Visualização computacional de música com suporte à discriminação de elementos de teoria musical

Gabriel Dias Cantareira

| SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP |
|--------------------------------------|
| Data de Depósito: |
| Assinatura: |

Visualização computacional de música com suporte à discriminação de elementos de teoria musical

Gabriel Dias Cantareira

Orientador: Prof. Dr. Fernando Vieira Paulovich

Dissertação apresentada ao Instituto de Ciências Matemáticas e Computação – ICMC-USP, como parte dos requisitos para a obtenção do título de Mestre em Ciências - Ciências de Computação e Matemática Computacional. VERSÃO REVISADA

USP – São Carlos Fevereiro de 2015

C229v

Cantareira, Gabriel Dias

Visualização computacional de música com suporte à discriminação de elementos de teoria musical / Gabriel Dias Cantareira; orientador Fernando Vieira Paulovich. -- São Carlos, 2015.

87 p.

Dissertação (Mestrado - Programa de Pós-Graduação em Ciências de Computação e Matemática Computacional) -- Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2015.

1. Música. 2. Visualização de Informação. I. Paulovich, Fernando Vieira, orient. II. Título.



Agradecimentos

Ao meu orientador, Prof. Dr. Fernando Vieira Paulovich, pelo incentivo e apoio no desenvolvimento do trabalho.

Ao Instituto de Ciências Matemáticas e de Computação (ICMC-USP) e seus professores e funcionários, por tornarem o trabalho possível.

A pesquisa descrita neste documento foi financiada pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) e pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Abstract

Information visualization is an expanding research field due to its offering of novel approaches to analyze data of great size or complexity, referring to many techniques and tools in order to offer ways to interact and explore data sets to find important or useful information. Music is a domain of high complexity and hard to study and analyze by computer due to its sometimes subjective features, dependant of human interpretation. Although many research initiatives have been published regarding this subject recently, most of the music-related information visualization applications tend to analyze datasets composed by many different musical pieces, aiming to classify or group the data according to certain criteria. Thus, visualization of the information contained in a single musical piece is an area that still could be better explored, especially regarding to the comprehension of the musical information involved – information extracted by a musician by reading musical scores. This document reports the development of a novel approach to musical data visualization based on electric guitar melodies, capable of showing elements such as harmony, melody and timing variations, aiming to aid a musician in the task of understanding such data.

Resumo

A visualização computacional de informação é um campo em expansão por oferecer meios de se interpretar e analisar vários tipos de dados em grande quantidade e/ou de grande complexidade, compreendendo diversas técnicas e ferramentas para fornecer a um usuário formas de interagir e explorar conjuntos de dados a fim de se obter informações úteis ou importantes. A música, por sua vez, é um domínio complexo e de difícil estudo sob o ponto de vista computacional devido à análise de seu conteúdo possuir caráter muitas vezes subjetivo e dependente da interpretação humana. Embora vários trabalhos tenham sido publicados a respeito do assunto nos últimos anos, a maior parte das aplicações de visualização de informação relativas a música tende a analisar conjuntos de composições musicais a fim de agrupar ou classificar dados de acordo com algum tipo de critério. Assim, a visualização das informações contidas em uma única peça musical por si só é uma área que ainda pode ser melhor explorada, sobretudo visando compreender a informação musical envolvida - o conteúdo extraído por um músico a partir de partituras e tablaturas. Esta dissertação relata o desenvolvimento de uma abordagem para visualização de dados musicais referentes a melodias em guitarra, com a capacidade de exibir elementos como variações de harmonia, melodia e tempo, tendo como objetivo auxiliar um músico (ou aprendiz de músico) na tarefa de interpretar tais dados.

Sumário

| 1. Introdução | 1 |
|---|----|
| 2. Revisão Bibliográfica | 4 |
| 2.1. Conceitos de teoria musical | 5 |
| 2.1.1. Notas e tons | 5 |
| 2.1.2. Ritmo e tempo | 8 |
| 2.1.3. Escalas e intervalos | 10 |
| 2.1.4. Acordes e harmonia | 13 |
| 2.1.5. Estabilidade e Instabilidade | 16 |
| 2.1.6. Teoria musical e visualização de informação | 17 |
| 2.2. Processamento e análise de dados musicais | 17 |
| 2.2.1. Obtenção de descritores e características | 18 |
| 2.2.2. Decodificação de sinal em elementos musicais | 19 |
| 2.2.3. Extração de modelos | 20 |
| 2.3. Visualização de composições musicais | 22 |
| 2.3.1. Visualização de tons e notas | 22 |
| 2.3.1.1. Condensação de partituras | 26 |
| 2.3.2. Visualização de acordes e harmonia | 29 |
| 2.3.2.1. <i>Isochords</i> | 33 |
| 2.3.3. Visualização de melodia e contexto | 36 |
| 2.3.4. Visualização por grupos ou estruturas | 38 |

| | 2.4. Considerações Finais | 43 |
|----|---|------|
| 3. | Visualizando Elementos de Teoria Musical | 45 |
| | 3.1. Motivações e Objetivo | 45 |
| | 3.2. Visualizando Elementos da Teoria Musical | 47 |
| | 3.2.1. Visualização de dados multinível | 49 |
| | 3.2.2. Visualização de dados variantes no tempo | 50 |
| | 3.2.3. Visualização de texto | 51 |
| | 3.3. Estrutura do arcabouço de visualização | 53 |
| | 3.4. Representações visuais | 54 |
| | 3.4.1. Compassos | 55 |
| | 3.4.2. Acordes e Intervalos | 55 |
| | 3.4.3. Visão de detalhe | 57 |
| | 3.4.4. Visão geral | 59 |
| | 3.5. MOSHViz | 64 |
| | 3.5.1. Interagindo com a visualização | 67 |
| 4. | . Resultados | . 68 |
| | 4.1. Bases de dados e testes | 68 |
| | 4.2. Características e parâmetros | 69 |
| | 4.3. Comparação entre músicas | 75 |
| 5. | Conclusões e Trabalhos Futuros | 80 |
| R | eferências | 82 |

Lista de Figuras

| Figura 1: Notas musicais nas teclas de piano. | 7 |
|---|-----------|
| Figura 2: Relação entre as diferentes durações de nota | 9 |
| Figura 3: Escala de dó maior (C-major). | 12 |
| Figura 4: Escala de lá menor (A-minor). | 13 |
| Figura 5: Visualização Piano Roll usando a ferramenta MIDITrail | 24 |
| Figura 6: Music Animation Machine | 25 |
| Figura 7: Visualização de partitura usando o modelo BRASS | 27 |
| Figura 8: Colorscore | 28 |
| Figura 9: Círculo de terças associado a uma roda de cores | 30 |
| Figura 10: Representações de cores para tons | 31 |
| Figura 11: Espaço bidimensional de tons de Lerdahl | 32 |
| Figura 12: Visualização de progressões e distribuições tonais | |
| Figura 13: Círculo de quintas com símbolos no grid e grid Tonnetz | 35 |
| Figura 14: Exemplos de acordes usando Isochords | 35 |
| Figura 15: ImproViz. Melodic Landscapes e Harmonic Palettes | 37 |
| Figura 16: Layer Braid e Theme Fabric | 40 |
| Figura 17: Shape of Sound | 41 |
| Figura 18: Visualização do Comp-i | 42 |
| Figura 19: Exemplo de visualização de detalhe | 59 |
| Figura 20: Exemplo do modelo de segmentos usado na visualização da min | iatura 61 |
| Figura 21: Pontos de interesse na miniatura | 62 |
| Figura 22: Barras de complexidade, repetição e variação de intervalos | 63 |
| Figura 23: Diagrama representando o pipeline de visualização da ferrament | a 65 |
| Figura 24: Janela da ferramenta | 66 |

| Figura 25: Diferentes modelos de coloração de notas | . 71 |
|---|------|
| Figura 26: Equivalência entre miniatura e visão em detalhe | . 72 |
| Figura 27: Equivalência entre miniatura e visão em detalhe 2 | . 72 |
| Figura 28: Mudança de escala para agrupamento de compassos | . 75 |
| Figura 29: Visualização miniatura com identificação de segmentos | . 76 |
| Figura 30: Visualização da estrutura de composições musicais usando a miniatura | . 77 |
| Figura 31: Visualização de guitarras solo | . 77 |
| Figura 32: Diferenças entre estruturas de composições | . 78 |
| Figura 33: Janela de visualização exibindo um conjunto de dados | . 79 |

Lista de Tabelas

| Tabela 1: Resumo de intervalos. | 15 |
|--|----|
| Tabela 2: Comparação de técnicas de visualização de música | 44 |

Lista de Siglas

BPM Beats per Minute

MAM Music Animation Machine

MIDI Music Instrument Digital interface

SVM Support Vector Machine

Capítulo 1. Introdução

A Visualização de Informação é uma subárea dentro do escopo da grande área de visualização computacional que lida com métodos e técnicas para processar e exibir visualmente grandes conjuntos de dados a fim de prover meios de se analisar e extrair informações importantes a partir deles [1]. O fato de o tamanho e a complexidade de tais conjuntos se mostrarem cada vez maiores, devido à modernidade da tecnologia e consequente maior facilidade de obtenção, geração e armazenamento de informações, faz com que a demanda por novas formas de se analisar esses dados seja expressiva [2].

Com o intuito de atender esta demanda e possibilitar a análise das informações reunidas, vem sendo desenvolvidas e estudadas diversas novas abordagens para permitir a um usuário a exibição e exploração de espaços de dados, usando-se de aspectos diferentes de nossa capacidade cognitiva visual (distribuição de cores, posicionamento geométrico, percepção de padrões ou distinção de grupos por proximidade são exemplos). Assim, a Visualização de Informação pode ser considerada uma área em expansão e com constante desenvolvimento científico.

Entre os vários tipos de dados e informações analisados usando técnicas de visualização de informação, a música (tanto em formato de notação musical quanto em gravações de áudio) é um domínio de informação que desperta interesse. Analisar composições musicais e exibir informações nelas contidas que sejam pertinentes a algum tipo de usuário se mostra uma tarefa complexa devido ao caráter frequentemente subjetivo da interpretação da mente humana ante à percepção da música e à dificuldade de se associar elementos encontrados numa análise computacional aos elementos mais perceptíveis e identificáveis por seres humanos ao ouvir uma composição musical [3,4].

É possível observar avanços expressivos realizados na área de visualização e mineração de dados para grandes conjuntos referentes a gravações musicais, com diversos objetivos: a identificação de similaridades (sob vários critérios, como melodia, ritmo, tempo ou timbre de instrumentos), formação de padrões, localização de elementos de importância, etc. A análise de grandes bases de dados de música é realizada por vários programas comerciais encontrados no dia-a-dia, como o SoundHound[5], o MusicBox[6] ou mesmo visualizações para coleções do iTunes, como o Planetary[7]. Estas ferramentas são capazes de interagir com bases de tamanho considerável, tanto para visualização e navegação como para busca, ainda que muitas das análises visuais encontradas para coleções de música lidem com metadados ao invés da informação musical propriamente dita.

Porém, conforme descrito acima, estas aplicações são focadas principalmente na análise de relações entre diferentes músicas, tratadas como instâncias de um grande conjunto de dados. A observação de padrões, fluxos e similaridades contidos dentro de uma única peça ou composição musical não é abordada com tanta frequência pela literatura, embora um grande grupo de usuários, como os músicos e instrumentistas (experientes ou em treinamento), poderiam fazer uso desse ferramental. Para estes usuários, uma partitura ou gravação por si só pode ser considerada um conjunto de dados, que geralmente corresponde a uma grande quantidade de informações sequenciais.

Ainda que o campo de conhecimento musical possua regras estabelecidas para interpretar o conteúdo de uma partitura ou gravação musical, o conhecimento atrelado, como harmonia, organização rítmica e estrutura de fraseamento, nem sempre possui fácil acesso para usuários que precisam de uma interpretação rápida dos dados ou que não possuam treinamento adequado para realizar a análise necessária. Portanto, a ideia de utilizar auxílios visuais para a extração de conhecimento a partir de um conteúdo que normalmente necessita da análise minuciosa de cada instância pode se encaixar neste contexto.

Tendo em mente este domínio de aplicação, muitas abordagens podem ser utilizadas no que diz respeito a representar visualmente informações pertinentes ao conhecimento musical. Relações melódicas e harmônicas entre notas podem ser exploradas computacionalmente, de forma a identificar padrões ou fluxos não tão facilmente notados à primeira vista. Ao organizar elementos musicais em estruturas,

também é possível usar níveis diferentes de abstração, oferecendo uma visão geral do comportamento dos dados ao longo do tempo, permitindo ao usuário a compreensão de vários aspectos importantes da sonoridade de uma composição mesmo com uma observação superficial.

Este projeto de mestrado desenvolveu uma nova abordagem de visualização musical aplicando conceitos e técnicas de visualização de informação, como a visualização multinível e hierárquica de dados temporais, sobre o campo da teoria musical, com foco na interpretação de composições musicais e melodias em guitarra elétrica. Foi elaborado um modelo de visualização que fornece uma interpretação dos dados contidos em uma partitura ou arquivo MIDI, visando facilitar ao usuário a compreensão de elementos como harmonia, melodia, ritmo e tempo, também como as relações que estes elementos possuem entre si.

É importante notar que este projeto não propõe formas de substituir métodos de representação tradicionais, como partituras e tablaturas, mas sim fornecer ferramentas visuais que dêem suporte a tais métodos, auxiliando o usuário a compreender seu conteúdo e tirar conclusões a respeito da informação fornecida, com o embasamento do conhecimento musical.

Nos capítulos seguintes, é descrito em detalhes o curso de ação tomado, bem como os conceitos envolvidos, pesquisas realizadas e discussão da metodologia empregada. O Capítulo 2 apresenta trabalhos relacionados e conceitos relevantes ao projeto descrito neste documento. O Capítulo 3 descreve em detalhes as abordagens, técnicas e ferramental desenvolvido no projeto. O Capítulo 4 discute resultados obtidos e testes realizados e o Capítulo 5 apresenta conclusões, bem como direções futuras de pesquisa.

Capítulo 2. Revisão Bibliográfica

Neste capítulo, são abordados conceitos e técnicas relacionados ao projeto realizado, necessários para seu entendimento, bem como técnicas que definem o estadoda-arte em visualização de composições musicais. Sua divisão é realizada em três seções, lidando com diferentes aspectos do problema estudado.

A primeira seção faz uma descrição de conceitos básicos de teoria musical importantes ao desenvolvimento do projeto: harmonia, acordes, escalas, estabilidade, entre outros. A familiarização com tais conceitos é fundamental para o entendimento do trabalho realizado e para a interpretação dos dados tratados pelas técnicas de visualização estudadas. A segunda seção fará uma breve discussão sobre o processamento de dados referentes a composições musicais, sob diversos pontos de vista: busca de similaridades, análise de elementos melódicos importantes ou a identificação de descritores que podem ser usados em técnicas de visualização. A terceira seção descreve técnicas para visualização de uma composição musical, sob diversas visões e com diversos objetivos, juntamente com discussões sobre quais características ou elementos destas técnicas foram usados no contexto do problema atacado neste projeto.

2.1. Conceitos de teoria musical

Entende-se por "teoria musical" o estudo de como a música funciona: a identificação de padrões e estruturas em composições musicais, a observação de como tais elementos interagem com nossa percepção dos sons e como eles podem ser usados para criar novos sons (ou conjuntos de sons) no processo de criação e composição.

Ao lidar com sons, é possível observar, na análise de ondas sonoras, diversas relações matemáticas entre os elementos de estruturas definidas em teoria musical. Isto implica na existência de modelos que podem nos fornecer algumas informações sobre as relações entre diferentes sons e as sensações que temos ao ouví-los simultaneamente ou em série, ainda que não sejam o suficiente para representar com exatidão a compreensão humana de uma música.

Devido à importância de certos conceitos ao trabalho a ser realizado e de sua distância da área da computação, esta seção fará uma breve descrição de alguns elementos importantes de teoria musical que serão trabalhados no restante da revisão bibliográfica e na descrição de metologia e resultados deste projeto.

O conteúdo elaborado para esta seção foi escrito a partir do estudo realizado sobre três livros, [8], [9] e [10]. Exceto se especificado o contrário, as afirmações realizadas adiante são discutidas nestes trabalhos.

2.1.1. Notas e tons

Há um seleto grupo de vibrações e ondas físicas que podem ser transmitidos por sólidos ou fluidos que chamamos de *sons*. Estas vibrações são captadas por nosso sistema auditivo, transformadas em sinais elétricos em nosso sistema nervoso e então processadas pelo cérebro. Graças a isso, somos capazes de perceber o mundo ao nosso redor, nos comunicar uns com os outros e produzir música, entre outras habilidades.

A principal característica que diferencia o que percebemos como música dos outros tipos de sons é a presença de ordem: a onda sonora emitida por um instrumento

musical não-percussivo é regular e constante, ao invés de caótica; quaisquer sons, quando emitidos em intervalos regulares e padronizados, estabelecem ritmo. Ondas regulares e periódicas geradas por instrumentos ou mesmo por nossa voz podem ser definidas como tons musicais. Grande parte do que entendemos por música está nas combinações de tons, que compõem melodia e harmonia.

Musicalmente, os três elementos mais importantes de uma onda sonora correspondente a um tom musical são *frequência*, *amplitude* e *forma de onda*. A frequência, usualmente medida em hertz (Hz), determina o quão grave ou agudo é o som, o que também pode enquadrá-lo em uma nota musical. A audição humana é em média capaz de perceber frequências de 20Hz a 20KHz, sendo grande parte deste intervalo utilizado normalmente em composições musicais.

As notas musicais são determinadas por frequências específicas em ondas sonoras, sendo agrupadas por relações de multiplicidade. Tons musicais cujas frequências tem relações equivalentes a potências de 2 possuem sons similares, o que identificamos como notas equivalentes. Assim, a progressão de notas musicais é um ciclo que retorna a uma nota similar cada vez que a frequência do som é dobrada: na música ocidental, a este ciclo se dá o nome de *oitava*, devido à presença de oito notas musicais entre a primeira nota de um ciclo até a primeira do próximo ciclo. Uma oitava também pode ser dividida em seis *tons* ou doze *semitons*: desta forma, as frequências dos semitons podem ser relacionadas a uma progressão geométrica de razão $2^{1/12}$. A afinação de instrumentos musicais também é feita em função destas frequências: eles são configurados de forma que, quando tocados em conjunto, sejam capazes de produzir sons similares para as mesmas notas. A nota musical mais comumente usada como padrão de afinação é identificada pela letra A (Lá), sendo sua frequência usualmente dada por 440Hz e quaisquer multiplicações deste valor por potências de 2 (110Hz, 220Hz, 880Hz, etc).

A nomenclatura utilizada para notas musicais é baseada na escala de C (Dó) maior (descrita mais adiante), que pode ser visualizada nas teclas de piano: cada tecla corresponde a um dos doze semitons de uma oitava. Uma letra é atribuída a cada uma das sete notas musicais: C, D, E, F, G, A, B (Dó, Ré, Mi, Fá, Sol, Lá, Si, respectivamente). Estas notas são representadas pelas teclas brancas de piano. As teclas pretas compreendem os cinco semitons restantes: a distribuição desigual de teclas brancas e pretas se dá devido ao modelo observado na escala mencionada, significando que a

diferença entre as notas descritas pode ser de um ou dois semitons. Os semitons representados pelas teclas em preto são expressados adicionando símbolos às notas já definidas: o símbolo "#" indica uma nota um semitom mais aguda (maior frequência), enquanto o símbolo "b" indica uma nota um semitom mais grave (menor frequência). Por exemplo, a tecla preta entre as notas C e D pode ser representada tanto como C# quanto como Db. Nos casos em que não há teclas pretas, a transição é direta: a representação E# é equivalente à nota F e, de forma análoga, Fb é equivalente a E.

A **Figura 1**, obtida em [8], faz uma descrição das notas musicais e seus respectivos símbolos. Cada tecla branca está marcada pela letra de sua nota correspondente, enquanto as teclas pretas são identificadas em função das notas imediatamente anterior e posterior. Este padrão representa uma oitava, e é repetido tanto para a esquerda quanto para a direita uma vez que a nota C é alcançada novamente.

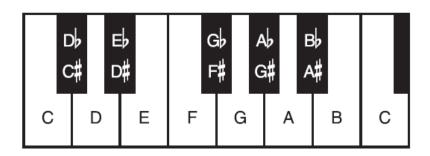


Figura 1: Notas musicais nas teclas de piano. São mostrados os doze semitons, organizados em sete notas da escala de C maior em branco e os cinco semitons restantes em preto. Imagem obtida em [8].

A amplitude da onda sonora, outra característica importante, indica a intensidade (ou "volume") com que o som em questão será absorvido por nossa audição. A intensidade de sons é medida em decibéis (Db), sendo esta uma escala logarítmica. Uma vez que o volume de sons é normalmente alterável em dispositivos de reprodução de música, a aplicação deste elemento no contexto musical diz respeito principalmente a como os volumes de cada som emitido durante uma música se relacionam, como por exemplo a diferença de amplitude entre dois instrumentos diferentes ou variações de intensidade em momentos específicos de uma composição.

A forma de onda, por sua vez, determina o timbre do instrumento ou a qualidade do tom musical (também chamada de *cor* do tom). Esta característica nos permite reconhecer as diferenças entre sons com a mesma frequência, como os provenientes de diferentes instrumentos. Isto se deve ao fato de, além da frequência equivalente à nota emitida (chamada *frequência fundamental*), instrumentos também emitem outras frequências, chamadas parciais (ou *modos de vibração*). Sintetizadores eletrônicos são capazes de gerar sons similares a diversos tipos de instrumentos por meio da combinação de diferentes frequências em torno de um mesmo tom.

2.1.2. Ritmo e tempo

Todos os sons emitidos em uma música podem ser associados a informações temporais: em que momento são iniciados e qual sua duração. A combinação destas informações para uma sucessão de sons ocorrendo ao longo do tempo de música compõe o que chamamos de ritmo.

O ritmo de uma música muitas vezes é definido por instrumentos de percussão, como uma bateria. Conforme descrito anteriormente, o conceito de música é intimamente ligado a organização e repetição, portanto há formas muito bem definidas de se estabelecer marcações de tempo: praticamente toda música possui padrões temporais muito nítidos, mesmo quando se trata de composições repletas de mudanças estruturais e variações sonoras.

O aspecto mais básico da marcação de tempo é a definição de compassos: a métrica temporal de música é definida por um ciclo com um número determinado de batidas, realizadas em intervalos de tempo iguais (*tempo*, geralmente expressado em batidas por minuto – BPM). Este padrão se repete, muitas vezes por toda a música, embora não seja difícil observar composições que alterem o que se entende por sua *fórmula de compasso* em momentos específicos: como a repetição de um padrão gera expectativa, tal mudança frequentemente confunde ou surpreende o ouvinte, podendo ser usada como uma ferramenta artística.

Numa partitura, a indicação de quais notas devem ser tocadas também vem acompanhada de por quanto tempo elas devem ser tocadas. A representação usa uma notação relativa que identifica a duração de uma nota quando comparada às outras, sempre em potências de 2: uma nota inteira (semibreve) dura o equivalente a duas meias notas (mínimas), quatro quartos de nota (semínimas) e assim por diante. A **Figura 2**, obtida em [8], mostra a relação entre as diferentes durações de notas, juntamente com os seus nomes e símbolos usados na notação musical.

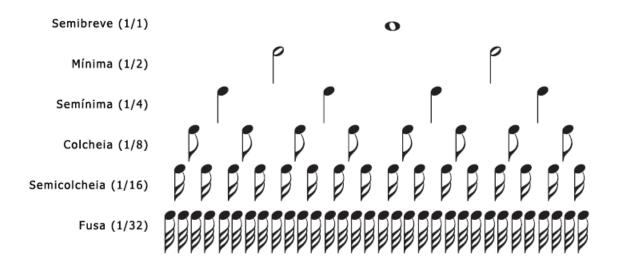


Figura 2: Relação entre as diferentes durações de nota. A duração de cada tipo de nota é expressa em função da duração das demais. Adaptação de imagem encontrada em [8].

O modo como os ciclos são calculados é descrito por uma fórmula de compasso. A notação geralmente usada para expressá-la corresponde a dois números, um numerador e um denominador. Eles podem ser representados tanto um sobre o outro como usando um sinal de divisão. O numerador indica quantas batidas ou tempos serão usadas por ciclo (podendo ser qualquer número) e o denominador indica qual a duração de cada batida (ele deve corresponder a uma subdivisão de compasso, sendo portanto uma potência de 2). Por exemplo, a notação 3/4 indica que cada ciclo possui três tempos, sendo que cada um deles equivale a um quarto de nota (semínima).

Para um dado número de tempos contidos em um compasso, ocorre uma subdivisão entre tempos *fortes* e *fracos*. Esta classificação indica a acentuação que cada

nota irá possuir durante a execução musical, recebendo certo destaque de acordo com sua posição dentro do compasso. Tradicionalmente, o primeiro tempo de cada compasso é sempre considerado forte, indicando uma ênfase nas primeiras notas a serem tocadas. Os tempos seguintes são então classificados conforme o número total de tempos no compasso, sendo considerados meio-fortes ou fracos. Geralmente, para um número par de tempos, o primeiro tempo da segunda metade do compasso é dito meio-forte, enquanto a compassos com tempos ímpares são analisados por meio da subdivisão de seus tempos em compassos pares e compassos ternários (três tempos), onde apenas o primeiro tempo é meio-forte.

2.1.3. Escalas e intervalos

Uma escala é uma sequência de notas musicais, geralmente dentro de uma mesma oitava e repetindo-se nas oitavas subsequentes. Composições musicais são construídas em torno de uma ou mais escalas, extraindo a partir delas as notas a serem usadas em acordes ou melodias. Todo tipo de música faz uso de escalas, ainda que elas possam não ser convencionais.

É importante para um músico conhecer uma grande variedade de escalas, tanto por implicar numa maior gama de possibilidades no momento de criar uma composição quanto por possibilitar a identificação e classificação de frases e acordes quando ele analisa uma peça criada por outro compositor. As escalas também são muito importantes para improvisação: de modo geral, pode-se dizer que o conhecimento da escala usada em uma música implica na habilidade de se adicionar sons consoantes a ela.

Escalas são comumente descritas por uma nota sobre a qual a escala é construída (chamada de raiz ou tônica) e uma sequência de distâncias em semitons que determinam quais serão as notas subsequentes. As magnitudes destas distâncias e a ordem em que são colocadas determinam as características sonoras de uma escala. Uma classificação frequentemente usada para escalas diz respeito à quantidade de passos ou classes de tons nela presentes, como Heptatônicas (sete tons por oitava, notação na qual o próprio nome

oitava é baseado e a mais comumente utilizada na música ocidental) ou Pentatônicas (cinco tons por oitava, utilizada frequentemente em *blues* ou *jazz*).

A escala mais básica que se pode definir na música ocidental é a chamada escala cromática, que é constituída por uma sequência de doze notas com apenas um semitom de distância entre si (listando a distância em semitons necessária para avançar cada passo da escala, pode-se representá-la por 1-1-1-1-1-1-1-1-1-1-1). Basicamente, esta escala compreende todos os semitons de uma oitava, cobrindo portanto toda a gama de notas musicais. Embora a escala cromática muitas vezes não seja considerada uma escala propriamente dita, ela oferece todos os tons musicais que podem ser selecionados para uso em outras escalas.

Uma das escalas mais comuns na música ocidental é a escala maior, uma heptatônica. A escala maior é definida por uma nota tônica e uma progressão de semitons na forma 2-2-1-2-2-1, totalizando os doze semitons de uma oitava e gerando portanto um ciclo. A denominação *maior* é dada devido à interação harmônica entre suas notas, que será descrita em detalhes mais adiante. Esta escala, quando construída com base na nota C (dó), é a sequência mais tradicional na música e a sequência na qual o teclado do piano é baseado: conforme descrito anteriormente, as teclas brancas são as notas pertencentes à escala, enquanto as notas pretas são os tons restantes dentre os doze existentes na oitava, estando ausentes entre as notas cuja distância é de apenas um semitom.

A **Figura 3**, obtida em [8], traz uma representação da escala maior em relação à frequência sonora de suas notas. O estilo de notação, conhecido como *piano roll*, será descrito mais adiante. A imagem é um gráfico simples cujo eixo vertical corresponde à frequência dos sons (*pitch*), que é a variação responsável por separar as notas musicais. Desta forma, usando a escala logarítmica adequada, o eixo pode ser representado como um teclado de piano, em que cada valor de frequência pode ser associado a uma nota. Nesta visualização, o eixo horizontal corresponde ao tempo, sendo ilustrada portanto uma escala de C-maior tocada uma nota por vez.

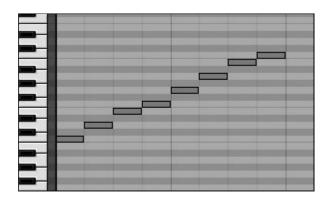


Figura 3: Escala de dó maior (C-major). Cada linha do diagrama representa um semitom. As notas da escala são tocadas em ordem, havendo saltos nos semitons não pertencentes a ela (teclas pretas). Imagem obtida em [8].

Outra escala comum é a chamada escala menor, também heptatônica. Ela se baseia nos mesmos conceitos harmônicos em que a escala maior é montada, mas sua terceira, sexta e sétima notas são colocadas um semitom abaixo, resultando numa progressão na forma 2-1-2-2-1-2-2. Devido ao fato de as escalas se repetirem ciclicamente a cada oitava, as escalas maior e menor podem possuir as mesmas notas se forem construídas em torno das tônicas certas: por exemplo, a escala maior em C e a escala menor em A possuem a mesma sequência de notas quando repetidas em todas as oitavas. Uma outra versão comumente usada é a chamada escala menor harmônica, onde a sétima nota permanece na mesma posição da escala maior, gerando uma progressão na forma 2-1-2-2-1-3-1.

A **Figura 4**, obtida em [8], mostra um exemplo gráfico de escala menor. De modo similar à imagem anterior, o eixo vertical representa as notas (e, portanto, a frequência dos sons) e o eixo horizontal representa o tempo em que as notas são tocadas. É possível observar a sobreposição das escalas maior e menor nesta imagem, em relação à imagem anterior: começando a partir da nota A, mesmo usando intervalos diferentes, a escala menor usa apenas as teclas brancas, o que significa que possui as mesmas notas da escala de C-maior.

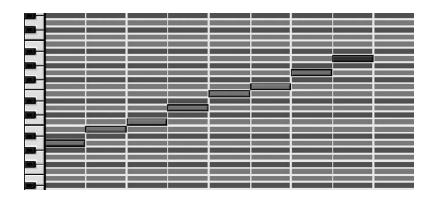


Figura 4: Escala de lá menor (A-minor). Verticalmente, cada barra representa um semitom. Imagem obtida em [8].

Usando a notação baseada em semitons, a distância entre uma nota e outra determina a diferença de frequência entre o som das duas notas. Esta distância, chamada de intervalo, é frequentemente usada com base na escala para determinar a relação harmônica ou melódica entre notas de uma composição musical: saber quais as posições ocupadas por cada nota dentro da escala utilizada auxilia no entendimento de interações entre os sons em termos de consoância e dissonância, entre outros aspectos.

2.1.4. Acordes e harmonia

Enquanto a melodia diz respeito a notas musicais tocadas em sequência, *harmonia* é o nome dado à consoância resultante do uso simultâneo destas notas. O estudo de propriedades harmônicas envolve a construção e progressão de acordes, assim como os princípios que levam à sua ligação. Desta forma, um dos focos principais deste estudo é a observação de intervalos harmônicos, ou seja, intervalos entre notas tocadas ao mesmo tempo.

Um ponto importante para a observação de intervalos é que, sejam eles harmônicos ou melódicos, a escala a ser usada como base é a mesma. Isto significa que, numa música, as notas usadas em acompanhamentos, bases e melodias principais são selecionadas a partir de um mesmo conjunto, proveniente da escala usada naquela seção da composição.

Os intervalos geralmente são referenciados em função das posições ocupadas pelas notas na escala maior, incluindo a nota inicial. Por exemplo, na escala maior de C, há cinco posições entre as notas B e F, contando com a inicial: B, C, D, E e F. Portanto, o intervalo entre B e F é chamado de *quinta*. Intervalos maiores que uma oitava são chamados compostos, mas devido à natureza cíclica das notas sua relação é similar: um intervalo de onze posições pode ser chamado de *quarta composta*.

Quando duas notas são tocadas em conjunto, o ouvido humano percebe uma relação única entre elas: devido a propriedades em suas frequências, nosso cérebro pode identificá-las como similares ou harmoniosas, sendo chamadas de consoantes ou concordantes. Caso contrário, transmitirão alguma sensação de irregularidade e são chamadas dissoantes ou discordantes. Todo o estudo de harmonia se baseia nestas relações.

A relação entre as frequências dos sons e a interpretação deles por nosso sistema auditivo é, de certa forma, simples: quanto mais as ondas sonoras se assemelham ou se complementam, mais concordantes os sons parecem. Desta forma, a relação mais consoante possível é entre uma nota e ela própria (primeiro intervalo), seguida pelo intervalo de uma oitava inteira (o dobro da frequência). De modo geral, quanto mais simples for a razão entre as frequências das duas notas, maior a concordância.

Em seguida, podem ser observados dois intervalos com razões consideravelmente simples. Enquanto oitavas possuem uma razão de frequência de 2:1, quintas e quartas possuem razões de 3:2 e 4:3, respectivamente. Estes intervalos são chamados de *perfeitos* e conhecidos na música ocidental como um padrão de consoância. Os intervalos de quinta e oitava são particularmente conhecidos pela sonoridade forte e, junto com a nota tônica, compõem os chamados *power chords*, popularizados por composições famosas de *rock*.

Além das consoâncias perfeitas, tradicionalmente, os intervalos de terças e sextas são considerados concordantes devido à razão entre frequências ser considerada simples o bastante. Ao colocar em conjunto uma nota e sua respectiva terça ou sexta em todas as posições da escala maior, porém, é perceptível que há variações na distância entre elas: isto ocorre devido à irregularidade da progressão, ora de dois semitons, ora de apenas um, o que não ocorre com quintas e quartas (com a exceção de intervalos aumentados, que não serão discutidos aqui). O resultado são terças com ora três, ora quatro semitons de

distância e sextas com oito ou nove. Graças a esta variação, é possível existir o que chamamos de terças e sextas *maiores* e *menores*. A própria escala menor harmônica é obtida a partir da escala maior, reduzindo terças e sextas em um semitom. Estas concordâncias são denominadas *imperfeitas*.

Por fim, segundas e sétimas são intervalos caracterizados como muito discrepantes e portanto discordantes. De forma similar às consoantes imperfeitas, também ocorrem em maiores e menores. A **Tabela 1**, obtida a partir de [8], mostra um resumo dos intervalos observados, compreendendo as doze variações de semitom correspondentes a uma oitava.

Tendo em mente tais intervalos, um acorde é definido com uma combinação de duas ou mais notas tocadas ao mesmo tempo. A classificação do acorde e sua denominação são obtidos a partir dos intervalos nele contidos, sempre no contexto de uma escala. Um dos tipos de acorde mais comumente utilizados são as tríades compostas pela nota principal, uma terça e uma quinta. A relação entre os sons é vista como concordante devido ao fato de terças e quintas também possuírem intervalos concordantes entre si (o que não acontece com terças e quartas, por exemplo). Quando a terça usada na tríade for uma terça menor, define-se o acorde como menor. De forma análoga, acordes maiores possuem terças maiores.

| | | | | Número de |
|-----------|------------|-----------|----------------|-----------|
| Intervalo | Tipo | Modo | Qualidade | Semitons |
| Primeira | Perfeita | Perfeita | Concordância | 0 |
| Segunda | Imperfeita | Menor | Discordância | 1 |
| Segunda | Imperfeita | Maior | Discordância | 2 |
| Terça | Imprefeita | Menor | Concordância | 3 |
| Terça | Imperfeita | Maior | Concordância | 4 |
| Quarta | Perfeita | Perfeita | Conc. Perfeita | 5 |
| Quarta | Cromática | Aumentada | Discordância | 6 |
| Quinta | Perfeita | Perfeita | Conc. Perfeita | 7 |
| Sexta | Imperfeita | Menor | Concordância | 8 |
| Sexta | Imperfeita | Maior | Concordância | 9 |
| Sétima | Impefeita | Menor | Discordância | 10 |
| Sétima | Imperfeita | Maior | Discordância | 11 |
| Oitava | Perfeita | Perfeita | Conc. Perfeita | 12 |

Tabela 1: Resumo de intervalos

Acordes maiores e menores são complementares: de acordo com a nota principal do acorde e a escala que está sendo empregada, será necessário utilizar uma versão maior ou menor da tríade para que a harmonia da música naquela escala seja mantida (para a obtenção de sons consoantes, as notas do acorde precisam sempre corresponder às notas presentes na escala). A este conjunto de acordes se dá o nome de *campo harmônico*, que varia com a escala.

2.1.5. Estabilidade e Instabilidade

Conforme discutido anteriormente, a ideia de organização e estruturação engloba vários dos conceitos centrais à percepção de música. Praticamente toda composição musical estabelece certos padrões, como escalas ou fórmulas de compasso a serem seguidas. Conforme a composição é executada, estes padrões tendem a gerar expectativas, que podem ou não ser atendidas.

Na literatura, muitas vezes combinações de notas ou padrões temporais podem ser descritos como *estáveis* ou *instáveis*. Nossa percepção da estabilidade de um trecho de música é dada por como ele manipula as expectativas e o quão coesa é a sua sonoridade. Este conceito é frequentemente usado para induzir sensações em uma composição musical, como no caso de trilhas sonoras.

A noção de estabilidade ou instabilidade pode ser transmitida por diversos fatores. A estabilidade é gerada pela manutenção de ciclos, cumprimento de expectativas, transições com sons consoantes, características que reforcem a idéia de que há uma estrutura no som a ser respeitada. A instabilidade é gerada por aspectos que apresentem idéias contrárias: quebra de expectativas e ciclos, mudanças súbitas de andamento ou fórmula de compasso, número desigual de repetições, sons dissoantes em geral [56].

Usualmente, intervalos de consoância forte são ditos estáveis. Porém, muitas vezes, a estabilidade percebida em uma composição musical é relativa: intervalos tipicamente dissoantes são considerados instáveis, mas quando aparecem num contexto em que sua presença é esperada, como a harmonia usada em *jazz*, se tornam parte da norma, acostumando o ouvinte a percebê-los como relativamente estáveis.

Conforme dito anteriormente, a idéia de estabilidade é obtida a partir da percepção de estrutura e organização, em todos os níveis: instrumentos desafinados ou desregulados, produzindo ondas sonoras irregulares, também podem contribuir para tornar a percepção de um trecho musical como instável.

2.1.6. Teoria musical e visualização de informação

O conteúdo abordado nesta seção corresponde a uma parcela do que foi estudado para a realização do projeto em relação a teoria musical, tendo como objetivo prover algum conhecimento básico sobre a área, especialmente no que diz respeito a conceitos utilizados na elaboração deste trabalho. Espera-se que, com o conteúdo aqui descrito, termos usados e metodologias discutidas possam ser melhor compreendidos.

A gama de conhecimentos englobados por teoria musical que podem ser aproveitados em ferramentas de visualização é muito mais extensa do que o conteúdo aqui discutido, sendo um domínio ainda passível de estudo. Por exemplo, outras características que poderiam ser exploradas neste contexto incluem temas, outros tipos de escalas, campos harmônicos, harmonia funcional, etc.

A seguir, será realizada uma discussão sobre técnicas de visualização referentes à representação de música com o uso de computadores, apresentando ferramentas e abordagens para fornecer auxílios visuais na tarefa de interpretar dados musicais, relacionando-os com os métodos desenvolvidos neste projeto.

2.2. Processamento e análise de dados musicais

Com o objetivo de se extrair informações adicionais ou características intrínsecas importantes, dados referentes a composições musicais são processados de diversas

formas, tanto para aplicações de visualização quanto de várias outras áreas. Isto é especialmente aplicável no caso de dados musicais em forma de sinal de áudio, ou seja, gravações de performances executadas por músicos obtidas usando receptores de som. Sinais de áudio são processados para obter vários tipos de resultados, como a redução de ruído, identificação de melodias predominantes, definição de cadências melódicas, entre outros.

Muitas vezes, visualizações são geradas sobre o resultado de tais métodos de processamento, como, por exemplo, a visualização de consoância, tonalidade ou fraseamento em composições musicais. Embora o modelo específico de processamento usado varie de acordo com o propósito e peculiaridades da técnica de visualização que o utiliza, a observação de abordagens existentes pode trazer informações importantes.

2.2.1. Obtenção de descritores e características

Um descritor (*feature*) pode ser definido como um atributo (ou conjunto de atributos) de uma instância de um conjunto de dados relevante ao problema sendo estudado, construído com uso de processamento a partir de dados de entrada [11]. Descritores geralmente contém informações que podem ser comparadas ou analisadas a fim de se obter conclusões a respeito dos dados, por exemplo fazendo uso de medidas estatísticas.

No que diz respeito a processamento de dados, a maior parte dos usos associados a descritores e características de dados musicais corresponde a aplicações que comparam músicas de alguma forma, como projeções ou classificadores. Estes descritores são geralmente usados em cálculos matemáticos que fornecem ao usuário determinados tipos de informação pertinentes à análise sendo realizada. Adicionalmente, mesmo a visualização ou análise da estrutura musical de uma única música pode ser facilitada com o auxílio de descritores: eles podem denotar o grau de importância de determinados elementos do conjunto de dados ou mesmo exibir informações adicionais que facilitem a sua compreensão.

Uma ferramenta de fácil acesso para a extração de descritores é a jAudio[12], desenvolvida em Java, que permite a extração de descritores tais como fluxo espectral harmônico, soma de batidas, momentos, entre outros. Estas características geralmente são usadas para alimentar sistemas de aprendizado de máquina, como classificadores ou algoritmos de agrupamento [13,29], mas também podem ser usadas em visualizações, como sistemas que utilizam técnicas de projeção. A ferramenta possui grande embasamento técnico, oferecendo em sua documentação estudos sobre cada descritor e seus métodos de extração.

No entanto, os descritores encontrados normalmente dizem respeito a aspectos técnicos da gravação digital de sons. A análise e obtenção de informações mais complexas, sobretudo referentes à estrutura musical dos dados, é muito difícil. Descritores como a fórmula de compasso de um trecho de uma gravação musical raramente são calculados com precisão, embora técnicas supervisionadas ofereçam certo grau de segurança [14]. Em muitas técnicas que realizam processamento de áudio, estes descritores não são calculados e acabam sendo presumidos, como ocorre em [15]. Desenvolvedores que necessitam destas informações tendem a optar por trabalhar com formatos com interpretação musical melhor definida, como partituras digitais ou instruções para sintetizadores, como ocorre neste trabalho.

2.2.2. Decodificação de sinal em elementos musicais

Por possuir análise complexa, existem várias linhas de pesquisa com o objetivo de decodificar o sinal de áudio e estruturar os elementos nele contidos sob algum tipo de visão. Quando se trata de música, a idéia de se obter uma estrutura bem definida de instruções para notas e instrumentos ao longo do tempo a partir de um arquivo de áudio (permitindo portanto extrair uma partitura a partir de uma gravação) implica numa técnica de processamento muito útil, mas que ainda parece um tanto distante de ser implementada, especialmente para músicas complexas.

Em computadores, o formato mais utilizado para representação de instruções musicais é o MIDI (*Musical Instrument Digital Interface*[16]), uma especificação para

sintetizadores que determina vários tipos de instruções que permitem desde o controle de equipamentos de som até a reprodução de uma composição musical. Um arquivo MIDI pode funcionar da mesma forma que uma partitura, provendo ao computador as instruções necessárias para reproduzir uma música usando sintetizadores e amostras de uma biblioteca de sons de instrumentos (contida no sistema e não no arquivo) em momentos específicos. Um outro exemplo de formato é o MusicXML [17], que codifica instruções musicais com linguagens de marcação.

Normalmente, composições no formato MIDI são codificadas no computador ou geradas a partir de um instrumento eletrônico, como um teclado. Porém, alguns avanços foram realizados no intuito de transmitir informação de um formato não processado (o sinal captado por dispositivos eletrônicos a partir de ondas sonoras) para uma estrutura organizada (no caso, o arquivo MIDI). Um exemplo desta funcionalidade pode ser observado no artigo escrito por Maezawa et al [18], que demonstra uma técnica capaz de gerar a transcrição de uma partitura de violino a partir de uma gravação de áudio. Embora o modelo apresentado tenha se mostrado eficiente e os resultados tenham sido positivos, é importante notar que uma gravação de uma peça de violino como as descritas no artigo é constituída de apenas um instrumento com uma pequena quantidade de notas sendo tocadas ao mesmo tempo, o que facilita a análise.

Neste projeto, propôs-se usar um formato voltado à instrução (partituras e MIDI) para o desenvolvimento do sistema de visualização devido ao foco em elementos musicais que necessitariam ser encontrados a partir de sequências de notas. Embora o uso de software de conversão e equipamentos para captação de instruções MIDI a partir do sinal proveniente de instrumentos musicais tenha sido cogitado, optou-se por utilizar apenas MIDIs criados em ferramentas de composição devido à maior praticidade na geração e obtenção de conjuntos de dados para testes.

2.2.3. Extração de modelos

Uma das tarefas mais difíceis na área de processamento de músicas em computador é realizar a análise musical dos dados, ou seja, encontrar nos conjuntos de

dados processados informações que possam ser encaixadas em modelos estabelecidos por uma visão em particular. Um dos motivos para esta dificuldade é o caráter muitas vezes subjetivo da análise musical, que pode não se mostrar exata e sofrer variações de acordo com o especialista responsável por ela. No entanto, existem diversas pesquisas na área buscando obter este tipo de informação. Esta seção faz uma breve descrição de algumas delas a fim de estabelecer certo contexto na tarefa de análise musical por computador.

Chew e Raphael [19] fazem uma análise sobre diferentes aspectos da música e como analisá-los usando um computador, identificando em vários casos elementos semelhantes a problemas de pesquisa operacional (programação dinâmica), sugerindo que esta área pode ter muito a oferecer no processamento de dados musicais. Entre eles podese citar estimativa de ritmo, análise de acordes ou mesmo otimização de uso das mãos para performance de composições.

Finkelstein [20] discute sobre a realização de segmentação de áudio usando métodos de cadeias de Markov e técnicas de agrupamento de dados. As ondas senoidais do sinal de áudio são convertidas em um espectrograma, que então é segmentado com algoritmos de agrupamento. O resultado final, embora sujeito a falhas, consegue identificar vários tipos de segmentação, como fraseamento ou refrãos.

Shao et al[15] propõem um sistema de sumarização automática capaz de encontrar e separar introdução, versos, refrãos e encerramento a partir de conjuntos de dados MIDI. A separação encontrada é usada para gerar o sumário tendo como foco o refrão e melodias recorrentes. No entanto, vários elementos da estrutura e do ritmo são fixados, o que pode tornar a análise inviável para vários tipos de música.

Stefik et al[21] apresentam um tradutor de linguagens musicais, como por exemplo notações específicas usadas em músicas étnicas ou históricas. Embora trate simplesmente de representações escritas, é necessária uma estrutura de dados intermediária que consiga armazenar a música e se comunicar com todos os tipos de notações.

Endo et al [22] propõem um método de análise de músicas baseado em melodia, considerando gêneros musicais. São extraídos diversos descritores para um número de intervalos de tempo, como timbre, *highest pitch*, volume médio, entre outros. As informações são passadas a um modelo de aprendizado de máquina *Support Vector*

Machine (SVM)[23] modificado de modo a poder realizar comparações *um-contra-o- restante* até que uma classe seja obtida.

Estabelecido um contexto para as técnicas de processamento de música estudadas neste trabalho, serão discutidas a seguir algumas abordagens para tentar visualizar composições musicais em computador e como elas se relacionam com o modelo desenvolvido.

2.3. Visualização de composições musicais

Mesmo quando analisada sob a forma de instruções estruturadas, uma música ainda corresponde a uma grande quantidade de dados e informações complexas, fazendo com que a interpretação por parte do usuário seja, em muitos casos, mais lenta do que se gostaria.

Algumas abordagens baseadas em Visualização de Informação foram desenvolvidas com o intuito de facilitar a interpretação de composições musicais, oferecendo informações mais acessíveis acerca de diversos aspectos de uma música, com diferentes áreas de atuação. Esta seção apresenta e discute algumas destas técnicas, suas aplicações, vantagens, desvantagens e seu contexto na análise musical.

2.3.1. Visualização de tons e notas

O nível de representação mais fundamental na visualização de composições musicais é a exibição de notas e as relações entre estas ao longo do tempo. O objetivo é permitir ao usuário uma visão geral de progressões e evoluções de melodias e acordes, podendo também visar transmitir informações adicionais de tons como intensidade do som ou timbre de instrumento.

A notação estilo *Piano Roll* denota o tipo mais simples e clássico de visualização de uma composição musical, derivada dos antigos rolos de programação que eram lidos automaticamente por pianos. As notas são marcadas num plano 2D, onde o eixo X representa instantes de tempo e o eixo Y representa a altura da nota, em tons (*pitch*).

Um exemplo desta visualização, usando a ferramenta MIDITrail[24], pode ser encontrado na **Figura 5**. Na seção anterior, foi mencionado o uso do eixo vertical como um teclado de piano, onde as notas são organizadas de baixo para cima de acordo com sua frequência sonora. Como a frequência de cada semitom é $2^{1/12}$ vezes maior que a frequência do semitom anterior, o eixo deve seguir uma escala logarítmica. Nesta ferramenta, as instruções em MIDI são então colocadas ao longo do eixo horizontal conforme o momento em que devem ser chamadas, podendo fazer uso de diferentes cores ou símbolos para diferentes faixas ou instrumentos. É possível aproximar ou afastar a imagem em qualquer nível de *zoom*, assim como ver a composição ser tocada em tempo real (a posição horizontal do teclado de piano representa o instante atual e se desloca ao longo da linha do tempo).

A notação permite que a previsão de notas a serem tocadas, assim como a percepção de progressões e observação de padrões de tempo, sejam realizadas de uma forma mais visual e intuitiva, ainda que a discriminação exata de notas seja difícil à primeira vista. Adicionalmente, também é possível notar certos tipos de repetição e fluxos numa composição de forma diferente em relação a uma partitura.

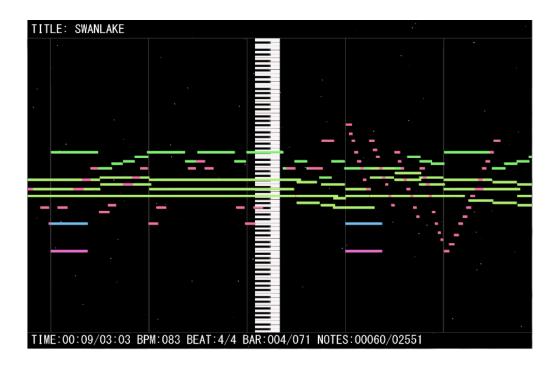


Figura 5: Visualização Piano Roll para trecho de "Swan Lake", de Tchaikovsky. Visualização obtida usando a ferramenta MIDITrail[24]. A ferramenta também informa na parte inferior aspectos relevantes sobre o progresso da música, como tempo, ritmo e número total de notas tocadas num dado instante de tempo.

Assim como na partitura, são exibidas apenas as notas e a distribuição de tempo, deixando totalmente a cargo do usuário a interpretação de interações entre notas próximas ou o contexto por trás da sequência sendo tocada.

Muitas das visualizações desenvolvidas para composições musicais, mesmo quando lidam com diferentes características dos dados, tendem a seguir estes mesmos princípios, tendo como exemplo o constante uso de uma linha do tempo ao longo da qual são marcadas as notas.

Uma das propostas mais antigas para se visualizar uma composição musical em computador foi feita por Smith e Williams [25], seguindo o modelo descrito acima, consistindo num mapeamento das instruções contidas em um arquivo MIDI num espaço tridimensional. A visualização compreendia o uso de formas esféricas para as notas, cujo tamanho indicava o volume ou amplitude da nota em questão. Os eixos X e Y continuam sendo usados de forma igual à *piano roll*, colocando ao longo do eixo Z as diferentes faixas de instrumentos contidas no arquivo, também separáveis por cor. Embora seja um

dos primeiros trabalhos da área, o artigo original não disponibiliza nenhuma imagem da visualização proposta.

Outra abordagem similar e relativamente conhecida é a ferramenta *Music Animation Machine* [26], que contém diversas alterações visuais à mesma fórmula citada no começo desta seção. Os símbolos usados para representar notas podem variar entre retas, retângulos, elipses ou círculos ligados por retas, com diversas versões diferentes disponíveis para visualização.

A **Figura 6** mostra um exemplo da utilização da ferramenta. Ao invés de usar uma linha vertical ou um teclado de piano para marcação de tempo, as notas brilham conforme são tocadas, gerando uma animação agradável. Diferentes faixas são representadas por objetos de diferentes formatos, como círculos e quadriláteros, de forma a tentar ilustrar a diferença entre os timbres de cada instrumento.

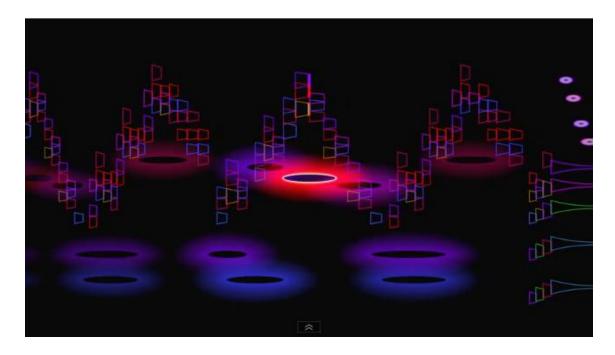


Figura 6: "Reflets dans l'eau", de Claude Debussy, na Music Animation Machine. As notas mais graves e lentas são representadas por círculos, enquanto as notas rápidas e sucessivas de outro instrumento são representadas por quadriláteros coloridos. As cores usada s variam de acordo com as notas tocadas. Imagem obtida em [26].

Além de prover este formato de visualização, a Music Animation Machine também conta com outros modos de exibição (como identificação de acordes), usando conceitos que serão descritos mais adiante.

2.3.1.1. Condensação de partituras

Ainda que as técnicas citadas anteriormente sejam capazes de fornecer algumas informações sobre a estrutura musical de uma composição e certa noção de seu cadenciamento, elas frequentemente são focadas em fornecer uma exibição temporal ou local dos dados (nota a nota), o que pode resultar em uma quantidade muito grande de conteúdo a ser explorada, especialmente em composições longas. Desta forma, a observação de uma composição musical como um todo usando estes modelos de visualização pode acabar não trazendo muita informação adicional ao usuário, quando comparada à partitura clássica (que possui uma representação semelhante).

Uma solução proposta para o problema das restrições geradas pelo caráter local da visualização de notas foram as técnicas de condensação de partituras, que visam reduzir a extensão da representação gerada para uma composição a ser analisada de modo a ainda manter informações importantes.

Uma abordagem para tratar o problema foi proposta por Watanabe et al [27], conhecida como o modelo BRASS (*Browsing and Administration of Sound Devices*). O modelo faz uso de uma notação reduzida para partituras, aliada à aplicação da técnica *fish eye* [28], para permitir a exploração de uma composição musical ao mesmo tempo em que a estrutura geral dos dados ainda pode ser observada, compreendendo assim uma visão *focus+context* [1]. Este modelo foi utilizado em conjunto com uma aplicação de rostos de Chernoff [30] para formar um sistema destinado ao aprendizado e treinamento de músicos [31].

A representação para a versão condensada da partitura agrupa as notas de compassos inteiros em uma barra, cujo tamanho, brilho e cor indicam características como quantidade e intensidade das notas presentes. As barras são então dispostas em ordem temporal no restante da composição, com o objetivo de transmitir ao usuário uma

idéia geral do que ainda vem a seguir, no contexto de fraseamento musical, tempo e dinâmica. Um exemplo da visualização com o sistema BRASS pode ser visto na **Figura** 7, obtida em [27], onde é possível observar uma seção intermediária em execução sendo exibida integralmente e o restante da composição condensado. As seções iniciais são mantidas a fim de prover informações básicas sobre a partitura, como as claves.

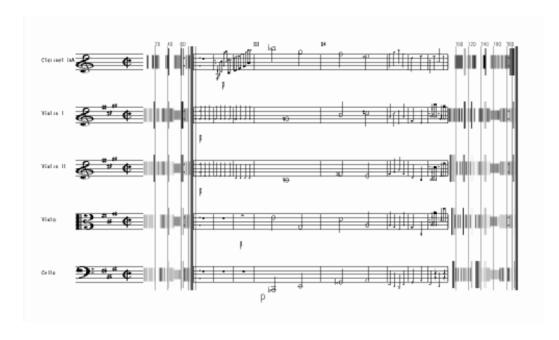


Figura 7: Visualização de partitura usando o modelo BRASS proposto por Watanabe et al [27]. "Clarinet Quintet A-major, K.V. 581", por W. A. Mozart. A seção em execução aparece exibida normalmente, enquanto a visualização condensa o restante da partitura.

Outra técnica que visa reduzir a representação de partituras, mas com abordagem e finalidade diferentes, é a *Colorscore*, desenvolvida por Hayashi et al [32]. O objetivo desta técnica é tentar reduzir o tempo levado para se analisar e interpretar uma obra de música clássica para orquestra, contendo as partituras de cada instrumento. Justamente por possuir uma grande quantidade de faixas para instrumentos que muitas vezes tem papel complementar, a técnica tenta encontrar o papel de cada instrumento em cada região da composição, identificando para cada uma os instrumentos que compõem a melodia dominante (em oposição a acompanhamentos ou decorações) e gerar uma partitura reduzida usando os principais aspectos da melodia.

A composição é dividida em blocos semânticos, a princípio usando momentos de silêncio dentro de cada faixa e posteriormente usando cálculos baseados na comparação

ponderada com um padrão dado pelo usuário. Papéis são então atribuídos aos blocos ainda usando o padrão definido anteriormente, até que toda a informação obtida é usada para se condensar a partitura tanto verticalmente (faixas com conteúdos considerados não tão importantes ou totalmente ausentes são eliminadas) quanto horizontalmente (regiões são reduzidas quando não há mudanças em seus papéis após uma repetição).

A **Figura 8**, obtida em [32], mostra um exemplo da utilização do Colorscore. Na seção superior da imagem são mostradas todas as faixas, em linhas: os blocos coloridos são partes identificadas como de grande importância para a composição, e portanto selecionadas. Na seção intermediária, é mostrado um novo grupo de faixas, em menor quantidade, contendo condensações dos conteúdos selecionados anteriormente. Por fim, na seção inferior, há um número ainda mais reduzido de faixas, contendo as informações dos blocos considerados de maior importância.

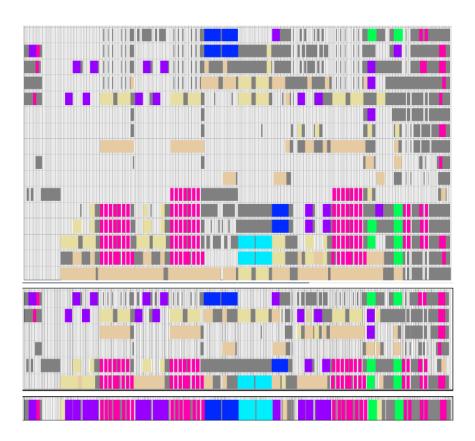


Figura 8: Colorscore. "Valse des fleurs", de Tchaikovsky, representada integralmente (cima), condensada em 5 faixas (meio) e apenas uma faixa (baixo). As cores representam os papéis dos blocos, as cores mais vivas descrevendo blocos com melodia mais dominante. Imagem de [32].

2.3.2. Visualização de acordes e harmonia

Enquanto algumas visualizações são focadas apenas em mostrar as notas a serem tocadas da forma mais clara possível, outras almejam exibir de forma prática informações que podem estar contidas numa partitura mas exigem certa interpretação das instruções descritas, tentando mostrar ao usuário determinados aspectos e peculiaridades de uma composição musical. Entre elas, pode-se citar as visualizações baseadas em harmonia, que tentam mostrar ao usuário as relações harmônicas entre notas tocadas simultaneamente.

Uma técnica que permite a visualização de uma destas informações é a desenvolvida por Ciuha et al [3] como parte do projeto *Colour Visualization of Music* [33]. Esta técnica se baseia na visualização de tons concorrentes usando cores, a fim de determinar consoância ou dissonância.

A técnica faz uso de um círculo de quintas, um círculo sobre o qual estão distribuídas todas as notas de uma oitava e sua consoância é expressada pelo tamanho do arco entre duas notas: em sentido horário, a próxima nota corresponde à quinta da nota atual. Este círculo é então associado a cores, de forma a atribuir valores diferentes de matiz a cada nota. A distribuição de notas no círculo pode ser observada em letras maiúsculas na **Figura 9**, obtida em [3].

As notas da composição, também lidas em formato MIDI, são então representadas como vetores radiais que saem do centro do círculo na direção do tom correspondente, com magnitude de acordo com sua intensidade. Quando mais de uma nota é tocada ao mesmo tempo, os vetores são somados. O vetor resultante irá possuir grande magnitude se as duas notas forem consoantes, resultando em uma cor mais viva no círculo (próximo às extremidades). Caso as notas sejam dissoantes, os vetores irão possuir direções opostas e o vetor resultante terá um valor baixo de magnitude, resultando numa cor acinzentada (próximo ao centro do círculo). Como o círculo de quintas faz uso de apenas um tipo de relação de consoância, foi sobreposto um segundo círculo ligeiramente deslocado, com as terças de cada nota do círculo original, buscando representar a consoância de acordes

maiores e menores com cores mais vivas. O conjunto passa a ser chamado de círculo de terças. A **Figura 9** mostra a distribuição de terças, representadas com letras minúsculas e em branco.

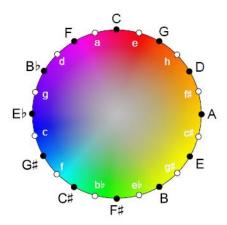


Figura 9: Círculo de terças associado a uma roda de cores. Quintas são identificáveis por letras de mesma cor, enquanto terças, de cores diferentes, são utilizadas para a definição de acordes menores e maiores. Imagem obtida em [3].

Esta notação é então aplicada sobre uma visualização *piano roll* simples como as mostradas no começo desta seção. O resultado final é uma visualização que permite não só a observação de tons predominantes em cada seção da composição usando um código de cores como também a identificação de regiões de alta consoância ou dissonância, baseando-se na saturação da cor em questão.

A **Figura 10**, obtida em [3], mostra uma composição visualizada com a técnica. A cada instante, é calculado o vetor resultante das notas sendo tocadas naquele ponto da linha do tempo e essas notas são então pintadas com a cor obtida no círculo de quintas a partir deste vetor. Assim, é possível observar o tom predominante em cada instante de tempo por sua cor e também o quão consoantes são as notas tocadas, por sua saturação.



Figura 10: "Clair de Lune", de Debussy, com representações de cores para tons. É possível observar, pelo uso de cores, a presença de partes dissoantes, além da evolução das tendências da composição de usar certos tons (do azul para o amarelo). Imagem obtida em [3].

A visualização conta com um sistema focado na identificação e associação de cores similares para acordes menores e maiores pela percepção de terças: a partir de uma nota mais forte considerada fundamental em um dado instante, o sistema tenta encontrar a terça mais apropriada no círculo para aquele conjunto de notas, resultando em cores mais saturadas. Embora seja efetivo para acordes comumente utilizados, esta funcionalidade pode impor certa limitação ao funcionamento da visualização se usada em composições com acordes e harmonia mais complexos.

Ainda visando mostrar relações entre notas, seja em intervalos harmônicos ou melódicos, Mardirossian e Chew [34] propuseram uma ferramenta de visualização de música capaz de mostrar progressões e distribuições de tons numa composição musical usando o espaço bidimensional de tons (*Bidimensional Tonal Pitch Space*) de Lerdahl [35], cujas coordenadas são baseadas em relações de consoância entre quintas e terças, usando escalas maiores e menores. A **Figura 11**, obtida em [34], oferece uma visão deste espaço.

O objetivo da ferramenta é gerar uma visualização mais intuitiva que revele aspectos da música que possam não estar disponíveis a pessoas destreinadas. A técnica consiste em mostrar frequências de tons (*keys*) ao longo de uma música, de uma forma mais informativa que um histograma: por mostrar acordes maiores e menores num plano 2D, é possível capturar relações entre eles. A composição musical é dividida em segmentos de tempo de comprimento uniforme aos quais são atribuídos acordes de acordo com um algoritmo de identificação (no caso, *Spiral Array Center of Effect Generator*

[36]). Os acordes são marcados no espaço bidimensional com discos coloridos, que aumentam de tamanho a cada vez que aquele acorde é detectado.

| d♯ | g# | c# | f# | b | e | a |
|-----------|----|-------------------|----------|----|-------------------|----|
| F♯ | В | E | A | D | G | C |
| $f\sharp$ | b | e | a | d | g | c |
| A | D | \mathbf{G} | C | F | $\mathrm{B}\flat$ | Εb |
| a | d | g | c | f | bb | еb |
| C | F | $\mathrm{B}\flat$ | $E\flat$ | Ab | $D\flat$ | G۶ |
| c | f | bb | eb | ab | db | gb |

Figura 11: Espaço bidimensional de tons de Lerdahl, com linhas para tons (*keys*) maiores e menores. As relações de consoância são dadas por quintas, sempre à esquerda, e terças, maiores e menores, respectivamente acima e abaixo. Imagem obtida em [34].

Esta técnica demonstra a visualização de características que não são visíveis no conjunto de dados à primeira vista (frequências de acordes), oferecendo uma representação sucinta de informações que normalmente necessitariam a análise visual ou auditiva de toda a composição. O modelo apresentado se baseia em uma visão animada, que mostra círculos surgindo e aumentando de tamanho com o passar do tempo.

A **Figura 12** mostra uma captura de tela da ferramenta, obtida em [34]. Por utilizar um espaço cíclico, é possível ver a mesma informação se repetindo ao longo da imagem. Cada ponto representa uma combinação diferente de tom e intervalo (maior ou menor), com sua posição dada pelo espaço de tons e sua cor atribuída a partir de um modelo similar ao que é observado em *Colour Visualization of Music*. A barra ao lado esquerdo indica o número de segmentos em que a composição será dividida: a cada segmento, é definido o acorde predominante e esta informação é incluída na visualização aumentando o círculo correspondente ou criando um novo círculo, caso não haja nenhum.

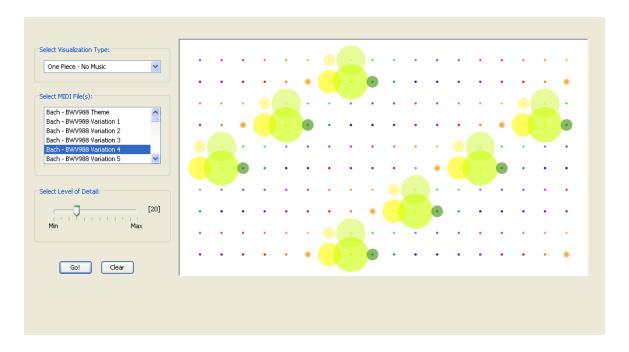


Figura 12: Visualização de progressões e distribuições tonais aplicada sobre uma composição de Bach. A composição é segmentada e os acordes predominantes detectados são mostrados em círculos, que aumentam conforme a sua incidência em diferentes segmentos. Imagem obtida em [34].

2.3.2.1. *Isochords*

Na área de identificação de tons, uma representação relativamente conhecida é a usada pela visualização *Isochords*, proposta por Bergstrom et al [37]. A visualização destaca os intervalos consoantes entre notas e acordes comuns em música, compreendendo informações como qualidade de intervalo, qualidade de acordes e progressão de acordes de modo síncrono durante a execução de uma música. Assim como outras abordagens vistas nesta seção, esta visualização busca proporcionar ao usuário meios de compreender a estrutura intrínseca da música que só seria acessível a músicos treinados.

A técnica faz uso de um espaço bidimensional isométrico de coordenadas triangulares denominado Tonnetz, proposto por Euler, de modo a fazer uma aproximação visual da consoância e dissonância de tons usando a representação de distância. O espaço, mostrado na **Figura 13**b (obtida em [37]), é um *grid* composto pelos 12 semitons de uma

oitava, colocados de forma cíclica onde tríades são representadas por triânglulos. A quinta de uma nota se encontra sempre à sua direita, enquanto as terças maior e menor se encontram entre as duas, acima e abaixo, respectivamente.

O conjunto de notas a serem tocadas é lido a partir de um arquivo MIDI e a visualização tenta montar acordes com as notas ativas em tempo real. Acordes são representados por um grupo de pontos correspondentes às notas que o compõem, juntando-os de modo a formar linhas e triângulos. Esta organização faz com que as três notas de acordes chamados maiores formem um triângulo apontando para cima e as notas dos acordes menores formem um triângulo apontando para baixo. As formas vão se tornando maiores e mais complexas com o aumento da dissonância das notas sendo tocadas, seguindo o círculo de quintas apresentado na **Figura 13**a, obtida em [37].

A visualização exibida é dinâmica e necessita ser apresentada em conjunto com a música, pois a imagem gerada compreende apenas o acorde identificado a partir das notas sendo tocadas em um instante de tempo. As coordenadas se repetem pelo *grid* de forma cíclica, como pode ser visto na **Figura 13**b.

Embora a informação contida na visualização não seja totalmente assimilada por um usuário de forma trivial, é possível identificar elementos dissoantes à primeira vista. A visualização pode fornecer suporte à interpretação detalhada de uma composição musical a um usuário que seja treinado em seu uso.

A **Figura 14**, obtida em [37], mostra uma captura de tela da visualização. Em tempo real, as notas contidas nas instruções MIDI são processadas e os símbolos são desenhados no *grid* conforme a música é tocada. Desta forma, é possível saber qual a combinação de notas sendo tocada em um dado instante a partir do símbolo formado pela visualização. Quando algum polígono é fechado pelos traços obtidos a partir das relações entre notas, a sua região interior é pintada em azul.

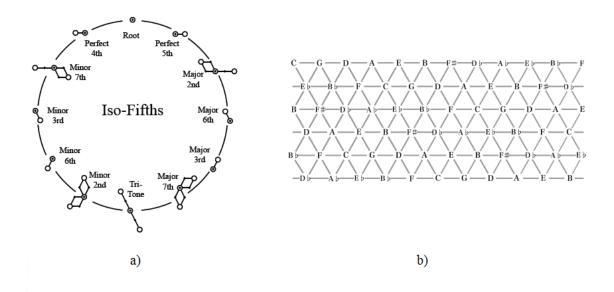


Figura 13: a) Círculo de quintas com símbolos no grid; b) Grid Tonnetz. Combinações de intervalos são representadas no *grid* de formas de diferentes complexidades. Imagens obtidas em [37].

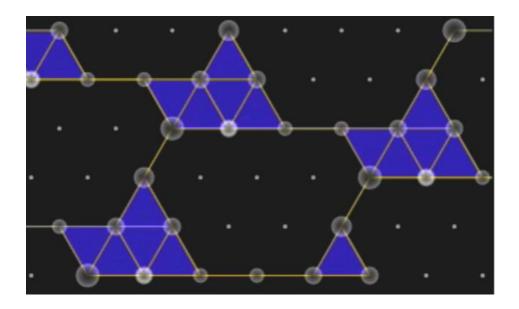


Figura 14: Exemplos de acordes usando Isochords. O símbolo correspondente ao conjunto de notas se repete com a repetição das notas pelo *grid*. Imagem obtida em [37].

2.3.3. Visualização de melodia e contexto

Em muitos casos, um ponto importante a se analisar em uma música é a melodia em destaque (como voz ou uma parte *solo* em guitarra), não só em como ocorre sua progressão mas também em como ela se relaciona com linhas de base e ritmo. Embora a simples visualização de notas possa fornecer uma visualização satisfatória de melodia, explorá-la mantendo em vista seu contexto pode ser muito importante mesmo quando ela não é um elemento fixo, como pode ser observado em improvisações musicais ou composições desenvolvidas em tempo real.

A técnica *ImproViz* [38] é uma técnica de visualização para diagramação de música capaz de revelar padrões característicos do estilo improvisacional de um músico de Jazz. Tais músicos possuem tendências a usar elementos melódicos similares ou sequências de notas semelhantes, mesmo durante improvisos: a ocorrência de notas é identificada e encaixada em padrões que podem ser observados pelo usuário. A técnica usa como entrada transcrições de performances realizadas por músicos.

A visualização é realizada por meio de duas abordagens: a primeira, denominada superfície ou paisagem melódica (*melodic landscape*) mostra os contornos gerais do freaseamento musical, oferecendo fácil observação de momentos de silêncio, entrada de solos complexos ou regiões em tons mais altos ou mais baixos. Esta superfície é gerada com um mapa de frequências (*pitches*), de maneira similar às visualizações discutidas no começo desta seção, mas baseando-se nas notações usadas em partitura e traçando as notas como uma linha contínua. A segunda, chamada de paleta harmônica (*harmonic palette*), é mais complexa e tenta marcar a ocorrência de notas tocadas pelo músico em relação ao acorde base em cada ponto da escala de tempo (calculada usando as batidas e o compasso da música). Desta forma, analisando improvisos realizados por um músico, é possível observar a preferência do mesmo por certos tipos de intervalos.

Um exemplo da visualização oferecida pelo ImproViz pode ser observado na **Figura 15**, obtida em [38]. Nela, são apresentadas quatro versões de improvisações sobre a música "All Blues", feitas por músicos diferentes e usando instrumentos diferentes. Na parte superior, podem ser vistas as *melodic landscapes*, sendo possível observar diferenças entre os momentos de silêncio, as densidades de melodias e as tendências da

progressão de notas para cada versão. Embora as posições das notas sejam as mesmas de uma partitura, o eixo de tempo é fixo, de modo que a comparação visual entre notas longas e curtas (ou a ausência de notas) seja mais fácil.

Na parte inferior da **Figura 15** estão as *harmonic palettes*: as notas do acorde sendo tocado naquele momento são representadas por círculos, nas mesmas posições que estas notas ocupam na partitura. Os círculos preenchidos representam o uso daquela nota em particular na improvisação, com a adição de mais círculos caso a ocorrência seja mais frequente. Assim, é possível observar a tendência dos músicos em usar certas notas em suas performances, sob o contexto dos acordes e da escala em uso. A segmentação é feita com o tempo da música, de forma que cada paleta represente o uso de notas em um compasso.

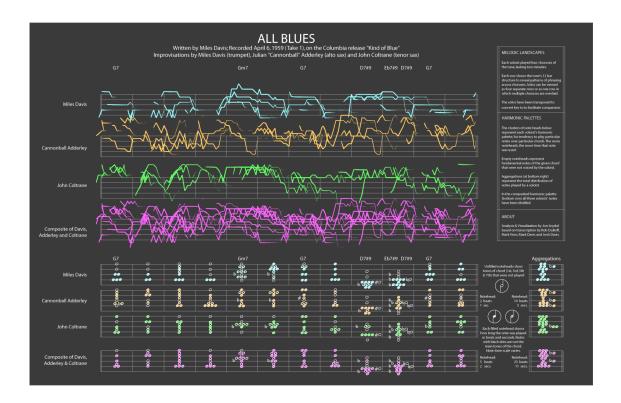


Figura 15: ImproViz. *Melodic Landscapes* (cima) e *Harmonic Palettes* (baixo) para quatro interpretações de "All Blues". O eixo X (linha do tempo) é o mesmo para todas as visualizações da imagem, de modo que seções das partes de cima e de baixo podem ser comparadas. Imagem obtida em [38].

2.3.4. Visualização por grupos ou estruturas

Outro aspecto que pode auxiliar na compreensão de uma composição musical é o entendimento de como ela pode ser subdividida e agrupada em áreas semelhantes. Existem vários meios de se caracterizar partes de uma música, seja por intervalos de tempo, instrumentos usados, variações na sequência de acordes e campo harmônico, entre outros.

Um ponto relevante, principalmente quando se trata de obras de música clássica a serem tocadas por orquestras, é a importância de cada instrumento ao longo da composição, qual sua atuação e qual sua funcionalidade dentro do contexto da apresentação. Foi mencionada a relevância de uma análise da melodia principal levando em consideração aspectos secundários, mas é importante notar que o papel de cada instrumento ou faixa pode variar conforme o andamento da música. Enquanto técnicas mencionadas anteriormente podem atribuir papéis a fim de selecionar regiões de maior importância dentro de partituras, essa alternância de papéis também pode ser visualizada em conjunto com outros elementos da estrutura da composição.

Uma abordagem proposta por Chan et al [39] propõe uma visualização para observar esta alternância. O modelo considera uma composição dividida em faixas ou camadas (*layers*) para cada instrumento (que podem assumir diferentes papéis dentro da harmonia com o passar do tempo) e seus temas, que compreendem melodias principais e idéias musicais persistentes na música, que possuem pequenas variações a cada ocorrência. Os dados de entrada são providos por análises manuais de partituras das composições.

São apresentadas estruturas de cor e símbolos específicas para representar a informação desejada, fazendo uso de dois protótipos: o trançado de camadas (*layer braid*) e o tecido de temas (*theme fabric*). O primeiro corresponde a uma visualização do papel exercido por cada grupo de instrumentos ao longo de diferentes temas da composição com o objetivo de exibir como as diferentes camadas se correlacionam e interagem entre si ao longo de uma linha do tempo. O segundo é focado num nível mais detalhado da representação, a fim de revelar microrrelações entre ocorrências de temas em camadas.

Usando uma notação específica para a representação de temas, a visualização permite observar quando um mesmo tema é tocado por camadas diferentes, por exemplo.

Imagens das visualizações podem ser observados na **Figura 16**, obtida em [39]. A imagem superior mostra o *layer braid*, onde linhas que representam instrumentos passam por diferentes partes da composição (exposição, desenvolvimento e recapitulação), subindo e descendo conforme a presença do instrumento naquele momento (dominante, acompanhamento, contraste, etc). A imagem inferior mostra o *theme fabric*, em que seções onde instrumentos participam de temas em comum são entrelaçadas para estabelecer maior conexão semântica.

Esta técnica se mostra pertinente dentro do contexto deste projeto por fornecer dois elementos importantes: visão baseada em *focus* + *context*, com o uso de ferramentas de *zoom* local, e o fornecimento de uma abstração para componentes musicais de destaque, como os temas.

Conforme discutido anteriormente, grande parte da estrutura intrínseca a uma composição musical diz respeito a repetições temporais. O que se conhece como música *pop*, em particular, frequentemente possui estruturas de repetição muito bem definidas, como refrãos. A análise destas estruturas pode ser muito útil para transmitir a um usuário um senso de fluxo dos elementos de uma música, principalmente caso este usuário seja um músico aprendendo a executar a composição em questão em seu instrumento.

Encontrar estruturas de repetição em uma música no formato de sinal sonoro é muito difícil por vários fatores, dos quais pode-se destacar dois: primeiro, os sinais de trechos da composição com repetições nunca serão exatamente iguais, devido a elementos como ruído, variações pequenas no timbre dos instrumentos ou variações na performance dos músicos; segundo, muitas vezes uma estrutura considerada como refrão não deve ser tocada exatamente da mesma forma, podendo possuir pequenas variações na letra, velocidade ou mesmo nas notas de alguns instrumentos, o que não necessariamente resultará em variações de mesma magnitude no sinal.

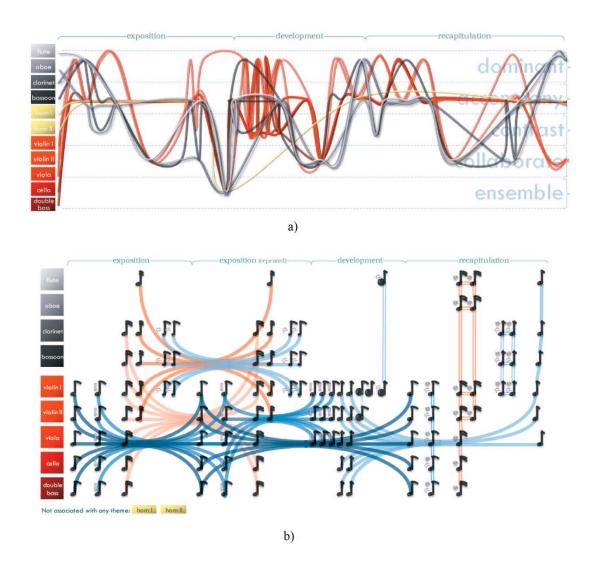


Figura 16: Estrutura semântica no primeiro movimento de Symphony No. 40, de W. A. Mozart.. a) Layer Braid; b) Theme Fabric. A imagem superior mostra as variações no papel de cada instrumento, enquanto a imagem inferior entrelaça instrumentos que participam de um mesmo tema. Imagem obtida em [39].

Portanto, a análise de estruturas de repetição também costuma ser realizada em arquivos no formato MIDI ou em partituras, nos quais é possível identificar com exatidão elementos que se repetem e mesmo estabelecer certa tolerância a pequenas variações. Uma das aplicações mais conhecidas desta análise é a *Shape of Song* [40], uma ferramenta Java que representa composições musicais como linhas do tempo em que arcos são erguidos para identificar regiões que se repetem. A análise é feita buscando sequências diretas de notas que aparecem em outras regiões da música, usando como base uma

comparação computacional de *strings* de caracteres [41]. Uma imagem desta aplicação pode ser vista na **Figura 17**, obtida em [40].

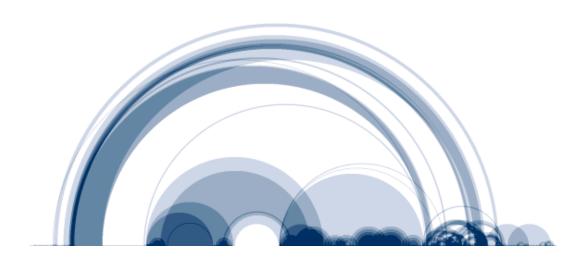


Figura 17: Representação de "Like a Prayer", de Madonna, usando a ferramenta Shape of Sound. Arcos ligam seções que se repetem, de tamanhos e durações variados. Imagem obtida em [40].

Outro modo eficiente para se visualizar uma composição musical longa é dividíla em estruturas hierárquicas e permitir a navegação por estas estruturas. Porém, embora seja mencionada nesta seção a identificação de elementos com repetições, encontrar elementos como passagens rítmicas, divisões em introdução, solos ou encerramento e frases de modo geral é uma tarefa complexa.

A ferramenta Comp-i, proposta por Miyazaki et al [42], é um visualizador 3D para arquivos MIDI capaz de explorar o conjunto de dados de diversas formas: um *grid*, uma notação comum baseada em linha do tempo (com faixas diferentes colocadas ao longo do eixo Z e tamanho de objetos correspondente ao volume da nota) e uma representação hierárquica baseada em *ConeTrees* [43], que mostra uma música vista por vários níveis hierárquicos diferentes, baseados em fraseamento musical.

O desenvolvimento da ferramenta é descrito como voltado à exploração *focus* + *context*, com o objetivo de produzir uma visualização que permita definir e entender estruturas globais, acessar informações de interesse dentro das estruturas e por fim ter acesso às instruções (definidas em MIDI como eventos) com possibilidade de edição.

Embora a visualização baseada em estrutura hierárquica tenha se mostrado eficiente, ela ainda é completamente definida pelo usuário, não havendo processo automático para extração de hierarquias ou de agrupamento de eventos. A visualização de níveis diferentes de hierarquia de forma radial pode facilitar a navegação e a identificação de elementos específicos, dependendo da tarefa a ser realizada.

Alguns dos conceitos usados na ferramenta, como a atribuição e visualização de estruturas hierárquicas, serviram como referência ao trabalho realizado neste projeto, por possibilitarem uma navegação mais fácil por conjuntos de dados complexos. A **Figura 18**, obtida em [42], tem exemplos de visualizações oferecidas pelo Comp-i. A imagem à esquerda apresenta a visão tradicional da música com linha do tempo, além de uma visão 3D da organização hierárquica radial. A imagem à direita apresenta uma visão *top-down* da abordagem hierárquica, em que níveis diferentes de hierarquia são mostrados em diferentes cores.

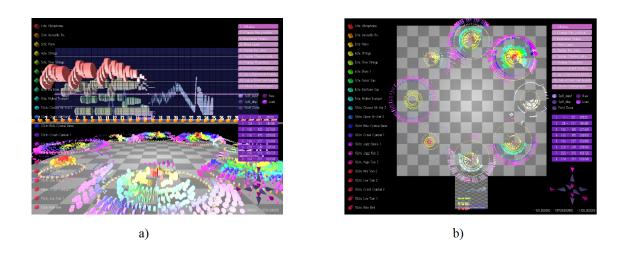


Figura 18: Visualização do Comp-i. a) Visão geral. b) Visão top-down da abordagem hierárquica. Imagem obtida em [42].

2.4. Considerações Finais

As abordagens vistas na seção 2.4 promovem várias possibilidades no que diz respeito à visualização computacional de música. No entanto, todas elas possuem áreas específicas de atuação e diferentes limitações. Entre as principais, pode-se citar o fato de que nenhuma das técnicas fornece auxílios para visualizar uma música por todos os aspectos básicos do domínio musical, focando-se numa área específica. Algumas também exigem interações com o usuário a fim de calibrar o processo ou mesmo realizar classificações, enquanto outras são capazes de gerar seus resultados usando apenas os arquivos MIDI.

Na maior parte dos casos, não há grande interseção entre as técnicas que permitem a análise e observação de elementos musicais e as que organizam e agrupam os dados contidos na música, oferecendo visões *top-down* ou *focus+context* que facilitem a exploração do conjunto de dados em alto nível. Neste aspecto, o princípio de visualização multinível poderia ser empregado a fim de tentar fornecer maneiras mais completas de interpretação do conteúdo a usuários.

Poucas das visualizações estudadas realizam algum esforço adicional para visualizar ritmo e métricas de tempo além de dividir a linha temporal e marcar nela as instruções de percussão, seja por não haver necessidade de tal auxílio ou pela dificuldade de seu desenvolvimento. No entanto, a identificação de repetições ao longo do tempo e o destaque de temas e frases frequentes, mesmo que por simples ligação como demonstrado no aplicativo *Shape of Song* (seção 2.3.4), pode ser usada tanto para diminuir o tamanho de uma representação visual (por meio da omissão de repetições) quanto para ressaltar elementos importantes.

Com a exceção do projeto *Colour Visualization of Music*, as técnicas estudadas voltadas para visualização de relacionamentos entre notas (vistas na seção 2.3.2) possuem a característica de serem dinâmicas, exigindo que o usuário acompanhe a visualização enquanto a música é reproduzida. Isto é, de certa forma, conflitante com o foco deste projeto pois não há como fornecer uma visão geral do conjunto de dados sem uma representação estática. Estes tipos de visualização poderiam, no entanto, ser usados paralelamente ao modelo de alto nível.

A técnica ImproViz fornece uma visão interessante sobre certos aspectos da melodia, que dificilmente seriam reconhecidos com as outras abordagens. Adicionalmente, simplesmente por representar uma plotagem direta como o *piano roll* por meio de linhas contínuas, as *Melodic Landscapes* são capazes de transmitir de forma mais intensa picos de tonalidade e momentos de som ou silêncio.

Dentre as técnicas estudadas e no contexto deste projeto, o modelo proposto por Ciuha et al em *Colour Visualization of Music* se mostrou um dos mais promissores, por permitir a visualização de elementos de harmonia por meio do uso de cores, o que libera o restante dos recursos visuais para a representação de outros aspectos importantes. A sua maior limitação sob o ponto de vista de identificação de harmonia é usar esquemas de cores diferentes para classificar tipos diferentes de acordes, não possuindo uma métrica absoluta e podendo dificultar o reconhecimento de certos elementos harmônicos.

O estudo destas técnicas permitiu um entendimento melhor da situação atual da visualização aplicada em música por parte do aluno, assim como possibilitou a definição de certas linhas para o desenvolvimento do projeto, que serão descritas posteriormente no Capítulo 3. A Tabela 2 faz um resumo das principais técnicas estudadas e algumas formas de compará-las, ainda que tenham propósitos diferentes.

| Ferramenta / Técnica | Visualização nota-a-nota | Visualização de harmonia | Visualização de ritmo | Processamento de dados | Permite visão global? | Organiza os dados? |
|-------------------------|-----------------------------|-----------------------------|--------------------------|---------------------------|-----------------------------|--------------------|
| Piano Roll | Sim | Não | Não | MIDI | Parcial | Não |
| BRASS | Sim | Como em partitura | Como em partitura | Partitura | Sim | Não |
| Comp-i | Sim | Não | Não | MIDI – parcial | Sim | Sim |
| Colorscore | Não | Não | Não | Partitura | Sim | Sim |
| Colour Visualization | Sim | Sim | Não | MIDI | Parcial | Não |
| Layer Braid | Não | Não | Não | MIDI | Sim | Sim |
| Theme Fabric | Não | Não | Não | MIDI | Sim | Sim |
| Isochords | Não | Sim | Sim | MIDI | Não | Não |
| Improviz | Sim | Parcial | Não | Transcrição | Sim | Parcial |
| Shape of song | Não | Não | Parcial | MIDI | Sim | Sim |

Tabela 2: Comparação de técnicas de visualização de música

Capítulo 3. Visualizando Elementos de Teoria Musical

As técnicas de visualização de composições musicais descritas no capítulo anterior visam oferecer auxílios visuais para facilitar a compreensão de uma música por parte de um usuário com diferentes graus de treinamento. Devido à quantidade de elementos e conceitos envolvidos neste domínio, esta tarefa pode ser realizada de diversas formas, focando-se em diferentes aspectos de uma música, como a análise de eventos concorrentes ou a identificação de padrões temporais. Com base no conteúdo estudado, pode-se dizer que a maior parte destas visualizações é especializada, de modo que poucas das abordagens observadas apresentam a possibilidade de se atacar o problema de mais de uma forma ao mesmo tempo.

Neste capítulo, será descrita em detalhes a abordagem de visualização desenvolvida neste projeto de mestrado, analisando seu impacto no processo de interpretação e estudo de composições musicais por parte de usuários experientes ou em treinamento. Serão apresentados os métodos de visualização elaborados para a discriminação visual de cada elemento músical explorado e o sistema de visualização desenvolvido que aplica tais métodos.

3.1. Motivações e Objetivo

De forma geral, as técnicas de visualização de música discutidas anteriormente possuem três limitações importantes. A primeira delas diz respeito à interação entre os diferentes aspectos da música: o foco na visualização de apenas um aspecto musical por

vez pode dificultar a compreensão da relação entre elementos musicais diferentes, além de possivelmente omitir informações potencialmente úteis. Por exemplo, embora existam técnicas que possibilitem visualizar estruturas de repetição de notas e técnicas que mostrem padrões harmônicos, a detecção de repetições de padrões harmônicos não chega a ser explorada.

A segunda limitação se refere ao fato de que a maior parte das técnicas estudadas não é multiescala, ou seja, não trabalha em níveis diferentes de refinamento e abstração. Isto significa que é difícil relacionar elementos observados em técnicas para visualização geral de músicas com elementos vistos em técnicas que trabalham localmente, nota a nota. Fazer estas relações, no entanto, pode ser fundamental para um músico que tenta criar conteúdo novo sobre uma composição existente.

A terceira limitação corresponde à dificuldade de realizar uma análise musical sem o auxílio de um usuário. Isso se deve ao fato das classificações e decisões a serem realizadas não serem absolutas. Por exemplo, um mesmo conjunto de notas tocado no mesmo instante pode ser classificado como diferentes acordes, dependendo do contexto. Logo, a maior parte das técnicas que trabalham com harmonia acaba assumindo certas condições ou pedindo informações ao usuário.

Definir quais elementos musicais seriam o foco do trabalho e o que exatamente se planejava visualizar foi uma decisão importante tendo em vista a quantidade de aspectos musicais diferentes que podem ser explorados e a diversidade de técnicas que os contemplam. Por fim, optou-se por trabalhar com melodias em guitarra elétrica, por ser um instrumento de fácil acesso e com certa popularidade. Assim, o objetivo central consistiu em criar uma representação visual da performance de um guitarrista, com suporte a análise de relações melódicas e harmônicas entre as notas tocadas e a detecção de frases e sequências usadas com frequência.

Originalmente, a visualização a ser desenvolvida possuía um foco adicional em solos e improvisos, mas a observação de padrões e repetições ao longo de composições musicais de forma geral se mostrou interessante o suficiente para que a meta do trabalho se tornasse mais abrangente. A seguir serão discutidos os detalhes da visualização produzida.

3.2. Visualizando Elementos da Teoria Musical

A visualização e navegação no espaço de instruções referentes às notas contidas numa composição musical pode ser implementada de várias formas, de acordo com as necessidades e objetivos da aplicação. Contudo, compreender os nuances de uma melodia e entender como ela é tocada em um determinado instrumento não é uma tarefa simples. Uma partitura determina uma sequência de notas e o momento específico em que cada uma deve soar, caracterizando assim uma forma exata de representar uma composição musical. No entanto, a análise de uma partitura com o intuito de executar a peça nela descrita, realizar improvisos ou simplesmente estudar as técnicas envolvidas em sua composição pode ser dificultada por alguns aspectos:

- A compreensão mental do som que a partitura representa nem sempre é feita pelo leitor em tempo real, especialmente quando se trata de um músico ainda em treinamento;
- A partitura deve ser lida sequencialmente, o que pode dificultar a previsão do que vem a seguir;
- As metáforas visuais empregadas na partitura nem sempre refletem da melhor maneira características de instrumentos específicos;

Além de auxiliar o estudo em conjunto com a partitura, acreditou-se ser importante prover ferramental que forneça uma boa visão geral da composição musical, de modo a dar suporte ao usuário na própria decisão de quais partituras estudar mais a fundo. Esta visão, além de garantir que uma idéia geral dos dados analisados já se encontre em mente, também possibilita a comparação visual imediata de duas ou mais composições musicais.

Assim, a maneira escolhida para se atacar o problema foi criar um arcabouço de representações gráficas que ajude o usuário a compreender e imaginar a sequência de sons que forma uma música, associando-os a instruções num instrumento que conheça ou esteja praticando (no caso, a guitarra elétrica), mas que ao mesmo tempo mantenha-o ciente do contexto que envolve esta sequência.

Dentre as técnicas de visualização musical relevantes a este projeto descritas no Capítulo 2, três delas se relacionam intimamente com os conceitos discutidos nesta seção: a abordagem descrita no projeto *Colour Visualization of Music*, a ferramenta *ImproViz* e o sistema de visualização BRASS (seção 2.3). A primeira técnica, que faz uso de cores para visualizar harmonia, pode possuir um papel importante em determinar a situação da melodia tocada em relação a um acompanhamento ou base, o que é essencial em solos e improvisos. A *ImproViz*, por sua vez, oferece a tendência de um dado músico a utilizar certas posições na escala, o que pode fornecer uma espécie de perfil da performance em questão. Por fim, o sistema BRASS realiza a condensação de partituras, característica importante no processo de definir a visão geral mencionada anteriormente.

Estas visualizações também possuem limitações: as *melodic landscapes* exibidas no *ImproViz* fornecem informações gerais sobre o que acontece na melodia, mas em nível essencial não são muito diferentes de um *piano roll* ou mesmo de uma partitura clássica. As *harmonic palettes* não oferecem uma análise profunda o suficiente para identificar elementos maiores que uma nota, além de possivelmente se tornarem congestionadas quando usadas para analisar composições inteiras ou trechos grandes. A visualização de tons com cores poderia ser voltada a operar não só em harmonia, mas em melodia (ao longo do tempo), além de possivelmente se beneficiar de um modelo absoluto e independente de acordes para a distribuição de consoância. A condensação de partituras observada no sistema BRASS possibilita a visão de contexto em uma composição musical, mas não permite a visualização de progressões melódicas e harmônicas. Adicionalmente, a união dos conceitos vistos nestas três abordagens poderia levar à obtenção de uma visualização mais informativa.

Após um processo iterativo de estudo e elaboração de modelos de visualização, a estratégia escolhida para o projeto foi a de adotar um sistema multinível que representa informações melódicas, harmônicas e temporais em detalhes e contexto simultaneamente. As relações entre notas a serem tocadas são representadas por metáforas visuais tanto a nível local quanto global, buscando exibir o máximo de informações possível de uma forma coesa, facilitando a identificação de pontos de interesse. A aplicação de conceitos da visualização de modelos multinível, de dados variantes no tempo e de documentos baseados em texto representam áreas da Visualização de Informação intimamente

envolvidas no desenvolvimento do projeto, sendo discutidas brevemente nos parágrafos seguintes.

3.2.1. Visualização de dados multinível

Muitos conjuntos de dados possuem diversos tipos de informações associadas que podem ser analisados ou visualizados usando mais de uma abordagem: observados por pontos de vista diferentes, organizados e agrupados de mais de uma forma, etc. Nestes casos, pode ser vantajosa a utilização de uma visualização que ataque os aspectos diferentes do problema e mostre-os em diferentes níveis de representação.

Técnicas de visualização multinível são técnicas que oferecem visão de diferentes aspectos ou organizações dos dados ao mesmo tempo ou, caso funcionem alternadamente, permitem a correlação entre elementos e características pertencentes a níveis diferentes. Estes níveis podem representar visões diferentes dos dados (como camadas de abstração) ou mesmo visualizações com objetivos específicos variáveis [44]. Dentro do contexto deste projeto, um tipo de visualização que pode ser citada é a visualização multinível de dados hierárquicos, por possuírem diferentes camadas que podem ser exploradas a fim de se obter visões gerais e específicas simultaneamente (focus+context). Frequentemente, dados hierárquicos são exibidos como estruturas menores contidas em estruturas superiores, como árvores ou treemaps (representação hierárquica de dados na forma de retângulos aninhados) [59,60]. As estruturas hierárquicas podem ser definidas de acordo com várias características, como herança de classes em software, chamadas de processos ou eventos, relações lógicas ou mesmo divisões de tempo. Quando aplicada para explorar organizações hierárquicas, a visualização multinível pode muitas vezes também servir como uma visualização multiescala, permitindo a observação do conjunto de dados em diferentes graus de aproximação.

As idéias observadas em visualização multinível se encaixam na perspectiva do sistema de visualização desenvolvido devido ao foco adotado em estuturas hierárquicas e visualização de contexto. Adicionalmente, as informações contidas numa partitura podem ser analisadas por diversos ângulos, sendo observadas tendo como guia o ritmo,

harmonia ou mesmo semântica, fornecendo várias possibilidades de interpretação e elementos de destaque para um mesmo conjunto de dados. A representação final adotada no projeto agrupa diferentes facetas dos dados em uma mesma imagem com o uso de diferentes metáforas visuais, além de dividir a visualização em duas imagens a fim possibilitar a observação de características gerais e locais simultaneamente.

3.2.2. Visualização de dados variantes no tempo

A visualização de dados variantes no tempo é um campo complexo de ser tratado devido ao fato de a passagem do tempo ser um aspecto que deve ser assimilado integralmente para um bom entendimento de uma técnica de visualização. Na maior parte das abordagens de visualização, isto significa o uso de um dos eixos em uma visualização de duas ou três dimensões para a passagem do tempo, como ocorre com a maior parte dos gráficos temporais [45]. Em outras abordagens, também pode ser observado o uso de uma progressão de cores, usando posicionamento espacial apenas para descrição dos dados.

Este tipo de representação é, de certa forma, similar às observadas para tratamento de música, principalmente devido ao uso da linha do tempo. Porém, em alguns casos, não é possível usar tais representações devido a limitações do conjunto de dados, como grande extensão temporal (resolução de imagem insuficiente para a exibição do eixo de tempo com a definição adequada ou paleta de cores insuficiente para representar variações de tempo de modo distinguível) ou uma necessidade muito grande de se representar outros atributos do conjunto usando estas características em detrimento do tempo.

Desta forma, muitas das técnicas recentes de visualização de dados variantes no tempo visam obter informações decorrentes de tal variação ao invés de visualizar a passagem do tempo propriamente dita [46,47]. Como por muitas vezes as interações entre as instâncias de dados com o passar do tempo são sutis e indiretas, muitos autores buscam representar os padrões encontrados na análise destas informações de formas alternativas. Um exemplo é encaixar as variações temporais em padrões reconhecíveis, como funções matemáticas inspiradas em movimentação de grupos de animais [48,49]. Outra

abordagem para a visualização de variações e fluxos temporais é a ilustração por meio de texturas ou superfícies [50,51].

As técnicas de visualização de dados variantes no tempo são pertinentes ao contexto deste projeto por motivos práticos: uma música corresponde a uma série de informações que variam ao longo do tempo, seja ela representada por um sinal de áudio cuja onda sofre variações ou uma lista de eventos de um arquivo MIDI. Embora o modelo temporal escolhido para visualizar os dados neste trabalho faça uso de uma linha do tempo simples, a observação destas técnicas foi importante para determinar quais informações exibir e como representá-las adequadamente em uma visão em escala.

3.2.3. Visualização de texto

A visualização de documentos contendo textos é uma área de destaque em visualização de informação: prover informações a respeito do conteúdo de um ou mais textos sem a necessidade da análise detalhada (leitura) por parte do usuário é uma tarefa difícil, mas que pode auxiliar dramaticamente o trabalho de pesquisadores, editores e quaisquer outros usuários que precisem trabalhar uma grande quantidade de textos em uma quantidade limitada de tempo.

É possível traçar paralelos entre conjuntos de dados referentes a texto e referentes a instruções musicais: ambos possuem uma grande quantidade de informações que são expressas a partir da combinação de certos elementos básicos (letras e notas) e ambos possuem uma sequência natural em que os dados são distribuídos (leitura e performance). De forma similar, muitas aplicações de visualização focadas em texto possuem semelhanças com àquelas focadas em música, como a identificação de repetições ou a frequência da ocorrência de determinados elementos.

Assim como ocorre com música, uma grande parcela dos desenvolvimentos recentes na área de visualização de textos diz respeito à observação de coleções de documentos, podendo analisar milhares de textos como parte de um mesmo conjunto de dados. Um exemplo deste tipo de visualização são as *tag clouds*, que representam palavras ocorrentes em uma coleção de textos em posição e tamanhos variados conforme a

frequência com que aparecem. Uma aplicação gratuita desta visualização pode ser vista em [52].

No entanto, embora existam similaridades, dados musicais e textuais não podem ser visualizados ou processados exatamente da mesma forma. Uma das principais diferenças se dá na dependência maior da informação sequencial (temporal) por parte da música: ainda que um texto necessite ser lido em ordem para ser compreendido, suas palavras podem ser examinadas fora de contexto em grande parte das aplicações, como agrupamento de dados, identificação de assuntos e tópicos de destaque, entre outras. Grupos de notas musicais, como as contidas em um compasso, tendem a ser mais dependentes de contexto, sendo muitas vezes necessário analisar o trecho da composição em que eles se encontram para compreender melhor qual o seu papel.

A visualização de um único texto pode ser explorada de diversas formas. Muitas das abordagens citadas para coleções podem ser usadas também neste caso, como a identificação da ocorrência de palavras-chave. Existem também abordagens orientadas à visualização de elementos textuais específicos, como eventos ligados a personagens em textos literários [54]. Em alguns casos, técnicas de visualização são aplicadas no intuito de comparar os resultados de sua utilização em dois ou mais textos distintos.

Um conceito que inspirou uma funcionalidade do sistema de visualização desenvolvido neste projeto é o *literature fingerprinting*, por Keim e Oelke [53]. Esta técnica consiste na geração de descritores para segmentos de um documento de texto extenso, que são posteriormente unidos para gerar uma imagem que representa o documento. A subdivisão em segmentos pode ser realizada em diferentes magnitudes, de acordo com a necessidade: palavras, frases, parágrafos ou mesmo capítulos inteiros. As características usadas para determinar estes descritores também são variáveis, o que permite a geração de vários *fingerprints* para um mesmo texto: comprimento dos segmentos, presença de palavras importantes, riqueza de vocabulário, entre outros.

Uma vantagem deste modelo é a possibilidade de usar diferentes descritores, possibilitando não só a comparação de resultados como também a utilização do mesmo tipo de representação para aplicações com características distintas. Oelke et al [55] fazem uso da mesma visualização para interação com o usuário num processo iterativo de análise

de textos, baseando a escala de cores na positividade ou negatividade das expressões usadas no texto.

Um conceito similar foi utilizado no sistema de visualização desenvolvido neste trabalho, em que a composição musical analisada é dividida em compassos que recebem então valores correspondentes a diferentes descritores, que podem ser observados em conjunto. Desta forma, o comportamento dos descritores pode atrair a atenção do usuário para a observação de certas áreas do conjunto de dados, pertinentes à análise visual sendo realizada no momento.

3.3. Estrutura do arcabouço de visualização

O sistema de visualização desenvolvido se baseia na observação *focus* + *context* dos elementos musicais identificados por meio do uso de duas frentes de visualização: uma de detalhe, visando a exibição das notas de forma local, e uma geral, exibindo a composição musical como um todo. São representadas as notas existentes na faixa de um dado instrumento (guitarras elétricas), escolhida pelo usuário, e sua comparação com uma faixa escolhida como acompanhamento (como baixo ou uma guitarra base), a fim de se obter informações de harmonia. Estas faixas serão chamadas de *faixa principal* e *faixa base*, respectivamente. Aspectos importantes da faixa principal, assim como da relação entre as duas faixas, devem ser exibidas de forma global, com o intuito de dar ao usuário a capacidade de identificar rapidamente padrões, regiões de alta ou baixa complexidade, mudanças em ritmo e harmonia, entre outras características.

Além de visualizar a sequência de notas propriamente dita, a visão de detalhe é responsável por fornecer informações que a complementem: a marcação de cada compasso, com indicações de qual o acorde ou tom predominante, andamento e tempo; a consoância de cada nota tocada com relação ao tom base identificado, tambem expressa por meio do intervalo equivalente; a força de cada nota dentro do compasso, dada por sua localização temporal, duração e intensidade. Este modelo foi projetado com o intuito de

prover suporte ao usuário na tarefa de observar o papel de cada nota na sonoridade de um dado trecho da composição.

A visão geral, também chamada de miniatura, corresponde a uma imagem que resume toda a composição musical. A linha do tempo ainda é utilizada da mesma forma, mas as notas são condensadas em compassos. Cada compasso é representado por uma caixa colorida cujas variações de tamanho e opacidade representam as variações das notas nele contidas. Nesse ponto, compassos que contém instabilidade, seja por notas dissoantes, repetições com quebra de expectativa ou mudanças de andamento são pintados de vermelho para assinalar possíveis pontos de interesse. Mudanças importantes, como trocas de tom, tempo e andamento são assinaladas ao longo da miniatura para fácil referência. Abaixo desta imagem, são exibidas outras características do compasso na forma de um mapa de densidade: complexidade da sequência de notas, nível de repetição e variação de intervalos usados.

3.4. Representações visuais

Definidos os elementos musicais a serem visualizados nos dados, esta seção descreve em detalhes as abordagens desenvolvidas para a obtenção e exibição de cada característica envolvida no arcabouço de visualização. Um dos aspectos mais importantes da abordagem elaborada é a presença de uma estrutura de dados básica referente a um compasso, que possibilita muitas das comparações e classificações necessárias à visualização.

3.4.1. Compassos

O compasso é a unidade básica utilizada em todos os aspectos da visualização do sistema desenvolvido, representando um intervalo de tempo específico definido pela fórmula de compasso naquele trecho da composição musical. Além das notas nele contidas, são armazenados descritores referentes ao modo como estas notas interagem entre si e com um compasso equivalente da faixa base, mudanças e transições ocorridas, entre outros aspectos. Desta forma, é possível comparar diferentes compassos, encontrar similaridades ou mesmo classificar grupos de compassos de acordo com diferentes critérios.

A um compasso são atribuídos vários descritores: densidade de notas, nota mais alta, nota mais baixa, valor de estabilidade, valor de complexidade, variação harmônica, primeiro intervalo, presença de mudanças de tom ou de fórmula de compasso e um acorde associado. Adicionalmente, também é associado a cada compasso o estado da execução musical no momento de seu início: quais notas estão ativas e quais são o tempo, tom e fórmula de compasso atuais. Ao objeto que contém estas informações de estado foi dado o nome de *snapshot*.

Muitos dos descritores de compasso podem ser obtidos apenas utilizando os dados do próprio compasso, como a densidade e variação de notas. Outros, como o valor de estabilidade ou complexidade, dependem de informações referentes aos compassos próximos. Descritores de natureza local podem ser calculados assim que os compassos são convertidos pelo interpretador MIDI, enquanto os demais necessitam ser calculados após a observação de todo o conjunto de dados, juntamente com descritores do conjunto de dados como um todo.

3.4.2. Acordes e Intervalos

Alguns descritores de compassos são calculados de maneiras simples: basta examinar a lista de notas associada para identificar características como densidade de notas ou nota mais alta. Outros descritores necessitam de cálculos mais complexos, por

vezes necessitando de informações a respeito de outros compassos ou da composição como um todo. As seções seguintes detalham o processo de obtenção de descritores relacionados a partes específicas da visualização, mas aqui serão discutidas duas informações atribuídas a notas e compassos que são responsáveis por influenciar todos os níveis da visualização: acordes e intervalos.

Grande parte da interpretação de harmonia realizada pela visualização depende da identificação e comparação de acordes e escalas. Originalmente, o protótipo inicial da ferramenta calculava a consoância entre notas de maneira absoluta: para cada instante (ou em todos os momentos em que ocorria uma mudança nas notas ativas), todas as notas ativas eram comparadas duas a duas e um valor médio de consoância era determinado. No entanto, este modelo não se mostrou particularmente eficiente. Além de computacionalmente custoso, o método não levava em consideração certos aspectos da harmonia, como o fato da percepção de consoância de intervalos maiores ou menores ser dependente do contexto.

Assim, optou-se por um modelo de identificação de acordes. A cada compasso é associado um tom, que essencialmente representa a sonoridade predominante naquele segmento de tempo. É possível ocorrer casos em que uma composição mude drasticamente de sonoridade no decorrer de um mesmo compasso, mas neste caso o sistema simplesmente identificará a mudança como um ponto de instabilidade a ser observado (mais detalhes sobre a visualização de instabilidade serão vistos adiante). O algoritmo utilizado identifica o tom em cada compasso a partir da faixa base de duas maneiras diferentes: a primeira, mais simples, procura a nota mais grave tocada num tempo forte do compasso (normalmente a primeira nota). Esta nota é considerada o tom base naquele compasso, que será usado como referência na análise das notas tocadas na faixa principal. Este modelo foi chamado de "harmonia sem sinal" (unsigned harmony), por não fazer distinções entre tons maiores e menores. A segunda maneira, mais complexa, tenta encontrar acordes propriamente ditos: a nota encontrada no método anterior é considerada fundamental e, estimando o campo harmônico sobre o tom atual da composição musical, atribui-se a ela a propriedade maior ou menor. A segunda abordagem apresenta resultados mais precisos para grande parte dos casos, mas não é flexível o suficiente para avaliar uma peça com harmonia mais complexa, justificando a necessidade de uma abordagem mais geral.

O método descrito para identificar acordes foi uma dentre quatro alternativas testadas com diferentes tipos de música. A concepção original consistia em comparar as notas tocadas em um compasso com um dicionário de acordes e identificar a correspondência mais provável. Porém, além de custoso, tal método poderia se mostrar incerto e difícil de configurar devido ao grande número de combinações possíveis. Por fim, optou-se pela simplicidade: as outras alternativas consistiam em usar a nota mais grave ou a primeira nota do compasso, abordagens que acabaram sendo unidas no modelo descrito no parágrafo anterior.

Uma vez que o tom ou acorde predominante em cada compasso da base foi identificado, pode-se calcular um valor de consoância de cada nota da faixa principal em relação a ele, expresso pelo intervalo. No modelo *unsigned*, as consoâncias são determinadas de forma mais simples: detectam apenas categorias de consoância de intervalos (quintas são mais consoantes que segundas, por exemplo). Já no modelo completo, os intervalos são calculados em função da escala dada pelo tom da composição, resultando em consoâncias baixas para notas fora da escala. Embora as escalas usadas sejam essencialmente maiores e menores, é possível configurar o sistema para admitir escalas e valores de consoância diferentes, caso tendências particulares dos conjuntos de dados analisados sejam conhecidas.

3.4.3. Visão de detalhe

A visão de detalhe se baseia no modelo *piano roll* para representar as notas tocadas conforme a passagem do tempo. As notas são posicionadas no eixo Y de acordo com a sua frequência (*pitch*), sendo o alcance da visão de detalhe da ferramenta desenvolvida igual a seis oitavas: as quatro normalmente contidas no braço de guitarras e uma adicional tanto para sons mais graves quando para agudos, a fim de ilustrar variações como efeitos especiais aplicados aos sons ou mudanças na afinação de instrumentos, além de uma parcela das notas contidas nas faixas de baixo. Também é possível alterar o alcance de frequências da visualização por meio de um parâmetro.

As notas são representadas por pontos verdes seguidos de uma linha que indica a duração da nota, sendo desenhadas sobre um *grid* que divide o espaço horizontalmente em compassos e verticalmente em oitavas. Dentro do espaço de cada compasso são também descritos a fórmula de compasso e o andamento. Outras duas informações também são adicionadas às notas: sua consoância em relação à base, expressa pelo brilho, e sua potência na sonoridade daquele compasso, representada por um círculo brilhante ao redor da nota.

A consoância de uma nota qualquer n (C_n) é representada de forma simples: valores reais entre 0 e 1 são linearmente atribuídos aos diferentes intervalos, do mais dissoante ao mais consoante. Este valor é então utilizado como componente de brilho ao desenhar as notas.

A potência de uma nota tenta representar o quão importante ela é na sonoridade de um trecho da composição e é atribuída baseando-se na força do tempo em que é tocada, sua intensidade (expressa por *velocidade* no formato MIDI, corresponde a quão rápida ou fortemente uma tecla de piano é pressionada) e duração. Desta forma, a potência da nota n é dada por:

$$P_n = S_n * V_n * D_n \tag{1}$$

Os três valores são dados como reais entre 0 e 1. A duração D_n e a intensidade V_n são linearmente normalizadas de acordo com a duração do compasso (notas mais longas que a duração do compasso recebem valor $V_n = 1$) e o valor máximo de intensidade. Em faixas de guitarra, é comum V_n permanecer constante em toda a composição, sendo efetivamente descartada o cálculo. Uma aproximação f é usada para obter a força do tempo S_n , dada por:

$$f(S_n) = \begin{cases} 1 & \text{se } n \text{ está no primeiro tempo do compasso} \\ 0.5 & \text{se } n \text{ está no início da segunda metade do compasso} \\ 0.25 & \text{caso contrário} \end{cases}$$
 (2)

Conforme o conjunto de dados, os valores finais de P_n podem ser normalizados novamente, atribuindo $P_i=1$ à nota i com o maior valor de potência.

Ambos os descritores foram configurados para serem gerais e se encaixarem em diferentes contextos, mas é possível alterar os números ou colocar pesos em diferentes componentes para refletir propriedades de tipos específicos de música. A **Figura 19**

mostra um exemplo da visão de detalhe da ferramenta para um dado conjunto de notas. A interação entre cada uma delas e a faixa base é dada pelo modo como seus símbolos são desenhados.



Figura 19: Exemplo de visualização de detalhe. A consoância das notas é dada pelo brilho e saturação dos símbolos e sua força é dada pelo brilho do círculo desenhado à sua volta. Informações de cada compasso são descritas ao fundo.

3.4.4. Visão geral

A visão geral contém informações a respeito da composição musical por inteiro. Todas as notas da faixa principal do arquivo MIDI são condensadas ali, junto com informações sobre complexidade, estrutura e mudanças importantes realizadas a partir de metaeventos MIDI. Todas as fórmulas listadas nesta seção presumem compassos com ao menos uma nota tocada; compassos vazios simplesmente não são representados.

Compassos são representados por segmentos de mesmo tamanho (a despeito do andamento ou fórmula de compasso observados), onde são desenhadas sobre um fundo negro caixas coloridas para representar o conjunto de notas contido em cada um. A altura do segmento é igual ao alcance das frequências das notas contidas no compasso, da mais baixa até a mais alta. Desta forma, dado um compasso *Mi* pertencente a um conjunto de

K compassos que correspondem a uma composição musical, a altura $H(M_i)$ e o comprimento $L(M_i)$ padrão da representação do compasso na miniatura são dadas por:

$$L(M_i) = L(W)/K$$

$$H(M_i) = maxPitch(M_i) - minPitch(M_i)$$
(3)

Onde L(W) é o comprimento da janela de visualização, $maxPitch(M_i)$ e $minPitch(M_i)$ são os valores máximo e mínimo de frequências atingidas pelas notas contidas em M_i , que podem assumir valores de 0 a 127 de acordo com o padrão MIDI.

O brilho do segmento (ou nível de transparência, neste caso) é proporcional ao número de notas tocadas durante aquele compasso, de modo que áreas de alta densidade sejam brilhantes e áreas com baixa ocorrência de notas fiquem mais escuras. Assim, o valor de brilho de M_i , um número real entre 0 e 1, é dado por:

$$B_{Mi} = m/\alpha \tag{4}$$

Onde m é o número de notas ativadas em Mi e α é um parâmetro que determina o valor limite de notas para o cálculo do brilho. Se $m > \alpha$, assume-se $B_{Mi} = 1$.

Em seguida, são desenhadas linhas verdes e brilhantes dentro de cada segmento, a fim de complementar a informação dada pelo alcance de frequências a cada compasso. As linhas são totalmente horizontais, posicionadas em uma altura igual à média ponderada das notas tocadas em relação à sua duração. Possuindo o mesmo comprimento de um segmento referente a compasso, a posição no eixo y da linha (Li_v) é dada por:

$$Li_{y} = \frac{\sum_{n \in Mi} F_{n} D_{n}}{\sum_{n \in Mi} D_{n}}$$
 (5)

Onde F_n e D_n são o valor de frequência a duração da nota n, respectivamente.

A **Figura 20** mostra um exemplo de visualização usando os passos mencionados até agora. É possível observar variações de densidade e progressões melódicas, assim como a ocorrência de repetições.

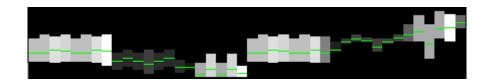


Figura 20: Exemplo do modelo de segmentos usado na visualização da miniatura. Sendo cada barra referente a um compasso, é possível perceber variações de densidade e alguns aspectos da progressão melódica. A linha verde tem a função de tornar o usuário ciente da distribuição das notas pelo alcance de frequências.

Neste ponto, todos os segmentos estão em tons de cinza. Então, para cada compasso, é calculado um valor de instabilidade, que tenta estimar o quão previsível é o som naquele trecho da composição. O modo como o sistema desenvolvido estima valores de instabilidade é definido por muitos fatores: quais notas são tocadas e quando são tocadas, se ocorreram mudanças no ritmo ou fórmula de compasso recentemente. A estimativa pode ser descrita como

$$I_i = \left(\frac{\sum_{n \in M_i} D_n P_n}{\sum_{n \in M_i} P_n}\right) + a_i \varphi + b_i \omega \tag{6}$$

 $a_i = 1$ se não há notas no primeiro tempo do compasso, 0 caso contrário

 $b_i = 1$ se a fórmula de compasso foi alterada, 0 caso contrário

Onde D_n é a dissonância da nota n, dada por $D_n = (1 - C_n)$, e P_n é a potência da nota n. Os parâmetros φ e ω determinam o nível de influência de mudanças na fórmula de compasso e momentos de silêncio não usuais na estimativa de instabilidade. De forma similar, mudanças nas definições de consoância usadas no programa também afetam o modo como a instabilidade é calculada. Valores de I_i utilizados pela ferramenta também são dados por números reais entre 0 e 1, exigindo uma normalização de acordo com os parâmetros φ e ω escolhidos.

Esta fórmula possui o objetivo de destacar aspectos musicais normalmene relacionados à instabilidade e quebra de expectativa em música, como sons dissoantes, mudanças repentinas na organização temporal e silêncio em tempos fortes de compassos. É importante observar que o valor calculado é subjetivo e pode não coincidir com a percepção de instabilidade do usuário em todos os casos – seu intuito é ser apenas um guia para indicar possíveis pontos de interesse. Os segmentos com valores altos de instabilidade são então pintados de vermelho, dando ao usuário acesso a uma barra que

permite selecionar um limiar (*threshold*) de instabilidade a partir do qual os compassos serão pintados.

Por fim, são adicionados símbolos à parte inferior da miniatura. Letras indicando trocas de tom, uma letra "X" em azul indicando trocas na fórmula de compasso e símbolos ">" e "<", verdes e vermelhos, indicando uma aceleração ou desaceleração no andamento, respectivamente. A **Figura 21** apresenta as metáforas visuais discutidas nos últimos parágrafos, sendo possível observar momentos em que ocorrem mudanças expressivas no trecho exibido.

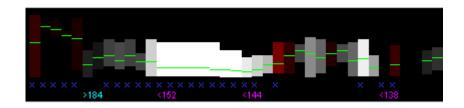


Figura 21: Visualização de instabilidade, trocas na fórmula de compasso e mudanças de andamento. Os símbolos na parte inferior assinalam possíveis pontos de interesse.

Abaixo da miniatura, são exibidas três barras coloridas. Estas barras também são divididas de acordo com os compassos e medem outras informações a respeito dos dados: complexidade, variação de intervalos e nível de repetição. Cada barra possui um valor para cada compasso, mostrando-os lado a lado de forma semelhante à visualização de descritores observada em *literature fingerprinting* (Seção 3.2.3). A coloração de cada uma é feita com uma escala de cinza a branco para valores baixos e de branco a uma cor saturada (verde, azul e ciano, respectivamente) para valores altos. A estimativa de complexidade é dada por:

$$CP_i = (ct_i * bpm)\beta + cm_i \gamma + b_i \theta \tag{7}$$

 $b_i = 1$ se a fórmula de compasso foi alterada, 0 caso contrário

Onde bpm é o valor atual do andamento em batidas por minuto, ct_i é o número de instantes diferentes em que ocorre a ativação de notas e cm_i o número de notas diferentes tocadas no compasso i. Os parâmetros β , γ e θ são responsáveis não só por controlar a influência de cada elemento na estimativa final de complexidade como também por

normalizar os seus valores, visto que a multiplicação de ct_i pelo valor em bpm do andamento pode tornar este termo dezenas ou centenas de vezes maior que cm_i , mesmo que ele não seja necessariamente mais importante.

A estimativa de complexidade tenta prover ao usuário alguma noção sobre o quão complexa é a execução daquele trecho da composição, baseando-se na velocidade em que as notas são tocadas e quantas notas diferentes são tocadas, além de possíveis alterações no ritmo. É importante notar que, assim como ocorre com a estimativa de instabilidade, os valores correspondem a apenas uma estimativa e que muitas vezes podem possuir interpretação subjetiva, sendo que a percepção de trechos de execução difícil ou complexa pode variar de acordo com o usuário.

Os outros dois valores são de natureza mais simples: o valor de repetição RP_i é igual ao número de vezes que a mesma sequência de notas observada no compasso i se repetiu em outros compassos, enquanto a variação de intervalos IV_i conta quantos intervalos diferentes foram utilizados no compasso i (até um máximo de 7, utilizando o modelo de harmonia unsigned, ou 12, com o modelo completo). O primeiro pode ser útil para determinar áreas propícias a possuírem padrões e o segundo fornece mais uma possibilidade de identificar padrões harmônicos. Na **Figura 22**, é possível observar a variação das cores nas três barras ao longo de um trecho de uma composição musical visualizada.

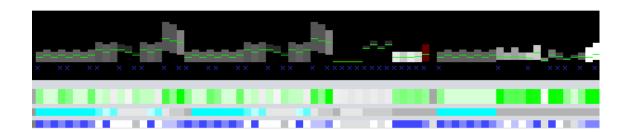


Figura 22: Visualização das barras de complexidade, repetição e variação de intervalos. A primeira indica, em verde, trechos potencialmente difíceis; a segunda marca compassos que se repetem muitas vezes em ciano; a terceira, usando a cor azul, marca compassos que utilizam muitos intervalos diferentes.

A última característica da miniatura é a capacidade de agrupar compassos a fim de simplificar a visão. Para composições muito grandes, é possível agrupar compassos de modo a exibir a miniatura em um espaço de tamanho reduzido e ainda oferecer informações a respeito de fluxos e padrões nas sequências de notas. Uma barra controlada pelo usuário determina o grau de agrupamento, sempre variando em potências de 2, até um máximo de 16 compassos por segmento. Todos os descritores de compasso passam a ser obtidos a partir de todo o grupo, mas é possível configurar a ferramenta para conservar certas informações, como atribuir a um grupo apenas o valor mais alto de instabilidade.

3.5. MOSHViz

A ferramenta protótipo que implementa o arcabouço de visualização desenvolvido neste projeto, denominada *Music Overview, Stability and Harmony Visualization* (MOSHViz), foi estruturada da seguinte forma: a entrada, recebida no formado MIDI, é lida e interpretada por um módulo responsável por converter os eventos MIDI em um padrão para instruções musicais interno ao programa (abrindo assim a posibilidade de futuramente estender o suporte do sistema a outros formatos, como partituras ETF). As instruções musicais são então processadas por um módulo de lógica que retorna conclusões a respeito dos dados, como a classificação de intervalos e identificação de momentos de instabilidade. Por fim, um módulo gráfico é encarregado de transformar a informação obtida em representações visuais e exibí-las em uma janela, por onde ocorre a interação com o usuário. A **Figura 23** apresenta um diagrama com o *pipeline* de visualização existente na ferramenta desenvolvida.

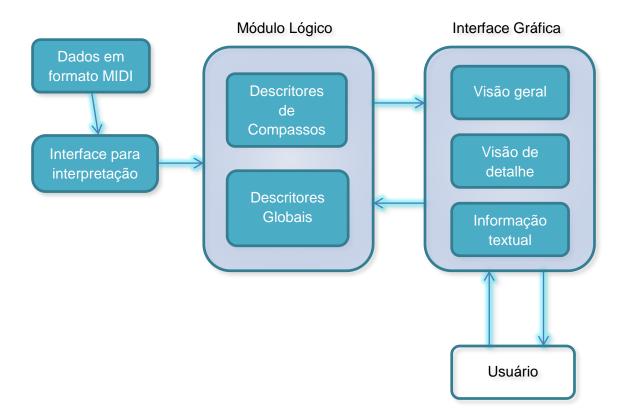


Figura 23: Diagrama representando o *pipeline* de visualização da ferramenta. A interação com o usuário ocorre por meio da interface gráfica e pode vir a requisitar novos cálculos por parte do motor lógico.

O sistema de visualização é composto por três seções principais: na maior parte da tela, é exibida a visão de detalhe, onde as notas são representadas uma a uma conforme o passar do tempo. Na parte inferior, está localizada a visão geral do conjunto de dados, permitindo ao usuário selecionar o trecho a ser exibido na visão em detalhe. Por fim, a região lateral contém informações adicionais sobre o conjunto de dados e o trecho selecionado, em forma de texto. A **Figura 24** mostra a janela de visualização da ferramenta, destacando as três seções descritas, assim como um painel de opções. A ferramenta MOSHViz foi desenvolvida em linguagem Java, utilizando a JDK 7.0.45.



Figura 24: Janela da ferramenta. 1: Painel de opções e configurações da ferramenta; 2: Visão em detalhe das notas; 3: Visão geral da composição analisada; 4: Caixas de texto com outras informações pertinentes.

A seção lateral direita da janela de visualização contém informações adicionais em formato de texto. Dados úteis sobre a composição incluem número total de compassos, número de compassos que contém notas ativas, número de compassos únicos, intervalos mais frequentemente utilizados e intervalos mais frequentemente usados no início de um compasso. O usuário também pode selecionar um compasso na miniatura e receber informações adicionais sobre ele, como índice, número total de notas, valores de complexidade e estabilidade, primeiro intervalo e índices de outros compassos com conjuntos de notas similares. Selecionar um compasso também irá ressaltar compassos com a mesma sequência de notas na miniatura. Exibir estas informações por escrito pode ser importante caso o usuário deseje ver exatamente quais fatores geraram a representação visual que está sendo exibida.

3.5.1. Interagindo com a visualização

A ferramenta desenvolvida possibilita a navegação no conjunto de dados de forma simples. Clicar em um ponto da miniatura selecionará um compasso, resultando em quatro efeitos: o instante de execução atual da composição musical será alterado para a região escolhida, a visão de detalhe será redirecionada para exibir aquele compasso em particular, estatísticas a respeito do compasso serão exibidas na caixa de texto lateral e outros compassos semelhantes ao selecionado serão destacados na miniatura.

A janela de visualização possui um painel superior onde estão localizadas opções de visualização, sendo possível desligar a exibição do *grid* na visualização de detalhe e inverter as cores da visualização, passando a utilizar um fundo branco. Além destas opções, também se encontram os botões por onde o conteúdo a ser visualizado é controlado: o arquivo a ser aberto, quais faixas são utilizadas na visualização, o modelo de detecção de tons usado no cálculo. A miniatura possui duas barras com *sliders* interativos, um para controlar o limiar em que compassos instáveis são pintados de vermelho e um para determinar o nível de agrupamento dos segmentos.

Neste capítulo foram descritos em detalhes a as metas do projeto, as abordagens utilizadas e o desenvolvimento do arcabouço de visualização. O capítulo seguinte apresenta resultados obtidos com a utilização da ferramenta apresentando exemplos que demonstram sua utilidade.

Capítulo 4. Resultados

Neste capítulo, são apresentados resultados da visualização de diferentes composições musicais usando o arcabouço desenvolvido neste projeto. É discutida a influência de cada parâmetro da visualização e como eles podem ser configurados. Por fim, são mostradas as conclusões obtidas, avaliando possíveis direções de pesquisa a serem seguidas no futuro.

4.1. Bases de dados e testes

Um dos aspectos positivos na realização do projeto foi a relativa facilidade de se obter conjuntos de dados. Por se tratar de uma única música em formato MIDI, o custo de armazenamento de cada conjunto de dados é relativamente reduzido e sua disponibilidade, considerável. Existe uma quantidade muito grande de transcrições e composições originais em formato MIDI disponíveis gratuitamente na internet. Foi utilizada como base para testes comparativos uma pequena coleção com cerca de trinta arquivos MIDI contendo faixas de guitarra em composições de $rock^{I}$, acompanhadas de partituras. Adicionalmente, foram realizados testes pontuais com outros arquivos referentes a diferentes tipos de música.

IIDIs obtidos om http://midi orghiya.com/_ http://

¹ MIDIs obtidos em http://www.free-midi.org/

De forma geral, um arquivo MIDI é composto por um cabeçalho (que contém informações como número de faixas e marcação de tempo) e uma série de informações estruturadas em forma de eventos: estes eventos são agrupados em blocos para as diferentes faixas, geralmente uma para cada instrumento, e informam onde sons começam, onde terminam, qual sua intensidade, qual sua nota (valor de *pitch*), sinais de controle, etc. Para alguns dos testes realizados, arquivos MIDI foram criados e editados usando as ferramentas MidiEditor [57] e Cakewalk Music Creator 6 [58].

Os testes consistiam na observação do comportamento da ferramenta com a variação de parâmetros e conjuntos de dados utilizados, visando identificar padrões que facilitassem a discriminação de elementos musicais contidos nos dados. A ferramenta também foi testada com diferentes usuários, a fim de indentificar facilidades ou aspectos mais informativos da visualização elaborada, mas tais testes foram realizados apenas em caráter informal.

Os testes realizados com composições musicais citadas neste capítulo fazem referência ao uso de transcrições não oficiais destas composições para arquivos MIDI, que reproduziam, sob análise inicial, fielmente a composição original. Em muitos casos, não existem versões oficiais destes arquivos, ou o acesso a eles é controlado.

4.2. Características e parâmetros

A ferramenta desenvolvida conta com muitos parâmetros diferentes a serem configurados e visa fornecer uma quantidade grande de informações. A influência de cada parâmetro e os valores a ele atribuídos são importantes no que diz respeito à comparação de diferentes conjuntos de dados.

A **Figura 25** mostra testes com diferentes modelos de coloração de notas, sendo a **Figura 25**a correspondente ao modelo descrito no capítulo anterior. Uma versão inicial do sistema representava a consoância utilizando a saturação e a potência utilizando o brilho, como pode ser visto na **Figura 25**b. No entanto, percebeu-se que isso poderia ser

ambíguo em alguns aspectos: a variação de saturação também provoca variação na luminância, o que pode provocar uma percepção errônea de brilho.

A **Figura 25**c mostra uma versão da mesma visualização, porém com cores invertidas para se aplicar a um fundo branco. Ao invés de diminuir o brilho de notas dissoantes, o valor é aumentado para que elas se mesclem ao fundo.

Outro fator importante a ser analisado é a correlação de informações descritas nas visualizações geral e de detalhe. Um dos requisitos de projeto é ser possível que pontos de interesse sejam facilmente identificados pelo usuário e que a causa do interesse na visão de detalhe seja facilmente observada. A **Figura 26** mostra um exemplo da identificação de pontos de instabilidade. Uma grande área foi marcada em vermelho na miniatura, indicando uma seção aparentemente instável. A observação daquele trecho na visão de detalhe mostra a razão: uma seção repleta de mudanças na fórmula de compasso, sendo que as notas mais fortes de cada compasso estão fora da escala (descritas na visualização como um intervalo igual a -1). O contrário também ocorre. Trechos que permanecem totalmente brancos mesmo com trocas na fórmula de compasso geralmente implicam em intervalos de consoância forte, como um primeiro intervalo, na abertura de cada compasso.

Um novo exemplo é mostrado na **Figura 27**, onde é possível observar a correspondência entre um trecho identificado como possuindo alta complexidade e suas notas: sequências muito rápidas, com 16 ou mais notas tocadas em instantes diferentes num mesmo compasso, combinados a uma variação considerável de categorias de *pitch*, fazem este segmento ser percebido como complexo. A visualização opera de forma similar para outras características, como grandes variações de notas, progressões observadas na miniatura e indicadores de repetição.

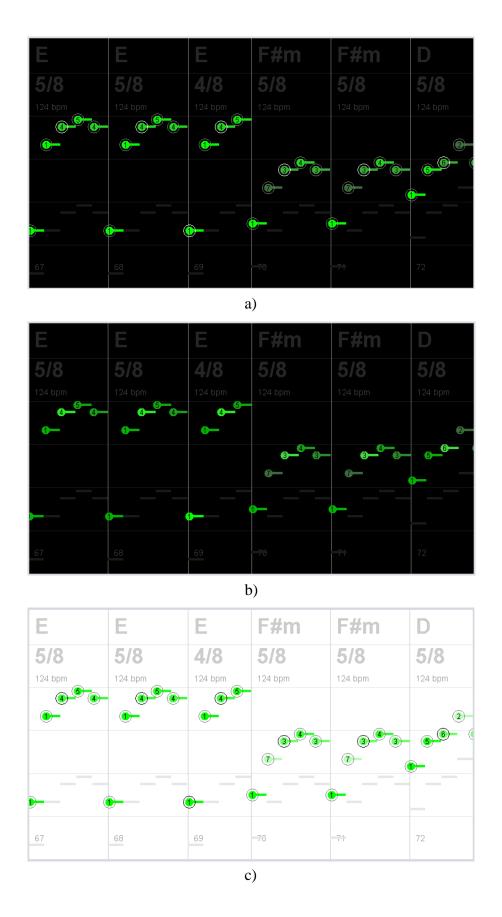


Figura 25: Diferentes modelos de coloração de notas. a) padrão; b) sem círculos, potência e consoância exibidos por brilho e saturação; c) fundo branco

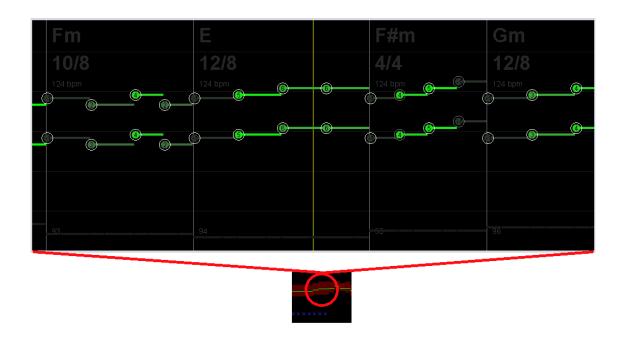


Figura 26: Equivalência entre miniatura e visão em detalhe. Um trecho determinado como instável apresenta muitas trocas na fórmula de compasso e primeiras notas do compasso sempre dissoantes.

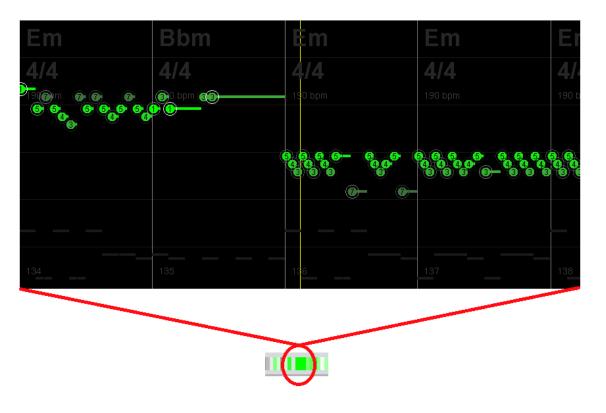


Figura 27: Equivalência entre miniatura e visão em detalhe. Uma região determinada como de alta complexidade possui grande quantidade de notas diferentes tocadas em instantes diferentes.

A ocorrência destas características é regulada pelos parâmetros das equações que determinam cada um dos descritores correspondentes. Eles podem ser regulados de duas formas, relativa ou absoluta, de acordo com a comparação desejada entre diferentes conjuntos de dados. Uma regulagem relativa implica na normalização de todos os parâmetros para atender cada conjunto de dados em particular, garantindo que sempre haja ao menos um instante da composição em que sejam atingidos valores máximos para cada descritor. Esta regulagem é útil quando se quer determinar variações que ocorrem dentro da própria composição. O modelo absoluto implica no estabelecimento de um padrão com o qual cada conjunto de dados é comparado, de modo que eles também possam ser comparados entre si.

Nos testes realizados, a ferramenta foi configurada usando o modelo absoluto, com parâmetros definidos de acordo, de modo que exista um padrão de comparação. Para tanto, este padrão precisou ser determinado por meio da observação de diferentes conjuntos de dados. É importante notar que os valores atribuídos aos parâmetros foram escolhidos de forma empírica baseando-se apenas na observação do comportamento da visualização sobre o conjunto de composições musicais, de tamanho limitado, utilizado para testes.

Os parâmetros do arcabouço de visualização essencialmente informam a importância de cada característica em cálculos e quais são suas medidas absolutas, ou seja, métricas de comparação para definir se um valor observado é alto ou baixo. Um estudo ainda precisa ser realizado no que diz respeito à observação de conjuntos de dados e definição de conjuntos de parâmetros específicos para avaliar determinados tipos de música, mas foi deixado como trabalho futuro devido à alta demanda de tempo e experiência na análise de composições musicais de diferentes estilos.

O parâmetro α , usado para determinar a densidade de um segmento na miniatura, representa um valor máximo de notas a ser observado em um compasso. Nos testes realizados, adotou-se $\alpha=24$ por ser um valor raramente ultrapassado mesmo em composições com grande densidade de notas. De modo a tolerar a presença de eventuais valores espúrios, todos os descritores foram calculados restritos ao intervalo [0,1], sendo atribuídos os valores das extremidades caso fossem ultrapassados.

Os parâmetros φ e ω , responsáveis por determinar a influência de momentos de silêncio no início de compassos e trocas na fórmula de compasso no cálculo de instabilidade, receberam valores $\varphi=0.15$ e $\omega=0.10$. Esta escolha se deu ao levar em conta o valor da estimativa de instabilidade inerente mesmo a compassos relativamente consoantes e como a soma destes valores no descritor final seria afetada.

Os parâmetros β , γ e θ , utilizados no cálculo de complexidade, passaram por um processo mais exaustivo para sua determinação. Nesse caso, o intuito foi atribuir importância semelhante aos aspectos temporais e espaciais das notas a serem tocadas (ou seja, o quão rápidas as notas são tocadas e quantas posições diferentes no braço da guitarra seriam utilizadas). Empiricamente obteve-se os valores $\beta = 0.0012$, $\gamma = 0.25$ e $\theta = 0.15$, fornecendo uma boa estimativa de complexidade. Estes valores também podem ser alterados de acordo com o desejo do usuário.

Por fim, as duas barras que complementam interação do usuário com a miniatura são responsáveis por alterar o agrupamento de compassos e o limiar de instabilidade. A Figura 28 mostra um exemplo do agrupamento de compassos. Nessa figura, uma faixa de guitarra contida numa composição musical de 23 minutos e 633 compassos ("A Change of Seasons", por Dream Theater) é mostrada em duas visões, uma com um compasso por segmento e outra com quatro compassos por segmento. A visão agrupada mantém todos os símbolos e grande parte dos contornos visuais referentes à progressão melódica são mantidos. O sistema de visualização permite a condensação de até 8 compassos num mesmo segmento, possibilitando que a composição inteira seja representada em um espaço de 80 pixels.

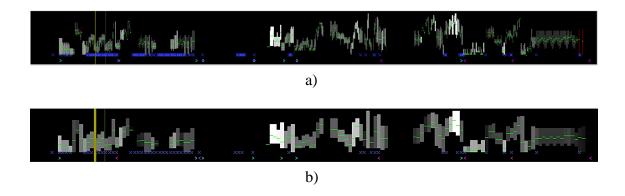


Figura 28: Mudança de escala para agrupamento de compassos: faixa de guitarra de composição de 23 minutos e 633 compassos, representada em a) um compasso por segmento e b) 4 compassos por segmento

4.3. Comparação entre músicas

Uma possibilidade interessante do sistema de visualização desenvolvido é a geração e comparação de perfis de composições musicais a partir da miniatura. Embora uma das funcionalidades originalmente planejadas para a visualização fosse a estruturação da composição a ser visualizada em seções com repetições bem definidas, a miniatura permite que a presença de padrões ou repetições seja percebida e interpretada pelo usuário de forma quase imediata.

A **Figura 29** exemplifica esta afirmação. As duas primeiras imagens, referentes às duas faixas de guitarra contidas na composição "The Evil that Men Do", de Iron Maiden, são muito bem estruturadas e com grupos facilmente discerníveis. A terceira imagem, **Figura 29**c, identifica seções contidas na segunda faixa. Com uma rápida observação, é possível ter certa noção da localização de elementos como refrão, ponte, versos e solo (áreas em verde, vermelho, azul e amarelo, respectivamente).

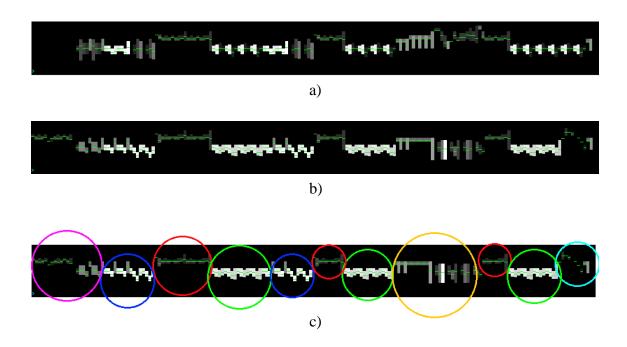


Figura 29: Visualização miniatura de "The Evil that Men Do", de Iron Maiden. As imagens a) e b) são referentes às duas faixas de guitarra contidas na composição. Em c), são destacados trechos com repetições divididos por cores.

É possível observar padrões e tendências semelhantes em diversas composições musicais. A **Figura 30** mostra a visualização de outras duas composições, "Harvester of Sorrow", por Metallica, e "Paranoid", por Black Sabbath, com regiões de repetição bem definidas, assim como regiões únicas (possivelmente referentes a partes solo) facilmente detectáveis. Além de repetições, também é possível notar áreas limitadas a um mesmo intervalo de *pitch* e um mesmo padrão de variações de densidade ao longo das composições, o que reforça a percepção da estrutura envolvida.

Não limitadas ao posicionamento das notas, as outras informações contidas na miniatura podem descrever várias outras características importantes, como a localização de pontos de alta complexidade ou concentrações de compassos instáveis. Desta forma, mesmo em faixas solo de composições focadas em guitarra, onde os padrões de repetição tendem a sumir, ainda é possível extrair muitas informações úteis. A **Figura 31** mostra duas composições focadas em melodias de guitarra e suas diferentes organizações. "Summer Song" de Joe Satriani, e "For the Love of God", de Steve Vai, são composições facilmente distinguíveis em relação aos modelos vistos anteriormente. Além de possibilitar a identificação de pontos de interesse, o estudo da miniatura fornece um perfil

de cada composição que permite identificar certas características desejadas (ou indesejadas), de acordo com os objetivos do músico.

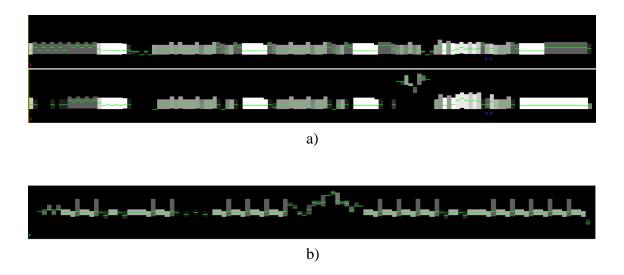


Figura 30: Visualização da estrutura de composições musicais usando a miniatura. a) as duas faixas de guitarra de "Harvester of Sorrow", de Metallica; b) a guitarra de "Paranoid", de Black Sabbath. É possível observar claramente estruturas de repetição, assim como localizar o momento onde se iniciam solos.

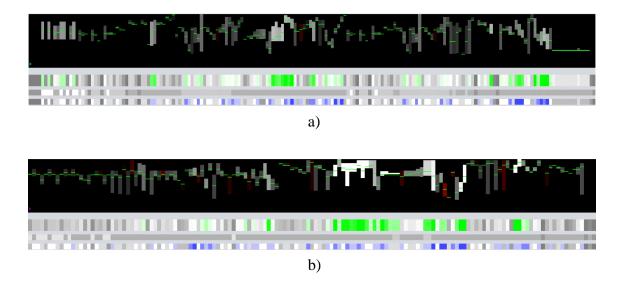


Figura 31: Guitarras solo de a) "Summer Song", de Joe Satriani; b) "For the Love of God", de Steve Vai; É possível notar que composições focadas em guitarra possuem bem menos padrões de repetição. As duas composições possuem perfis diferentes no que diz respeito a variações melódicas, de complexidade e estabilidade.

As informações exibidas podem representar conjuntos de dados muito diferentes, mesmo quando se trata de gêneros não tão distantes. Na **Figura 32**, é possível observar as faixas de guitarra de duas composições com características bem distintas: enquanto uma delas ("The Trooper", por Iron Maiden), mais curta, apresenta um formato estruturado, com repetições definidas, a outra ("Learning to Live", por Dream Theater), mais longa, possui muitas variações e mudanças ao longo de sua extensão. A observação destas diferenças pode ser fundamental a um usuário que deseja antecipadamente ter uma idéia geral do esforço necessário para executar uma composição ou algumas das técnicas envolvidas.

Todas essas representações visuais são agrupadas na ferramenta MOSHViz e compõe ferramental útil para análise e interpretação de músicas. A **Figura 33** mostra todo o modelo de visualização contido na tela principal da ferramenta. Nela é possível observar todos os elementos musicais discutidos anteriormente exibidos em conjunto. A miniatura permite observar padrões e progressões de notas, possibilitando também que pontos de interesse sejam selecionados e examinados nota-a-nota pelo usuário.

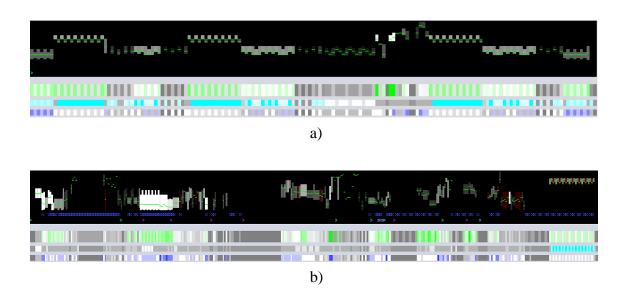


Figura 32: Diferenças entre estruturas de composições. a) faixa de guitarra contida em "The Trooper", de Iron Maiden; b) faixa de guitarra de "Learning to Live", de Dream Theater.

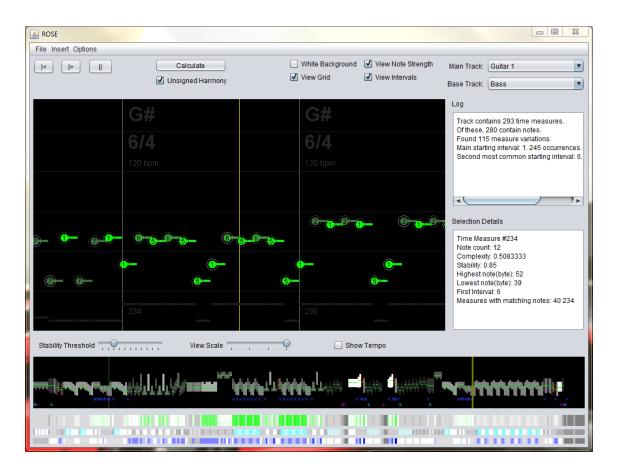


Figura 33: Janela de visualização exibindo um conjunto de dados referente a uma faixa de guitarra de "Stream of Consciouness", de Dream Theater. Todos os elementos musicais identificados pelo sistema exibidos em conjunto a fim de fornecer auxílios ao usuário que pretende interpretar uma composição musical.

Capítulo 5. Conclusões e Trabalhos Futuros

O desenvolvimento deste projeto exigiu uma aquisição de conhecimento de magnitude considerável a respeito do domínio da aplicação. No entanto, acredita-se que é possível ir muito além no que diz respeito a conceitos de teoria musical e suas possíveis representações utilizando visualização de informação, o que ilustra a complexidade do campo da música de forma geral.

A ferramenta desenvolvida se mostrou capaz de exibir uma gama de informações a respeito de uma composição musical analisada. Embora testes informais com músicos tenham obtido retornos positivos, ainda não foram realizados testes formais para determinar o grau de utilidade e funcionalidade do arcabouço de visualização descrito neste projeto. No entanto, as imagens obtidas com a ferramenta, sobretudo na visualização das miniaturas, identificam claramente estruturas, padrões e mudanças que ocorrem ao longo de uma composição musical, ilustrando a melodia a ser tocada e portanto demonstrando seu potencial para análise e entendimento de peças musicais.

O modelo de visualização de acordes, usado para cálculo de intervalos, é baseado na identificação de apenas acordes maiores e menores, aspecto que foi criticado em visualizações como a oferecida em *Colour Visualization of Music*. O modelo *unsigned* foi elaborado justamente para prover uma alternativa, que embora menos eficiente consegue operar de modo mais geral.

Novamente, deve-se notar que alguns dos descritores usados na visualização, como os valores de complexidade e instabilidade, não denotam de maneira alguma características absolutas ou referentes a alguma métrica musical precisa, servindo apenas

como guia para que o usuário explore o conjunto de dados. Nenhuma delas é infalível, podendo haver divergência entre os valores informados e a percepção do usuário.

As direções futuras de pesquisa a partir deste trabalho são diversas. É possível estudar melhores representações de características discutidas, como repetições ou harmonia (ou mesmo combinar o modelo desenvolvido com outros já existentes), ou ir mais a fundo e representar características musicais mais complexas. Adicionalmente, o uso de instâncias de dados com a estrutura criada para armazenar informações de compassos em aplicações diferentes, como técnicas de agrupamento ou projeção, poderia gerar resultados interessantes.

Uma análise mais criteriosa dos parâmetros também se faz necessária. Apesar dos bons resultados alcançados, um melhor entendimento das métricas envolvidas nas características muiscais observadas precisa ser provido para ser possível um maior controle sobre as informações gráficas geradas, possibilitando assim um melhor enquadramento do modelo de visualização às necessidades de diferentes usuários. De forma similar, um teste mais formal e completo com diferentes tipos de usuários com variados níveis de proficiência em música ajudaria a informar a utilidade do ferramental de visualização desenvolvido.

Referências

- [1] Stuart K. Card, Jock D. Mackinlay, and Ben Shneiderman (Eds.). 1999. Readings in Information Visualization: Using Vision to Think. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
- [2] Daniel A. Keim. 2002. Information Visualization and Visual Data Mining. IEEE Transactions on Visualization and Computer Graphics 8, 1 (January 2002), 1-8.
- [3] Peter Ciuha, Bojan Klemenc, and Franc Solina. 2010. Visualization of concurrent tones in music with colours. In Proceedings of the international conference on Multimedia (MM '10). ACM, New York, NY, USA, 1677-1680.
- [4] Fourney, D.W.; Fels, D.I., "Creating access to music through visualization," Science and Technology for Humanity (TIC-STH), 2009 IEEE Toronto International Conference, vol. 939, no. 944, pp. 26-27 Sept. 2009
- [5] "SoundHound", 2014. Acessado em 01 de Novembro de 2014. http://www.soundhound.com
- [6] Lillie, A. "MusicBox", 2008. Acessado em 01 de Novembro de 2014. http://thesis.flyingpudding.com
- [7] "Planetary", 2011. Acessado em 01 de Novembro de 2014. http://planetary.bloom.io/>
- [8] Hewitt, M. 2008. Music Theory for Computer Musicians. Course Technology.
- [9] Schmeling, P. 2011. Berklee Music Theory. Berklee Press.

- [10] Lacerda, O. 1967. Compêndio de Teoria Elementar de Música. Ricordi Brasileira S.A.E.C.
- [11] Isabelle Guyon and André Elisseeff. 2003. An introduction to variable and feature selection. J. Mach. Learn. Res. 3 (March 2003), 1157-1182.
- [12] McEnnis, D., Fujinaga, I. 2006. "jAudio: Improvements and additions". International Symposium/Conference on Music Information Retrieval ISMIR, pp. 385-386, 2006
- [13] Scardapane, S.; Comminiello, D.; Scarpiniti, M.; Uncini, A., Music classification using extreme learning machines, Image and Signal Processing and Analysis (ISPA), 2013 8th International Symposium on , vol., no., pp.377,381, 4-6 Sept. 2013
- [14] Gainza, M., Coyle, E. 2007. Time Signature Detection by Using a Multi Resolution Audio Similarity Matrix. 122nd Audio Engineering Society Convention.
- [15] Xi Shao; Maddage, M.C.; Changsheng Xu; Kankanhalli, M.S., "Automatic music summarization based on music structure analysis," Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). IEEE International Conference on, vol.2, no., pp.ii/1169,ii/1172 Vol. 2, 18-23 March 2005
- [16] MIDI Manufacturers Association. "MIDI.ORG", 2014. Acessado em 01 de Novembro de 2014. <www.midi.org>
- [17] MakeMusic.inc. "musicXML", 2014. Acessado em 01 de Novembro de 2014. www.musicxml.com>
- [18] Akira Maezawa, Katsutoshi Itoyama, Kazunori Komatani, Tetsuya Ogata, and Hiroshi G. Okuno. 2012. Automated violin fingering transcription through analysis of an audio recording. Comput. Music J. 36, 3 (September 2012), 57-72.
- [19] Chew, E., Raphael, C., 2010. Music and Operations Research. Wiley Encyclopedia of Operations Research and Management Science. John Wiley & Sons, inc.
- [20] Finkelstein, P. 2011. Music Segmentation Using Markov Chain Methods.
- [21] Stefik, A; Stefik, M; Curtiss, M, "An Automatic Translator for Semantically Encoded Musical Languages," Computer Music Journal, vol.31, no.4, pp.33,46, Dec. 2007.

- [22] Endo, T.; Ito, S.; Mitsukura, Y.; Fukumi, M., The music analysis method based on melody analysis, Control, Automation and Systems, 2008. ICCAS 2008. International Conference on , vol., no., pp.2559,2562, 14-17 Oct. 2008.
- [23] Hearst, M.A.; Dumais, S.T.; Osman, E.; Platt, J.; Scholkopf, B., Support vector machines, Intelligent Systems and their Applications, IEEE, vol.13, no.4, pp.18,28, Jul/Aug 1998.
- [24] "MIDITrail", 2012. Acessado em 01 de Novembro de 2014. http://sourceforge.jp/projects/miditrail>
- [25] Smith, S.M.; Williams, G., A visualization of music, Visualization '97., Proceedings , vol., no., pp.499,503, 24-24 Oct. 1997
- [26] Malinowski, S. "Music Animation Machine", 2010. Acessado em 01 de Novembro de 2014. http://musanim.com>
- [27] F. Watanabe, I. Fujishiro, and R. Hiraga: BRASS --An Interactive Interface for Supporting Score Reading--, 02-MUS-46, pp. 49-54, July 2002.
- [28] G. W. Furnas. 1986. Generalized fisheye views. SIGCHI Bull. 17, 4 (April 1986), 16-23.
- [29] Jingxuan Li; Bo Shao; Tao Li; Ogihara, M., Hierarchical Co-Clustering: A New Way to Organize the Music Data, Multimedia, IEEE Transactions on , vol.14, no.2, pp.471,481, April 2012
- [30] H. Chernoff. 1968. The Use of Faces to Represent Points in k-dimensional Space Graphically. Journal of the American Statistical Association, Vol. 68, No. 342, pp. 361-368 Jun., 1973.
- [31] Rumi Hiraga, Fumiko Watanabe, and Issei Fujishiro. 2002. Music learning through visualization. In Proceedings of the Second international conference on Web delivering of music (WEDELMUSIC'02), Christoph Busch, Michael Arnold, Paolo Nesi, and Martin Schmucker (Eds.). IEEE Computer Society, Washington, DC, USA, 101-108

- [32] Hayashi, A.; Itoh, T.; Matsubara, M., Colorscore -- Visualization and Condensation of Structure of Classical Music, Information Visualisation (IV), 2011 15th International Conference on , vol., no., pp.420,425, 13-15 July 2011
- [33] Klemenc, B., Ciuha, P., Solina, F. 2011. Educational Possibilities of the Project Colour Visualization of Music. Organizacija (Organization Journal of Management, Information Systems and Human Resources) vol. 44, no. 3, pp. 67-75 May 2011.
- [34] Mardirossian, A., Chew, E. 2007. Visualizing Music: Tonal Progressions and Distributions. In Proceedings of the 8th International Conference on Music Information Retrieval. Vienna, Austria. September 23-27. pp. 189--194.
- [35] Lerdahl, F. 2001. Tonal Pitch Space. Music Perception: An Interdisciplinary Journal. Vol. 5, No. 3, Cognitive and Perceptual Function (Spring, 1988), pp. 315-349. Oxford University Press.
- [36] Chew, E. 2001. Modeling Tonality: Applications to Music Cognition.
- [37] Tony Bergstrom, Karrie Karahalios, and John C. Hart. 2007. Isochords: visualizing structure in music. In Proceedings of Graphics Interface 2007 (GI '07). ACM, New York, NY, USA, 297-304.
- [38] Jon Snydal and Marti Hearst. 2005. ImproViz: visual explorations of jazz improvisations. In CHI '05 Extended Abstracts on Human Factors in Computing Systems (CHI EA '05). ACM, New York, NY, USA, 1805-1808.
- [39] Wing-Yi Chan; Huamin Qu; Wai-Ho Mak, Visualizing the Semantic Structure in Classical Music Works, Visualization and Computer Graphics, IEEE Transactions on , vol.16, no.1, pp.161,173, Jan.-Feb. 2010
- [40] Wattenberg, M. "The Shape of Song". 2002. Acessado em 01 de Novembro de 2014. http://www.turbulence.org/Works/song/index.html
- [41] Wattenberg, M., Arc diagrams: visualizing structure in strings, Information Visualization, 2002. INFOVIS 2002. IEEE Symposium on , vol., no., pp.110,116, 2002
- [42] Miyazaki, R., Fujishiro, I., and Hiraga, R.: comp-i: a system for visual exploration and editing of MIDI datasets, ICMC 2003, ICMA, Nov. 2004.

- [43] George G. Robertson, Jock D. Mackinlay, and Stuart K. Card. 1991. Cone Trees: animated 3D visualizations of hierarchical information. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '91), 189-194.
- [44] Peter Eades and Qing-Wen Feng. 1996. Multilevel Visualization of Clustered Graphs. In Proceedings of the Symposium on Graph Drawing (GD '96), Stephen C. North (Ed.). Springer-Verlag, London, UK, UK, 101-112.
- [45] Havre, S.; Hetzler, B.; Nowell, L., ThemeRiver: visualizing theme changes over time, Information Visualization, 2000. InfoVis 2000. IEEE Symposium on , vol., no., pp.115,123, 2000
- [46] Younesy, J.; Moller, T.; Carr, H., Visualization of time-varying volumetric data using differential time-histogram table, Volume Graphics, 2005. Fourth International Workshop on , vol., no., pp.21,224, 20-21 June 2005.
- [47] Ali, A.S.; Hussien, A.S.; Tolba, M.F.; Youssef, A.H., Visualization of large time-varying vector data, Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on , vol.4, no., pp.210,215, 9-11 July 2010
- [48] Yun Jang; Ebert, D.S.; Gaither, K., Time-Varying Data Visualization Using Functional Representations, Visualization and Computer Graphics, IEEE Transactions on , vol.18, no.3, pp.421,433, March 2012
- [49] Moere, A.V., Time-Varying Data Visualization Using Information Flocking Boids, Information Visualization, 2004. INFOVIS 2004. IEEE Symposium on , vol., no., pp.97,104, 0-0 0
- [50] Krishnan, H.; Garth, C.; Joy, K.I., Time and Streak Surfaces for Flow Visualization in Large Time-Varying Data Sets, Visualization and Computer Graphics, IEEE Transactions on , vol.15, no.6, pp.1267,1274, Nov.-Dec. 2009
- [51] Caban, J.J.; Joshi, A.; Rheingans, P., Texture-based feature tracking for effective time-varying data visualization, Visualization and Computer Graphics, IEEE Transactions on , vol.13, no.6, pp.1472,1479, Nov.-Dec. 2007
- [52] Feinberg, J. 2013. "Wordle". Acessado em 01 de novembro de 2014.

http://www.wordle.net

- [53] Keim, D.A.; Oelke, D., Literature Fingerprinting: A New Method for Visual Literary Analysis, Visual Analytics Science and Technology, 2007. VAST 2007. IEEE Symposium on , vol., no., pp.115,122, Oct. 30 2007-Nov. 1 2007
- [54] Vuillemot, R.; Clement, T.; Plaisant, C.; Kumar, A., What's being said near "Martha"? Exploring name entities in literary text collections, Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on , vol., no., pp.107,114, 12-13 Oct. 2009
- [55] Oelke, D.; Bak, P.; Keim, D.A.; Last, M.; Danon, G., Visual evaluation of text features for document summarization and analysis, Visual Analytics Science and Technology, 2008. VAST '08. IEEE Symposium on , vol., no., pp.75,82, 19-24 Oct. 2008
- [56] Bigand, E. 1997. Perceiving Musical Stability: The Effect of Tonal Signature, Rhythm, and Musical Expertise. Journal of Experimental Psychology: Human Perception and Performance, 808-22. June 1997.
- [57] Schwenk, M. 2013. "MidiEditor". Acessado em 01 de novembro de 2014.

http://midieditor.sourceforge.net

[58] Cakewalk, Inc. 2014. "Cakewalk Music Creator 6 Touch." Acessado em 01 de novembro de 2014.

https://www.cakewalk.com/Products/Music-Creator

- [59] Vliegen, R.; van Wijk, J.J.; van der Linden, E.-J., Visualizing Business Data with Generalized Treemaps, Visualization and Computer Graphics, IEEE Transactions on , vol.12, no.5, pp.789,796, Sept.-Oct. 2006
- [60] Muelder, C.; Kwan-Liu Ma, A Treemap Based Method for Rapid Layout of Large Graphs, Visualization Symposium, 2008. PacificVIS '08. IEEE Pacific , vol., no., pp.231,238, 5-7 March 2008