

O uso de descritores acústicos como ferramenta na análise, criação e performance musical

Gabriel Rimoldi

Universidade Estadual de Campinas – gabriel.rimoldi@gmail.com

Ivan E. Y. Simurra

Universidade Estadual de Campinas – ieysimurra@gmail.com

Adriano Claro Monteiro

Universidade Estadual de Campinas – monteiro.adc@gmail.com

Resumo: Discutimos neste artigo sobre o uso de ferramentas computacionais de recuperação de informação musical (MIR) como proposta de sistematização à criação, análise e performance musicais. Apresentamos um panorama metodológico a partir de trabalhos desenvolvidos pelos autores nos últimos anos que abordam a sonoridade como um aspecto estrutural do discurso musical. As propostas discutidas confluem numa abordagem sistemática da sonoridade que, reduzida a um conjunto de dimensões mensuráveis, possibilite uma nova exploração metodológica nos mais diversos campos de pesquisa musical.

Palavras-chave: Recuperação de informação musical (MIR). Descritores acústicos. Análise musical. Criação musical. Performance musical.

The use of audio features as tool for musical analysis, composition and performance

Abstract: We discuss in this paper the use of computational tools of Music Information Retrieval (MIR) as methodological approach to music analysis, composition and performance. We present an overview of works developed by the authors in recent years that conceive the sonority as structural aspect of musical discourse. These proposals converge to a systematic perspective of sonority that reduced to a set of measurable dimensions allows a new methodological exploration in various fields of music research.

Keywords: Music Information Retrieval (MIR). Audio Features. Music Analysis. Music Creation. Music Performance.

1. Introdução

Descrever metodologicamente a sonoridade a partir de um conjunto de características capazes de circunscrevê-la num espaço observável e replicável é um aspecto desafiador à criação, análise e performance musical. Torna-se ainda mais preponderante quando tratamos de práticas musicais dos séculos XX e XXI, em que nota-se uma crescente preocupação na obtenção de resultados timbrísticos mais refinados e particulares. A progressiva expansão dos dispositivos de criação no campo da música instrumental, seja pela utilização de técnicas estendidas ou por modelos prescritivos não convencionais de acabo, bem como a produção e manipulação sonora através de suporte eletroacústico, reforçam a emergência da sonoridade em sua

natureza complexa e multifacetada como aspecto elementar na criação contemporânea. A ideia de sonoridade como uma “metadimensão” (GUIGUE, 1996) traduz, deste modo, o interesse de considerá-la não mais como uma simples “cor”, mas como um espaço para integrar os parâmetros musicais e, até mesmo, como categoria central do discurso sonoro. Seja no campo da escritura instrumental ou na criação com suporte eletroacústico, observa-se que as múltiplas dimensões que se interpõem podem alterar sensivelmente as características sonoras que, em consonância à ideia discutida por Solomos (2013), torna assim a sonoridade uma entidade global, um “artefato” articulado e construído a partir do “interior” do som, uma propriedade emergente da interação de uma diversidade de elementos em jogo.

No campo da engenharia, a área Recuperação de Informação Musical (com acrônimo MIR, do inglês, *Music Information Retrieval*) tem proposto a extração de características a partir do sinal de áudio com o objetivo de analisar e organizar arquivos digitais a partir de seus conteúdos musicais. Alguns dos principais objetivos específicos em MIR são: desenvolver métodos para transcrição automática de música (HERRERA-BOYER; KLAPURI; DAVY, 2006), organizar grandes bases de dados com arquivos musicais (TZANETAKIS; COOK, 2002), desenvolver tecnologia musical, como automatização de processamento computacional de áudio e de equipamentos relacionado a música (RAFII; PARDO, 2009), interação homem-máquina ligada à música (BROSSIER, 2006), novas abordagens musicológicas (COLLINS, 2010), dentre outros.

Um dos aspectos que deve se ter em mente quando empregam-se técnicas de MIR é que boa parte dessas técnicas são orientadas para reconhecer no áudio padrões musicais pré-estabelecidos, de acordo com teorias musicais coerentes ao repertório analisado (*e.g.* harmonia tonal, contraponto tradicional, canções com repetição de refrão, etc.). A generalidade de algumas das ferramentas dessa área permite-nos, entretanto, ampliar seu escopo às praticas musicais que exploram aspectos outros daqueles apontados pelo repertório tradicional, sobretudo àquelas que trazem a sonoridade como um aspecto estrutural do discurso sonoro. Assim, trataremos neste artigo de resultados de sete anos de pesquisa desenvolvida no Núcleo Interdisciplinar de Comunicação Sonora (NICS – UNICAMP)¹ sobre a utilização das ferramentas MIR nos campos da criação, análise e performance com suporte

¹ www.nics.unicamp.br

computacional. As propostas discutidas confluem numa abordagem sistemática da sonoridade que, reduzida a um conjunto de dimensões mensuráveis, possibilite uma nova exploração metodológica nos mais diversos campos de pesquisa musical.

2. A biblioteca PDescriptors

PDescriptors é um conjunto de algoritmos de análise de áudio usuais na pesquisa em MIR que foram implementados como abstrações no software Pure Data (PD) e tem como intuito fundamentar pesquisas em performance, criação e análise musical (MONTEIRO; MANZOLLI, 2011). Na configuração atual a biblioteca PDescriptors contém implementações de descritores de áudio de baixo nível divididos em três sub-grupos: a) descritores de características temporais, nos quais os procedimentos de extração são aplicados a representação temporal digitalizada do som; b) descritores de características espectrais, nos quais os procedimentos de extração são aplicados após a Transformada Discreta de Fourier (TDF) do sinal sonoro digitalizado; c) descritores de características psicoacústicas, nos quais as extrações são calculadas de acordo com modelos da percepção auditiva humana. Na T, descrevemos todas os descritores presente na biblioteca PDescriptors.

Tabela 1: Subgrupos da biblioteca PDescriptors de acordo com a tipologia de descritores.

Características Temporais	Centróide Temporal, Taxa de Cruzamento por Zero, Proporção de Quadros com Pouca Potência, Auto-correlação, Média Energética, Desvio Médio de energia, Variância de Energia, Desvio Padrão de Energia, Obliquidade de Energia, Curtose de Energia.
Características Espectrais	Média Energética, Desvio Médio de energia, Variância de Energia, Desvio Padrão de Energia, Obliquidade de Energia, Curtose de Energia, Centróide Espectral, Variância Espectral, Desvio Médio Espectral, Desvio Padrão Espectral, Obliquidade Espectral, Curtose Espectral, Largura de Banda, Medida de Crista Espectral, Fluxo Espectral, Irregularidade Espectral, Conteúdo de Altas Frequências, Desvio de Fase, Domínio Complexo, <i>Spectral Flatness</i> , <i>Spectral Noisiness</i> , <i>Spectra Decrease</i> , <i>Spectral Slope</i> , <i>Spectral Smoothness</i> , <i>Roll-off</i> , <i>Cesptrum</i> , <i>MFCC (Mel-Frequency Cepstral Coeficients)</i> , <i>BFCC (Bark-Frequency Cepstral Coeficients)</i> .
Características Psicoacústicas	Curva de <i>Loudness</i> , Energia por Banda da Escala <i>Bark</i> , Energia por Banda da Escala MEL.

3. Análise, criação e performance via descritores acústicos

3.1 Análise da sonoridade no repertório instrumental e eletroacústico

No âmbito da musicologia, a “estética da sonoridade” apresenta peculiaridades em relação a outros repertórios, visto que os registros gráficos remanescentes dos processos criativos musicais, ou das partituras, não representam diretamente as qualidades sonoras apreensíveis na audição musical, mas representam os procedimentos físicos para a geração do resultado sonoro desejado pelo músico. Tal aspecto pode ser observado, por exemplo, pelo uso de técnicas estendidas, que expandem os preceitos da técnica instrumental historicamente consolidados, no intuito de se obter sonoridades de natureza complexa e que dão destaque a outros aspectos do som que não aqueles privilegiados pela escritura tradicional; ou ainda, nos resultados de técnicas de síntese sonora instrumental em que o percepto musical é emergente da interação entre os componentes espectrais dos sons instrumentais que são anotados na partitura muitas vezes de maneira convencional, ou em conjunto a notações prescritivas de ações. A problemática torna-se ainda mais incisiva quando se trata da análise de material musical gerado por meios eletroacústicos, pois nesses casos os suportes de criação e de performance convergem reduzindo significativamente a necessidade de representações gráficas das obras, salvo rascunhos e registro de planejamento dos processos composicionais, alguns casos de partituras de difusão sonora, e partituras de escuta.

A cisura entre notação e percepto auditivo coloca, assim, à musicologia uma relação diferente daquela antes estabelecida na musicologia dedicada ao repertório tradicional da música de concerto quanto à análise do chamado nível neutro da obra (NATTIEZ, 1990), isto é, de um estágio da análise musical voltada à descrição das características da obra de acordo com uma lógica de construção musical intrínseca e extraída de relações observáveis apenas no material de registro. A aplicação de tecnologia MIR provê à pesquisa musicológica informações objetivas concernentes às características sonoras e às possíveis estruturas musicais salientes no áudio das obras analisadas que podem ser confrontadas às análises dos demais registros gráficos, bem como às análises auditivas (KLIEN; GRILL; FLEXER, 2012).

Em linha com as pesquisas de aplicação de tecnologia MIR para a musicologia de obras recentes de concerto, desenvolvemos trabalhos de utilização da biblioteca *PDescriptors* para esse propósito. Um deles trata de uma análise comparativa entre as peças *Jonchais*, para orquestra, e *La Légende d'Eer*, para meios eletroacústicos, ambas do compositor Iannis Xenakis. No texto introdutório da

partitura de Jonchaies, Xenakis relata que a peça orquestral é baseada em resultados obtidos com a criação de *La Légende d'Eer*. Em nosso trabalho (BONDUKI; MONTERIO, 2015), utilizamos gráficos gerados a partir de dados numéricos extraídos por três descritores de áudio: centróide espectral, desvio padrão espectral, e RMS (a média quadrática da energia do sinal), como meio de identificar padrões de similaridade dos resultados sonoros nas gravações das duas peças, e que foram investigados como indícios das relações composicionais compartilhadas entre as peças.

O desenvolvimento do trabalho fez-se no reconhecimento visual de características semelhantes desses gráficos (vide figura 1), tomadas como indícios, que dirigiam o foco de investigações mais aprofundadas sobre as similaridades entre as peças, que por sua vez foram realizadas através de comparações dos espectrogramas em regiões localizadas, ou ainda recorrendo à análise da partitura de Jonchaies em comparação a estruturas identificadas nos espectrogramas. Dessa maneira, o uso dos descritores propiciou a comparação das duas obras sobre uma mesma base de representação visual, salientando características globais e livre de interpretações subjetivas do fenômeno sonoro (auditivo), que vieram apenas em uma etapa seguinte de identificação de estruturas de semelhança nas representações. Outra característica importante foi a possibilidade de redução integral do fenômeno sonoro a uma única representação estática, que ajudou os analistas à identificar e relacionar elementos de difícil apreensão pela escuta, devido à longa duração das peças, principalmente de *La Légende d'Eer*, bem como às diferenças significativas em suas proporções (*Jonchaies* com aproximadamente 15' ; e *La Légende d'Eer* com 45'31").

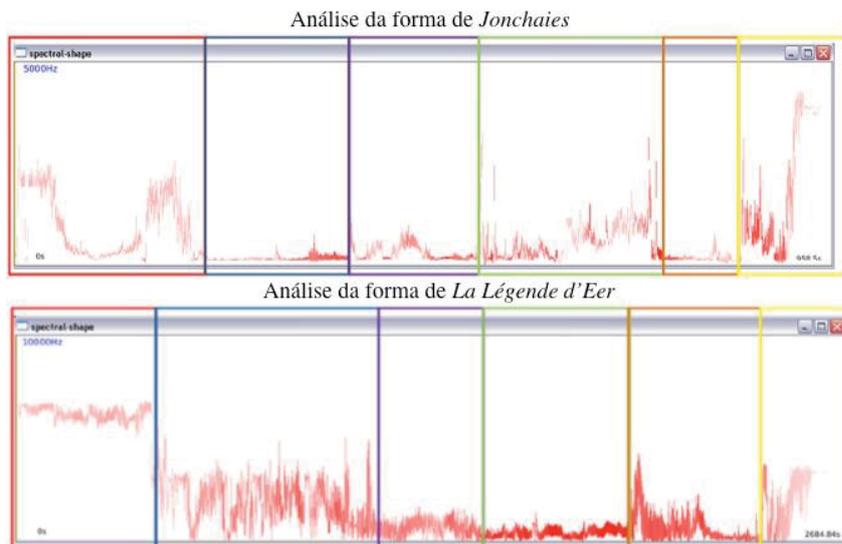


Figura 1: Comparação formal entre as obras *Jonchaies* e *La Légende d'Eer*, ambas de I. Xenakis via descritores acústicos.

Em outro trabalho, utilizamos o *PDescriptors* aplicado à análise à série *Audible Ecosystems*, do compositor Agostino di Scipio, também para suporte eletroacústico (RIMOLDI; MANZOLLI, 2016). A abordagem composicional tratada em *Audible Ecosystems* coloca a sonoridade como um aspecto emergente das interações locais e de menor complexidade entre os componentes colocados em jogo pelo compositor. Para isto, utilizamos a extração de características do sinal de áudio que dialogam com as parametrizações apontadas pelo próprio compositor (DI SCIPIO, 2003), a dizer: energia quadrática média (RMS), centroide espectral e fluxo espectral. No intuito de observar a formação de padrões globais e a dinâmica de comportamento entre estados subsequentes do sistema, que sugerissem um certo design composicional da obra, utilizamos então gráficos de recorrência e medidas de quantificação de recorrência a partir dos vetores de características extraídos do áudio das obras. A utilização de tais ferramentas propiciou o delineamento de características mais globais da série de obras analisadas que, como instâncias de um mesmo processo criativo, permitiu-nos observar aspectos de parametrização apontados pelo compositor. Na Figura 2 demonstramos a presença de diferentes estados intermediários na obra *Audible Ecosystems n. 2* através dos mapas de recorrência extraídos a partir dos descritores acústicos.

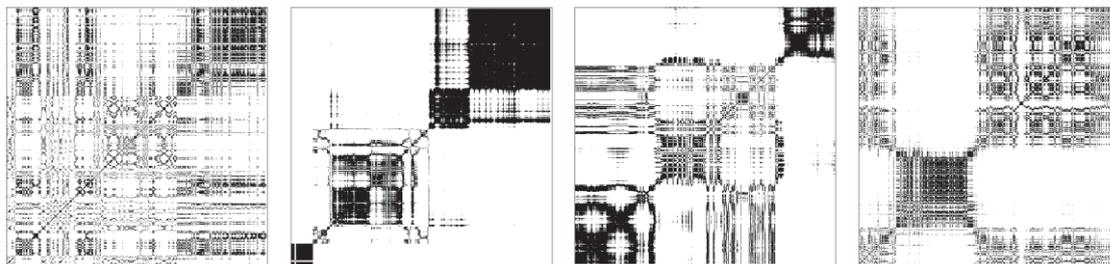


Figura 2: Estados Intermediários na obra *Audible Ecosystems n.2*, de Agostino di Scipio, através de Mapas de Recorrência.

3.2 Composição musical de sonoridades orquestrais

As ferramentas computacionais são recursos tecnológicos que amplificam e expandem as possibilidades e as escolhas do material composicional sem, entretanto, substituir o trabalho particular do próprio compositor. O planejamento composicional aqui apresentado, foca em sonoridades musicais que possuem dois universos distintos, mas não disjuntos dos quais a) representa as técnicas de execução instrumental estendidas e b) representa as ferramentas computacionais para analisar e descrever estatisticamente o conteúdo espectral do material gerado por essas técnicas. Para tanto, desenvolvemos um ambiente computacional, denominado *Sound Shizuku Composition – SSC*, para analisar sonoridades do ponto de vista de seu conteúdo espectral. Em nosso protótipo, as sonoridades resultam da interação das análises dos descritores de áudio com a própria percepção de suas características espectrais. O escopo do desenvolvimento do nosso ambiente de orquestração assistida por computador centra-se nas particularidades contrastantes de variações de sonoridades, a partir da análise sonora via descritores de áudio. A Figura 3 ilustra o ambiente de pesquisa e exploração das misturas sonoras, que estão dispostas em quatro espaços bidimensionais, representando o conjunto de descritores de áudio. O primeiro espaço é representado pelos eventos de cor amarela. O segundo espaço é representado pelos eventos de cor verde. O terceiro espaço é representado pelos eventos de cor roxa. Por último, o quarto espaço é representado pelos eventos de cor vermelha. Todos os pontos representam as mesmas sonoridades dispostas em espaços de análise distintos pelos descritores de áudio.

Composição e análise são as duas frentes de trabalho fundamentais neste estudo. Construímos espaços colaborativos de trabalho entre os dados objetivos da análise sonora junto ao processo de escrita composicional para ampliar a reflexão artística, estética e composicional. Num escopo mais generalizado, o processo criativo

Nosso intuito foi estabelecer uma função entre qualidades timbrísticas, que são representadas pelo espaço dos descritores, e a posição aural e ordenamento temporal dos grãos sonoros no espaço de difusão em suporte eletroacústico. A posição de cada grão é determinada por uma combinação ponderada dos descritores associados pelo próprio usuário do sistema a cada dimensão do espaço de difusão sonora (bi ou tridimensional). A partir da seleção de um conjunto de amostras de áudio previamente selecionadas, o usuário pode controlar então o sequenciamento e espacialização dos segmentos sonoros pela manipulação de trajetórias no espaço paramétrico dos descritores.

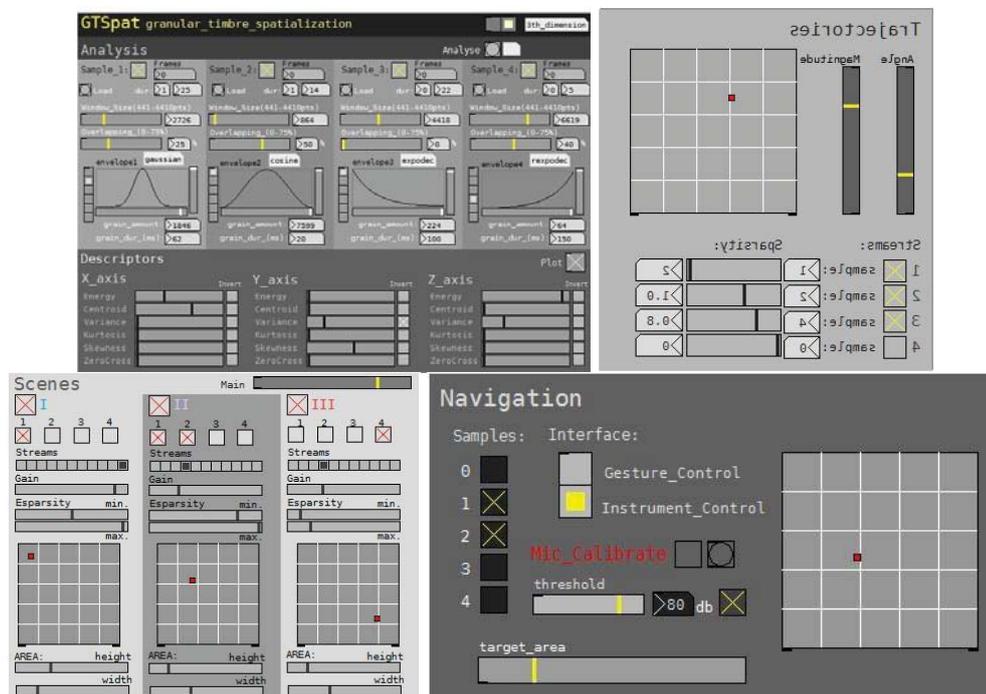


Figura 4: Interface do sistema GTSpaT - Granular Timbre Spatialization, que associa informações obtidas via descritores acústicos para o sequenciamento e espacialização de síntese sonora.

3.4 Estratégias de controle em instrumentos aumentados via descritores

O uso de descritores acústicos tem se configurado também como uma ferramenta útil na captura gestual indireta de instrumentos acústicos (TRAUBE; DEPALLE; WANDERLEY, 2003). Tal metodologia baseia-se na extração de características de baixo nível a partir do sinal de áudio e na utilização de modelos que associem estes dados a informações relacionadas à gestualidade do instrumentista. Em nossa pesquisa, temos investigado a utilização de modelos híbridos de captura gestual associado à flauta transversal em ambientes de improvisação. Nosso protótipo de

instrumento aumentado, o qual denominado *Metaflauta*, emprega tanto a aquisição direta de movimentos do instrumentista através de sensores acoplados ao instrumento, bem como a aquisição indireta via descritores de áudio (RIMOLDI; MANZOLLI, 2015). A aquisição indireta relaciona-se a informações gestuais efetivas, ou seja, ações diretamente à produção do som. Em nosso modelo, temos sobretudo empregado o uso dessas ferramentas na detecção de técnicas estendidas para flauta e nas possibilidades de controle e interação através do mapeamento das mesmas. A recuperação de informações gestuais efetivas foi realizada a partir da extração de descritores temporais e espectrais do próprio som do instrumento. A partir do vetor de características extraído, aplicamos um conjunto de técnicas de aprendizado não supervisionado que tem como propósito a redução de dimensionalidade dos dados. O processamento dos dados obtidos via descritores permiti-nos explorar aspectos da sonoridade do instrumento, sobretudo pelo uso de técnicas estendidas, no controle de processamento de imagens e sons em ambientes de improvisação assistida por computadores.

4. Discussão e Projeções Futuras

Neste artigo apresentamos um conjunto de pesquisas desenvolvidas no Núcleo Interdisciplinar de Comunicação Sonora que estudam a aplicação de tecnologias MIR como metodologias sistemáticas para a pesquisa em música que exploram a sonoridade como elemento central do discurso musical. Os trabalhos expostos abordam os temas: musicologia e análise musical; composição; e performance e desenvolvimento de novos instrumentos. Todos eles foram realizados com a biblioteca de descritores acústicos PDescriptors, implementada na plataforma Pure Data como subproduto e suporte metodológico da pesquisa aqui reportada. No campo da musicologia, resumimos dois trabalhos que utilizaram os dados obtidos de descritores acústicos para a criação de representações gráficas que ressaltaram propriedades acústicas específicas das gravações das obras analisadas. No primeiro desses dois trabalhos os gráficos serviram como intermediação visual comum para comparação das obras *Jonchaies* e *La Légende d'Eer* ambas de Iannis Xenakis. No segundo trabalho, que analisa a obra *Audible Ecosystems* de Agostino Di Scipio, houve o acréscimo de uma etapa de mensuração de propriedades dos gráficos (medidas de quantificação de recorrência extraídas das séries temporais resultantes da aplicação

dos descritores acústicos) que salientaram aspectos formais macroestruturais da obra musical analisada. No âmbito da composição musical, apresentamos uma estratégia de composição para explorar o universo sonoro das técnicas estendidas de execução musical, a partir das sonoridades orquestrais cujas características espectrais são analisadas pelos descritores de áudio. O trabalho de pesquisa relacionou a análise musical do ponto de vista objetivo, utilizando ferramentas computacionais e estatísticas, com o próprio planejamento conceitual, estético e particular do compositor. Desenvolvemos diversos experimentos composicionais utilizando os procedimentos metodológicos, tanto aplicado à escritura instrumental, como para o controle de síntese e espacialização sonora. Por fim, aplicamos essas ferramentas como potencial de expansão de instrumentos acústicos aplicada a contextos de improvisação com suporte computacional.

Como se nota nos trabalhos apresentados, uma importante característica das técnicas MIR é seu potencial de integração da representação das características sonoras micro-temporais (da ordem dos milissegundos) em elementos sonoro/musicais médio ou macro-temporais (respectivamente da ordem de segundos e minutos), o que é confluyente com o ideário técnico-musical de obras que têm a sonoridade como elemento de integração do discurso, pois nelas a construção e variação de objetos perceptivos é fundamentada pela manipulação de características micro-temporais do som. Em relação a esse aspecto, umas das projeções de desenvolvimento futuro da pesquisa aqui reportada está justamente a possibilidade de integração metodológica entre esses domínios temporais do fenômeno sonoros/musical. Reportamos apenas em um dos trabalhos, especificamente na análise da obra *Audible Ecosystems* de Agostino Di Scipio, metodologias de uma mesma natureza (sistemática e objetiva) para a extração de características micro-temporais e posterior análise de padrões macro-temporais; enquanto que nos demais trabalhos esse salto se deu por uma mudança de paradigma metodológico em que a subjetividade do analista, compositor ou performer, visa a integração dos dados micro-temporais em elementos musicais de níveis médio ou macro-temporais.

Entretanto, é importante notar que esse tipo de desenvolvimento metodológico não é trivial. A definição de um arcabouço metodológico universal e generalista, pois se trata de um repertório musical em que os estilos e abordagens divergem em grande medida na qual a inovação e diferenciação em relação ao

repertório existente é um dos valores de juízo artístico. O que se tem na prática são ferramentas desenvolvidas para um grupo bem definido de obras ou, muitas vezes, para casos particulares. Sendo assim, a análise e a interpretação computacional dos dados extraídos via descritores acústicos, com intuito de identificar e representar de elementos musicais, é o principal aspecto de continuidade da pesquisa aqui reportada.

Referências

- BONDUKI, S. A.; MONTERIO, A. Compositional Influences in Jonchaies from La Légende d'Eer1. *Iannis Xenakis, la musique électroacoustique: The electroacoustic music*, p. 53, 2015.
- BROSSIER, P. M. The aubio library at mirex 2006. *MIREX 2006*, p. 1, 2006.
- COLLINS, N. *Computational Analysis of Musical Influence: A Musicological Case Study Using MIR Tools*. ISMIR. Anais...2010
- DI SCIPIO, A. "Sound is the interface": from interactive to ecosystemic signal processing. *Organised Sound*, v. 8, n. 3, p. 269–277, 2003.
- GUIGUE, D. Une étude "pour les sonorités opposées". *Villeneuve d'Asq: Presses Universitaires du Septentrion*, 1996.
- HERRERA-BOYER, P.; KLAPURI, A.; DAVY, M. Automatic classification of pitched musical instrument sounds. In: *Signal processing methods for music transcription*. [s.l.] Springer, 2006. p. 163–200.
- KLIEN, V.; GRILL, T.; FLEXER, A. On automated annotation of acousmatic music. *Journal of New Music Research*, v. 41, n. 2, p. 153–173, 2012.
- MONTEIRO, A.; MANZOLLI, J. *A Framework for Real-time Instrumental Sound Segmentation and Labeling*. Proceedings of IV International Conference of Pure data-Weimar. Anais...2011
- NATTIEZ, J.-J. *Music and discourse: Toward a semiology of music*. [s.l.] Princeton University Press, 1990.
- RAFII, Z.; PARDO, B. *Learning to Control a Reverberator Using Subjective Perceptual Descriptors*. ISMIR. Anais...2009
- RIMOLDI, G.; MAIA JR, A. *GTSpat: um sistema interativo de síntese e espacialização sonora*. Anais do 13º Simpósio Brasileiro de Computação Musical. **Anais**...Brasília: 2011
- RIMOLDI, G.; MANZOLLI, J. *Metaflauta : design e performance de instrumento aumentado via suporte computacional*. Proceedings of the 15th Brazilian Symposium on Computer Music. Anais...Campinas: 2015
- RIMOLDI, G.; MANZOLLI, J. *Medidas de quantificação recorrência: uma proposta de análise para Audible Ecosystems de Agostino Di Scipio*. XXVI Congresso da Associação Nacional de Pesquisa e Pós-Graduação em Música - Belo Horizonte. Anais...2016
- SOLOMOS, M. *De la musique au son: l'émergence du son dans la musique des XXe-XXIe siècles*. [s.l.] Presses universitaires de Rennes, 2013.
- TRAUBE, C.; DEPALLE, P.; WANDERLEY, M. M. Indirect Acquisition of Instrumental Gesture Based on Signal, Physical and Perceptual Information. *Proceedings of the 2003 International Conference on New Interfaces for Musical Expression (NIME-*

03), p. 42–47, 2003.

TZANETAKIS, G.; COOK, P. Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing*, v. 10, n. 5, p. 293–302, 2002.