

Universidade Federal de Santa Catarina
Centro Tecnológico
Departamento de Informática e Estatística
Curso de Bacharelado de Ciências da Computação

EDSON EDGAR DA SILVEIRA

UMA CONTRIBUIÇÃO À COMPARAÇÃO ENTRE ARQUIVOS DE MÚSICA
PARA UMA IMPLEMENTAÇÃO COMPUTACIONAL

Florianópolis, 29 de setembro de 2013
Edson Edgar da Silveira

UMA CONTRIBUIÇÃO À COMPARAÇÃO ENTRE ARQUIVOS DE MÚSICA PARA UMA IMPLEMENTAÇÃO COMPUTACIONAL

Trabalho de Conclusão de Curso,
elaborado como parte dos requisitos para
a obtenção do grau de Bacharel em
Ciências da Computação.

Orientação: Prof. Dr. João Candido Dovicchi

Universidade Federal de Santa Catarina
Florianópolis, 2013.2

Edson Edgar da Silveira

Uma Contribuição À Comparação Entre Arquivos De Música Para Uma Implementação
Computacional

Trabalho de Conclusão de Curso apresentado ao
curso de Ciências da Computação como parte
dos requisitos para a obtenção do grau de
Bacharel em Ciências da Computação.

Orientador: Prof. Dr. João Candido Dovicchi

Comissão Examinadora

Prof. Dr. João Candido Dovicchi – Orientador
Departamento de Informática e Estatística – UFSC

Prof. Dr. Luiz Fernando Jacintho Maia

Prof. Dr. Wemerson Delcio Parreira

Florianópolis, SC _____ de _____ de _____

DEDICATÓRIA

À minha família, por todo apoio durante a graduação

AGRADECIMENTOS

À minha família e amigos,
ao professor e orientador Dr. João Candido Dovicchi pela orientação,
e aos membros da banca avaliadora, os professores Dr. Luiz Fernando Jacintho Maia e
Dr. Wemerson Delcio Parreira.

Resumo

O compartilhamento e compra de músicas online vem crescendo a medida que o meio digital se torna cada vez mais presente na vida das pessoas. Assim como outras atividades, este novo meio possibilita e muitas vezes exige formas mais eficientes de lidar com a distribuição musical. Uma dessas formas envolve como melhor sugerir e buscar produtos semelhantes.

Na área musical foram incorporadas técnicas provindas de outros mercados como análise do histórico do consumidor e também o uso extenso de labels para melhor indicar artistas ou músicas semelhantes. Em outros casos, diversas técnicas foram desenvolvidas para que se possibilite a busca de uma música com base em um pequeno trecho.

Se por um lado todas essas formas são bem-vindas, convenientes e muitas vezes eficientes, elas não abrangem todas as possibilidades e possuem claras limitações. Poder sugerir e comparar arquivos de música com o auxílio de conceitos musicais pode abrir um universo de novas possibilidades para que, ao lado das técnicas já bem estabelecidas, atenda melhor atender as necessidades desse mercado.

Este trabalho busca apresentar uma forma para comparação de arquivos de música baseados em conceitos da teoria musical. A proposta apresenta os passos a serem tomados para que se tornem possíveis assim como suas dificuldades e limitações.

Palavras chaves: música, áudio digital, DSP

Abstract

The sharing and selling of music online has been growing steadily as the digital medium becomes more present in everyone's everyday life. Much in the same way as other activities, this new medium enables and in many cases requires new and more efficient ways of dealing with the distribution of music. One of them consists in how to better suggest and search similar products.

In music, techniques from different markets have been incorporated such as using consumer's purchase history or the use of labels to better suggest new artists or music the consumer may potentially be interested in. In other cases techniques have been developed with the aim of searching a particular music by inputting a small fragment for example.

If on one hand all those methods are welcome, convenient and often efficient, on another they do not offer the whole breadth of possibilities as well as having clear limitations. Being able to suggest music files by the very music nature in them can open a whole new set of possibilities to better deal with this market with the aid of the previous well established methods.

This study aims to present a method to approach the comparison of music files based musical theory. Such method takes into account the steps required to make them possible as well as their problems and limitations.

Key words: music, digital audio, DSP

Sumário

1	Introdução.....	1
1.1	Motivação.....	5
1.2	Objetivos.....	6
1.2.1	Objetivo Geral.....	6
1.2.2	Objetivos Específicos.....	7
1.3	Metodologia.....	7
2	Fundamentação Teórica.....	9
2.1	Som e suas características.....	10
2.2	Observação do valor de frequência como nota musical.....	11
2.3	Modulações.....	13
2.4	Notas Não-harmônicas e Ornamentação.....	17
2.5	Transformadas de Fourier.....	18
2.6	Centroide.....	22
3	Proposta.....	23
3.1	Motivação.....	23
3.2	Resumo.....	23
3.3	Desenvolvimento.....	25
4	Resultados e Discussão.....	50
4.1	Vantagens.....	50
4.2	Desvantagens.....	51
5	Conclusão.....	55
5.1	Considerações Finais.....	55
5.2	Trabalhos futuros.....	56
6	Referências Bibliográficas.....	57
7	Apêndice.....	62
7.1	Apêndice A – Artigo.....	62

Lista de Figuras

Figura 1: Venda de música em diferentes formatos p.3

Figura 2: Círculo de quintas para tons maiores e menores p.15

1 Introdução

Durante as últimas duas décadas a indústria fonográfica tem passado por grandes e significativas mudanças. Entre elas a maneira de como produzir, vender e atrair consumidores. Um dos fatores mais relevantes para esta mudança de paradigma neste setor comercial vem do avanço tecnológico. A popularização de mídias como o CD criado pela *Sony* e *Philips*, o advento de novas mídias, como o DVD também pela *Sony* e *Philips* em conjunto com a *Panasonic* e *Toshiba*, assim como o sucesso do computador pessoal somado ao importante surgimento da Internet promoveram mudanças bruscas para a indústria (DOLATA, 2011). Mudanças quais afetaram não só as empresas envolvidas nesta área como os consumidores. Uma nova cultura de como consumir música surgiu. Ouvir música se tornou algo muito mais acessível e se tornavam mais fáceis de compartilhar. Consumidores passaram a ser bombardeados por um mercado gigantesco. Produtos como o *Discman*®, que se tornou o primeiro reprodutor de CD da linha *Walkman*® da *Sony*, se tornavam muito comum entre a população mais jovem para ouvir música fora de casa, a caminho da escola ou do trabalho (BBC, 2004). Isso também promoveu a produção e uma maior popularização de músicas que servissem melhor este propósito e que em consequência popularizava ainda mais esta nova cultura de consumo.

Paralelamente cresceu a pirataria devido as facilidades tecnológicas (BBC). Se por um lado a indústria expandia seus horizontes, com a globalização promovida pela Internet, por outro a pirataria surgiu como uma ameaça cada vez mais preocupante ao

modelo de negócios ainda muito tradicional apesar das mudanças que o setor havia passado. Cada vez mais se pensava como ela poderia ser controlada. Tentativas diversas de segurança das mídias pouco efeito faziam. Propagandas conscientizadoras não conseguiam acompanhar o aumento da pirataria.

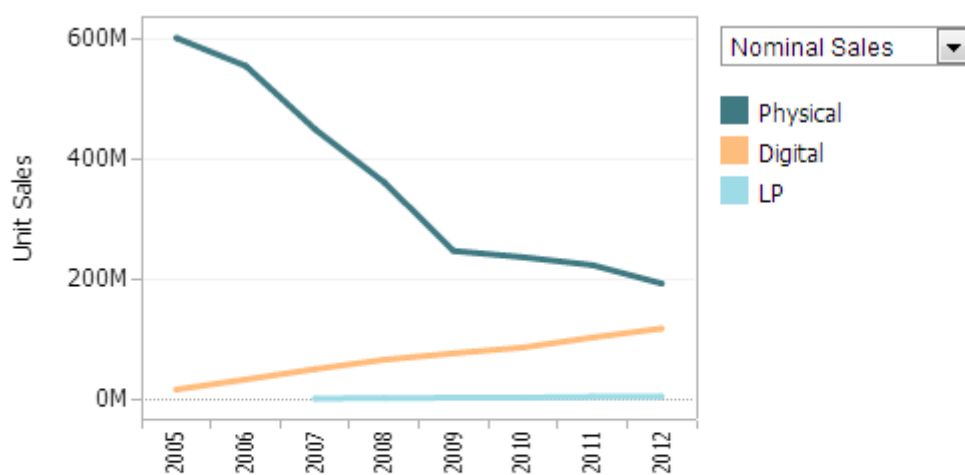
O surgimento das redes *peer-to-peer* no final da década de noventa elevou a situação a um patamar ainda mais sério levantando diversas questões quanto a validade daquela forma de compartilhamento de arquivos assim como o papel da indústria fonográfica atual (BBC, 2004). Empresários, artistas e consumidores pareciam divididos. Independente disso o acesso à música por meios não considerados regulares pelas leis dos países aumentava à velocidades exorbitantes (CNN, 2010).

Pouco depois a popularização dos aparelhos de reprodução de arquivos MP3 (*MPEG-1/2 Audio Layer 3*) garantia que o consumo de música se tornava cada vez maior, trazendo grandes lucros para as empresas que se aproveitavam bem do novo mercado. A criação de sites de venda de músicas no formato MP3 como o *iTunes™* da *Apple* foi um dos grandes sucesso por parte das empresas de se adaptar à nova realidade. Passaram a oferecer os produtos de maneira barata a ponto de atrair parte do público que buscava de maneira ilegal além de usar a Internet que se tornava o meio mais popular para venda de músicas como através do site de compras *Amazon*.

Atualmente o foco para venda online de músicas tem sido grande com altos investimentos. A *Apple*, por exemplo, gasta em média 3,75 bilhões de dólares por ano com o *iTunes™* segundo pesquisas além de ter uma renda anual em torno de 13,5

bilhões de dólares (MACNN, 2013). São números significativos que mostram o poder desse meio de venda quando bem utilizado por empresas.

É relevante ainda observar que mais do que um novo meio de comercializar música, o meio digital tem se tornado quase obrigatório. Hoje em dia é através da internet que boa parte dos futuros consumidores tem contato com as lojas. Não possuir um sistema online de vendas pode vir a prejudicar significativamente a visibilidade da empresa se tornando para muitos consumidores virtualmente inexistentes uma vez que as vendas por meios físicos vem caído consideravelmente como observado pelos *reports* anuais da companhia *Nielsen*, através das quais o site *tableausoftware* geral gráfico abaixo.



Data: The Nielsen Company & Billboard
LP album sales unreported before 2007. Overall music sales includes all albums, singles, music videos and digital tracks. Album sales do not include digital track-equivalent

Figura 1: Venda de música em diferentes formatos

Fonte: [http://www.tableausoftware.com/public/blog/2013/01/rise-digital-music-](http://www.tableausoftware.com/public/blog/2013/01/rise-digital-music-1773)

[1773](http://www.tableausoftware.com/public/blog/2013/01/rise-digital-music-1773)

Possuir um sistema online é então, para todos os efeitos, “entrar no mapa”. É um primeiro passo imprescindível. Mas não basta porém apenas isso. O cuidado com a imagem e a reação do consumidor para com o site deve ser o maior possível assim como em uma loja física. Como então concorrer nesse novo meio se tornou um novo desafio aberto para grandes inovações das mais diversas origens. Novas possibilidades de integração com *players* de música e outros aplicativos logo surgiram.

Dados porém, mostram que as vendas de música estava até recentemente em um longo período de declínio (IFPI, 2013). Enquanto isso a pirataria se tornava cada vez mais difundida com o crescimento da população com acesso à tecnologia e conhecimento para usá-la comparado com anos anteriores..

Com muitas discussões e controvérsias, a indústria passou a tomar medidas para concorrer com a pirataria. Preços de música baixaram consideravelmente e a possibilidade de baixar faixas individuais em vez de um álbum completo de um artista cada vez mais comum.

Como essa mudança de foco para vendas online, que é o estado que a sociedade se encontra há alguns anos, a busca torna-se cada vez maior por maneiras de usar o meio, Internet, para melhor atrair um consumidor em potencial. Diversas propostas foram levadas adiante para tirar o máximo desta tecnologia e mercado. Softwares que além de oferecer maneiras de comprar também são reprodutores de mídia, como o já citado *iTunes*[™], foram uma maneira de chamar novos consumidores através dessa integração e simplicidade o que também desestimula a pirataria por parte dos usuários. A possibilidade de compra através de aparelhos móveis, que ruma

à primeira posição em termos de aparelhos mais usado para acesso à Internet no futuro próximo, foi outra maneira.

Apesar disso, dentro deste período de grandes transformações que se faz presente, a necessidade de adaptações para o cenário atual e de novas ideias ainda é grande. Ideias e conceitos que foram previamente utilizados para outros fins podem agora ser vistos com outros olhos, com foco nessa indústria.

Em uma era com a crescente tendência de dar liberdade ao usuário de usar a tecnologia de acordo com seu perfil individual não há nada mais natural que personalizar a experiência do sistema online de vendas de música. Uma das formas comumente usadas é sugestão de produtos com base no histórico do consumidor. Outra forma de facilitar o consumir a encontrar a música desejada é por meio de *tags* indicado características ou gênero da música. Devido as limitações dessas formas, outras formas de análise de músicas e comparação de arquivos de música se tornam cada vez mais estudadas.

1.1 Motivação

Há muitas formas nas quais um indivíduo poderia julgar duas músicas semelhantes. Apesar de serem subjetivos muitos dos critérios costumam ser recorrentes entre as pessoas. Ainda que em muitos casos exista uma dificuldade de explicar os motivos de semelhança ou características de uma música elas costumam ter uma explicação musical e física do fenômeno. É comum ainda, que muitas e muitas

peessoas tentam com dificuldade buscar novas músicas de acordo com tais parâmetros. Ainda que sistemas como o de *tags* seja comum ele pode não representar de fato características buscadas por um indivíduo. Por este motivo se faz necessário uma análise musical do conteúdo presente na música comumente disponível pelo meio digital. Através de uma compreensão musical de muito desses critérios subjetivos pode-se buscar fazer uma análise de um arquivo de música para encontrar semelhanças, ou não, em relação a outro.

Uma das grandes motivações de buscar a comparação desses arquivos de música é a possibilidade de então poder oferecer a um consumidor ou mesmo um indivíduo qualquer a possibilidade de buscar músicas semelhantes através de critérios baseados na fundamentação teórica música. Através disso a pessoa pode então ter acesso a músicas semelhantes a alguma outra que dificilmente se apresentaria por outras formas de sugestão como análise do histórico de compras ou o sistema de *tags*. De todas as questões teóricas musicais uma se destaca e motiva o estudo que é a questão da análise harmônica.

1.2 Objetivos

1.2.1 Objetivo Geral

- Compreender e identificar uma forma de comparação de arquivos de música com base em questões musicais tais como escalas e campos harmônicos que

possa oferecer uma nova forma de abordar a questão de sugestão de músicas similares.

1.2.2 Objetivos Especificos

- Criar uma proposta assim como os passos gerais a ser seguidos para tornar possível uma futura realização da proposta apresentada no trabalho.
- Em conjunto dos passos deve entender as limitações, dificuldades, desvantagens e possíveis formas de melhorias da abordagem proposta..

1.3 Metodologia

A parte central do trabalho consiste em uma proposta para comparação de arquivos de áudio. Sua elaboração contempla não somente sua explicação como também sugestões de propostas a serem abordados em futuros trabalhos assim como dificuldades presentes.

Após a revisão teórica segue-se um resumo da proposta e também a motivação para escolha da mesma. É então apresentado a proposta propriamente dita com seus os passos que a compõe. Será apresentado como poderia ser aplicada e seus passos usando linguagem comum. Momentos que se fizerem necessário podem conter linguagem matemática.. Será analisado vantagens e desvantagens da proposta em

foco considerando o que foi apresentado, novas formas e técnicas que possam eventualmente contribuir para como o desenvolvimento da ideia proposta.

Concluído a proposta considerações finais serão apresentadas comparando impressões iniciais da proposta com as finais após o estudo extensivo da mesma.

O desenvolvimento da proposta será apresentado de forma a buscar compreender o escopo geral da mesma. Ainda que diversos passos envolvam questões possivelmente técnicas o foco será mantido no todo, como melhor abordar a comparação de arquivos de áudio e não em soluções técnicas para cada um dos passos que a proposta contém.

2 Fundamentação Teórica

Para o escopo deste trabalho será muitas vezes referenciados conceitos da música ocidental. Porém, devido ao vasto número de conceitos da área musical a ser utilizado não cabe aqui fazer um estudo para cada conceito musical. Alguns porém são de suma importância para o trabalho e serão aqui devidamente revisados. É importante portanto estar a par do vocabulário básico musical já que será utilizado ao longo do estudo.

Por outro lado alguns conceitos básicos serão abordado com o intuito de como compreendê-los de forma científica e mais especificamente no ramo computacional. Conceitos como notas podem ser explicados fisicamente e se pode buscar enxergar esses elementos através de arquivos de áudio. Poder extrair esse tipo de informação da estrutura dos arquivos é de grande importância para manipulação dos mesmos e para poder derivar outras informações a partir destas.

Isso também significa que técnicas que envolvem manipulação de dados referentes a música e arquivos de áudio também podem merecer esclarecimento. Conceitos como transformada de Fourier que são muito decorrentes na compreensão da análise de sinais digitais merecem uma introdução.

2.1 Som e suas características

Os teóricos musicais Bruce Benward e Marilyn Saker (SAKER;BENWARD, 2008, pg. XIII-XV) definem som como sendo a sensação percebida pelos órgãos quando se ouve vibrações, isto é, ondas de som, através do ouvido. Segundo os mesmos o som possui quatro características identificáveis: sua altura, intensidade, duração e timbre.

A altura está diretamente ligada com a frequência. Sendo que as variações de frequências é o que se identifica como variações de altura. Isto significa que quanto maior o número de ondas sonoras produzidas por um corpo elástico mais agudo e, no caso contrário, mais graves. Uma nota por exemplo se refere a uma altura específica.

A intensidade, também conhecida como amplitude se refere a altura no sentido de volume.

A duração por sua vez diz respeito a, como o próprio nome sugere, a duração de uma nota por exemplo. Para padrões de duração se usam dos termos “métrica” e “ritimica”. O primeiro termo descreve os recorrentes pulsos de duração igual enquanto o segundo o padrão de durações irregulares.

Finalmente, o timbre se refere a qualidade ou cor de um som. É o que torna possível diferenciar o som de diferentes instrumentos. A qualidade do som é resultado da forma de um corpo e como este é colocado em um estado de vibração e então como o ouvido humano percebe as séries harmônicas provindas dessas vibrações. As séries harmônicas por sua vez contém muitas alturas sendo produzidas ao mesmo

tempo por um corpo em vibração. Múltiplas vibrações ocorrem ao mesmo tempo, não somente no corpo como um todo, mas também, ao imaginar uma corda, ao longo de sua metade, terça parte, quarta parte, etc. (SAKER;BENWARD, 2008).

2.2 Observação do valor de frequência como nota musical

Como já indicado, este estudo toma com base conceitos da música ocidental. Nela observa-se a escolha de 12 semitons para representar frequências todas em torno de um elemento central definido em 440 Hz por razões históricas. As notas, de maneira ascendente são as seguintes *dó* (C), *dó#* (C#), *ré* (D), *ré#* (D#), *mi* (E), *fá* (F), *fá#* (F#), *sol* (G), *sol#* (G#), *lá* (A), *lá#* (A#) e *si* (B). O 440 Hz corresponde a uma nota A. É importante ressaltar porém que essas 12 notas são se referem a uma única frequência. Na realidade cada uma delas podem representar inúmeras frequências porém seguem um padrão. Cada nota representa uma frequência e todos os seus harmônicos tornando assim cíclico o que significa que a nota após o *si* é novamente *dó* porém com uma frequência diferente do *dó* anterior. Isso pode ser facilmente compreendido observando o conceito de séries harmônicas já apresentado.

Ao analisar um arquivo de áudio em busca de uma nota o retorno vem como frequência. Dificilmente os valores retornados serão exatamente os que representam as notas e então é portanto necessário fazer um cálculo para descobrir a nota representada.

Antes porém, é necessário compreender como acontece a divisão da frequência em notas. Por convenção o sistema utilizado atualmente é de temperamento igual. Esta forma da divisão, ou afinação, separa os doze semitons em intervalos iguais, cada um representando $2^{1/12}$, o que se define como 100 cents. Também por convenção toma-se como referência o A4, que equivale a 440 Hz. Com estas informações se pode facilmente calcular uma nota para uma frequência seguindo a seguinte fórmula:

$$f = 2^{12k/12} \times 440 \text{ Hz} = 2^k \times 440 \text{ Hz}$$

em que f representa a frequência enquanto o k , a quantidade de oitavas acima do A4 mais 1, já que $k = 0$ representa a oitava atual. A fórmula fornece portanto a posição que a frequência se encontra dentro de uma oitava com base no A4 representado pelo valor 440 Hz. Se a frequência resultar em $k = 0$ então tem-se simplesmente o mesmo A, enquanto um valor 1 ou outro inteiro também representará o A porém em alturas diferentes. Uma vez que uma oitava é dividida em 12 semitons o valor de $k = 1/12$ fornece o próximo semitom após o A, isto é, A#. $k = 11/12$ pela mesma lógica fornece G#.

Ao seguir a fórmula acima é natural que haverá situações nas quais o cálculo gerará valores fracionários para k . Para compreender a parte não inteira do k é necessário lembrar que como mostrado anteriormente a distância entre semitons equivale a 100 cents. Já que o k representa a quantidade de oitavas de distância ao 440 Hz, uma oitava inteira então corresponde a $12 \times 100 = 1200$. Portanto multiplicando a parte fracionária por 1200 obtem-se a quantidade de cents que se encontra acima do A da oitava que se encontra. Uma vez que a distância para cada semitom é de 100

cents é possível dividir o valor resultante por 100 e também aplicar mod. Enquanto a parte inteira da divisão resulta em um número de 0 a 11, representando os semitons, o resultado do mod fornece um valor de 0 a 99 representando a os 100 cents entre um semitom e outro. Com isso se pode finalmente definir qual a nota mais próxima da frequência dada. De 0 a 49 permanece-se da nota atual enquanto um valor superior a 49 ate 99 indica que está mais próxima ao próximo semitom (University of New South Wales).

2.3 Modulações

"The process of modulation, meaning change o key, is one of the most important resources of variety in music."(PISTON, 1959. p.77)

Assim Piston descreve a modulação e segue explicando que a mudança para outro tom significa mudar o centro tonal, se referindo a tônica uma vez que este possui essa característica. Essa mudança é feita por uma adoção de um outro tom no qual todos os outros notas são relacionados. Conclui explicando que qualquer acorde ou grupo de notas pode ser interpretado em qualquer tom e que é essa ambiguidade presente num único acorde qualquer que forma a base de modulação (PISTON, 1959).

Saker e Benward definem de forma similar dizendo que:

"Modulation is a process that results in a shift of tonal center. The term applies to those occasions in music when one established tonal center gives way to another."(SAKER;BENWARD, 2008. p. 315)

E também explica que a maior parte das modulações ocorre entre tons relacionados, isto é, aqueles que possuem no máximo um acidente de diferença em suas notas. Como exemplo dado se tem que, se o tom original é de *dó* maior, os "tons vizinhos" são *G* maior e *fá* maior e as relativas menores dessas tonalidades, *lá* menor, *mi* menor e *ré* menor respectivamente. É explicado ainda a existência de outras duas formas comuns de modulação. A primeira sendo a modulação modular que ocorre no ponto de uma progressão cromática. A segunda é a modulação direta, também chamada de modulação de frase, que ocorre entre frases ou grande seções onde uma frase musical é feita em um tom sendo que a seguinte começa imediatamente em outro tom (SAKER; BENWARD, 2008).

Uma ferramenta de visualização da relação entre diferentes tonalidades é o círculo de quintas. Ele consiste em um círculo dividido em doze no qual cada um dos pontos onde há divisão é indicada uma nota. Na direção horária as notas estão em uma ordem que representa que a seguinte está há um intervalo de quinta justa em relação a anterior. De forma inversa, ou anti-horária, se observa uma ordem seguindo intervalos de quartas justas. A importância do círculo de quintas quando se pensa é modulação é explicada pelos intervalos que controlam a ordem que as tonalidades estão distribuídas. Uma escala maior natural com o centro tonal em uma das notas do círculo possui apenas um acidente musical de diferença para as escalas centradas nas

notas vizinhas que são a quarta justa à esquerda e quinta justa à direita. Elas são, pelos motivos explicados anteriormente boas candidatas para modulação pela diferença em apenas um grau da escala (SAKER; BENWARD, 2008).

Uma outra forma comum na qual o círculo de quintas é representado é com um círculo interno representando escalas menores. O círculo central busca explicitar outra relação comum entre escalas, a entre uma escala e sua relativa maior ou menor. Assim sendo, o círculo central também obedece a ordem dos intervalos mas é ajustado de tal forma que uma nota do círculo central esteja alinhada, ou paralela, com intervalo de terça menor em relação ao círculo externo. Vendo do círculo externo para o interno isso significa um intervalo de sexta maior. Uma vez que uma escala menor natural pode ser visto como uma escala maior iniciada no sexto grau, ou ainda como o modo eólio da escala diatônica fica claro que essa relação apresentado entre o círculo externo e interno é entre escala maior, no círculo externo, e sua tonalidade relativa menor, no interno.

O gráfico a seguir apresenta o círculo de quintas pensado para tonalidade maiores e menores.

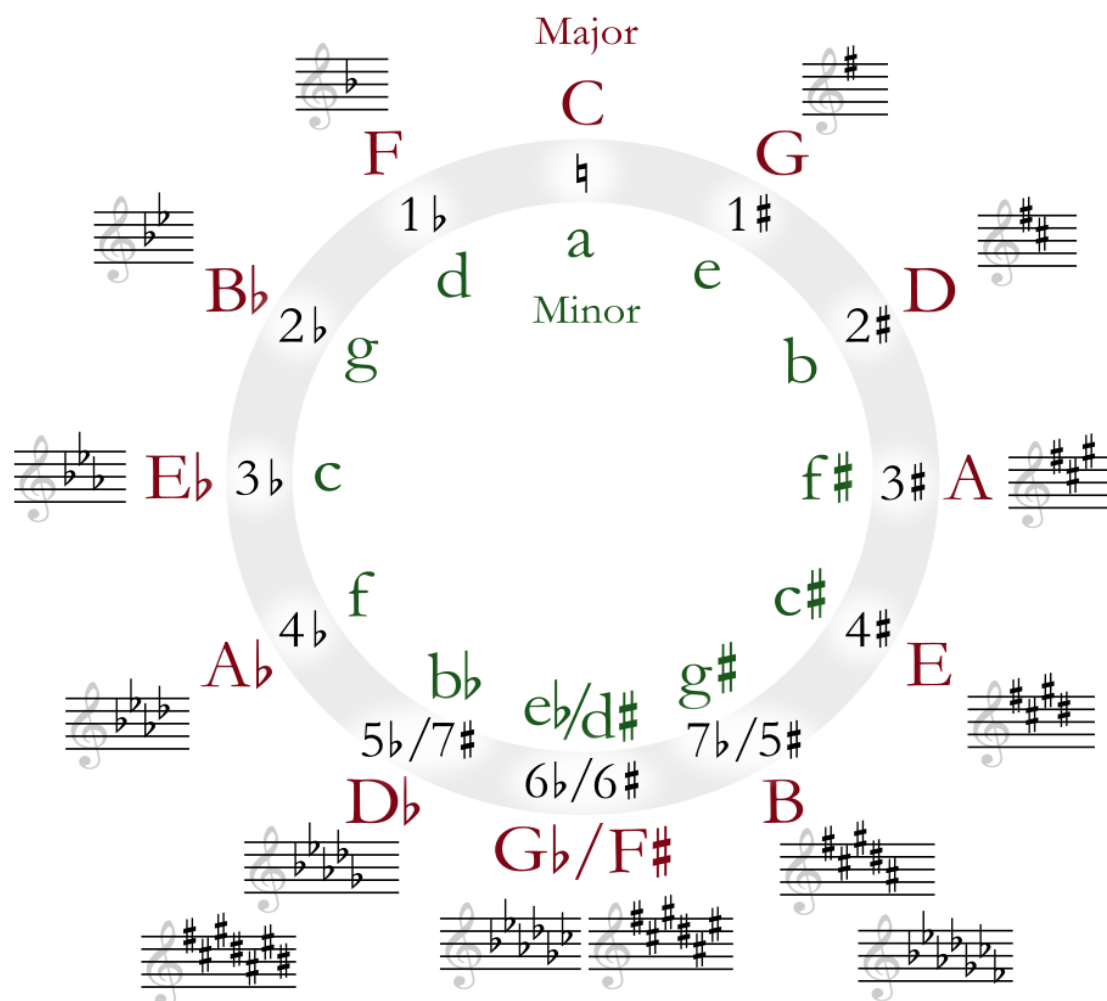


Figura 2: Círculo de quintas para tons maiores e menores

Fonte: http://upload.wikimedia.org/wikipedia/commons/thumb/3/33/Circle_of_fifths_deluxe_4.svg/500px-Circle_of_fifths_deluxe_4.svg.png

2.4 Notas Não-harmônicas e Ornamentação

Ao explicar a existência de notas não-harmônicas dentre as primeiras coisas que Piston deixa claro é:

"Literally, there is no such thing as a nonharmonic tone, since tones sounding together create harmony" (PISTON, 1959. p. 102)

A partir disso porém ele vai a explicar a evolução da linguagem musical na qual notas não-harmônicas vem a representar aquelas que não aparecem como fatores de acordes. Ele ainda busca evitar a má interpretação de que notas não harmônicas não derivam suas características do fato de serem estranhas à harmonia presente e ainda escreve:

"Their true nature is inherently melodic and can be discovered by study of the melodic line alone" (PISTON, 1959. p. 102)

Saker e Benward (2008. p. 102), trata da mesma maneira falando de forma simples que são notas que soam juntas a um acorde porém não fazem parte das notas dos mesmos. Indica ainda que os casos mais comuns são dissonantes e criam intervalos de segunda, quarta e sétima sendo que intervalos diminutos e aumentados também são considerados dissonantes. A dissonância dessas notas não-harmônicas é explicado por eles como sendo calculados contra a nota de altura mais baixa de um

acorde, isto é, a mais grave, independentemente de quantas outras vozes se fazem presente com a exceção do caso dela mesma ser a mais grave. É indicado ainda a forma como mais são encontradas que é em um padrão de três notas no qual a nota precedente e a que sucede a nota não-harmônica não notas presentes no acorde.

A literatura ainda indica uma grande variedade de notas não-harmônicas possuindo diversas funções tais como serem notas de passagem (SAKER; BENWARD, 2008. p. 104-111).

Pode-se pensar que esses representam as notas de ornamentação porém é possível obter notas de ornamentação que fazem parte de acordes (HURON, 2002).

2.5 Transformadas de Fourier

A série de Fourier que recebe o nome do matemático francês Jean Baptiste Joseph Fourier (21 de março de 1768 – 16 de maio de 1830) representa uma função periódica como soma de funções periódicas. A forma segue, sendo $f(t)$ uma função periódica em t :

$$f(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cdot \cos(k\omega t) + b_k \cdot \text{sen}(k\omega t))$$

em que ω representa a frequência da seguinte forma:

$$\omega = 2\pi/T$$

em que T representa o período, medido em segundos.

Se $f(t + nT) = f(t)$, então $f(t)$ é periódica então a_n e b_n são os coeficientes de expansão da série.

A série pode ainda ser vista de forma complexa:

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{in\omega t}$$

Com,

$$c_n = \frac{1}{T} \int_c^{c+T} f(t) e^{-in\omega t} dt$$

Mas sendo que $c = -T/2$ se tem:

$$c_n = \frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{-in\omega t} dt$$

É dessa série que se baseia uma série de transformações matemáticas conhecidas como Transformadas de Fourier que possuem aplicações em diversas áreas e uma frequentemente empregada é na área de análise de sinais digitais (KUC, 1988).

A transformada clássica de Fourier consiste em transformar funções contínuas produzindo como retorno uma função contínua de frequência que é conhecida como distribuição de frequência e vice-versa já que esta operação pode ser reversível. É possível ainda aplicar em funções descontínuas caso o número de descontinuidade seja finito. No caso da entrada ser uma função de tempo e a saída uma função de frequência a função de transformação pode ser vista como:

$$S(f) = \int_{-\infty}^{\infty} s(t) \cdot e^{-i2\pi f t} dt$$

em que S é a função, f se refere a função de tempo a ser transformada e f a frequência. O resultado é um número complexo. Ao avaliar para todos os infinitos pontos se produz a função do domínio da frequência. De forma inversa tem-se que a recombinação de todas as possíveis frequência resultará em $S(f)$:

$$s(t) = \int_{-\infty}^{\infty} S(f) \cdot e^{i2\pi f t} df$$

Observa-se porém que ela se aplica para domínios contínuos. De forma prática, como em análise de sinais digitais é necessário trabalhar com uma versão discreta. É então que se fazem necessárias o uso da transformada de Fourier determinística (DFT) que é calculada comumente pela Transformada Rápida de Fourier (FFT). Há ainda a transformada de tempo determinístico de Fourier (DTFT) que exige que a entrada seja discreta porém a DFT tanto a entrada como saída é discreta.

Já a transformada rápida de Fourier é um algoritmo que permite calcular a transformada discreta de Fourier e sua inversa. Enquanto a transformada clássica de Fourier permite também converter tempo ou espaço como o domínio da frequência e ao contrário, a transformada rápida permite que isso seja feito com uma maior velocidade (BATENKOV, 2005).

"A essência do cálculo da Transformada Rápida de Fourier (FFT) é uma série de operações matemática conhecidas como Transformação Discreta de Fourier (DFT) que é um conjunto de m variáveis no domínio da frequência a partir de um conjunto n de amostras no domínio do tempo." (DOVICCHI; PARREIRA, 2003)

As transformadas discretas de Fourier (DFT) são de grande importância para análise de sinais funcionando como um filtro de espectro, isolando os componentes do sinal periódico.

Pode-se compreender essas transformadas como rotações no espaço de uma função, de um domínio ao outro. Uma vez que se usa funções senoidais como bases que estão localizadas no domínio da frequência que podem ser analisados através de espectros é possível isolar e calcular suas distribuições de potência.

Elas consistem em uma combinação linear em um vetor contendo $\log_2 n$ elementos que transforma em um vetor de $(1, 2^n)$, com n pertencendo aos reais. Deduz-se então que a matriz transformada inversa de uma DFT é a transporta da matriz original (DOVICCHI; PARREIRA, 2003).

Os coeficientes complexos da DFT é dado pela seguinte equação:

$$X_k = \sum_{n=0}^{N-1} x_n \cdot \omega^{-kn}$$

com $\omega = e^{2\pi i/N}$ que costuma a ser definido como a n -ésima raiz de unidade.

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} X_k \cdot e^{i2\pi k n/N}$$

com $e^{2\pi i/N}$ sendo a n -ésima raiz de unidade. Como se nota, está é a versão inversa da DNF (BRIGHAM, 1988).

Assim sendo, a função X_k é uma função periódica complexa em k , com N pontos em cada período, dos quais somente metade são únicos. A função X_k gera m variáveis

no domínio da frequência com incremento de $F = 1/tp$. Para k_n real com N amostras, um único espectro pode ser computador para $N/2$ pontos da frequência.

Com uma complexidade razoável na ordem de $O(N \cdot \log_2(N))$ essa transformada possui uma grande vantagem em questão de desempenho e conseqüentemente aplicações práticas (DOVICCHI; PARREIRA, 2003).

2.6 Centroide

Mais especificamente centroide espectral que é o tipo de centroide tipicamente empregado em análise de sinais. A centroide espectral indica o centro de massa de um espectro. Através da Transformada de Fourier é possível ela é calculada através de uma média ponderada das frequências presentes em um sinal.

Sua fórmula segue:

$$\text{Centroide} = \frac{\sum_{n=0}^{N-1} f(n)x(n)}{\sum_{n=0}^{N-1} x(n)}$$

Com $x(n)$ representando o valor ponderado de frequência ou magnitude ou o número bin e $f(n)$ representando a frequência central desse *bin*.

Há duas maneira principais de abordar esse cálculo (DOVICCHI; ALVES), uma definida por valores absolutos como sugerido por Kendall e por *Hertz*, como sugerido por Sandell que segue conforme a fórmula acima.

3 Proposta

3.1 Motivação

Uma das muitas maneiras que se pode pensar de como comparar uma música à outra, sem pensar diretamente na questão técnica da computação, é pela natureza musical de ambas. Culturas diferentes possuem interpretações diferentes da música assim como a forma que dividem as frequências em grupos. A mais comum e globalmente usada vem da escola europeia de música clássica. Tomando suas teorias como referência se pode analisar músicas diferentes com o propósito de buscar similaridades e dissimilaridades. Baseando-se na possibilidade de análise musical automatizada, esta proposta busca trazer estes elementos para o universo da ciência da computação assim como entender suas limitações.

3.2 Resumo

Esta proposta tem como foco central extrair informações quanto a harmonia para que seja possível estabelecer uma comparação entre músicas diferentes e identificar suas similaridades.

Para comparar uma música com a outra inicialmente será necessário extrair dados de frequência do arquivo. Antes porém deverá ser decidido o que será considerado silêncio para que trecho que foram considerados como tanto, não afetem as análises feitas com sons considerados válidos pela proposta.

Com tais dados já obtidos a música será dividida em intervalos discretos com uma duração mínima a ser decidida pela implementação a ser feita. Para cada um desses intervalos será identificada sua centroide se usado métodos relacionados à Transformada de Fourier.

Armazenando estas informações se obtêm os dados das mudanças de frequência em relação ao tempo para uma determinada música de forma discreta. Isso porém não é adequado ainda para uma comparação já que para fazer a análise musical desejada será preciso se usar das teorias musicais que neste caso exige que se trabalhe com apenas 12 semitons. Se faz necessário então estabelecer um método para que se tenha apenas frequências correspondentes a esse número. Uma forma possível de resolver este caso pode simplesmente ser trocar cada um dos valores de frequência por um símbolo representando a nota mais próxima daquela frequência. Isso pode resultar em valores frequências muito diferentes para uma mesma nota devido aos harmônicos das notas mas isso não foge do interesse desta proposta.

Como resultado obtido se tem as mudanças de notas, independente de frequência exata, ao longo do tempo. Alguns ajustes ainda se fazem necessários. Um deles é agrupar notas iguais consecutivas considerando-as uma só e tendo como duração a quantidade de notas agrupadas. Outro passo é avaliar estes dados até o

momento e eliminar notas extremamente infrequentes. Para este passo funcionar de forma mais eficiente há algumas dificuldades que serão analisadas.

Com esta fase de preparação encerrada pode-se finalmente partir para técnicas de análise harmônica. Várias são possíveis e algumas inclusive podem se somar para melhorar a análise. Buscar identificar possíveis escalas dentre os dados obtidos, notas frequentes, clichês musicais e outras são algumas das opções.

Último passo é então comparar os dados analisados de uma música para com a outra. As possibilidades são inúmeras e podem ser diferentes em situações diversas uma vez que é um tanto subjetivo. Esta proposta tenta se utilizar de métodos que vão de acordo com a natureza das análises feitas.

3.3 Desenvolvimento

- **Identificação da região de silêncio**

Antes de seguir aos próximos passos para dividir a música em partes, identificar notas dentre outras coisas se deve considerar um fator muito importante presente na música: o silêncio.

Já no século 13 dizia o teorista musical alemão Franco of Cologne:

*“Time is the measure of actual sound as well as of the opposite,
its omission.”*(apud CROSBY, 1997, p.152)

Identifica-se a omissão no qual o teorista alemão se refere como a ausência do som, isto é, o silêncio. Contudo, por muitas vezes muitos dos momentos que são considerados como silêncio apresentam som. Ainda que em uma partitura esteja explicitada pausas raras são os momentos nos quais o som se faz ausente. Isto significa que para uma determinada situação tem-se uma certa tolerância ao som presente e se considera silêncio até que o limite desta tolerância seja ultrapassado. Em muitos desses casos existem ruídos externos que não fazem parte do foco da nossa atenção e portanto são ignorados. O mesmo se mostra real na música. Se por um lado uma música pode ter sido bem definida como em geral são representadas através de partituras, por outro a execução invariavelmente apresenta ruídos fora do controle de quem está executando a peça ou até mesmo pequenos erros cometidos pelo mesmo. Ao analisar e comparar músicas costuma-se ignorar tais elementos uma vez que originalmente deveriam ser silêncio.

Ao considerar que este estudo busca fazer análises automatizadas com auxílio de computadores dificuldades se apresentam ao tratar estes ruídos acima comentados. Enquanto o ser humano tem a capacidade de, ao ser treinado, distinguir muitos dos elementos com o objetivo de identificar o que é relevante e o que não, o computador pode não estar em posição de fazer o mesmo. Com a natureza do problema aqui exposto de comparar músicas, não será considerado que será desenvolvido um sistema com conhecimento prévio de como seriam as músicas originalmente em ambiente ideal. Isto significa que não há conhecimento prévio sobre ontem possivelmente existem pausas em uma determinada música. Considerando isso a maneira de evitar problemas de ruídos externo é limitada.

Dentre outras possíveis de serem desenvolvidas esta proposta considera que haverá uma capacidade de distinção do silêncio. Isto significa analisar a música e marcar da maneira desejada os espaços de tempo que serão considerados como silêncio. A dificuldade encontrada nesse momento é saber como e a partir de que momento se pode considerar um momento como um provável silêncio.

Métodos de identificação de silêncio são comuns e muito usados na análise de silêncio em conversas como o VAD (*Voice Activity Detection*). Comumente os testes para detecção de silêncio consiste na análise da energia de um sinal e o ZCR (*Zero-Crossing Rate*) que podem ser aplicados para arquivos de áudio (PAPAJ, 2008). Pode-se assim fazer uma distinção de possíveis trechos de silêncios. Assim sendo, outras formas de detecção de silêncio podem ser incorporadas de áreas que lidem com identificação de silêncio em conversas dentre outras porém nem sempre podem oferecer os resultados esperados para a análise de músicas (PAPAJ, 2008). Pode modificar ou buscar métodos que considerem a questão de silêncio do contexto de arquivos de música.

- **Divisão do sinal em frames**

Este primeiro passo busca dividir a música em uma certa quantidade de trechos, conhecidos comumente como *frames*, de mesmo tamanho. Isto porque uma música terá uma certa quantidade de notas. É importante que os trechos sejam curtos o suficiente para que nos próximos passos, ao identificar as frequências de cada trecho,

evite-se o problema de obter uma nota não presente originalmente. Isto é, um trecho ser longo o suficiente para possuir variação de frequência o suficiente para conter duas ou mais notas resultando no próximo passo considerar este grupo uma nota diferente de todas as envolvidas ou ignorar a presença de alguma que pode dificultar a análise mais a frente.

Uma forma de fazer essa divisão é se usar das transformadas de Fourier. Obtendo sua transformada em partes consecutivas do áudio criando assim um espectograma. Para esses casos é comum o uso de Transformada de Fourier de Tempo Reduzido, conhecida como STFT (*Short-Time Fourier Transform*). Uma vez que ela também é inversível é possível reconstruir o sinal quando se faz necessário (JONT, 1977).

Definir o número de trechos ideais pode ser uma tarefa complexa. Uma vez que o tempo pode variar entre músicas, ao longo de uma única música ou ainda apresente polirritmia¹ é difícil prever o trecho de tamanho ideal para minimizar o problema explicado anteriormente de obtenção de notas erradas.

Poderia se desenvolver um método para obtenção aproximada do tempo mas devido à complexidade e dificuldade para um possível método não será aqui abordado.

Tomando isso em consideração, a definição de um trecho mínimo fica a critério do implementador. Pode-se usar um método de obtenção de tempo se houver um a disposição e não for complexo o suficiente para comprometer o funcionamento desta proposta. Ou ainda, alternativamente, definir um trecho que se considere adequado na

¹ uso de uma ou mais estruturas rítmicas contrastantes que não são prontamente identificadas como sendo derivadas uma da outra ou sendo manifestações de uma mesma métrica.

grande maioria das vezes. Neste último caso seria necessário uma pesquisa e estudo dos tempos mais comuns encontrados nas músicas comparadas. É razoável imaginar porém, que este seja um valor baixo já que valores grande grandes causam grandes alterações de resultados nos próximos passos. Trechos de duração curtos demais por outro lado não causam este problema já que se torna mais próximo da música original. A desvantagem porém nestes casos é ter como resultado uma quantidade elevada de trechos a ser trabalhadas a partir desde passo e que pode eventualmente comprometer o desempenho. É importante levar em consideração os valores máximos e mínimos para o tamanho dos trechos. O limite superior é definidor pela duração total do áudio enquanto o limite inferior vem da frequência de amostragem de dados por uma unidade de tempo, medida geralmente em segundos chamada normalmente de *sampling rate* (taxa de amostragem) ou *sampling frequency* (frequência de amostragem) representada em *hertz* (BURK et al, 2011). Considerando o *sampling rate* do arquivo de áudio é conveniente que o número escolhido divida a música igualmente porém diferenças muito pequenas não provocam diferenças significativas.

Com a definição da quantidade de trechos a dividir a música encerra-se este passo e se inicia o passo de extração das frequências destes trechos. É importante que ao final deste passo todos os trechos tenham o *mesmo sampling rate*. Caso isso não ocorra será necessário calcular o *sampling rate* de cada um deles. Isto significa que arquivos como VBR (*Variable Bitrate*) possam ter maiores dificuldades.

- **Extração de frequência**

Com o passo anterior concluído é necessário extrair dados referentes a frequência de cada um dos trechos formados pela divisão da música. Isto tem como objetivo identificar as notas que tais frequências representam no próximo passo.

Obter a frequência de um trecho de áudio é algo já muito conhecido e usado em diversas aplicações relacionados com áudio há muitos anos (GERHARD, 2003). Isto significa também que existem já métodos consolidados para esta tarefa que podem aqui ser usados.

Muitos dos métodos usados baseiam-se na Transformada de Fourier e suas muitas variantes como a STFT (*Short Time Fourier Transform*) ou a FFT (*Fast Fourier Transform*). As Transformadas de Fourier e podem ser encontradas em três diferentes variedades como clássica, como Série de Fourier e versões para tempo discreto conhecidas como Transformada de Fourier de tempo discreto, ou DTFT (HARDESTY, 2009). É possível também usar outros métodos como a Transformada de Laplace e a Transformada Z. A transformada Z porém exige continuidade entre dois pontos de descontinuidade (VILLATE, 2001), isto é, ser parcelarmente contínua ao passo que a Transformada Z considera o tempo discreto (PICARD, 2003).

Independente do método escolhido este passo é concluído com a obtenção das frequências dos trechos.

- **Transformação em notas**

Este passo meramente consiste em transformar os dados obtidos anteriormente referentes a frequência em uma representação de notas musicais. Pode-se escolher pular este passo neste momento e fazer as transformações dos valores de frequência para notas, sempre que se fizer necessário nos próximos passos mas, considerando o quanto serão usadas as notas para análises é conveniente que este passo exista neste momento.

O motivo desta transformação é simplificar e facilitar a análise musical que seguirá. Tendo referência a teoria musical se pode definir valores possíveis de serem assumidos pelas frequências como notas musicais. Como um mesmo som pode ser representado de diversas formas é importante lembrar que na análise algumas notas poderão ser eventualmente consideradas como sendo outra. Exemplificando, um C# (*dó sustenido*) é equivalente a Db (*ré bemol*). Isto em teoria musical é conhecido como enarmonia (SAKER;BENWARD, 2008). Uma enarmonia na notação musical é uma nota equivalente à outra mais escrita de forma diferente (SAKER;BENWARD, 2008).

Como não se possui informação sobre qual notação será adequada numa determinada musical não é possível decidir com antecedência se, assumindo o exemplo dado, deve-se usar C# ou Db ou ambos. Com isso considerado deve-se escolher tentativamente mas considerar as possíveis enarmonias no momento da análise. Por simplicidade e por ser uma das bases da música (PISTON, 1959) no que

se refere as notas a escala cromática ascendente é uma boa escolha. A escala cromática possui com todos os doze semitons. Partindo do A (*lá*) até G# (*sol sustenido*) tem-se: A, A#, B, C, C#, D, D#, E, F, F#, G, G# (PISTON, 1959).

Tomando as notas da escala cromática ascendente como base se pode então fazer a conversão dos valores de frequência obtidos para as notas buscando a nota mais próxima à frequência retornada. Para isso é importante considerar os cents para poder encontrar a nota mais próxima com precisão. Como as frequências não necessariamente representaram exatamente uma das notas será preciso aproximar os valores para a nota mais próxima, o que faz parte do objetivo deste passo. É importante considerar que uma mesma nota representa infinitos valores de frequência. Isto ocorre porque considerando uma frequência fundamental todos os seus harmônicos são considerados a mesma nota. Isso faz com que neste processo de transformar as frequências em notas haverá uma mesma nota representando frequências muito distantes. Como para a análise feita por esta proposta não se faz necessário este tipo de distinção isto não se torna um problema.

Pode ser feita uma diferenciação de acordo com a altura. Isto significa atribuir um número após a nota para indicar sua altura. Ao escrever A4 indica-se que a nota *lá* indicada é aquela de 440 Hz seguindo padrão adotado atualmente (ISO 16:1975). Este tipo de distinção porém não é considerada nessa proposta e portanto se limitará a indicar qual nota é, independente de altura.

É importante lembrar dos trechos que foram considerados como silêncio anteriormente. Esses casos são aqui considerados como não havendo frequência,

razão pela qual é feita essa distinção antes de chegar nesse passo. A presença desses silêncios ainda que sejam ignoradas nesse passo devem permanecer para futuras manipulações em outros passos.

Terminado a análise de todos os trechos e associando as notas às frequências de acordo com o proposto encerra-se este passo.

- **Agrupamento de notas similares**

Este passo consiste no agrupamento de notas. Isto é, considerar notas iguais consecutivas como uma única. O objetivo para isso é auxiliar a análise no qual a informações relevante são as mudanças de notas.

Isto não significa porém que a duração não é um dado importante. Sugere-se que para cada grupo de notas agrupadas seja mantida a informação referente a quantos trechos foram agrupados em um só. Essa informação somada a duração conhecida de cada trecho nos gera a duração total do novo grupo de notas obtidas e será importante nos próximos passos.

Da mesma forma que as notas os trechos contidos nas regiões de silêncio devem ser agrupados e a quantidade agrupada obtida.

O resultado final deste passo é um grupo muito menor de notas e silêncios mas que preserva informação necessária de como foi formada uma vez que é

importante não perder a relação com o áudio original nos momentos que se fizer necessário comparar os dados obtidos com os dados de entrada originais.

- **Eliminação de notas infrequentes**

Ainda que em passos anteriores exista preocupação de buscar eliminar informações consideradas irrelevantes, como no caso dos trechos considerados como silêncio, é esperado que outras persistam. Muitos ruídos altos e outras distorções sonoras podem existir, propositalmente ou não. Este passo busca minimizar esses casos para facilitar a análise musical do arquivo.

Considerando que muitos ruídos resultam em uma mudança brusca de frequência e de muito baixa duração se pode buscar trechos que possam indicar que isso tenha ocorrido. Definindo uma duração máxima para esses casos se deve considerar ruído todos os trechos abaixo ou igual a esta duração. As informações referente a frequência desses trechos passaram a ser consideradas para a análise.

O estudo de redução de ruídos porém, abrange de forma mais avançada muitos desses problemas e algoritmos específicos e melhor definidos podem ser uma alternativa melhor para redução dos ruídos presentes no áudio.

É importante observar que se tais trechos não possuem um valor relevante por serem ruídos é necessário observar os erros que provocaram nos trechos vizinhos até este passo. É possível que em muitos casos uma nota tenha interrompido uma nota por

um curto período de tempo. Isso significa que durante o passo anterior a nota não foi agrupada e sim mantida como dois trechos separados. É preciso então refazer o passo anterior para que, nos momentos em que o trecho anterior e posterior a um ruído corresponderem à mesma nota, poder agrupá-las.

Devido essa inconveniência se pode preferir fazer esse passo junto com o passo de identificação de silêncio. Ambas os passos possuem similaridades e em alguns casos são tratados de forma semelhantes.

Alguns algoritmos como o usado no *Audacity*® reduz grandemente o volume de ruídos (AUDACITY). Isso significa que em períodos que só há ruídos o resultado da redução pode oferecer algo próximo do silêncio. Se isto for seguido da identificação de silêncio pode resultar em uma maior quantidade de trechos considerados como silêncio. Resultados para cada abordagem pode variar e portanto a decisão fica para a implementação desejada.

- **Identificação de notas longas**

Ao alcançar este passo se inicia a análise musical. Esta parte pode ser abordada de diversas formas com resultados e sucessos diferentes. Considerando o foco que foi dado nos passos anteriores ao conceito de nota este será o tema central aqui sugerido. Informações sobre timbre e amplitude não serão consideradas aqui.

Ao chegar aqui o resultado dos passos anteriores fornece a mudanças de notas e silêncios e suas durações presentes no áudio com a eliminação ou redução de alguns possíveis ruídos. Dentre as notas restantes é de esperar uma grande variedade, mesmo em músicas harmonicamente e melodicamente simples. Ao considerar que qualquer trecho possui apenas 13 valores possíveis, sendo eles 12 semitons e o silêncio, e que uma música ainda que curta provavelmente conterà milhares de trechos é de esperar que todas essas possibilidades apareçam em algum momento, ainda que por distorções restantes e outros problemas mesmo após os cuidados dos passos anteriores. Ao lidar com uma situação, uma análise simples buscando quais possíveis escalas se encaixam na música tem uma possibilidade muito grande de falha. Uma música que oferece todas as possibilidades de notas resultará em boa parte dos casos com grande semelhança a escala cromática dificultando a comparação com outras músicas. Além disso, é de se esperar que quanto mais longo um arquivo de áudio maiores as chances de que apareçam ao longo de sua duração uma maior variedade de notas. Considerando estas dificuldades é necessário buscar maneiras de contorná-las para que seja possível uma análise a partir das informações extraídas até o momento.

É importante compreender que, mesmo além de defeitos no áudio, as músicas não necessariamente seguem regras claras. Mesmo as que seguem algum padrão estrito, como muitas músicas eruditas, possuem ainda muitas possibilidades e tais regras podem ser complexas. Além de muitas outras possibilidades, também é comum a existência de trechos cromáticos e notas de passagem. É possível imaginar inúmeros casos no qual uma única nota diferente presente resulte em uma análise falha uma vez

que há somente 12 semitons disponíveis. Um pequeno trecho com notas de passagem podem levar a uma interpretação muito diferente da música como um todo.

Para evitar algumas dessas situações é preciso saber como elas ocorrem e quais notas das muitas presentes possuem maior importância para sua compreensão musical. É de grande importância então entender quais dessas estão presentes numa possível escala sendo seguida naquele momento pela música.

As notas anteriores e posteriores às notas de ornamentação ou de passagem, por outro lado, podem nos fornecer grandes indicações uma vez assumido que as notas entre essas compõe um ornamento musical. Ao pensar desta forma é possível imaginar que essas notas divididas por um trecho que ornamentos sejam importantes para a compreensão da escala. Podendo separá-las, o número de notas analisadas totais, resultante dos passos anteriores, cai grandemente.

A dificuldade se apresenta porém, ao tentar distingui-las de outras. Como visto, elas costumam ser breves e é uma técnica que pode ser usada para diferenciá-las. Como isto não significa que outras notas não possam ser breves o problema se mantém. Ainda assim, uma vez que elas não costumam ser longas, é possível considerar as não breves como sendo de importância maior. O foco nas notas mais longas diminui a quantidade de notas para análise grandemente, porém ainda são muitas e ainda podem ser representativas da música como um todo ao assumir anteriormente a relativa menor importância das ornamentações e notas de passagem para a comparação. Assim sendo, ao deixar de lado notas mais breves, ainda que com

a perda de notas válidas pode facilitar a análise sem um prejuízo grande na compreensão da música.

O quão breve ou quão longa essas devem ser também pode afetar os resultados e deve ser definido. Para isso de maneira similar ao passo que o áudio foi dividido em trechos, o ideal seria definir esses valores com base numa análise do tempo da música. Como isso não está sendo levado em consideração nesta proposta devido à sua complexidade deve-se definir com base em um estudo das durações de tais tipos de nota. Seguindo o estudo deve-se definir uma duração mínima que comumente é considerada suficiente para ser considerada uma nota aqui julgada importante para a análise da música desta proposta, isto é, possui menos chances de ser uma nota de passagem ou ornamentações dos mais variados tipos. Implementações diferentes porém podem definir diferentes valores com base nos resultados obtidos usando diferentes valores mínimos.

- **Nova Divisão por Blocos**

Definido a diferença entre notas no passo anterior o foco passa a ser nas notas que tem duração acima ou igual a mínima definida. Como é comum mudança de tonalidade em certos trechos musicais (PISTON, 1959), fazer uma análise geral pode não ser ideal. Por outro lado identificar de forma algorítmica o início e fim dessas mudanças pode não ser algo trivial. Ainda assim algumas tentativas podem se fazer válidas.

É razoável pensar que após um período significativo de silêncio a direção da música pode potencialmente mudar. Seguindo este raciocínio se pode buscar dividir a música pelos seus trechos que silêncio mais longos. O resultado disso porém, são partes de tamanhos diversos e que pode potencialmente falhar em músicas que não apresentam trechos com silêncios longos o suficiente para esta técnica.

Pensando em práticas comuns para modulação é possível observar algumas tendências. Contudo, para essa proposta a identificação desses elementos podem ser de um grau de dificuldade muito elevado que pode inviabilizar sua aplicação. Uma das formas comuns que precede uma modulação em uma música costuma ser a escolha de um acorde pivô, sujeito a modulação devido a presença nos campos harmônicos atual e o que se deseja alcançar (PISTON, 1959). Esse tipo de análise requereria informações ainda desconhecidas sobre os campos harmônicos da música, além de ser uma de várias formas de modulação.

Alternativamente então, pode-se buscar dividir em partes de duração definida. É possível ainda a união de ambas as divisões extrapolando um valor máximo em mínimo possível para cada parte para evitar partes muito grandes que podem ocorrer ao separar pelos trechos de silêncio. De acordo com a implementação estas ou outras técnicas podem ser aplicadas. Com a música dividida nessas grandes partes é possível fazer análise individualmente em cada um desses blocos.

- **Análise dos Blocos**

Para cada um dos blocos é necessário fazer uma lista de possíveis escalas associando cada um grau de semelhança para com o respectivo bloco. Para isso é necessário listar todas as notas ali presentes, isto é, aquelas escolhidas previamente por serem consideradas mais significantes. Essas notas são então comparadas a várias escalas para analisar seu grau de semelhança. As regras para definir essa similaridade pode ser uma ou mais métodos escolhidas para uma determinada implementação. Contudo é importante considerar e compreender o que torna mais ou menos similar.

Um bloco no qual nenhuma das notas está presente em uma determinada escala pode significar que a similaridade daquele bloco para com a escala comparada é nula ao passo que um bloco que possui todas as notas de uma escala e somente estas tem um grau de similaridade alto. Alto somente, e não com um grau de certeza absoluto porque é possível encontrar diferentes escalas com as mesmas notas ou ainda, é possível que as notas presentes sejam parte de uma outra escala dentre outras razões.

Pelo mesmo raciocínio é importante considerar como um ponto de similaridade escalas que possuem muitas de suas notas no trecho, ainda que não todas, mas que ao mesmo tempo notas não pertencentes não estejam presentes. Isto é importante uma vez que é possível que aquela seja a escala sendo usada no momento porém naquele trecho não foram explorados todos os graus da escala. No caso de aparecer notas não pertencentes da escala pode-se definir uma tolerância máxima de um nota e ainda assim não atribuir um grau de semelhança elevado. Mais que uma nota

indesejada deve-se considerar a escala como não possuindo pontos de semelhança para aquele trecho. Uma nota foi aqui escolhida como tolerância devido a notas possivelmente emprestadas de tonalidades vizinhas que possuem apenas diferença de um semitom. Deve-se tratar isto caso a caso uma vez que escalas podem se comportar diferente. A escala menor melódica por exemplo possui diferença entre sua forma descendente e ascendente (PISTON, 1959) e deve-se levar em consideração.

Problemas maiores surgem justamente no caso citado de escalas diferentes com as mesmas notas. Como exemplo, tem-se que toda escala maior possui uma relativa menor, que se visto pelos modos gregos nada mais é que o modo eólio e assumindo a maior como modo jônico se tem que a escala menor é o mesmo que iniciar o modo jônico através do seu sexto grau (SAKER;BENWARD, 2008). O resultado disto são escalas com as mesmas nota porém com sonoridades diferentes já que possuem graus diferentes resultante da ordem diferente das notas. Isso também significa uma dificuldade grande para discernir entre modos maiores e menores que possuem as mesmas notas. Se os tamanhos de cada bloco analisados forem pequenos é preferível que o problema seja deixado para após a análise dos blocos. Assim sendo nesses casos todas as escalas com as mesmas notas terão o mesmo grau de similaridade para o bloco.

Após a análise dos blocos tem-se então uma lista de possíveis escalas para cada um dos blocos. Com isso em mãos se pode calcular para a música como um todo quais dessas foram mais frequentes de forma ponderada considerando tanto o grau de similaridade quanto a duração total de cada bloco.

- **O Problema das Escalas Com as Mesmas Notas**

A lista global de possíveis escalas gerado ao fim do passo anterior ainda pode possuir possivelmente o problema de escalas com as mesmas notas uma vez que isso não foi tratado anteriormente. Nesses casos pode-se tentar diversas técnicas para buscar aprimorar a exatidão e estabelecer quais dentre estas é a mais provável.

No caso das escalas maiores e menores naturais pode-se analisar todos os blocos nos quais essas escalas foram consideradas muito prováveis novamente para estabelecer uma lista das notas mais comuns. Essa lista pode nos dar informação de qual das escalas é mais comum durante a música uma vez que alguns graus costumam ser mais frequentes que outros e assim elevar o seu grau de similaridade. Para melhorar o uso dessa técnica exigiria o estudo de uma grande número de músicas em cada tipo de escala e que pode variar um pouco os resultados devido aos diferentes estilos musicais.

Se isso não for possível se pode partir da teoria musical para considerar quais os graus de maior importância de uma escala. No caso da escala maior ou menor a tônica, isto é, o primeiro grau é considerada de maior importância, é o centro o centro tonal e com grande frequência o grau de resolução da música (SAKER;BENWARD, 2008). Seguido a tônica tem-se o dominante que é o quinto grau e é considerado o

grau de maior importância após a tônica (SAKER; BENWARD, 2008) podendo em momentos indicar melhor a tonalidade de uma música mais que a própria tônica (PISTON, 1959). Portanto, espera-se que uma escala de *dó* maior tenha a nota *dó*, que é a tônica, e a nota *sol*, que é a dominante, como as mais frequentes ao passo que a escala de *lá* menor natural por sua vez tenha as notas *lá* e *mi*, tônica e dominante respectivamente, como as mais frequentes. Neste exemplo observa-se duas escalas que possuem as mesmas notas sendo diferenciadas pelas notas mais comuns em ambas. Para isso porém houve a necessidade de assumir que o par de notas mais comuns que formaram um intervalo de quinta justa representam de fato à tônica e dominante, uma suposição que se espera ser verdade mais vezes que o contrário.

Vale notar que as duas notas mais frequentes encontradas podem não corresponder a um intervalo de quinta justa para resultar em tônica e dominante. Outros pares são possíveis ou ainda mesmo, pares ambíguos, dependendo da nota que se tome como inicial. Um intervalo de quinta justa afinal pode ser visto de forma inversa como um intervalo de quarta justa gerando dúvidas como no exemplo a seguir.

Avaliar um par frequente de notas porém pode não bastar. Supondo que um caso hipotético onde as notas *mi* e *lá* se apresentem como mais comuns pode indicar que *lá* seja a tônica e *mi* a dominante ou ainda que *mi* seja a tônica e *lá* a subdominante. Se não for definido por suposição prévia, é preciso saber qual dos casos está sendo avaliado. Imaginando que o caso seja uma escala diatônica sem acidentes gera dúvidas se trata-se de uma escala de *lá* eólio, caso de *lá* como tônica, ou *mi* frígio, caso de *mi* como tônica. Para diferenciar é necessário considerar mais um

grau pelo menos. Pode-se aqui definir um grau como o provavelmente mais incomum. Assumindo o sétimo grau como o mais incomum se espera que a nota *sol*, sétimo grau no caso de *lá* tônica, ou *ré*, no caso de *mi* tônica, como a mais incomum. Uma vez que para *lá* eólio o *ré* é a subdominante e para *mi* frígio *sol* é a medianta resulta que a sétima (*sol*), no caso de considerar tônica (*lá*) e dominante (*mi*), seria a terça para o outro (*mi*) e que a sétima (*ré*), no caso de considerar subdominante (*lá*) e tônica (*mi*), seria a quarta para o outro (*lá*). Isso é importante porque se considerar que o quarto e o terceiro grau costumam ser comuns e assumir o sétimo grau como mais incomum ao analisar as notas há maiores chances de disparidade entre as notas *sol* e *ré*, no exemplo em questão. Se for definido que a nota menos comum entre ambas como o sétimo grau pode-se finalmente resolver a dúvida inicial entre *lá* como tônica ou *mi* como tônica e então encontrar qual é a escala mais prováveis de ambas. É assim necessário aplicar um método como da nota mais incomum da escala para discernir entre pares de notas frequentes.

Como observado acima, muitas suposições são necessárias para discernir entre escalas com as mesmas notas o que mostra a dificuldade de lidar com estes casos e portanto outras formas de diferenciá-las pode ser somadas a esta para uma melhor precisão. Uma possibilidade é limitar os tipos de escalas possíveis, forçando um limite de escolhas. Isto significa, por exemplo, pensar só em tonalidade maior e menor natural no campo diatônico ignorando os outros modos da escala. Assim basta diferenciar entre o modo jônico e eólio que pode ser mais fácil pela grande diferença de graus. Por exemplo, para uma escala de *dó* maior, *si* é o sétimo grau enquanto que para sua relativa menor, o *lá* menor, o *si* é o segundo grau. Ainda são suposições mas

ao limitar as possíveis tonalidades diatônicas para duas apenas dentre sete possibilita uma facilidade maior para diferenciar entre as possibilidades.

Outro ponto importante é que a maneira de ver os graus de maior importância e mais frequentes pode ainda variar de acordo com o tipo de escala. Para outras escalas como pentatônicas se pode adotar outras técnicas e o mesmo segue para qualquer outra situação no qual possa haver confusão devido a similaridade entre as escalas em relação às suas notas. Isso deve ser levado em consideração uma vez que escalas diferentes podem ser usadas com objetivos diferentes.

A análise de notas comuns para identificação de grau significa também a necessidade de não somente um estudo adicional para desenvolver estas técnicas como também uma camada maior de complexidade uma vez que é necessário analisar novamente todos os blocos. O resultado porém é vantajoso uma vez que pode diferenciar escalas aparentemente semelhantes mas que musicalmente possam soar significativamente diferente.

O resultado final dessa análise permite concluir esse passo com a lista de escalas potencialmente comuns para a música assim como seu grau de semelhança para com ela. Pode com isso partir para o próximo passo para comparar resultados de músicas diferentes, o objetivo inicial da proposta.

- **Comparação entre Resultados**

Supondo que duas músicas distintas tenham passado por todos os passos anteriores e tenham retornado como resultado suas listas de possíveis escalas é possível comparar tais resultados a fim de julgar sua semelhança segundo esta proposta.

Primeiramente, é importante entender que essas escalas indicam possivelmente qual o campo harmônico da música ainda que esta proposta não lide com o conceito de acordes. Uma música no qual a presença da escala menor melódica é a mais marcante indica que possivelmente a música esteja no campo harmônico menor melódico ou em grande parte uma vez que é possível haver modulações. O campo harmônico por sua vez nos dá o foco da comparação dessa proposta já que é grande responsável pela sonoridade da música. É possível pensar dessa forma uma vez que os acordes que formam um campo harmônico são derivados das notas pertencentes da escala de mesmo nome (BUCKINGHAM;PASCAL, 1997).

Partindo-se dessa informação, para cada um dos resultados deve-se buscar as escalas que tiveram o maior grau de similaridade, excluindo definitivamente as que não tiveram similaridade alguma. Pode-se adotar um padrão de buscar somente as cinco primeiras de cada músicas ou usar outra faixa de corte. Para cada uma das escalas da lista de resultados de uma música procura-se uma semelhante da na lista da outra música. De acordo com a posição que se encontre na sua lista e na lista ganha-se mais ou menos pontos de similaridade. Mais pontos para as mais a frente da lista e menos quando a escala estiver mais abaixo na lista de possíveis escalas.

Junto dessa forma de comparação é possível avaliar outros aspectos que possam pesar na avaliação da semelhança de uma escala presente na lista de uma música com a de outra. Uma delas é não somente buscar a escala correspondente na outra lista mas também dar valor ao tipo. Em outras palavras, duas escalas de mesmo tipo ainda que em tonalidades diferentes possuem um certo nível de similaridade. Isso pode ser facilmente compreendido ao considerar que uma mesma música pode ser transposta para diferentes tons desde que mantenha o tipo de escala (SAKER;BENWARD, 2008). Como exemplo se tem que uma canção em *si* bemól menor pode também ser executada em *si* menor. Pensando desta forma é possível dar pontos de similaridade nesses casos ainda que não da mesma forma que se fossem exatamente a mesma escala.

A segunda forma pode não ser intuitiva como a primeira mas é importante para que a comparação entre duas músicas baseado em escalas comuns em ambas não se torne simplística. Trata-se aqui de comparar a natureza de diferentes tipos de escalas. Isto é, por exemplo, analisar de uma escala dórica tem maior semelhança com uma escala menor natural ou com uma escala maior jônica. Sendo assim, deve-se definir para cada escala se ela tem uma sonoridade maior ou menor.

Como exemplo se observa o modo dórico das escalas diatônicas e o quão próxima ela está na escala menor natural, que também pode ser visto como o modo eólio das escalas diatônicas (SAKER;BENWARD, 2008). Em um tom qualquer tem-se que, de forma ascendente e em relação a nota tônica, os intervalos presentes no modo dórico são: segunda maior, terça menor, quarta justa, quinta justa, sexta maior, sétima

menor e finalmente uníssono. A escala menor natural por sua vez possui os seguintes intervalos: segunda maior, terça menor, quarta justa, quinta justa, sexta menor, sétima menor e então uníssono. Como se nota a diferença de intervalos apresenta-se apenas na subdominante. Ao pensar que as escalas consideradas menores possuem predominantemente uma terça menor e quinta justa formando uma tríada menor nota-se que a escala maior lícria segue essa forma. Entre os modos da escala diatônica apenas o modo lícrio pode causar confusão uma vez que possui uma quinta diminuta e, ao contrário dos outros modos com características menores, possui uma segunda menor. Ela pode ser vista porém como uma escala menor com uma segunda e quinta meio tom abaixo.

Com o auxílio de um profissional pode definir com maior precisão os tipos de sonoridade para as escalas e inclusive criar vários níveis para semelhança mesmo entre duas escalas com sonoridade menor ou maior.

A similaridade fica definida então, de forma decrescente de importância, pela presença de uma escala em ambas as listas, pelos tipos comuns e as suas sonoridades.

Adicionalmente alguns outros casos específicos também podem aqui ser considerados para contribuir com os a comparação. Um possível e relevante caso pode ser a questão da escala menor natural. Se por um lado a escala menor natural, menor harmônica e menor melódica serão consideradas semelhantes em relação a característica menor, pode-se tornar desejável considerá-las mais próximas que em outros casos comuns. Isso se deve porque as escalas menor harmônicas e melódicas

podem ser vistas como variantes da escala menor natural e compositores costumam usar os recursos dessas três escalas em suas composições (PISTON, 1959).

O resultado das aplicações dessas técnicas para comparação conclui este passo e a proposta oferecendo a semelhança entre dois diferentes arquivos de áudio com base no resultado da análise das escalas possivelmente presentes. Os resultados das comparações que obtiverem uma certa quantidade mínima de pontos de semelhança (a ser definido pela implementação) indicarão que duas músicas possuem entre si um grau elevado de semelhança do ponto de vista harmônico.

4 Resultados e Discussão

4.1 Vantagens

A abordagem oferecida neste estudo possui a grande vantagem de buscar semelhanças nas mais diversas músicas, muitas vezes em estilos diferentes do ponto de vista do campo harmônico. Isso significa possibilitar encontrar peças do período barroco que potencialmente possuem similaridades com uma canção popular como um exemplo. Isso permite sugerir músicas que possuem semelhanças a um público que talvez não tenha tido exposição ao outro estilo e portanto dificilmente teria encontrado as músicas sugeridas por conta própria.

Pode-se considerar uma vantagem a obtenção das possíveis escalas mais comuns para uma música uma vez que pode haver outras aplicações além da forma como este resultado é utilizado nesta proposta.

Alguns estilos musicais podem possivelmente possuir maiores facilidades de análise. Dentre eles músicas do período barroco devido as regras de composição da época exigir bastante rigidez. Este período possui um número menor de dissonâncias e de forma geral menor um uso menos sofisticado de acordes do campo harmônico ao comparar com o período romântico ou à bossa-nova.

4.2 Desvantagens

Primeiramente tem-se a questão da complexidade. Ainda que não possa ser aqui calculada, uma vez que o trabalho oferece a visão geral dos passos, observa-se que em cada um dos passos necessários são complexos exigindo diversos cálculos e manipulação dos dados presentes. Essa complexidade pode aumentar ao considerar estudos adicionais sugeridos em alguns passos como lidar com a identificação do tempo da música. Área que possui um grande volume de estudo porém ainda apresenta muitos problemas para identificação de sinais polifônicos mas que pode ser muito importante uma vez que muitos dados podem ser derivados da compreensão do tempo (ALONSO et al, 2003).

Eventualmente pode-se tornar necessário limitar as técnicas escolhidas em cada passo para buscar mais eficiência. Mas não só a forma técnica e na implementação surgem essa dificuldade. Toda proposta exige conhecimento de questões musicais que podem se tornar complexas. Em relação ao esforço necessário para comparar músicas entre si é possível fazer uma análise prévia dos arquivos de áudio e armazenar em um banco de dados os resultados das análises. A partir disso no momento de comparar arquivos só existe a necessidade de aplicar o último passo na proposta que consiste em comparar os resultados da análise das escalas feita até o passo anterior. Isso pode tornar a aplicação da proposta mais viável que apresenta ser inicialmente.

O fato de possibilitar encontrar similaridades em diversos estilos de música diferentes pode por um lado ser considerado um ponto negativo. Ao encontrar uma

música popular similar a uma peça clássica por esta proposta indica apenas a presença de escalas similaridades. Estilos e músicas diferentes podem usar uma mesma escala de formas muito diferentes atingindo resultados muito diversos. Isso significa que a análise aqui sugerida pode não ser suficiente de forma isolada em muitos casos uma vez que há outros elementos que chamem atenção do ouvinte que sugerem uma diferença entre músicas.

Da forma como está presente também não possui um passo para lidar com múltiplos canais de áudio assim como para buscar separar instrumentos dentro de uma faixa. Como consequência a análise se limita a estudar o resultado final observado no áudio em vez de compreender e estudar as partes que possam compor esse áudio.

Relacionado a este mesmo problema anterior se faria necessário um estudo sobre a análise de timbre presente no áudio que pode auxiliar na separação de instrumentos mas que possui dificuldades uma vez que a extensão da frequência dos instrumentos se sobrepõe em alguns casos. A análise do timbre também poderia contribuir para indicar possíveis estilos musicais da música ao buscar certos padrões de timbre em certas extensões da frequência. Outra possibilidade dessa análise seria tentar identificar notas de natureza rítmica e não melódica ou harmônica. Instrumentos como tambores podem trazer dificuldades para a análise musical uma vez que os sons provindos de tais instrumentos serão considerados notas válidas e poderá afetar o resultado final de forma negativa dependendo da sua presença.

Outra dificuldade enfrentada por essa proposta vem dos estilos de músicas analisadas. Um dos problemas inerentes dessa abordagem é a incapacidade de

analisar de forma confiável músicas que fogem do espectro da típica música ocidental. (PISTON, 1959) Se por um lado algumas dessas possuam semelhanças e até correspondências de escalas outras se usam de elementos diferentes para a composição. São assim efetivamente diferentes conceitos para abordar a atividade musical. Assim sendo músicas que trabalhem com, por exemplo, diferentes divisões do campo da frequência para definir suas notas, muito diferentes do sistema de igual temperamento utilizado pela música ocidental não podem ser avaliados por essa proposta. As dificuldades porém, não se limitam a isso. Alguns estilos muito conhecidos e que fazem parte da música ocidental podem possuir características problemáticas para o tipo de análise aqui feita. Estilos que possuem um alto grau de complexidade harmônicas, se usem de muitos elementos dissonantes, possuam muitas passagens cromáticas e modulações, incluindo portanto harmonias cromáticas, dentre outros elementos como a bossa-nova podem não oferecer resultados que bem representem as músicas. Músicas atonais também se situam na mesma classe de problemas das músicas fora do sistema tradicional e portanto analisar peças assim por essa proposta estaria conceitualmente incoerente.

Algo não abordado mas que poderia ser incluso é o cuidado ao analisar diferentes versões de uma mesma música. Ao considerar que o objetivo é encontrar músicas similares pode ser mais desejável evitar o caso de duplicatas ainda que possuam pequenas diferenças. Uma forma que poderia ser aplicada é ignorar casos que os resultados das análises sejam excessivamente semelhantes tanto nas escalas resultantes como no grau de possibilidade das mesmas.

Outro elemento que limita o funcionamento dessa proposta é na realidade citado durante os passos: a análise do tempo de música. A análise do tempo pode ser complexa e cheia de dificuldades mas que o não conhecimento do mesmo pode dificultar em diversos momentos em que se deseja contextualizar a duração de certos elementos ao longo de uma música.

Para finalizar talvez a maior dificuldade da proposta está na obtenção das escalas mais prováveis para a música. Ainda que diversos cuidados sejam tomados tais como selecionar só as notas julgadas mais importantes e dividir em blocos, já que pode haver modulações, ainda pode ser uma tarefa complicada. Uma única nota pode alterar completamente o tipo de escala. Ainda que sejam considerados graus de semelhança de cada escala para cada trecho ainda é possível haver ruídos ou distorções causando significantes mudanças. A mudança de um semitom para cima no quarto grau de uma escala de *dó* maior pode fazer que seja considerada como uma escala de *sol* maior. A divisão em blocos no passo que isso pode ocorrer limita um pouco a extensão desse problema evitando que ele se estenda para o resto da música mas a presença frequente de impurezas não tratadas pode fazer que essa medida não seja eficiente o suficiente.

5 Conclusão

5.1 Considerações Finais

Com a conclusão deste estudo se pode inferir algumas observações. Primeiramente é que oferece uma forma de comparar arquivos de áudio diferente da usualmente aplicada. Essa forma por si só pode em muitos casos não justificar sua aplicação de maneira isolada porém, junto de outras técnicas, mesmo as mais simples como usando rótulos para associar características à certa música, pode se tornar relevante. Permite encontrar similaridades muitas vezes não esperadas de outras formas. Assim sendo, o objetivo inicial do estudo foi atingido. A apresentação de diversos passos para tornar tal proposta aplicável também foi trabalhada como se desejava ainda que existam certas dificuldades a serem exploradas em trabalhos futuros e implementações.

Nota-se porém a grande quantidade de conceitos que são envolvidos tornando-a complexa, dificultando sua compreensão e possível implementação. Muitos dos passos contidos na proposta dependem de outros estudos que possuem suas próprias limitações e problemas. Assim sendo, este trabalho oferece uma proposta de análise e por este motivo não cobre a validação do método sugerido.

Contudo, a proposta oferecida pode abrir caminho para uma forma alternativa de abordar a comparação de arquivos de música.

5.2 Trabalhos futuros

Como trabalhos futuros se pode buscar estudar, desenvolver e integrar técnicas sugeridas durante a concepção da proposta. Questões como análise do tempo da música e identificação de diferentes instrumentos são questões que podem vir a contribuir com essa abordagem.

Além de questões teóricas, a implementação prática da proposta, completamente ou de passos específicos, pode contribuir não só para a realização prática da proposta mas também para sua melhor compreensão assim como possibilitar testes para analisar seu funcionamento.

Melhorias ou novas formas de abordar o problema das escalas com as mesmas notas da forma apresentada pode vir a enriquecer grandemente a identificação de escalas.

6 Referências Bibliográficas

DOLATA, Ulrich, 2011. "The Music Industry and the Internet: A Decade of Disruptive and Uncontrolled Sectoral Change". Disponível em http://www.uni-stuttgart.de/soz/oi/publikationen/soi2_dolata_music_industry.pdf>. Acesso em 31 outubro 2013.

DOWLING Stephen, 3 jul, 2004, "The music you could take anywhere". Disponível em <http://news.bbc.co.uk/2/hi/entertainment/3859903.stm>>, Acesso em 31 outubro 2013.

BBC, sem data, "Global Music Machine". Disponível em http://www.bbc.co.uk/worldservice/specials/1042_globalmusic/page8.shtml>, Acesso em 31 outubro 2013.

SIMMONS Dan, 10 dez 2004, "The rise and rise of file-sharing". Disponível em http://news.bbc.co.uk/2/hi/programmes/click_online/4085719.stm>, Acesso em 31 outubro 2013.

GOLDMAN David, 3 feb 2010, "Music's lost decade: Sales cut in half". Disponível em http://money.cnn.com/2010/02/02/news/companies/napster_music_industry/>, Acesso em 31 outubro 2013.

MACNN, 22 mar 2013, "Asymco: Apple apps turns iTunes Store into profit powerhouse". Disponível em <http://www.macnn.com/articles/13/03/22/music.video.still.low.margin.growth.and.softw.are.add.billions/>>, Acesso em 31 outubro 2013.

HOM Daniel, 9 jan 2013, "The Rise of Digital Music". Disponível em <<http://www.tableausoftware.com/public/blog/2013/01/rise-digital-music-1773>>, Acesso em 31 outubro 2013.

BUSINESSWIRE, 4 jan 2013, "The Nielsen Company & Billboard's 2012 Music Industry Report". Disponível em <<http://www.businesswire.com/news/home/20130104005149/en/Nielsen-Company-Billboard%E2%80%99s-2012-Music-Industry-Report>>, Acesso em 31 outubro 2013.

IFPI, 26 fev 2013, "IFPI publishes Digital Music Report 2013". Disponível em <http://www.ifpi.org/content/section_resources/dmr2013.html>, Acesso em 31 outubro 2013.

CROSBY Alfred W.. *The Measure of Reality: Quantification in Western Europe, 1250 1600*, Cambridge University Press, 13 dez 1997. p.152.

PAPAJ M., set 2008, "Silence/noise detection for speech and music signals". Disponível em <<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=5967594>>, Acesso em 31 outubro 2013.

BURK; POLANSKY; REPETTO; ROBERTS; ROCKMORE, 2011, "Music and Computers, A Theoretical and Historical Approach". Disponível em <<http://music.columbia.edu/cmcmusicandcomputers/>>, Acesso em 31 outubro 2013.

GERHARD David, nov 2003, "Pitch Extraction and Fundamental Frequency: History and Current Techniques". Disponível em <<http://www.cs.uregina.ca/Research/Techreports/2003-06.pdf>>, Acesso em 31 outubro 2013.

HARDESTY Larry, 25 nov 2009, "Explained: The Discrete Fourier Transform". Disponível em <<http://web.mit.edu/newsoffice/2009/explained-fourier.html>>, Acesso em 31 outubro 2013.

VILLATE Jaime E., dez 2001, "Equações Diferenciais e Equações de Diferenças". Disponível em <<http://people.ufpr.br/~jcvb/online/eqdiferenciais.pdf>>, Acesso em 31 outubro 2013.

PICARD R. W., 2003, "MIT MAS 160/510 Additional Notes, Spring 2003". Disponível em <<http://courses.media.mit.edu/2012spring/mas160/zl.pdf>>, Acesso em 31 outubro 2013.

BENWARD; SAKER. *Music in Theory and Practice, Vol. I*. McGraw-Hill, 8a ed., mai 2008. 407 p..

PISTON Walter. *Harmony*. Londres: Victor Gollancz LTD, 2a ed. revisada, 1959. 344 p..

BUCKINGHAM Bruce; PASCAL Eric. *Rhythm Guitar: The Complete Guide*. Musicians Institute Press, 1 out 1997. 144 p..

ISO 16. 1975 Acoustics -- Standard tuning frequency (Standard musical pitch). [S.I.]: International Organization for Standardization, 1975. Disponível em <http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=3601>, Acesso em 31 outubro 2013.

AUDACITY, "How Noise Removal Works". Disponível em <http://wiki.audacityteam.org/wiki/How_Noise_Removal_Works>, Acesso em 31 outubro 2013.

University of New South Wales, "Music Acoustics". Disponível em <<http://www.phys.unsw.edu.au/jw/basics.html>>, Acesso em 31 outubro 2013.

BATENKOV Dima, "Fast Fourier Transform: Key Papers in Computer Science Seminar 2005". Disponível em <<http://www.wisdom.weizmann.ac.il/~naor/COURSE/fft-lecture.pdf>>, Acesso em 31 outubro 2013.

COOLEY J. W.; TUKEY J. W.. "An algorithm for the machine calculation of complex Fourier series. *Mathematics of Computation*", 297-301, 1965.

DOVICCHI, J. C. L. e PARREIRA, W. D. "Combinacões Lineares de Vibracões Harmônicas Simples e Complexas", *Tech. Rep.*, Lab. Música Eletroacústica, NACSM/UFU, 2003.

BRIGHAM, E. Oran. "The fast Fourier transform and its applications". Englewood Cliffs, N.J.: Prentice Hall, 1988.

DOVICCHI, J. C. L. e ALVES, J. B. M. "Functional Programming as a Natural Choice for Spectral Centroid Analysis in Digital Signal Processing (DSP)" Enviado ao Journal of Functional Programming.

READ, Gardner. "Music notation: a manual of modern practice". Allyn and Bacon, 1969. 482 p..

HURON David, 1 fev 2002, "The role of embellishment tones in the perceptual segregation of concurrent parts.". Disponível em <<http://www.music-cog.ohio-state.edu/Huron/Publications/huron.embellish.html>>, Acesso em 31 outubro 2013.

JONT B. Allen, Jun 1977. "Short Time Spectral Analysis, Synthesis, and Modification by Discrete Fourier Transform". *IEEE Transactions on Acoustics, Speech, and Signal Processing*. ASSP-25 (3): 235–238. Disponível em <<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=1162950>>, Acesso em 31 outubro 2013.

Alonso, M.;David, B.; Richard G., March 31, April 1, 2003. "*A study of tempo tracking algorithms from polyphonic music signals*". Disponível em <<http://perso.telecom-paristech.fr/~grichard/Publications/cost03.pdf>>. Acesso em 31 outubro 2013.

KUC, Roman. *Introduction to Digital Signal Processing*. McGraw-Hill, 1 ago 1988. 492p..

7 Apêndice

7.1 Apêndice A – Artigo

Uma Contribuição À Comparação Entre Arquivos De Música Para Uma Implementação Computacional

Edson Edgar da Silveira

Departamento de Informática e Estatística – Universidade Federal de Santa Catarina
(UFSC)

88.040-970 – Florianópolis – SC – Brasil

edsonsilveira@inf.ufsc.br

***Abstract.** This paper discusses the main points of the author's bachelor's degree final project. The project [1] consists of a study of music files comparison and proposes an alternative way of approaching the problem. It focuses on the use of some music theory concepts through computer science means. It describes the general steps needed in order to make it possible a future computer implementation.*

***Resumo.** Este artigo discorre os principais pontos do trabalho de conclusão de curso do autor. O projeto [1] consiste em um estudo de comparação entre arquivos de música e propõe uma forma alternativa de abordar o problema. Ele foca no uso de alguns conceitos da teoria musical no âmbito computacional. Descreve ainda os passos gerais necessários para uma futura implementação computacional.*

1. O Problema Da Comparação

O compartilhamento e compra de músicas online vem crescendo a medida que o meio digital se torna cada vez mais presente na vida das pessoas. Assim como outras atividades, este novo meio possibilita e muitas vezes exige formas mais eficientes de lidar com a distribuição musical. Uma dessas formas envolve como melhor sugerir e buscar produtos semelhantes.

Este trabalho busca apresentar uma forma para comparação de arquivos de música baseados em conceitos da teoria musical. A proposta apresenta os passos a serem tomados para que se tornem possíveis assim como suas dificuldades e limitações.

2. A Proposta De Comparação

Uma das muitas maneiras que se pode pensar de como comparar uma música à outra, sem pensar diretamente na questão técnica da computação, é pela natureza musical de ambas. Culturas diferentes possuem interpretações diferentes da música assim como a

forma que dividem as frequências em grupos. A mais comum e globalmente usada vem da escola europeia de música clássica. Tomando suas teorias como referência pode-se analisar músicas diferentes com o propósito de buscar similaridades e dissimilaridades. Baseando-se na possibilidade de análise musical automatizada, esta proposta busca trazer estes elementos para o universo da ciência da computação assim como entender suas limitações.

3. Solução Proposta

A proposta tem como foco central extrair informações quanto a harmonia para que seja possível estabelecer uma comparação entre músicas diferentes e identificar suas similaridades.

Para comparar uma música com a outra inicialmente é necessário extrair dados de frequência do arquivo. Antes porém deve-se decidir o que será considerando silêncio para que trecho que foram considerados como tanto, não afetem as análises feitas com sons considerados válidos pela proposta.

Com tais dados já obtidos a música é dividida em intervalos discretos com uma duração mínima a ser decidida pela implementação a ser feita. Para cada um desses intervalos será identificada sua centroide se usado métodos relacionados à Transformada de Fourier.

Armazenando estas informações obtêm-se os dados das mudanças de frequência em relação ao tempo para uma determinada música de forma discreta. Isso porém não é adequado ainda para uma comparação já que para fazer a análise musical desejada será preciso se usar das teorias musicais que neste caso exige que se trabalhe com apenas 12 semitons. Se faz necessário então estabelecer um método para que se tenha apenas frequências correspondentes a esse número. Uma forma possível de resolver este caso pode simplesmente ser trocar cada um dos valores de frequência por um símbolo representando a nota mais próxima daquela frequência. Isso pode resultar em valores frequências muito diferentes para uma mesma nota devido aos harmônicos das notas mas isso não foge do interesse desta proposta.

Como resultado obtido tem-se as mudanças de notas, independente de frequência exata, ao longo do tempo. Alguns ajustes ainda se fazem necessários. Um deles é agrupar notas iguais consecutivas considerando-as uma só e tendo como duração a quantidade de notas agrupadas. Outro passo é avaliar estes dados até o momento e eliminar notas extremamente infrequentes. Para este passo funcionar de forma mais eficiente há algumas dificuldades que serão analisadas.

Com esta fase de preparação encerrada pode-se finalmente partir para técnicas de análise harmônica. Várias são possíveis e algumas inclusive podem se somar para melhorar a análise. Buscar identificar possíveis escalas dentre os dados obtidos, notas frequentes, clichês musicais e outras são algumas das opções.

Último passo é então comparar os dados analisadas de uma música para com a outra. As possibilidades são inúmeras e podem ser diferentes em situações diversas uma vez que é um tanto subjetivo. Esta proposta tenta se utilizar de métodos que vão de acordo com a natureza das análises feitas.

4. Conclusão

Com a conclusão deste estudo se pode inferir algumas observações. Primeiramente é que oferece uma forma de comparar arquivos de áudio diferente da usualmente aplicada. Essa forma por si só pode em muitos casos não justificar sua aplicação de maneira isolada porém, junto de outras técnicas, mesmo as mais simples como usando rótulos para associar características à certa música, pode se tornar relevante. Permite encontrar similaridades muitas vezes não esperadas de outras formas. Assim sendo, o

objetivo inicial do estudo foi atingido. A apresentação de diversos passos para tornar tal proposta aplicável também foi trabalhada como se desejava ainda que existam certas dificuldades a serem exploradas em trabalhos futuros e implementações.

Nota-se porém a grande quantidade de conceitos que são envolvidos tornando-a complexa, dificultando sua compreensão e possível implementação. Muitos dos passos contidos na proposta dependem de outros estudos que possuem suas próprias limitações e problemas. Assim sendo, este trabalho oferece uma proposta de análise e por este motivo não cobre a validação do método sugerido.

Contudo, a proposta oferecida pode abrir caminho para uma forma alternativa de abordar a comparação de arquivos de música.

Referências

- [1] da Silveira, Edson Edgar. (2013) “Uma Contribuição À Comparação Entre Arquivos De Música Para Uma Implementação Computacional”. UFSC.