emergent Timbre and Extended Techniques in Live-Electronic Music: An Analysis of
Desdobramentos do Contínuo Performed by Audio Descriptors
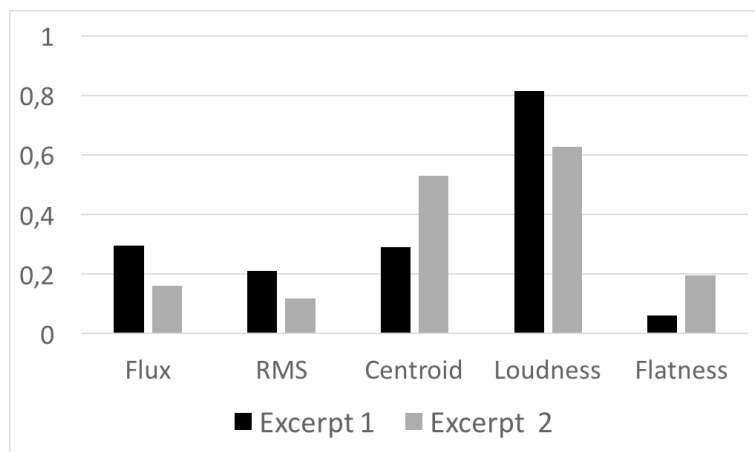Revista Música Hodie, Goiânia, V.18 - n.1, 2018, p. 16-30

Fig. 11: Descriptors average values of excerpts 1 and 2.

It is important to emphasize that the objective of this analysis is not to compare the absolute obtained mean values among different descriptors in order to describe the auditory results but to compare the behavior of each descriptor in different excerpts. In keeping with this approach, because of the high difference values of Spectral Flatness Descriptor, the obtained normalized arithmetic mean has relatively low values (around 0.1 and 0.2). On the other hand, in the Spectral Flux Descriptor higher absolute normalized values are found: around 0.3 and 0.2. The interpretation of the information extracted by these Descriptors could be that in excerpt 1 we have higher values of Spectral Flux than in excerpt 2, which means that the amount of spectral movement is higher in excerpt 1. When it comes to the Spectral Flatness Descriptor, higher values are found in, which means that in this excerpt the spectromorphology is closer to the noise configuration, holding a more inharmonic distribution.

Upon the other employed descriptors, RMS and Loudness have similar behaviors, even though their absolute values are very different. Both of them refer to intensity, but RMS is a physical measure, whereas Loudness is a psychoacoustic measure. In both average intensity measures the normalized average values are higher in excerpt 1 in relation to excerpt 2. Lastly, we find Spectral Centroid values higher for the second excerpt, which means that the brightness average timbre perception is concentrated in a higher region in this excerpt. A correlation between Spectral Centroid and Spectral Flatness average results is detected. By our observation, this can be explained because higher Spectral Flatness values are normally found in inharmonic electroacoustic textures composed by higher frequencies and not in instrumental harmonic or even inharmonic timbres. This kind of timbre quality is closer to the second excerpt where we have the electroacoustic tape in the first plan in some parts in combination with high-frequency *tremolo* artificial harmonics in the violoncello.

6. Conclusion

In this article, we intended to propose a methodology for analysis of live-electroacoustic music based on the utilization of Audio Descriptors; an attempt to contribute in the field of computer-aided musical analysis. In the analysis of the emergent timbre of *Desdobramentos do contínuo*, we found it is relevant the application of Spectral Flux, Energy Mean, Spectral Centroid, Loudness and Spectral Flatness Descriptors. In further works, our objective is to perform a more detailed research about the orthogonality between most Spectral Flux and Spectral Flatness values, in order to have a more nuanced understanding about the information that can be extracted from these Descriptors. From them, we presume that issues about spectral richness, inharmonicity and noise features can be extracted

ROSSETTI, D.; TEIXEIRA, M., WILLIAM J. Emergent Timbre and Extended Techniques in Live-Electronic Music: An Analysis of
Desdobramentos do Contínuo Performed by Audio Descriptors
Revista Música Hodie, Goiânia, V.18 - n.1, 2018, p. 16-30

from the analyzed audio.

From the analysis of the fixed tape sounds of the piece by these Descriptors, we extracted important information data in order to clarify how timbre features change in time. On a first listen, we tend to consider them similar to each other, due to their continuous time evolution. However, after the application of our Descriptors model, subtle variations become noticeable and our perception becomes more attentive to these nuances. It is also desirable during the performance that the interpreters are aware of these nuances. Thus, they can interact with them with more accuracy, in order to produce a more balanced performance, considering acoustic and electroacoustic parts.

These subtle timbre variations, in a certain way, complement the previously presented compositional hypothesis. The tape sounds have a global continuous evolution. However, for the phase-vocoder sound, after a certain point, there is a discontinuity perception demonstrated by higher levels of Spectral Flux and Loudness values. In relation to the five tape sounds of the second part, the variability of RMS and Spectral Centroid values characterizes different features of their global timbre perception. In addition, despite the tape sounds nuances, the main timbre differences in the global perception of the work (defined by the variations of the Spectral Flux and Spectral Flatness Descriptors applied to the entire piece) are related to the employed instrumental extended techniques and their real-time granulation. Considering other audio descriptors, these timbre variations reflect mostly on RMS and Spectral Centroid differences.

Finally, from this analysis of the emergent timbre, we verified that the change in the perceived timbre morphology is mostly guided by the instrumental part, especially in terms of spectral harmonic/inharmonic activity. Structures with an inharmonic and noisy configuration are normally provided from the electroacoustic parts (tape sounds or real-time processing). The fusion of these structures into one single perceived timbre is possible due to the level of their permeability and their different or complementary spectral qualities.

The emergent timbre arises from the interaction between instrumental and electroacoustic sounds. During the performance, there is a constant process of adaptation between the instrumental and electroacoustic interpreters guided by listening. This auditory feedback modulates the reactions of the interpreters with the aim of merging the different sound sources: the electroacoustic interpreter controls the diffusion of the tape sounds, the real-time granulation of the violoncello, its clean amplification and the general sound intensity, and the instrumental interpreter modulates the intention of his performance in terms of dynamics and musical time from this information. It is important to emphasize that in live-electronic music, the instrumental performer is the main responsible for musical time, since he or she dictates the succession of musical events. This temporality is dependent and is always in relation to the listening of the resonance of the electroacoustic sound events.

7. Acknowledgments

Note

1 <http://www.music-ir.org/mirex/wiki/MIREX_HOME>.

2 Musical discourse is a term adopted here to mean the whole relationship between musical agents and not as a synonym to musical work and even less to notation. *Cf*: Teixeira, 2016.

3   In this article, the band from 250 to 500Hz was chosen for the Spectral Flatness computation.

ROSSETTI, D.; TEIXEIRA, M., WILLIAM J. Emergent Timbre and Extended Techniques in Live-Electronic Music: An Analysis of
Desdobramentos do Contínuo Performed by Audio Descriptors
Revista Música Hodie, Goiânia, V.18 - n.1, 2018, p. 16-30

4   The performance took place in the SBCM 2017 at ECA-USP in 05th September 2017, by William Teixeira (violoncello) and Danilo Rossetti (live-electronics). This audio recording can be found at:
    <https://soundcloud.com/danilorossetti/desdobramentos-do-continuo>.

5   The *tapes* Seq. 3 and Seq. 5 can be listened at: <https://drive.google.com/drive/folders/1M4bLF8o77bEqetqoogcK3U- -XTU2tUNyH?usp=sharing> also: <https://emac.ufg.br/up/269/o/Seq3_main.mp3> <https://emac.ufg.br/up/269/o/ Seq5_main.mp3>.

6   Both analyzed excerpts can be listened at: <https://drive.google.com/drive/folders/1GdmSypSC56atebGr9oFWsC 5tbZTua-Lz?usp=sharing> also: <https://emac.ufg.br/up/269/o/Excerpt1.mp3> <https://emac.ufg.br/up/269/o/Excerpt2.mp3>.

7   In order to have the sonogram and both the Spectral Flux and Spectral Flatness curves in the same figure, this audio analysis was performed in *AudioSculpt* software.

8   The concept of permeability is defined by György Ligeti in his text "Évolution de la Forme Musicale" (1958), French translation. For more information, *Cf.* LIGETI, 1958, In LIGETI, 2010, p. 119-135 and Rossetti, 2017, p. 275-276.

References

AGOSTINI, Giulio; LONGARI, Maurizio; POLLASTRI, Emanuele. Musical Instrument Timbre Classification with Spectral Features. *EURASIP Journal on Applied Signal Processing*, New York, v. 2003, n. 1, p. 5-14, 2003.

GILLOT, Pierre. Les traitements musicaux en ambisonie. Dissertação de Mestrado. Centre de recherche Informatique et Création Musicale da Université Paris 8, 2012-2013. Saint Denis: Université Paris 8, 2012-2013. 104p.

LIGETI, György. Évolution de la forme musicale. In: LIGETI, György. *Neuf essais sur la Musique*. 2ª Ed. Genève: Contrechamps, 2010. p. 119-137.

MALT, Mikhail; JOURDAN, Emmanuel. Zsa.Descriptors: A Library for Real-Time Descriptors Analysis. In: SOUND AND MUSIC COMPUTING CONFERENCE, 5., 2008, Berlin. *Proceedings…* Berlin: Sound and Music Computing, 2008.

_____. Real-Time Uses of Low Level Sound Descriptors as Event Detection Functions Using the Max/MSP Zsa.Descriptors Library. In: BRAZILIAN SYMPOSIUM ON COMPUTER MUSIC, 12., 2009, Recife. *Proceedings…* São Paulo: USP, SBC, 2009, p. 45-56.

MANZOLLI, Jônatas. Interpretação mediada: pontos de referência, modelos e processos criativos. *Revista Musica Hodie*, Goiânia, v. 13, n. 1, p. 48-63, 2013.

MCADAMS, Stephen. Musical Timbre Perception. In: DEUTSCH, Diana. *The Psychology of Music*. 3ª Ed. San Diego: Academic Press, 2013. p. 35-67.

MENEZES, Flo. Fusão e contraste entre a escritura instrumental e as estruturas eletroacústicas. In: MENEZES, Flo. *Atualidade estética da música eletroacústica*. São Paulo: Editora UNESP, 1999. p. 13-20.

MONTEIRO, Adriano. Criação e performance musical no contexto dos instrumentos musicais digitais. Dissertação de Mestrado. Instituto de Artes da Universidade Estadual de Campinas, 2012. Campinas: UNICAMP, 2012. 139p.

PEETERS, Geoffroy. *A Large Set of Audio Features for Sound Description (Similarity and Classification) in the CUIDADO Project*.

<http://recherche.ircam.fr/anasyn/peeters/ARTICLES/Peeters_2003_cuidadoaudiofeatures. pdf>. Acessed 30 nov. 2016.

PEETERS, Geoffroy *et al*. The Timbre Toolbox: Extracting audio descriptors from musical signals. *Journal of the Acoustic Society of America*, Melville, v. 130, n. 5, p. 2902-2916, November 2011.

ROSSETTI, D.; TEIXEIRA, M., WILLIAM J. Emergent Timbre and Extended Techniques in Live-Electronic Music: An Analysis of
Desdobramentos do Contínuo Performed by Audio Descriptors
Revista Música Hodie, Goiânia, V.18 - n.1, 2018, p. 16-30

PEREIRA, Erica. Estudos sobre uma ferramenta de classificação musical. Dissertação de Mestrado. Faculdade de Energia Elétrica e de Computação da Universidade Estadual de Campinas, 2009. Campinas: UNICAMP, 2009. 57p.

ROSSETTI, Danilo. Processos microtemporais de criação sonora, percepção e modulação da forma: uma abordagem analítica e composicional. Tese de Doutorado. Instituto de Artes da Universidade Estadual de Campinas, 2016. Campinas: UNICAMP, 2016. 475p.

_____. The Qualities of the Perceived Sound Forms: A Morphological Approach to Timbre Composition. In: ARAMAKI, Mitsuko; KRONLAND-MARTINET, Richard; YSTAD, Sølvi (Eds.). *Bridging People and Sound:* 12th International Symposium, CMMR 2016, São Paulo, Brazil, July 5-8 2016, Revised Selected Papers. Cham: Springer, 2017. p. 259-283.

ROSSETTI, Danilo; MANZOLLI, Jônatas. De Montserrat às ressonâncias do piano: uma análise com descritores de áudio. *Revista Opus*, versão *online*, v. 23, n. 3, p. 193-221, 2017.

SIMURRA, Ivan; MANZOLLI, Jônatas. O azeite, a lua e o rio: o segundo diário de bordo de uma composição a partir de descritores de áudio. *Revista Música Hodie,* Goiânia, v. 16, n. 1, p. 101-123, 2016a.

_____. Sound Shizuku Composition: a Computer-Aided Composition Systems for Extended Music Techniques. *MusMat Brazilian Journal of Music and Mathematics*, Rio de Janeiro, v. 1, n. 1, p. 86-101, 2016b.

TEIXEIRA, William. O discurso musical. In: PRESGRAVE, Fabio; NODA, Luciana; MENDES Jean Joubert (Eds.). *Ensaios sobre a música do século XX e XXI*: composição, performance e projetos colaborativos. Natal: EDUFRN, 2016, p. 226-243.

_____. Por uma performance retórica da música contemporânea. Tese de Doutorado. Escola de Comunicações e Artes da Universidade de São Paulo, 2017. São Paulo: USP, 2017. 264p.

Danilo Rossetti – Doutor em Composição Musical pela UNICAMP, com período sanduíche no Centre de recherche Informatique et Création Musicale da Université Paris 8. É Mestre em Música e Bacharel em Composição (instrumental e eletroacústica) e Regência pela UNESP. Atualmente, realiza pesquisa de pós-doutorado junto ao Núcleo Interdisciplinar de Comunicação Sonora da UNICAMP, sob supervisão do Prof. Dr. Jônatas Manzolli, com apoio da FAPESP, e é professor de pós-graduação do Centro Universitário Senac SP, nas áreas de comunicação e artes. Foi um dos premiados do Prêmio Funarte de Música Clássica 2016, na categoria solos, música acusmática e mista. Atua principalmente nas seguintes áreas: composição musical (instrumental, eletroacústica e mista), teoria e análise musical.

William Teixeira William – É Bacharel em música com habilitação em violoncelo pela UNESP. Tem desenvolvido trabalho dedicado à interface entre aspectos teóricos e práticos da música contemporânea, tendo estreado dezenas de obras de diversas gerações de compositores brasileiros. Atualmente é Professor Adjunto da Universidade Federal de Mato Grosso do Sul - UFMS.

Jônatas Manzolli – É graduado em Matemática Aplicada Computacional (1983) e em Composição e Regência (1987) e é mestre em Matemática Aplicada (1988) ambos pela Unicamp. Desenvolveu seu doutorado (PhD) na University of Nottingham (1993) sobre Composição Musical. Atualmente é Professor Titular do Instituto de Artes da Unicamp e Coordenador do Núcleo Interdisciplinar de Comunicação Sonora (NICS). Compositor e matemático, pesquisa a interação entre Arte e Tecnologia em criação musical, computação musical e ciências cognitivas. Atua no programa de pós-graduação em Música com ênfase em Processos Criativos e Fundamentos Teóricos em Música e Tecnologia. Suas publicações focam, principalmente, os seguintes temas: composição musical, síntese de som, auto-organização e criatividade sonora, ambientes interativos para composição, modelos matemáticos e computação evolutiva aplicados a processos sonoros. Sua produção artística relaciona música instrumental, eletroacústica, obras multimídia para dança e instalações sonoras.