

MUSICA THEORICA 201802

SCIENTIFIC ARTICLE

Data do recebimento: 01/08/2018

Data da aprovação final: 08/10/2018

Análise Musical Assistida por Descritores de Áudio: um estudo de caso da obra *Reflexões* de Jônatas Manzolli

*Computer-Assisted Music Analysis: a case study of the work
"Reflections" by Jônatas Manzolli*

Ivan Eiji Simurra

Universidade Federal do ABC

ieysimurra@gmail.com

Resumo: Trata-se de um artigo de análise musical com suporte computacional que associa as informações simbólicas de uma obra sinfônica com dados gerados por descritores de áudio, a partir da gravação da estreia da obra. Esse método foi aplicado na análise de *Reflexões* (2011) de Jônatas Manzolli, pois trata-se de uma obra que foca variações de 95 sonoridades a partir de um único referencial e da mistura com a eletroacústica. O objetivo foi encontrar padrões estruturais por intermédio de diversas características sonoras e analisar as correlações entre as mesmas.

Palavras-chave: Análise Musical; Suporte Computacional para Análise; Descritores de Áudio; *Pure Data*; Música Orquestral.

Abstract: It is an article of musical analysis with computational support that associates the symbolic information of a symphonic work with data generated by audio descriptors, from the recording of the premiere of the work. This method was applied in the analysis of *Reflections* (2011) of Jônatas Manzolli, because it is a work that focuses on variations of 95 sounds from a single reference and the mixture with electroacoustics. The objective was to find structural patterns through different features and to analyze the correlations between them.

Keywords: Music Analysis; Computer-Aided Analysis; Audio Features; *Pure Data*; Orchestral Music.



MUSICA THEORICA

Revista da Associação Brasileira de Teoria e Análise Musical
Journal of the Brazilian Society for Music Theory and Analysis

@ TeMA 2018 - ISSN 2525-5541

1.1 - Introdução

O trabalho de análise musical realizado diretamente pela partitura disponibiliza dados consideráveis que fornecem evidências sobre aspectos estéticos, conceituais e musicais que podem se relacionar com o planejamento particular do compositor. Contudo, a análise realizada somente por essa fonte pode não ser suficiente para a compreensão macro e microestrutural do universo sonoro no qual o compositor está inserido em seu projeto composicional, principalmente quando a obra em questão faz uso de dispositivos de escrita que ampliam o escopo das transformações timbrísticas. Nesses casos, além dos tradicionais métodos de análise musical, tais como destacados por Cook (1994), os ambientes computacionais possibilitam novas técnicas de análise musical e recuperação da informação musical. Uma considerável quantidade de obras compostas a partir da segunda metade do século XX, possui características e objetivos composicionais os quais expandem a paleta timbrística por intermédio das misturas instrumentais e pelas técnicas instrumentais expandidas. No âmbito de várias obras do repertório produzido a partir dessa abordagem composicional, destacamos o compositor brasileiro Jônatas Manzolli. Trata-se de um compositor inserido no panorama da música brasileira contemporânea que utiliza recursos de análise e criação musical com suporte computacional. Como destacado em Le Groux et al. (2010), sua pesquisa criativa busca uma interação entre arte e tecnologia, computação musical e ciências cognitivas. Recentemente, a convite das comemorações do aniversário de 50 anos da UNICAMP, Manzolli compôs uma obra multimodal na qual integra a sonoridade da orquestra com dança contemporânea, coro cênico e tecnologia digital¹. No trabalho apresentado neste artigo analisamos *Reflexões* para orquestra sinfônica e audiovisual. Essa obra constrói um discurso que se relaciona com variações dinâmicas de sonoridades, intensidades e timbres orquestrais. As texturas sonoras, estabelecidas em *Reflexões*, relacionam-se com as “torrentes de ideias que germinam e povoam os processos de descoberta e transformações”².

Motivados por tal questão, desenvolvemos um projeto de análise que dialoga com a área de *Music Information Retrieval* ou MIR (Sturm 2014). Nela, utilizam-se recursos de processamento digital de sinais para extrair características sonoras do espectro do som. Essas medidas de características são denominadas “descritores de áudio” na literatura especializada da área. Segundo Peeters (2004, p. 1), os procedimentos para caracterizar um sinal sonoro foram propostos pela comunidade científica com o objetivo de reconhecer padrões da

¹ Para mais detalhes, ver: <http://www.nics.unicamp.br/eventos/descobertas-opera-multimodal-em-4-atos-de-jonatas-manzolli/>. Data de acesso: 30/07/2018

² Retirado da nota de programa do concerto de estreia da obra. Disponível em: <http://www.osmc.com.br/novo/noticias/400/autores-e-obras.aspx>. Data de acesso: 30/07/2018.



fala e para a classificação de sonoridades produzidas por instrumentos musicais. As próximas seções do artigo apresentam uma breve contextualização do objeto de pesquisa (Seção 2) e as técnicas utilizadas para a realização das análises musicais (Seção 3). As seções subsequentes descrevem o método de trabalho de pesquisa e os resultados obtidos (Seção 4); as análises e considerações a partir dos resultados (Seção 5); as discussões, conclusões e as projeções para trabalhos futuros (Seção 6).

1.2 – Sobre o Compositor da Obra Analisada

De acordo com Jônatas Manzolli, a área de pesquisa musical encontra-se “diante da possibilidade de utilizar diversos níveis de representação computacional como suporte para análise, composição e performance musical” (Manzolli 2013, p. 48). Inserido neste contexto, tais procedimentos proporcionam novas formas de escritura musical. Como destacado na Seção 1, o compositor e pesquisador Jônatas Manzolli insere-se no panorama da produção de música brasileira contemporânea cujas pesquisas relacionam-se com a interação entre arte e tecnologia em criação musical, computação musical e ciências cognitivas. Manzolli utiliza os recursos computacionais de análise e criação centrados nos métodos de recuperação da informação musical e, em especial, em descritores de áudio. Na Seção 3 apresentamos a revisão da literatura da pesquisa relacionada com os descritores de áudio.

No contexto particular da nossa pesquisa relacionada com orquestração, Jônatas Manzolli identifica que o som de uma orquestra sinfônica chama a atenção pela sua complexidade, equilíbrio e pelo efeito estético percebido pelos nossos sentidos. Manzolli relaciona a interação dos instrumentos sinfônicos com a percepção, envolvendo fatores de três naturezas diferentes: estética, sensorial e física (Manzolli 1988, p. iii). Para Manzolli, a questão da orquestração centra-se na criação de blocos timbrísticos que, ao serem contrapostos, caracterizam de maneira única sua linguagem sinfônica. A obra orquestral ou orquestrada, constitui-se de trechos dos quais descrevem características do seu compositor (Manzolli 1988, p. iv). De forma análoga, Manzolli elucida o procedimento de orquestração por intermédio de um exemplo emprestado das artes plásticas e, em especial, por intermédio do ofício do pintor de uma aquarela. Como apresentado nos parágrafos anteriores, uma grande parte das obras composicionais de Jônatas Manzolli centra-se na interação entre tecnologia e escritura instrumental. A produção do compositor envolve obras tanto para instrumentos solo [*Sufrágio* para piano e audiovisual (2014); *Escriba* para texto, piano e audiovisual interativo (2014)], música de câmara [*Transeuntes* para soprano, grupo de câmara e vídeo (2013); *Curto Circuito II* para cordas, percussão



e imagens (2011)], e ainda orquestras de câmara e sinfônica [*Reação em Cadeia* para orquestra de violoncelos (2014); *Reflexões* para orquestra sinfônica, tape e imagens (2011)]. Para o nosso trabalho de análise musical, utilizamos *Reflexões* por tratar-se de uma obra cujo discurso sonoro relaciona-se com variações de sonoridades, intensidades e timbres orquestrais. As texturas sonoras, estabelecidas em *Reflexões* relacionam-se com as “torrentes de ideias que germinam e povoam os processos de descoberta e transformação”³. A Seção 2.1 apresenta um recorte de análise da obra, na qual utilizamos o suporte dos descritores de áudio.

2 - Contexto e Ponto de Vista em *Reflexões*

A obra *Reflexões* para orquestra, tape e imagens (2011) utiliza um planejamento composicional cujo objetivo centra-se na construção de sonoridades a partir da relação das sonoridades instrumentais e orquestrais com os sons eletrônicos. A obra concatena momentos de variação em densidade sonora criados com princípios de síntese sonora aplicada à escritura sinfônica e à geração de sons com suporte computacional. A estreia da obra foi realizada pela Orquestra Sinfônica Municipal de Campinas, no dia 30 de outubro de 2011, sob a regência de Christian Karlsen. O projeto poético da obra dialoga com a sua data da estreia, 30 de outubro, pois, justamente se aproxima da data da celebração do dia 31 de outubro de 1517, a qual atribui-se o momento no qual as 95 teses foram afixadas em Wittenberg. A parte da eletrônica e a concepção estrutural da obra contou com a colaboração do Prof. Adolfo Maia cuja difusão foi realizada pelo Prof. José Augusto Mannis⁴. Os sons eletrônicos foram processados a partir de registros de sinos e de murmúrios, ressonantes no interior de uma catedral, os quais foram ‘pulverizados’ pelos naipes da orquestra. A gravação sonora que utilizamos para realizar a nossa análise foi obtida pela gravação ao vivo do concerto, cujo parâmetros são os seguintes: estéreo, 32 Bits e 44100 Hz. A obra estruturou-se em seis movimentos breves: *Prólogo*, *Fragmentos*, *Corais*, *Ressonâncias*, *Ícones* e *Epílogo*. O compositor organizou as diferentes sonoridades a partir de 95 entradas do tímpano, cada qual estabelecendo uma construção sonora distinta. A quantidade de entradas do tímpano faz alusão às 95 teses afixadas na porta da Igreja do Castelo de Wittenberg⁵. Dentre todos os

³ Retirado da nota de programa do concerto de estreia da obra. Disponível em: <http://www.osmc.com.br/novo/noticias/400/autores-e-obras.aspx>. Data de acesso: 30/07/2018

⁴ Retirado da nota de programa do concerto de estreia da obra. Disponível em: <http://www.osmc.com.br/novo/noticias/400/autores-e-obras.aspx>. Data de acesso: 30/07/2018

⁵ Extraímos as informações relacionadas com o contexto poético da obra de entrevistas realizadas com o próprio compositor. No que se refere às informações relativas às 95 Teses de Lutero, ver <https://pt.wikipedia.org/wiki/95#Teses>. Data de acesso: 30/07/2018.



movimentos presentes em *Reflexões*, Prólogo é o único do qual não consta a presença do tímpano. A Tabela 1 ilustra a quantidade de entradas do instrumento que apresentam sonoridades distintas entre si.

Movimento (duração)	Quantidade entradas de tímpanos
Prólogo (01' 40'')	0 Tímpano
Fragmentos (02'47'')	19 Tímpanos
Corais (02'26'')	20 Tímpanos
Ressonâncias (01'49'')	18 Tímpanos
Ícones (02'06'')	19 Tímpanos
Epílogo (01'58'')	19 Tímpanos

Tabela 1: Quantidade de entradas do tímpano

O fato do compositor ter explicitado os 95 pontos como fator de mudança na estrutura da obra, já sugeriu uma segmentação para a análise. Assim o trabalho de análise de *Reflexões*, com o suporte de descritores de áudio, centrou-se em identificar as principais características espectrais das sonoridades estabelecidas entre as 95 entradas do tímpano. Por não utilizar tal instrumento, a nossa proposta de análise não levará em consideração o primeiro movimento, *Prólogo*, etapa que abordaremos em trabalhos futuros.

3 - Ferramentas de Análise Computacional: os Descritores de Áudio

Nesta Seção 3, discutiremos o uso dos descritores de áudio como ferramenta de análise sonora. Apresentaremos os descritores de áudio que dialogam diretamente com o nosso estudo. Como apresentado na Seção 1, a área de conhecimento científico sobre o assunto denomina-se *Music Information Retrieval* ou simplesmente, MIR. Os estudos sobre MIR utilizam funções matemáticas, apoiadas em medidas estatísticas e modelos psicoacústicos para desenvolver os chamados descritores de áudio. Segundo Peeters (2004, p. 01), os procedimentos para descrever as características de um sinal sonoro foram propostos pela comunidade científica para o reconhecimento de padrões da fala e para a classificação de instrumentos musicais. Esses procedimentos também são ferramentas de relevância significativa no contexto da composição e orquestração musical assistidas por computador. Os métodos para analisar o conteúdo espectral, a partir de sinais de áudio digitalizados, são realizados via *Short-Time Fourier Transform* ou STFT, que é definida matematicamente por Sheh e Ellis (2003, p. 02) como a seguir:



$$STFT_{[k,n]} = \sum_{m=0}^{N-1} x[n-m]w[m]e^{-j2\pi km/n} \quad (\text{equação 1})$$

onde k indexa o eixo de frequência (*bin*) dentro do intervalo $0 < k < N-1$ do n ésimo quadro e $w[m]$ é uma janela de N amostras.

Do ponto de vista generalizado, os descritores de áudio são ferramentas de análise sonora que representam determinadas características unidimensionais do sinal sonoro. Como apontado por Rimoldi (2013, p. 1), embora tenham caráter reducionista em relação objeto analisado, os descritores de áudio são ferramentas úteis para uma taxonomia de características relacionadas com o conteúdo espectral do sinal sonoro analisado. Tais características podem ser correlacionadas com – mas não necessariamente equivalentes a – atributos subjetivos da percepção do sinal sonoro, tais como “brilho”, “opacidade”, “rugosidade”, “ruidosidade”, “maciez”, dentre outros. Nas subseções seguintes, apresentaremos um conjunto de descritores de áudio e as suas relações frente ao nosso escopo de trabalho, ou seja, os critérios com os quais os descritores podem relacionar-se com a análise musical. No total utilizamos quinze descritores de áudio.

3.1 - Centroide Espectral

Segundo Theodoridis e Koutrumbas (2008, p. 460-461), um dos parâmetros mais utilizados para a extração de características do sinal sonoro é o Centroide Espectral, que será denominado apenas de centroide no restante do nosso trabalho. Tal como apontado por Pereira (2009, p. 16), por intermédio do centroide é possível medir a percepção de brilho do sinal sonoro analisado. Analogamente, tal descritor pode ser definido como o centro de massa ou baricentro do espectro analisado. Dessa forma, quanto mais alto o valor do centroide, dado em Hz, maior será a concentração de energia nas componentes espectrais de altas frequências o que pode levar, em muitos casos, à percepção de um som mais brilhante. Cabe ressaltar que o registro e a altura musical não se relacionam diretamente com o centroide. Por exemplo, uma nota tocada no grave, pode apresentar um valor de centroide em frequências altas. Isto ocorrerá se as magnitudes de seus componentes espectrais de alta frequência tiverem mais energia que os de baixa. O uso de técnicas estendidas pode produzir transientes que elevam o valor da energia nas componentes de alta frequência. Analisamos duas sonoridades com características musicais distintas do ponto de vista do centroide espectral. Utilizamos o contrabaixo, executando a altura musical $Mi3$, em *mezzoforte* e com a técnica de arco denominada arco *sul ponticello*. A outra sonoridade utilizada é a da viola com a mesma altura $Mi3$, mesma dinâmica



mezzoforte mas com a utilização de *sordina*, em arco *ordinario*. A Figura 1 ilustra os contornos da análise das suas sonoridades e a média, em Hz, para cada uma delas. A Figura 1a, que representa a análise sonora do contrabaixo, aponta um valor maior de média em comparação com o valor da média da Figura 1b, que descreve o resultado da análise da viola. A análise dessas duas sonoridades, a partir do cálculo do centroide espectral, aponta que o contrabaixo obteve um valor médio de centroide mais alto que a sonoridade da viola.

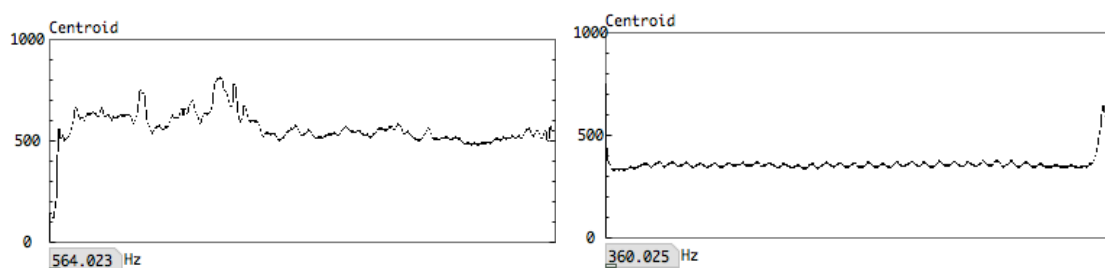


Figura 1: Análise do Centroide Espectral de dois instrumentos executando a mesma altura, com a mesma dinâmica
 - à esquerda, análise espectral do contrabaixo
 - à direita, análise espectral da viola

3.2 - Desvio Padrão Espectral

O Desvio Padrão Espectral representa o cálculo da raiz quadrada da variância do espectro de frequências sonoras, analisadas a partir do centroide espectral. Consideramos o desvio padrão espectral como uma medida calculada a partir do momento central da segunda ordem da distribuição do cálculo do centroide. Ela mede a espalhamento da distribuição de energia espectral, ou seja, a média de desvio padrão em relação ao centroide e que pode ser entendido como uma estimativa média da largura de banda espectral (Monteiro 2012, p. 125). Analisamos duas sonoridades distintas para demonstrar o uso da medida do desvio padrão espectral, na qual a) um conjunto de instrumentos de cordas, executando a nota Lá4, em *ordinario* e b) uma viola solo executando a mesma nota Lá4, em *ordinario*. A sonoridade a), formada pelos instrumentos contrabaixo, violoncelo, viola, violino II e violino I apresenta um valor médio maior de desvio padrão espectral que a viola solo. O descritor identificou que o conjunto instrumental analisado possui uma largura de banda maior que o instrumento solo. A Figura 2, ilustra a análise realizada com as duas sonoridades e os seus respectivos valores de média. Já Mikhail Malt e Emmanuele Jourdan utilizam a medida de desvio padrão espectral para identificar eventos musicais discretos ou técnicas de execução instrumental (Malt and Jourdan 2009). O desvio padrão



espectral é utilizado ainda para identificar eventos ruidosos inseridos em um contexto musical no qual há diferentes materiais sonoros misturados. Em sua análise, a técnica instrumental de *key clicks*, na flauta, pode ser facilmente discriminado de um contexto sonoro mais harmônico.

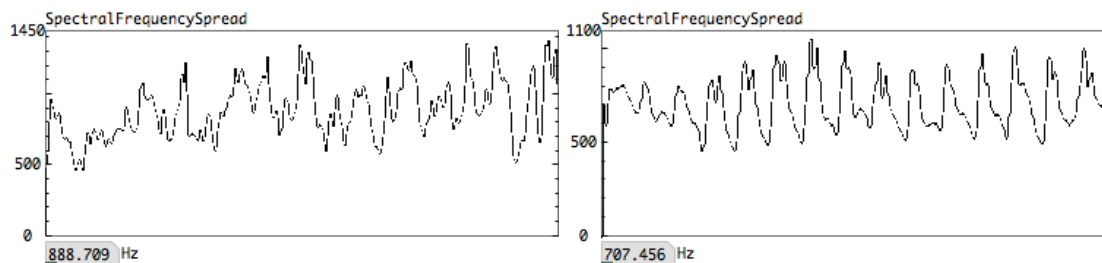


Figura 2: Análise do desvio padrão espectral de duas sonoridades executando a mesma altura e a mesma dinâmica
 - à esquerda, análise da orquestra de cordas
 - à direita, análise da viola

3.3 - *Skewness* ou Obliquidade Espectral

A medida de *Skewness* espectral, também conhecida como obliquidade espectral calcula a assimetria da largura de banda espectral com relação ao centroide. É computada a partir do momento central de terceira ordem da distribuição (Rimoldi 2013, p. 100). Quando o resultado da medida for igual a zero, a *skewness* indica uma distribuição simétrica em relação ao pico e à média da largura banda de frequência, aproximando-se de uma distribuição gaussiana. Valores maiores que zero indicam assimetria para a direita, com maior desvio em relação à média das frequências agudas. Já os valores menores que zero indicam assimetria para a esquerda, com maior desvio em relação à média das frequências graves (Esling and Agon 2013), (Rimoldi 2013). Na Figura 3, ilustramos dois momentos de análise da obliquidade espectral. À esquerda, a ilustração indica obliquidade negativa, no qual a curva do vale esquerdo é mais alongada. Inversamente, à direita, a curva do vale direito da distribuição é mais alongada, indicando uma obliquidade positiva.

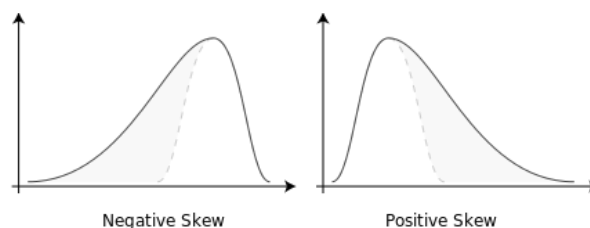


Figura 3: Dois momentos de análise a partir da obliquidade espectral



No seu trabalho de pesquisa, William Brent analisa sonoridades de instrumentos de percussão a partir do uso das ferramentas estatísticas dos descritores de áudio. Brent utiliza a medida de obliquidade para analisar as sonoridades do bongô, que resulta em um alto índice positivo de obliquidade e de instrumentos metálicos de percussão, tal como os gongos, que resultam em índices positivos consideravelmente baixos, comparando-se com os bongôs (Brent 2010, p. 39). Por tratar-se de apresentar uma maior predominância em frequências graves, a obliquidade positiva é associada com a percepção subjetiva de sonoridades arredondadas e quentes (Brent 2010, p. 130). Inversamente, a obliquidade negativa, por indicar uma predominância em frequência agudas, relaciona-se com a percepção de sonoridades brilhantes, a partir do cálculo de centroide (Brent 2010, p. 131).

3.4 - Curtose Espectral

O descritor de Curtose Espectral é calculado pela razão entre o quarto momento central e o quadrado da variância da magnitude. É uma medida de análise do formato da largura de banda relacionada com os picos e os vales da distribuição estatística. A curtose espectral é considerada como uma ferramenta estatística que pode indicar a presença de séries de transientes e a suas posições no domínio das frequências sonoras (Antoni 2006). Valores altos de curtose indicam uma distribuição com picos, caracterizando a presença de valores que se aproximam da média do desvio padrão e se relaciona com sonoridades sinusoidais. Inversamente, valores baixos de curtose indicam uma distribuição da energia mais achatada ou plana do que a gaussiana e são relacionadas com sonoridades mais ruidosas. Contudo, segundo Lawrence DeCarlo, há interpretações equivocadas junto à análise estatística a partir da curtose, onde não se consideram a presença e os contornos dos vales de uma distribuição (DeCarlo 1997). Analisamos dois pares de sonoridades distintas a partir da curtose espectral. Nosso objetivo foi ressaltar as diferenças entre as sonoridades com baixa quantidade de ruído sonoro instrumental e as sonoridades mais ruidosas. Utilizamos os instrumentos de flauta e violino. Na Figura 4a, analisamos a flauta executando a técnica de *whistle sound*, que apresenta sonoridade mais limpa e pura. Na Figura 4b, a flauta executa a técnica de *jet whistle*, que resulta em grande quantidade de ruído sonoro. A partir dos valores da média da análise, ilustrados em cada figura, a sonoridade em *whistle tone* possui maior índice de curtose, indicando sonoridades mais tonais.



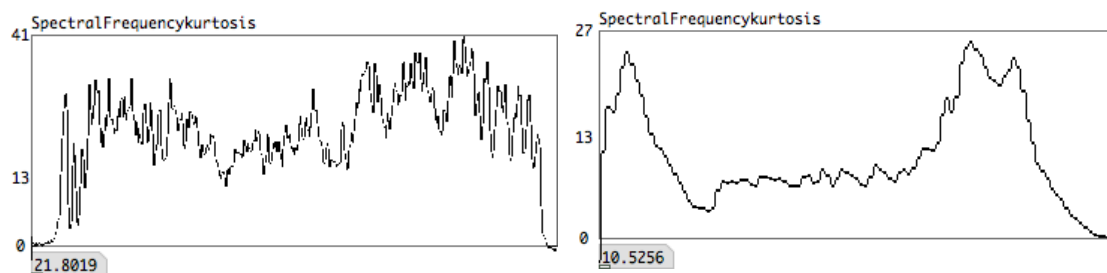


Figura 4: Análise da curtose espectral da flauta executando *whistle tone*, à esquerda, e *jet whistle*, à direita

A segunda análise que realizamos centra-se em duas sonoridades do violino, na qual destacamos que os harmônicos artificiais representam alto índice de curtose em relação à técnica de arco tocado atrás do espelho do violino. A Figura 5a representa o contorno e o valor da média da análise da técnica de harmônico artificial. Já a Figura 5b ilustra distribuição da análise e o valor da média da técnica de arco tocado atrás do espelho.

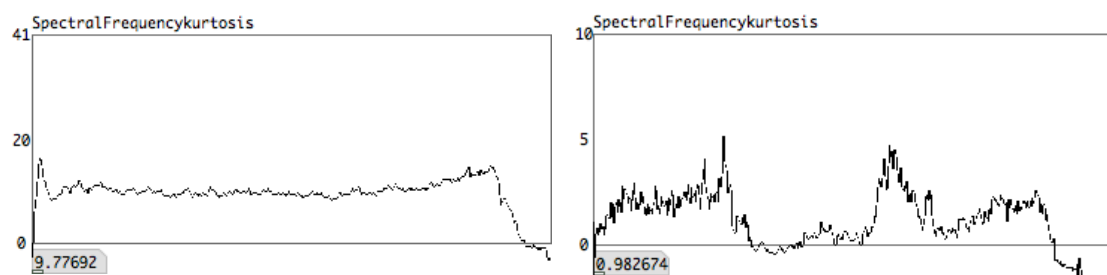


Figure 5: Análise da curtose espectral de duas sonoridades do violino executando harmônico artificial, à esquerda, e com o arco atrás do espelho, à direita

3.5 - Fluxo Espectral

O descritor de Fluxo Espectral é uma medida de análise temporal do espectro de frequências sonoras. O objetivo desse parâmetro é caracterizar quão rápidas são as mudanças no espectro do sinal ao longo dos quadros (Pereira 2009). O descritor, também denominado como ‘variação temporal de espectro’ é utilizado para identificar os tempos de ataques, i.e. *onsets* e representa a quantidade de variação do espectro no decorrer do tempo. Se próximo a zero, os espectros sucessivos são similares. Perto de um, os espectros sucessivos são dissimilares (Peeters 2004). A Figura 6 ilustra a análise da justaposição de um conjunto de sonoridades, que resultam de experimentos composicionais junto à obra *O azeite, a lua e o rio* (2014). O descritor de fluxo espectral resultou em altos



valores de índice nos momentos de ‘entradas’ das sonoridades, contrastando-as com os intervalos de ‘silêncio’ entre si.

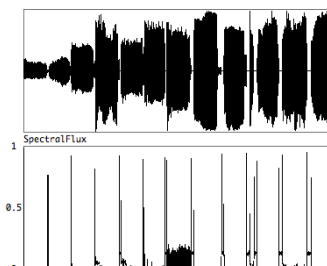


Figura 6: Análise do conjunto de sonoridades a partir do cálculo de fluxo espectral

O descritor de fluxo espectral é utilizado ainda para identificar sonoridades que possuem alto índice de variação de articulação. Analisamos a técnica instrumental de *tremolo* na viola. A Figura 7 representa a análise do fluxo espectral da técnica de arco *ordinario con poco vibrato* e, posteriormente, a técnica de *tremolo*. A análise do descritor destaca o momento da mudança de técnica instrumental, na qual o *tremolo* resulta em uma maior variação de fluxo espectral.

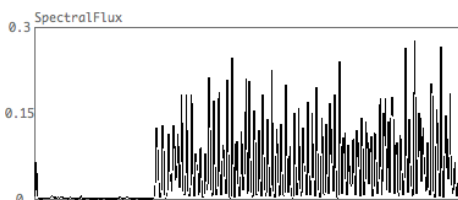


Figura 7: Análise sonora do fluxo espectral da viola
(arco *ordinario con poco vibrato* e *tremolo*)

3.6 - Nivelamento Espectral

Nivelamento Espectral ou *Spectral Flatness* define-se por se tratar de uma medida da relação ruidosidade/som tonal de um espectro. De acordo com Dubnov, o índice de tonalidade de um sinal caracteriza-se pela sua quantidade de picos espectrais (Dubnov 2004, p. 1). O nivelamento espectral é calculado pela razão da média geométrica com a média aritmética da energia do valor espectral (Peeters 2004) e (Jehan 2005). Quando normalizados, valores próximos de um indicam uma distribuição de energia homogênea, correlata a espectros mais ruidosos e sem formas definidas. Quando próxima a zero, a distribuição da energia no espectro de frequência é heterogênea, indicando a presença de picos de magnitude. Tal configuração correlaciona-se com sonoridades tonais ou senoidais (Rimoldi 2013, p. 99). Utilizamos o descritor de nivelamento para



analisar duas sonoridades distintas da flauta: a) executando a técnica de *jet whistle* e b) executando a nota Dó4, em *ordinario*. A Figura 8 ilustra os contornos dos dados de análise e os seus respectivos valores de média. A Figura 8a, que representa a flauta em *jet whistle*, possui um valor médio de nivelamento muito maior que a flauta em *ordinario*, como ilustra a Figura 8b.

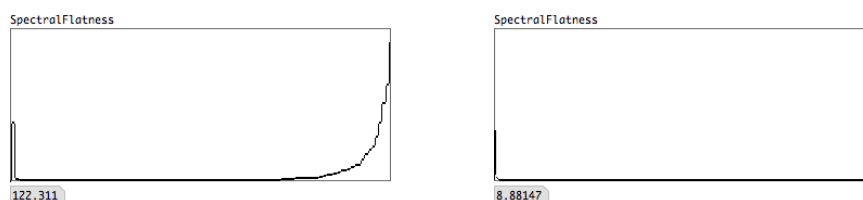


Figura 8: Análise da média de nivelamento espectral de duas técnicas instrumentais executadas pela flauta
 - à esquerda, figura 8a: análise da flauta em *jet whistle*
 - à direita, figura 8b: análise da flauta em *ordinario*

O nivelamento espectral é utilizado, ainda, para identificar e segmentar ‘momentos’ de silêncio daqueles que possuem informações sonoras musicais. A Figura 9 representa a análise do violoncelo tocando em *pizzicato Bartók*. No momento da execução da sonoridade, o valor de nivelamento é baixo, em comparação com o decaimento da sonoridade e o aumento do ruído.

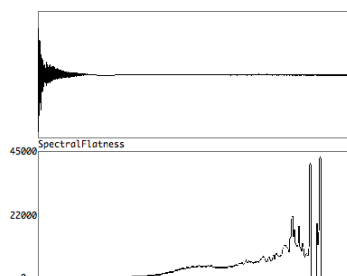


Figura 9: Análise do Nivelamento Espectral do violoncelo executando a técnica de *pizzicato Bartók*

3.7 - Irregularidade Espectral

Como apontado por Krimphoff et al. (1994), a Irregularidade Espectral apresenta-se como um atributo perceptivo fundamental para identificar sonoridades, separar e distinguir instrumentos musicais. Essa medida relaciona-se com o comportamento do envelope espectral que é uma curva que caracteriza a variação das magnitudes das componentes de frequência. Contornos suaves, com pouca alternância de picos espectrais e, conseqüentemente, com baixo índice



de irregularidade, caracterizam sonoridades aveludadas e lisas. Inversamente, contornos com maior presença de picos espectrais e com alto índice de irregularidade espectral, caracterizam-se por sonoridades rugosas e estriadas. Para detalhar as considerações expostas no parágrafo anterior, a Figura 10 apresenta dois instrumentos distintos, clarinete e violoncelo, executando a mesma nota Lá#, com a dinâmica musical *forte*. Nota-se que o clarinete possui o valor médio de irregularidade espectral maior (Figura 10-esquerda) que o violoncelo, o qual possui comportamento mais difuso, mas com valor médio menor de irregularidade espectral (Figura 10-direita). O clarinete é mais rugoso e estriado que o violino, que é mais aveludado e suave.

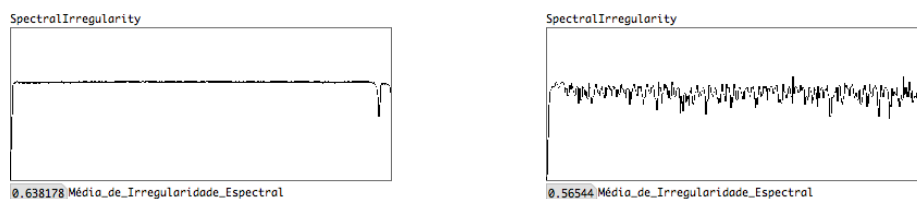


Figura 10: Análise da média de Irregularidade Espectral de dois instrumentos executando a mesma altura e a mesma dinâmica musical
 - à esquerda: análise do clarinete em Sib
 - à direita: análise do violoncelo

O maior valor da média de irregularidade espectral, extraído do clarinete, pode ser justificado pelo comportamento das componentes espectrais, o qual constitui o seu envelope espectral. Pelas características acústica e física do instrumento, o clarinete apresenta maior energia entre os seus componentes ímpares. A Figura 11 apresenta o comportamento dos picos espectrais do instrumento.

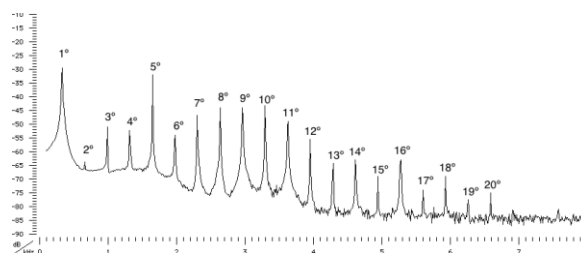


Figura 11: Envelope Espectral do clarinet

Há trabalhos que utilizaram o descritor de irregularidade espectral como ferramenta de análise da composição musical. Luka Mikula desenvolveu uma interface de controle de síntese sonora utilizando, dentre outros descritores, a irregularidade espectral (Mikula 2008). Bullock compôs uma obra para piano e



eletrônica, em tempo real, na qual as manipulações e os processamentos sonoros eram extraídos pelos descritores (Bullock, 2007).

3.8 - Roll-Off Espectral

O descritor de *Roll-off* está correlacionado com o ponto de declividade que antecede a frequência de corte ou o pico de frequência com a maior amplitude do Envelope Espectral. Por Envelope Espectral, entende-se uma persistência nos padrões de energia e de quantidade de componentes espectrais de um sinal sonoro analisado. A Figura 12, representa a análise do envelope espectral do trompete em duas intensidades: pp e ff (Manzolli 1988, p. 47). A terceira curva representa a diferença entre as frequências analisadas das duas sonoridades.

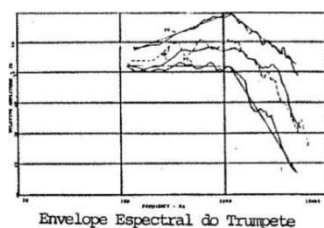


Figura 12: Análise do Envelope Espectral de duas sonoridades do trompete, em intensidade pp e ff

- o eixo y representa a intensidade sonora, mensurada em dB
- o eixo x representa a escala de frequência sonoras, em Hz

O trompete quando tocado em pp, curva inferior da Figura 12, exibe um *Roll-Off* próximo a zero, com declividade da reta à esquerda. De acordo com Manzolli, tal característica descreve um timbre no qual sua fundamental é predominante e, com decaimento rápido de seus componentes mais agudos, produz uma sonoridade menos penetrante e mais agradável. De modo contrário, para a sonoridade em ff, o ponto de declividade do *Roll-Off* é semelhante à declividade à direita da frequência de corte. A esta característica, Manzolli associa a percepção de sonoridades mais penetrantes (Manzolli 1988, p. 49). Contudo, de acordo com Pereira, os sinais musicais tendem apresentar valores maiores para este descritor que os sinais de voz. Isto porque os sons percussivos, os transientes de ataque, etc, tem mais energia nas altas frequências que os sinais de voz (Pereira 2009). Analisamos a mesma técnica instrumental de *pizzicato Bartók*, no contrabaixo, em duas cordas diferentes. Nosso objetivo foi identificar, a partir do *roll-off*, qual a frequência sonora de corte que representa 85% de toda a energia espectral analisada. As Figuras 13 ilustram o contorno da análise do *pizzicato Bartók* executado nas cordas três e quatro. O procedimento de análise demonstrou que a sonoridade realizada na corda 3, mais aguda que a corda 4,



possui uma frequência *roll-off* aproximada a 585Hz. Já a média de *roll-off* da corda 4 possui um valor de frequência menor.

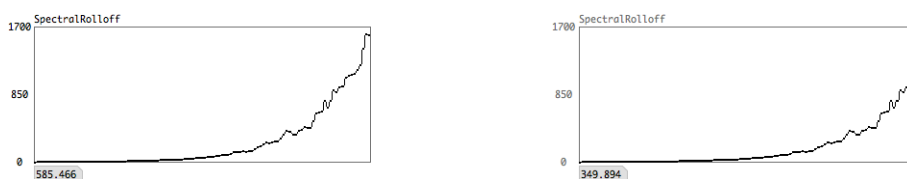


Figura 13: Análise da média da frequência *roll-off* de *pizzicato* Bartók, no contrabaixo, executado em duas cordas diferentes
 - a figura à esquerda ilustra a análise da corda 3
 - a figura à direita representa a análise da corda 4

3.9 - Razão entre Componentes Ímpares e Pares

Este descritor analisa a Razão entre os Componentes Espectrais Ímpares e Pares, que é utilizado para identificar sonoridades distintas, tais como o clarinete, com maior predominância de componentes ímpares. Índices maiores que um indicam a predominância de energia nos componentes ímpares. Inversamente, os índices menores que um indicam a predominância de energia em componentes pares. Utilizamos o cálculo da razão entre componentes ímpares e pares para analisar duas sonoridades distintas: a) clarinete, em Sib, tocando em *ordinario* e *poco vibrato*; b) fagote tocando em *ordinario* e *poco vibrato*. O clarinete, como esperado, resulta em valores altos junto à análise do descritor. Tal resultado relaciona-se com as características acústicas e espectrais do instrumento, ressaltando os componentes ímpares. A análise do fagote, em contrapartida, retorna valores baixos, próximos ou menores que um. O resultado indica que a sonoridade do fagote, em *ordinario* e *poco vibrato*, possui grande quantidade de energia em seus componentes pares. A Figura 14 ilustra a análise das suas sonoridades, na qual a Figura 14a representa a variação do contorno da análise da razão dos componentes ímpares e pares do clarinete. A Figura 14b representa a variação de índices baixos da análise do fagote.

3.10 - Energia RMS

O descritor cujo acrônimo RMS equivale a *Root Mean Square* ou raiz da média dos quadrados, calcula a potência média de curta duração do sinal analisado (Loureiro et al. 2008, p. 113). Correlaciona-se o RMS com o descritor de *loudness*, que será descrito na subseção 3.11. Segundo Tzanetakis, o RMS é muito utilizado no processo de segmentação sonora, já que as mudanças de amplitudes



são indicativos consideráveis para a identificar novos eventos (Tzanetakis & Cook 1999, p. 103-106).

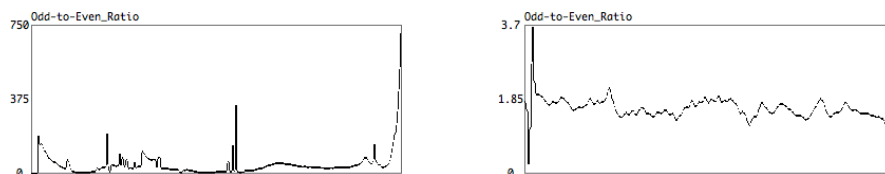


Figura 14: Análise da razão entre componentes espectrais ímpares e pares do clarinete, à esquerda, e do fagote, à direita, executando técnica instrumental em *ordinario* e *poco vibrato*

3.11 - Loudness

Segundo Fletcher e Munson (1933, p. 377), *Loudness* é um termo que descreve a magnitude da sensação auditiva de variação de intensidade sonora. Os atributos subjetivos para medir a intensidade sonora, tais como “muito forte”, “forte”, “moderadamente forte”, “moderadamente baixo”, “baixo” e “muito baixo” são correlacionados com termos musicais como fortissimo, forte, mezzoforte, mezzopiano, piano e pianíssimo. Mas esses termos não são absolutos nem tampouco unívocos para mensurar a capacidade de percepção humana de intensidade sonora. Diversos estudos propuseram modelos para descrever o comportamento do sistema auditivo com relação ao *Loudness* (Zwicker & Scharf 1965), (Florentine et al. 1978), (Miskiewicz & Rakowski 1994), (Thiede 1999) e (Plomp 2001). Os modelos da percepção de *Loudness*, foram utilizados em trabalhos de musicologia e análise, como Collins (2005), Einbond (2005) e Bullock (2007). A percepção de intensidade sonora varia de acordo com a frequência, ou seja, a percepção de intensidade varia com o registro. Esse comportamento psicoacústico é descrito pelas curvas de audibilidade, que são denominadas de “Curvas de Iso-Loudness” (Fletcher & Munson 1933). Dentro do âmbito da nossa pesquisa, utilizamos como referência o trabalho de Thiede (1999). O pesquisador destaca que a percepção sonora pode ser descrita a partir de cinco etapas, distribuídas em dois módulos distintos. O primeiro modelo é relacionado às estruturas fisiológicas do ouvido humano, nas quais as informações sonoras, do meio externo, são convertidas em impulsos elétricos, para comunicação neural. O segundo é relacionado às propriedades cognitivas que realizam o reconhecimento de diversos padrões para recuperar a informação sonora. A Figura 15 sintetiza as etapas do processamento auditivo, discutidas pelo pesquisador.



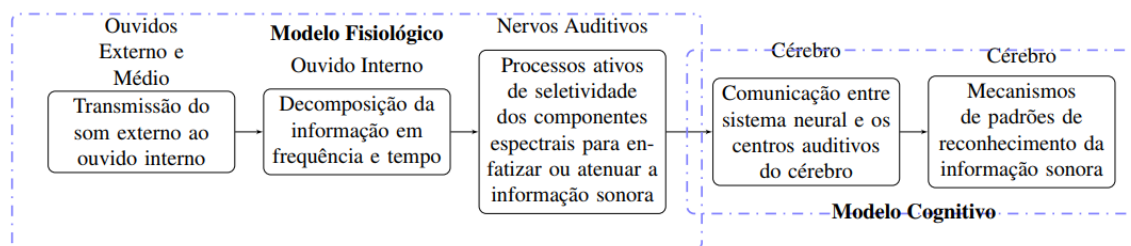


Figura 15: Síntese das etapas do processamento auditivo

3.12 - Taxa de Cruzamento por Zero

O descritor que calcula a Taxa de Cruzamento por Zero é uma medida que se relaciona com o ruído sonoro. Segundo Monteiro, para sinais simples, como ondas sinusoidais, o cálculo da taxa de cruzamento por zero é útil para identificar altura musical pois correlaciona-se com a frequência fundamental do sinal analisado. Aplicada em sinais complexos, o descritor fornece indicações sobre a ruidosidade do sinal (Monteiro 2012, p. 121). Os sinais de música apresentam uma taxa de cruzamentos na qual normalmente os valores não variam muito ao longo do sinal. No caso de sinais de voz, por exemplo, os valores variam intensamente de quadro para quadro de análise (Pereira 2009, p. 18). O descritor é utilizado para segmentar sinais sonoros e silêncio nos excertos analisados. Utilizamos três sonoridades para ressaltar os contornos de análise de três técnicas, em três diferentes instrumentos. Na Figura 16, à esquerda, a flauta executa a transição da técnica de *ordinario* para sons eólicos. O descritor de taxa de cruzamento por zero retorna maior variação nas sonoridades em sons eólicos, o que se caracteriza pela considerável presença de ruído sonoro. Na Figura 16, centro, analisamos o trompete executando a transição entre *ordinario* e *brassy* ou uma sonoridade mais metálica. Na nossa análise, o descritor não identificou nuances de ruído sonoro, demonstrando que o descritor não é associado com as mudanças de timbre ou de variação da magnitude dos componentes espectrais de um sinal sonoro. Por fim, a Figura 16, à direita, ilustra a análise da transição da técnica de *ordinario* para arco *overpressure*, no violoncelo. Apesar de maior variação quando o instrumento executa arco *overpressure*, a variação não é muito acentuada.



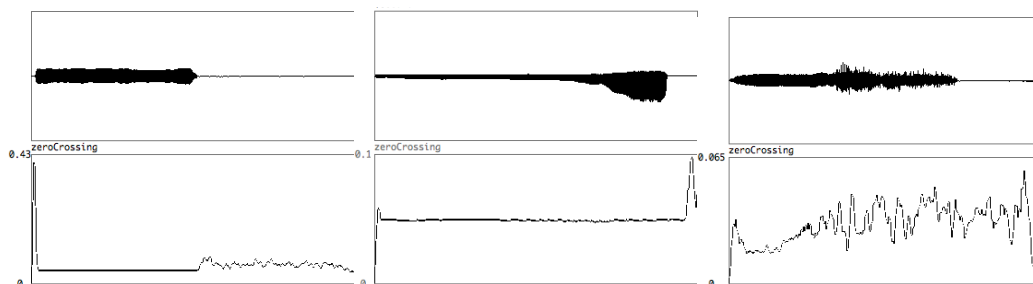


Figura 16: Análise da taxa de cruzamento por zero da flauta, à esquerda, do trompete, ao centro, e do violoncelo, à direita.

3.13 - Decrescimento Espectral

O Decrescimento Espectral é uma medida que representa a taxa decaimento da energia dos componentes espectrais de um sinal sonoro. O descritor de áudio de decrescimento espectral é relacionado por Malt e Jourdan (2008) como uma medida para analisar sons percussivos e sons sustentados. A energia dos componentes espectrais dos sons percussivos tende a decair rapidamente em relação a energia dos componentes espectrais de sons sustentados. Analisamos duas sonoridades distintas a partir do uso do descritor de decrescimento espectral, onde a) o contrabaixo com a técnica de *pizzicato Bartók*, caracterizando uma sonoridade percussiva do instrumento de corda e b) uma orquestração utilizando contrabaixo em *pizzicato Bartók* e um conjunto instrumental tocando notas sustentadas. A Figura 17, à esquerda, ilustra o contorno do comportamento espectral no qual logo após o ataque do *pizzicato Bartók*, há um decrescimento acentuado na energia dos seus componentes. Na análise da sonoridade orquestral, devido a presença do *pizzicato Bartók*, o momento do ataque ilustra um comportamento semelhante ao da Figura 17, à direita. Contudo, após esse ataque, a decrescimento espectral é estabilizado pelas outras sonoridades mais sustentadas.

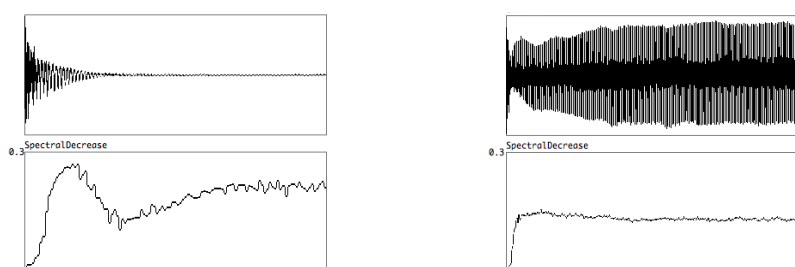


Figura 17: Análise do decrescimento espectral com notas percussivas, à esquerda e notas sustentadas, à direita



3.14 - Centroide Temporal

O Centroide Temporal é associado por Peeters como um importante descritor de análise de sonoridades percussivas ou sustentadas, contudo, a partir da análise temporal da energia do sinal analisado (Peeters et al. 2000). O descritor calcula a média ponderada da energia do sinal no tempo. Sons percussivos tendem a apresentar índices baixos de centroide temporal. Os sons sustentados resultam em índices mais altos. Analisamos duas versões de sonoridades, com a mesma duração, do ponto de vista do centroide temporal. Ambas sonoridades possuem a duração aproximada de 1,46 segundos. Na primeira, orquestramos um conjunto de sonoridades estabelecido por: violino em *pizzicato* seco; oboé e trompete em *crescendo*; trombone em *brassy*; harpa em *glissando*; saxofone alto em *crescendo súbito* e fagote em *molto vibrato*. O descritor calculou o tamanho do arquivo analisado e apontou que o centro de massa da energia da sonoridade encontra-se em aproximadamente 0,76 segundo. A outra sonoridade utilizou os mesmos instrumentos do experimento anterior com a adição do violão, contrabaixo e violoncelo em *pizzicato* e o tam-tam em deixar vibrando. Ao utilizar as técnicas de *pizzicato* e a percussão, o centroide temporal da segunda versão da sonoridade reduziu para aproximadamente 0,25 segundo, demonstrando que a média da energia do sinal encontra-se mais próxima ao ataque do som. A Figura 18 ilustra as análises realizadas para as duas versões de sonoridades.

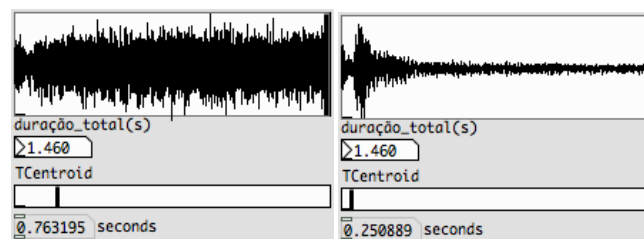


Figura 18: Centroide temporal de duas versões de uma mesma sonoridade em *pizzicato*

3.15 - Croma Espectral

O Croma Espectral, denominado apenas de croma no restante do trabalho, é caracterizado pela distribuição estatística acumulada da magnitude do espectro de Fourier, por bandas fixas de frequência. Essas, referem-se às classes de alturas da escala cromática. Originalmente, o cálculo do Croma foi descrito no trabalho de Fujishima (1999), que o denominou como perfil de classes altura, Pitch Class Profile ou PCP. A Figura 19 apresenta o esquema geral do algoritmo de croma, segundo o diagrama de Fujishima (1999).



A Tabela 2 resume os principais pontos destacados sobre os descritores de áudio. Nela, sintetizamos a lista de descritores com as suas possíveis aplicações.

Descritor	Definição	Correlato Sonoro	Aplicação Possível
Centroide Espectral Seção 3.1	Baricentro Espectral	Brilho / Opacidade	Deteção de variação sonora
Desvio Padrão Espectral Subseção 3.2	Largura de banda espectral do Centroide Espectral	Massa Sonora	Equalização de bandas espectrais
<i>Skewness</i> Espectral Subseção 3.3	Assimetria ou Obliquidade da distribuição estatística do espectro	Quentes e Arredondadas / Brillhantes e Penetrantes	Deteção de tipos de instrumentos de percussão
Curtose Espectral Subseção 3.4	Achatamento da distribuição estatística	Ruído	Deteção de transientes
Fluxo Espectral Subseção 3.5	Varição Espectral no decorrer do tempo	Ataque	Detectar tempos de ataque de eventos sonoros
Nivelamento Espectral Subseção 3.6	Razão entre Média Geométrica com Média Aritmética Espectral	Ruído/Som Tonal	Remoção de ruído sonoro
Irregularidade Espectral Subseção 3.7	Diferença da Magnitude Espectral	Aveludado e Liso / Rugoso e Estriado	Equalização de Banda Espectral
<i>Roll-Off</i> Espectral Subseção 3.8	Declividade do Envelope Espectral	Rugosidade	Masterização de voz e música
Razão entre componentes Ímpares e Pares Subseção 3.9	Quociente da Magnitude entre Componentes Espectrais	Nasal / Macio	Deteção de instrumentos musicais
Energia RMS Subseção 3.10	Raiz da média dos quadrados da energia	Forte / fraco	Deteção de Intensidade Sonora
<i>Loudness</i> Subseção 3.11	Sensação auditiva da variação de intensidade sonora	Forte / Fraco	Percepção de Intensidade Sonora
Taxa de Cruzamento por Zero - Subseção 3.12	Varição do Sinal no tempo	Ruído	Deteção de Ruído Sonoro
Decrescimento Espectral Subseção 3.13	Decaimento da Energia Espectral	Percussividade	Deteção de Sons Percussivos
Centroide Temporal Subseção 3.14	Baricentro Temporal	Percussividade	Deteção de Sons Percussivos
Croma Espectral Subseção 3.15	Análise Espectral por classes de alturas musicais	Tonalidade	Identificação de Harmonia

Tabela 2: Tabela com o resumo da apresentação dos descritores de áudio utilizados na Seção 3



Utilizando-se o HPCP é possível configurar outras bandas de frequência que extrapolam a escala cromática tais como quartos de tom e sextos de tom. Para desenvolver a nossa pesquisa em análise utilizamos um conjunto de dezesseis descritores de áudio. Na Seção 3, apresentamos os descritores de áudio como ferramentas potentes para a criação de uma taxonomia de características de conteúdo espectral. Essa taxonomia pode ser correlacionada e não necessariamente equivalente aos atributos subjetivos da percepção sonora.

4 - Procedimento de Análise

Esta Seção 4 apresenta as etapas do processo de análise relacionadas com a motivação estabelecida anteriormente.

A motivação principal do trabalho de análise de *Reflexões* centrou-se em identificar as principais características espectrais das sonoridades estabelecidas entre as 95 entradas do tímpano. A nossa proposta de análise não levará em consideração o primeiro movimento por não utilizar o tímpano. A primeira etapa do procedimento de análise centrou-se em extrair os dados objetivos a partir da gravação sonora da obra. Para tanto, utilizamos um vetor de características, constituído por um conjunto de 11 descritores de áudio: centroide espectral, *skewness*, desvio padrão, curtose, irregularidade, *flatness*, fluxo, *roll-off*, *odd-to-even*, *RMS* e *zero crossing rate*. Todos os dados coletados foram normalizados dentro do intervalo [0, ..., 1]. Para realizar a extração de características, utilizamos o ambiente *Pure Data* com a biblioteca de funções *PDescriptors* (Monteiro 2012). A Figura 21 ilustra as curvas dos dados de análise, extraídos a partir dos descritores de áudio. O eixo “x” representa o domínio do tempo, em cada movimento da obra.

A escolha dos descritores de áudio centrou-se nas características espectrais das sonoridades correlacionadas, mas não necessariamente equivalentes, com atributos perceptivos, tais como “brilho”, “volume”, “opacidade”, “rugosidade”, “ruidosidade”, “maciez” e “intensidade”. Depois da coleta de dados, realizada com o suporte dos descritores de áudio, segmentamos a obra por intermédio de procedimento semelhante utilizado em Simurra et al. (2013): a) consideramos a partitura musical, centrando-se nas entradas do tímpano e; b) desenvolvemos a análise estatística da obra. Na Seção 5, detalharemos as análises realizadas e os resultados obtidos pelas etapas do método descrito anteriormente.



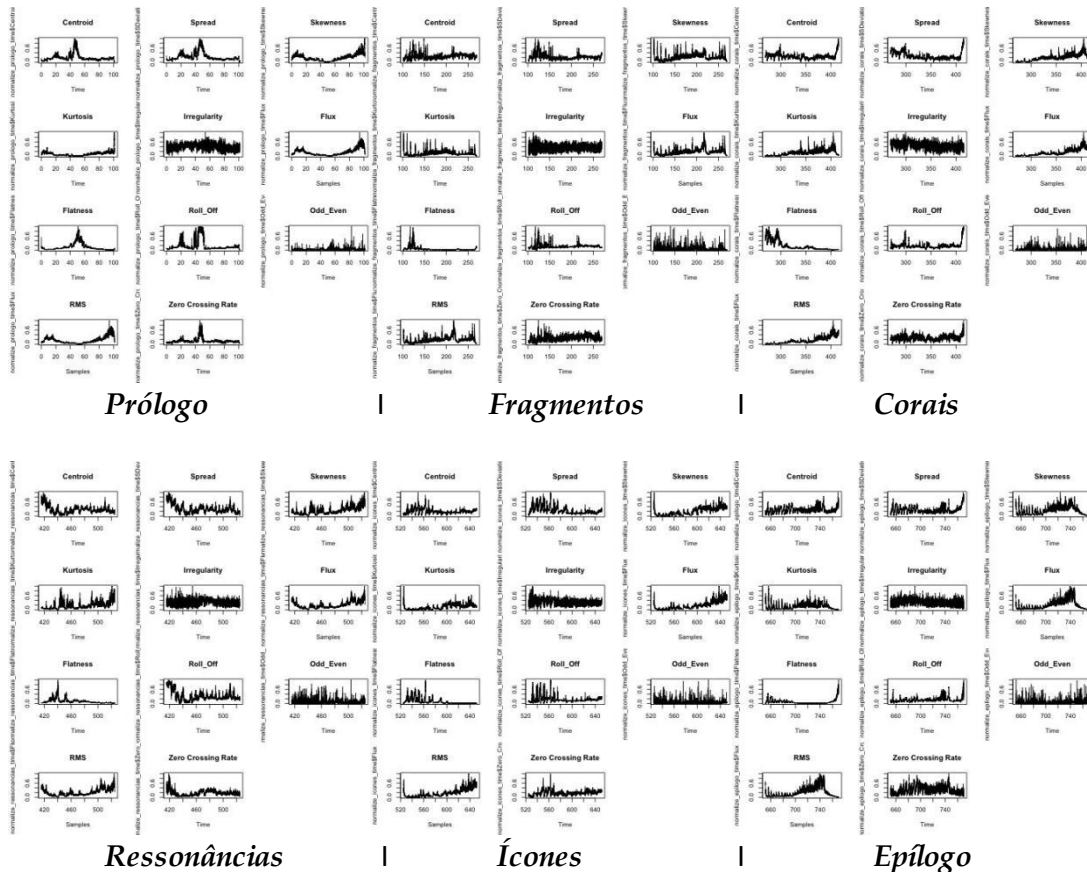


Figura 21: Extração das características, via descritores de áudio, dos seis movimentos de *Reflexões*

5 - Resultados e Análise dos Dados

A análise apresentada nos próximos parágrafos baseia-se em dois pontos principais: a) discutir como a interação dos dados coletados, pelos descritores de áudio relaciona-se com a complexidade planejada; b) analisar o comportamento das sonoridades de cada entrada do tímpano, do ponto de vista dos descritores de áudio. Calculamos a correlação dos descritores de áudio utilizados em cada movimento da obra. A correlação dos dados obtidos pelos descritores de áudio objetiva discutir a correspondência, a força e a direção do relacionamento linear entre as medidas utilizadas. Tal procedimento forneceu resultados para a compreensão global das características sonoras de cada movimento de *Reflexões*. Naturalmente, reduzimos os detalhes e as nuances dos procedimentos composicionais e sonoros estabelecidos pelo compositor. Contudo, ao elaborarmos tal etapa, iniciamos as primeiras conclusões do trabalho composicional. A Figura 22 apresenta a matriz de correlação de cada movimento da obra. Pela convenção da gradação de cores, utilizada para representar o índice



de correlação, quanto mais próximo à cor vermelha maior o índice de correlação negativa. De modo análogo, quanto mais próximo à cor azul maior o índice de correlação positiva. O índice de correlação nula é representado pela cor branca.

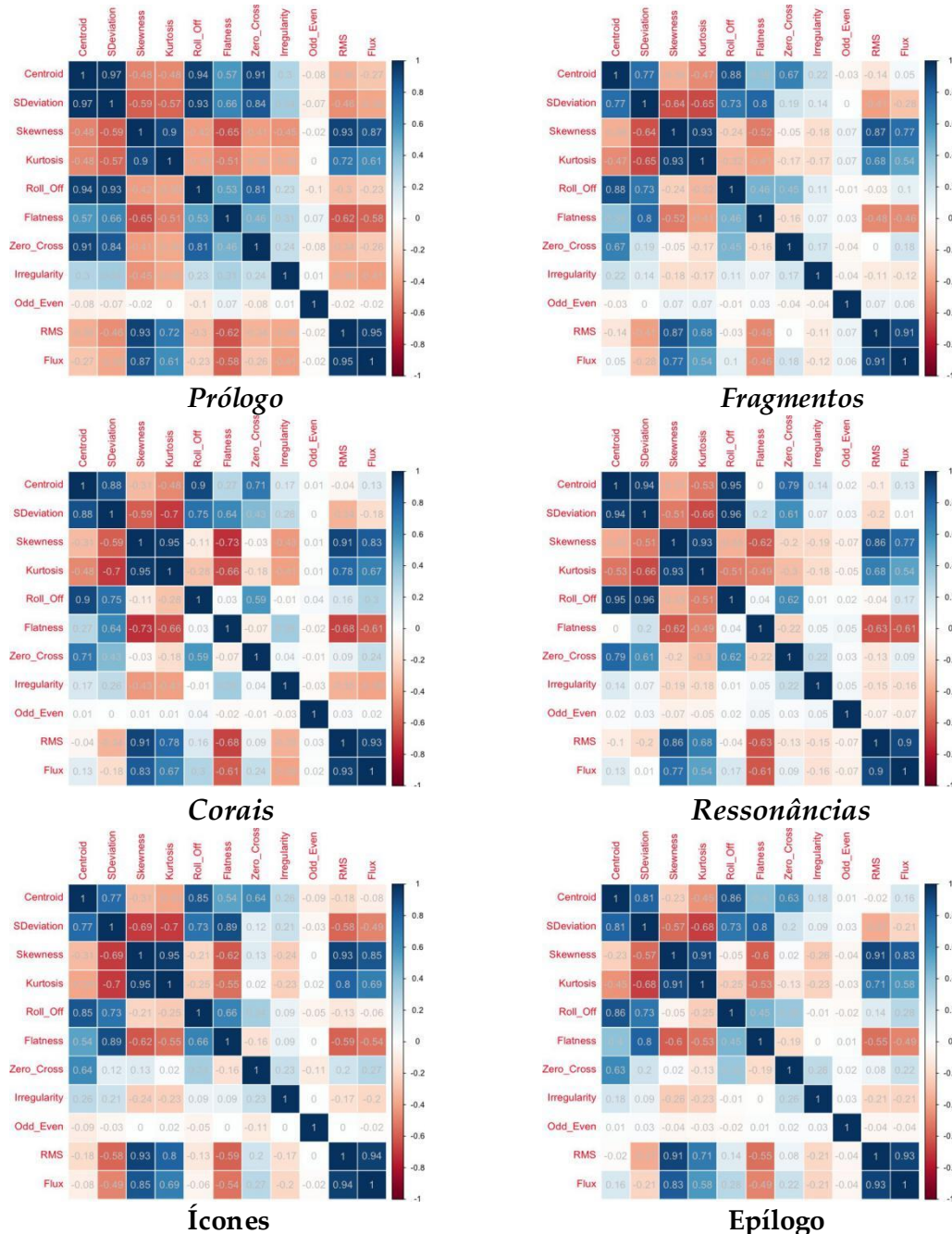


Figura 22: Matriz de correlação dos descritores de áudio utilizados em cada movimento de *Reflexões*

Os dados dispostos na Figura 22 apontam que o descritor de áudio *Odd to Even*, o qual se relaciona com a identificação de sonoridades anasaladas e aveludadas é a única medida de análise da qual não possui correlação com



nenhum outro descritor. Tal resultado indica que a variação da qualidade timbrística, explorada por tal descritor, não se relaciona com as outras características sonoras, exploradas pelo compositor. Em todos os movimentos da obra, o índice de correlação do descritor *Odd to Even* manteve-se dentro da escala entre 0,25 e 0,25. O intervalo abrange a correlação nula das medidas analisadas. Em contrapartida, o único movimento que não possui a participação do tímpano, *Prólogo*, possui a menor ocorrência de descritores não correlacionados. Tal resultado indica que as sonoridades elaboradas, em cada entrada do tímpano, são determinantes na correlação dos descritores de áudio. Tabela 4 ilustra o índice da média de ocorrências entre os descritores. A correlação positiva alta representa uma escala maior que 0,75; correlação positiva média abrange a escala menor que 0,75 e maior que 0,25; correlação positiva e/ou negativa muito baixa e não correlação abrange a escala maior que -0,25 e menor que 0,25; a correlação negativa média abrange a escala maior que -0,75 e menor que -0,25 e; correlação negativa alta abrange a escala menor que -0,75.

Frequência de Correlação	Prólogo	Fragmentos	Corais	Ressonâncias	Ícones	Epílogo
Positiva Alta	0,26	0,21	0,21	0,22	0,22	0,21
Positiva Média	0,15	0,12	0,17	0,07	0,13	0,15
Nula	0,21	0,5	0,4	0,53	0,46	0,5
Negativa Média	0,38	0,18	0,23	0,18	0,18	0,15
Negativa Alta	0	0	0	0	0	0

Tabela 3: Média dos índices de correlação de todos os descritores de áudio utilizados, em cada movimento.

A próxima etapa centrou-se em segmentar os movimentos a partir das entradas do tímpano. Para tanto, utilizamos o trabalho colaborativo entre a partitura musical e as curvas de análises via descritores de áudio. Relacionamos a Figura 21, que ilustra as curvas dos dados de análise, com a Tabela 1, que discrimina o número de entradas. Identificamos o descritor de áudio RMS como aquele que melhor ressalta tais entradas. A Figura 23 ilustra os picos dos índices de RMS com a presença do tímpano.



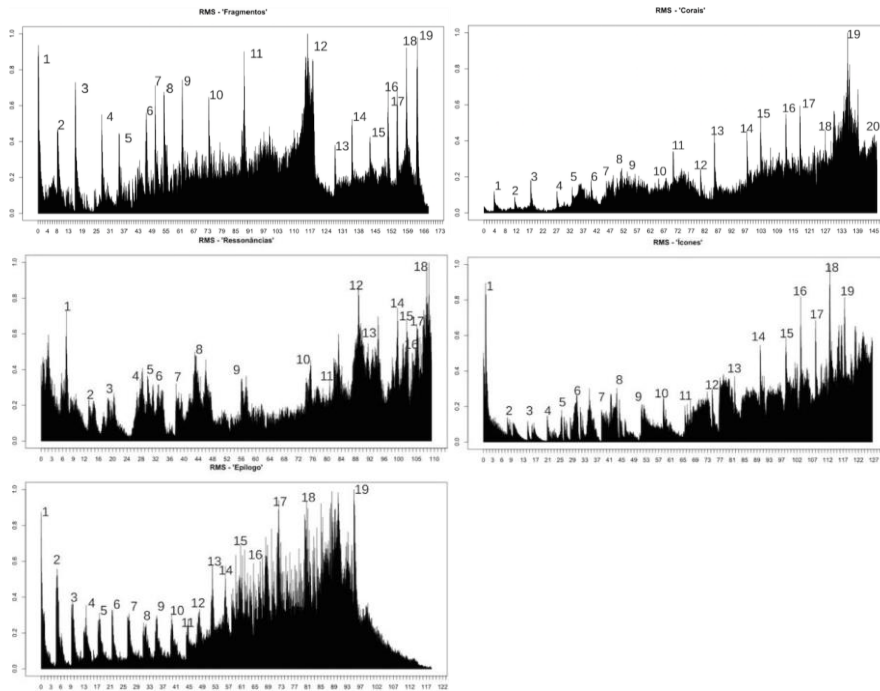


Figura 23: Picos de amplitude, extraídos pelo RMS, com as entradas do tímpano, em cada movimento de *Reflexões*

Para representar as sonoridades distintas entre os movimentos da obra distinguimos 25 agrupamentos com um vetor de 11 valores médios para cada descritor de áudio utilizado. Utilizamos o método denominado como *K-means*, para agrupar as médias das variáveis de análise⁶. A propriedade multidimensional das diversas características que são atribuídas ao timbre musical, sejam perceptivas ou espectrais, torna a análise sonora complexa e muito abrangente. Os agrupamentos elaborados a partir das médias dos descritores dialogam com a complexa característica multidimensional das sonoridades analisadas. O algoritmo *K-means*, implementado em ambiente de análise estatística *R*⁷, calculou automaticamente os 25 agrupamentos. O resultado correspondeu a aproximadamente 85% de efetividade para distinguir as sonoridades. A Figura 24 ilustra os 25 agrupamentos, a partir da média, normalizada no intervalo [0...1], de cada descritor de áudio. A escala de gradações de cinza, representam os valores da média de cada descritor de áudio. Tal procedimento permitiu encontrar relações entre todas as sonoridades estabelecidas por cada entrada do tímpano.

⁶ De uma maneira generalizada, o conceito de *K-means* relaciona-se com um método de agrupamento no qual objetiva-se particionar n observações dentre k grupos onde cada observação pertence ao grupo mais próximo da média. Para mais informações, ver: [https://en.wikipedia.org/wiki/K-means\\$clustering](https://en.wikipedia.org/wiki/K-means$clustering). Data de acesso: 30/07/2018

⁷ Para mais informações sobre o ambiente *R*, ver: <https://www.r-project.org>. Data de acesso: 30/07/2018



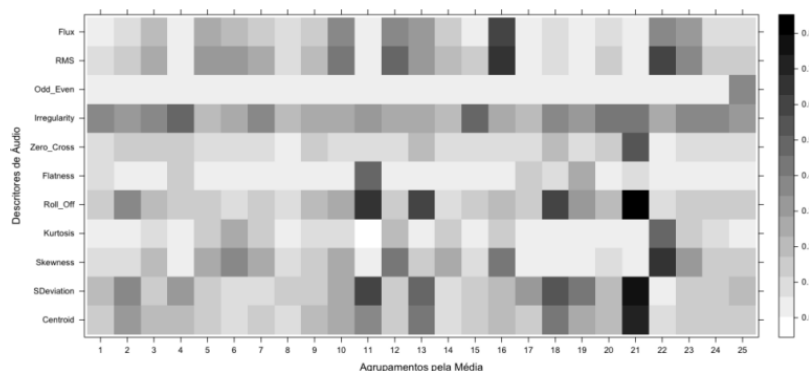


Figura 24: Representação da configuração dos 25 agrupamentos elaborados pela análise de K-means.

As gradações de cinza representam os valores, normalizados, das médias de cada descritor de áudio. As colunas ilustram os agrupamentos e as linhas representam os descritores de áudio utilizados.

A próxima etapa de análise relacionou os agrupamentos *K-means* com cada entrada do tímpano em cada movimento. A Figura 25 ilustra os 25 agrupamentos *K-means* relacionados com a identificação da entrada dos tímpanos, tal como apresentados pela Figura 23. Representamos cada sonoridade a partir de sua localização, em cada um dos movimentos da obra: “frag”, para as entradas do tímpano, em *Fragmentos*; “cora” para *Corais*; “res” para *Ressonâncias*; “ico” para “Ícones” e “ep” para *Epílogo*. O mapeamento da quantidade de agrupamentos, tal como ilustrado pela Figura 25, permitiu correlacionar as 95 entradas do tímpano. O procedimento permitiu encontrar similaridades entre as sonoridades estabelecidas pelo compositor. A Figura 26 ilustra a matriz de correlação entre todas as entradas do tímpano, do ponto de vista da média dos descritores utilizados⁸. A Figura 27 ilustra o diagrama com a frequência dos índices de correlação, entre as sonoridades analisadas. Pela escala, as sonoridades que apresentaram frequências iguais a zero, são representadas pela cor branca. Inversamente, as sonoridades que apresentaram frequências iguais a 95, são representadas pela cor preta.

⁸ Ressaltamos a presença da correlação das sonoridades entre si. Portanto, cada uma das sonoridades possui, pelo menos, uma correlação positiva alta.



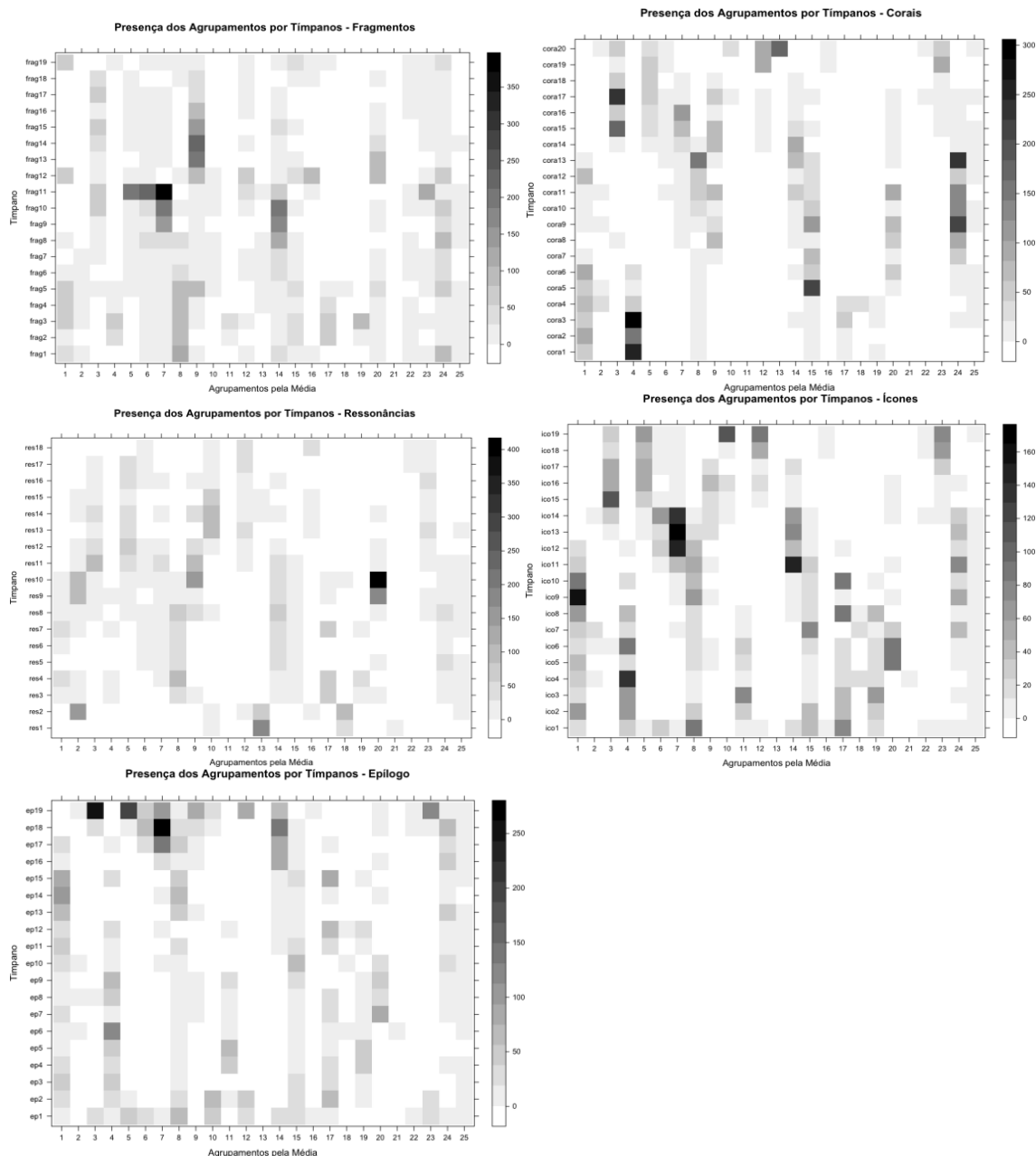


Figura 25: Correlação entre a frequência dos agrupamentos *K-means* e as entradas do tímpano.

Do ponto de vista global, cada sonoridade possui, em média, 65,31 sonoridades entre a escala de correlações baixas ou nulas. Ressalte-se a não presença de índices de correlação negativa alta. Contudo, identificamos algumas sonoridades que possuem índices de correlação positiva alta. Do ponto de vista dos descritores utilizados, o resultado indica que há similaridades entre algumas das sonoridades analisadas. Em média, para cada sonoridade analisada, há 5,82 sonoridades com índices de correlação positiva alta. Discriminamos, a seguir, em



perspectiva, as sonoridades que possuem maior frequência de correlação positiva alta.

- “ep11”, com total de 12 correlações positivas altas;
- “res4”, “ep3” e “ep19”, com total de 11 correlações positivas altas;
- “frag4”, “frag9”, “frag19”, “res11” e “ico16”, com total de 10 correlações positivas altas.

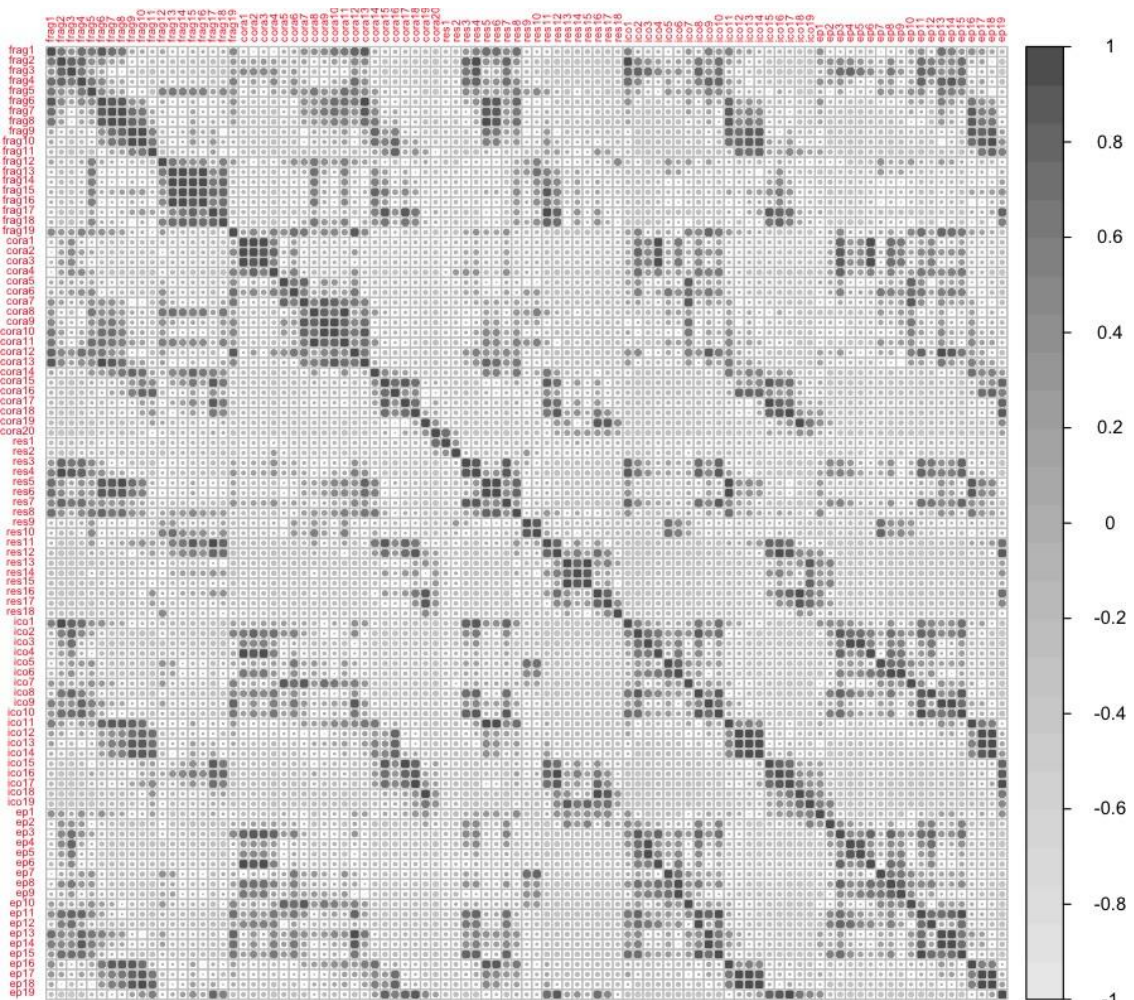


Figura 26: Matriz de correlação entre todas as 95 entradas do tímpano. Por convenção, os quadrados totalmente preenchidos representam a mais alta correlação positiva.



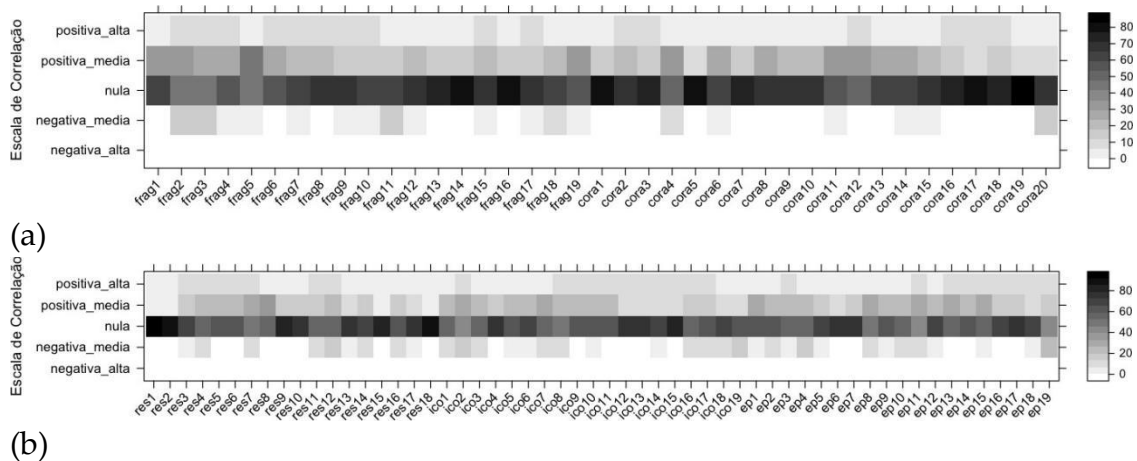


Figura 27: Frequência de correlação de cada uma das 95 sonoridades, estabelecidas pelas entradas do tímpano.

Identificamos as sonoridades das quais apresentam os maiores índices de correlação positiva alta, com a sonoridade “ep11”: “frag3”, “frag4”, “cora12”, “res4”, “ico2”, “ico8”, “ico9”, “ico10”, “ep11”, “ep13”, “ep14” e “ep15”. A sonoridade “ep11” possui características simbólicas das quais relacionam-se com as principais construções musicais, desenvolvidas pelo compositor, dentre elas: escalas ascendentes, acordes estáticos e as articulações rítmicas.

Figura 28: Excerto da partitura da sonoridade “ep11”, compasso 228, que possui a maior correlação positiva alta. Destacamos os instrumentos: flautas, oboés, clarinetes, fagotes, trombones, tímpano e vibrafone, piano e cordas.



A Figura 28 ilustra a sonoridade “ep11”, na partitura, que possui a maior frequência de correlação positiva alta. Subdividimos o excerto para destacar as construções musicais. Destacamos a análise das principais sonoridades, estabelecidas em *Reflexões*, por intermédio das 95 entradas do tímpano, dispostos nos cinco movimentos da obra. A seguir, discutimos toda a trajetória de análise realizada na obra. Por tratar-se de um trabalho ainda em andamento, apresentamos a nossa projeção de desenvolvimento de pesquisa, a ser realizada posteriormente.

6 - Discussão e Conclusão

A obra *Reflexões*, do compositor Jônatas Manzolli, possui características sonoras que foram destacadas a partir de recursos computacionais e de procedimentos de análise simbólica musical. O procedimento de análise da obra utilizou o mesmo sistema de análise empregado em trabalhos anteriores: o trabalho colaborativo entre a partitura musical e a gravação sonora (Simurra et. al 2013 e Simurra 2014). Para o estudo de caso particular da obra de Manzolli, utilizamos um vetor de características constituído por 11 descritores de áudio resultando em uma gama muito maior de dados coletados. A considerável quantidade de descritores utilizados ampliou as possibilidades de análise, do ponto de vista das características sonoras, extraídas pelas medidas. Contudo, dois descritores de áudio destacaram-se. Ao utilizar o *Odd to Even*, concluímos que o compositor ressaltou características que podemos relacionar a sonoridades anasaladas e aveludadas. Do ponto de vista do desenvolvimento do projeto composicional, em *Reflexões*, as sonoridades aveludadas e anasaladas não representam sonoridades específicas. Contudo, tais características fazem parte da sua macroestrutura. O descritor de áudio RMS destacou as 95 entradas do tímpano dispostas ao longo dos cinco últimos movimentos. Dentre todos os descritores utilizados, a variação dos picos de amplitude, extraída pelo RMS, permitiu relacionar o trabalho de segmentação das entradas com as informações simbólicas fornecidas pela partitura musical. Após a segmentação, analisamos as características sonoras de cada uma das sonoridades. Utilizamos um algoritmo de agrupamento, o qual estabeleceu 25 vetores com os valores médios das medidas. O procedimento permitiu calcular as correlações entre as sonoridades destacadas. Dentre as sonoridades analisadas, “ep11” possui a maior frequência de correlação positiva alta. No total, “ep11” correlacionou-se positivamente alto com 12 sonoridades distintas. Toda a trajetória de análise estabeleceu um interessante procedimento para identificar as características sonoras de construções sonoras distintas. Utilizar o ambiente computacional de análise



permitiu segmentar a obra e relacioná-la com a partitura musical. No caso particular de *Reflexões*, *Prólogo* é o único movimento que não possui a presença do tímpano. Por tal característica, não realizamos a etapa de segmentar o primeiro movimento. A projeção de trabalhos futuros de análise, que extrapola o escopo deste recorte de pesquisa, centra-se em identificar similaridades e correlações entre as 95 sonoridades com o primeiro movimento da obra. Todavia, ressaltamos que o uso das técnicas e ferramentas de análise musical é uma das contribuições originais do nosso trabalho. A obra de Jônatas Manzolli contrasta com os diversos projetos composicionais da música do século XXI cujo foco centra-se em estruturas formais para a organização sonora, tais como, apenas para citar alguns poucos dentro de um escopo muito diverso, Scelsi ou Lachenmann. Contudo, os compositores compartilham o mesmo objetivo: a variação de sonoridades como elemento estrutural de suas obras. O compositor brasileiro, diferentemente de Scelsi e Lachenmann, concatena suas sonoridades por intermédio de agrupamentos distintos de materiais composicionais. Manzolli condiciona as suas sonoridades pelas 95 entradas do tímpano. Manzolli integra as informações simbólicas com as suas construções sonoras. A partitura não representa somente uma interface de comunicação entre o compositor, o intérprete e o resultado sonoro. Para o compositor, a partitura também é um ambiente de assistência à composição musical. Por fim, Jônatas Manzolli explora as diferentes possibilidades sonoras a partir de construções musicais, claramente destacadas na partitura, além de explorar técnicas composicionais relacionadas às escalas musicais, às articulações rítmicas e às construções harmônicas.

Agradecimentos

O autor agradece o apoio da FAPESP – Fundação de Amparo à Pesquisa do Estado de São Paulo pela concessão de bolsas de doutorado e pós-doutorado, processos: 2011/23972-2 e 2015/25160-6, respectivamente.

Referências

1. Antoni, Jérôme. 2006. The spectral kurtosis: a useful tool for characterising non-stationary signals. *Mechanical Systems and Signal Processing*, 20(2), p. 282–307.
2. Brent, William. 2010. Physical and perceptual aspects of percussive timbre. PhD Thesis - UC San Diego Electronic Theses and Dissertations, 171 pages.
3. Bullock, Jamie. 2007. Libxtract: A lightweight library for audio feature extraction. In: *Proceedings of the International Computer Music Conference*, v. 43, p. 1-4.



4. Collins, Nick. 2005. An automated event analysis system with compositional applications. In: *Proceedings of the International Computer Music Conference*, Barcelona.
5. Cook, Nicholas. 1994. *A guide to musical analysis*. UK: Oxford University Press.
6. DeCarlo, Lawrence 1997. On the meaning and use of kurtosis. *Psychological methods*. Vol. 2(3), Sep 1997, p. 292-307.
7. Dubnov, Shlomo. 2004. Generalization of spectral flatness measure for non-gaussian linear processes. *Signal Processing Letters, IEEE*, 11(8), p. 698-701.
8. Einbond, Aaron. 2005. Composer as curator: Uncreativity in recent electroacoustic music. *Proceedings of the Electroacoustic Music Studies Network Conference Electroacoustic Music Beyond Performance*. Berlim.
9. Esling, Philippe e Agon, Carlos. 2013. Multiobjective time series matching for audio classification and retrieval. *Audio, Speech, and Language Processing, IEEE Transactions on*, 21(10), p. 2057-2072.
10. Fletcher, Harvey e Munson, Wilden. 1933. Loudness, its definition, measurement and calculation. *Bell System Technical Journal*, 12(4), p. 377-430.
11. Florentine, Mary e Bonding, Per. 1978. Loudness of complex sounds as a function of the standard stimulus and the number of components. *The Journal of the Acoustical Society of America*, 64(4), p. 1036-1040.
12. Fujishima, Takuya. 1999. Realtime chord recognition of musical sound: A system using common lisp music. In: *Proc. ICMC*, v. 1999, p. 464-467.
13. Gómez, Emilia. 2006. Tonal description of polyphonic audio for music content processing. *INFORMS Journal on Computing*, 18(3), p. 294-304.
14. Jehan, Tristan. 2005. *Creating music by listening*. 2005. Tese de Doutorado. Massachusetts Institute of Technology, School of Architecture and Planning, Program in Media Arts and Sciences.
15. Krimphoff, Jochen, McAdams, Stephen e Winsberg, Suzanne. 1994. "Caractérisation du timbre des sons complexes. II. Analyses acoustiques et quantification psychophysique." *Le Journal de Physique IV* (4), p. C5-625.
16. Le Groux, Sulvain et al. U. 2010. Disembodied and collaborative musical interaction in the multimodal brain orchestra. In: *Proceedings of the International Conference on New Interfaces for Musical Expression*. Sydney: NIME.



17. Loureiro, Mauricio e Magalhães, Tairone. 2008 Segmentação e extração de descritores de expressividade em sinais musicais monofônicos. *Seminário Música Ciência Tecnologia*, v. 1, n. 3.
18. Malt, Mikhail e Jourdan, Emmanuel. 2008. Zsa Descriptors: a library for real-time descriptors analysis. In: *5th Sound and Music Computing Conference, Berlin, Germany*, p. 134-137.
19. Malt, Mikhail e Jourdan, Emmanuel. 2009. Real-time uses of low level sound descriptors as event detection functions using the max/msp zsa descriptors library. *Proceedings of the 12th Brazilian Symposium on Computer Music*.
20. Manzolli, Jônatas. 1988. Um modelo matemático para timbre orquestral. Dissertação de mestrado. IMECC, Unicamp, 136 páginas.
21. Manzolli, Jônatas. 2013. Interpretação mediada: pontos de referência, modelos e processos criativos. *Música Hodie*, 13(1), p. 48-63.
22. Mikula, Luka. 2008. Concatenative music composition based on recontextualisation utilising rhythm-synchronous feature extraction. PhD Thesis. Institute of Electronic Music and Acoustics. University of Music and Dramatic Arts. Graz, Austria.
23. Miśkiewicz, Andrzej e Rakowski, Andrzej. 1994. Loudness level versus sound-pressure level: A comparison of musical instruments. *The Journal of the Acoustical Society of America*, 96(6), p. 3375–3379.
24. Monteiro, Adriano. 2012. Criação e performance musical no contexto de instrumentos musicais digitais. Dissertação de Mestrado Instituto de Artes, Unicamp.
25. Peeters, Geoffroy. 2004. *A large set of audio features for sound description (similarity and description) in the cuidado project*. Paris: IRCAM,
26. Peeters, Geoffroy; McAdams, Stephen e Herrera, Perfecto. 2000. Instrument sound description in the context of MPEG-7. In: *ICMC - International Computer Music Conference*, p. 166-169.
27. Pereira, Erica. 2009. Estudos sobre uma ferramenta de classificação musical. Campinas: Dissertação de Mestrado, Universidade Estadual de Campinas.
28. Plomp, Reinier. 2001. *The intelligent ear: On the nature of sound perception*. New York: Psychology Press, Reinier Plomp.



29. Rimoldi, Gabriel. 2013. Modelagem interativa aplicada à síntese e espacialização no domínio microtemporal. Mestrado Instituto de Artes, Unicamp.
30. Sheh, Alexander e Ellis, Daniel. 2003. Chord segmentation and recognition using entrained hidden Markov models. *ISMIR 2003*, p. 185–191.
31. Shepard, Roger. 1982. *Structural representations of musical pitch - The psychology of music*. New York: D. Deutsch Academic, p. 343–390.
32. Simurra, Ivan Eiji; Maia, Igor e Manzolli, Jônatas. 2013. Quattro pezzi per orchestra de Scelsi sob a ótica da análise musical assistida por computador. *14º Simpósio Brasileiro de Computação Musical*, p. 221–224.
33. Simurra, Ivan Eiji e Tavares, Tiago. 2014. A comprehensive approach in music and audio analysis: Similarities in Helmut Lachenmann's music compositions. In: *SysMus 2014 – 7th International Conference of Students of Systematic Musicology*, p. 1.
34. Sturm, Bob. 2014. The state of the art ten years after a state of the art: Future research in music information retrieval. *Journal of New Music Research*, v. 43, n. 2, p. 147-172.
35. Theodoridis, Sergios e Koutroumbas, Konstantinos. 2008. Pattern recognition. *IEEE Transactions on Neural Networks*, v. 19, n. 2, p. 376.
36. Thiede, Thilo. 1999. Perceptual Audio Quality Assessment Using a Non-linear Filter Bank. Vom Fachbereich Elektrotechnik der Technischen Universität Berlin zur Erlangung des akademischen Grades. Genehmigte Dissertation. 167 pages.
37. Tzanetakis, George e Cook, Perry. 1999. Multifeature audio segmentation for browsing and annotation. In: *Applications of Signal Processing to Audio and Acoustics, 1999 IEEE Workshop*, p. 103–106.
38. Zwicker, Eberhard e Scharf, Bertram. 1965. A model of loudness summation. *Psychological Review*, 72(1), p.3.

