

Sumário

1	INTRODUÇÃO	3
1.1	Fundamentação Teórica	4
1.1.1	A etapa de Pré-Processamento	4
1.1.2	A etapa de Processamento	6
1.1.2.1	O algoritmo Needleman & Wunsch (NW)	7
1.1.2.2	O algoritmo Dynamic Time Warping (DTW)	8
1.1.2.3	O algoritmo Smith Waterman (SW)	9
1.1.3	A etapa de Pós-Processamento	10
1.2	Questão de Pesquisa e Hipótese	11
1.3	Objetivo desta Pesquisa	12
2	ANÁLISE DO PROBLEMA	13
2.1	Contorno Melódico , Harmônico e Rítmico	14
2.2	A atitude do Músico	15
2.3	Convergências Conceituais	15
2.4	Conclusão da Análise	16
3	A MATRIZ MUSSIX (MU6)	17
3.1	Cálculos da Mu6	18
3.1.1	Definição de Match/Mismatch da Mu6	18
3.1.2	Definição de Penalidades da Mu6	18
3.2	O Algoritmo Mu6	20
3.3	O Attitude-Gram	21
3.3.1	As Características Attitude-Gram	23
4	PROTOCOLO EXPERIMENTAL	25
4.1	O Sistema de Busca	25
4.1.1	Pré-Processamento QBPLab	25
4.1.2	Processamento QBPLab	26
4.1.2.1	O Balanço de Precisão	27
4.1.2.2	O Balanço de Tamanho de Resultado	27
4.1.3	Pós-Processamento QBPLab	27
5	AValiação DOS RESULTADOS	28
5.1	Experimentos	28
5.1.1	Experimento 1	28

5.1.2	Experimento 2	28
5.1.3	Experimento 3	28
5.1.4	Experimento 4	28
5.2	Resultados Obtidos	28
5.3	Análise dos Resultados	28
6	CONSIDERAÇÕES FINAIS	29
6.1	Relevância do Estudo	29
6.2	Contribuições da Pesquisa	29
6.3	Limitações da Pesquisa	29
6.4	Trabalhos Futuros	29
	REFERÊNCIAS	30

1 Introdução

Music Information Retrieval (MIR) é um campo de pesquisa multidisciplinar que se baseia nas tradições, metodologias e técnicas de uma variedade notavelmente ampla de disciplinas incluindo Acústica, PsicoAcústica, Musicologia, Ciência da Computação, Aprendizagem de Máquina, entre outros.

O principal objetivo da pesquisa MIR, independentemente do paradigma disciplinar sob o qual é realizada, é proporcionar o acesso ao vasto conteúdo musical existente atualmente em um nível igual ou superior ao realizado pelos mecanismos de pesquisa baseados em texto.

Uma música por si só, é uma junção complexa de fenômenos acústicos, rítmicos, harmônicos, estruturais e culturais, onde o músico expressa sua arte através da sua atitude musical, ou seja, a maneira como o Músico se expressa através de seu instrumento musical, seja ele qual for.

O grande desafio da pesquisa MIR é o desenvolvimento de sistemas de recuperação que sejam capazes de interpretar este conteúdo por si só, Ou seja, os pesquisadores do MIR se esforçam para criar sistemas de recuperação em que a própria música, representada na forma de áudio (por exemplo, MP3s, WAV, etc.) ou de maneira simbólica (por exemplo, MIDI, partituras etc.) , seja o principal mecanismo pelo qual os usuários/músicos interagem com tais sistemas.

Em termos simples, a pesquisa do MIR quer desenvolver sistemas que permitam aos usuários pesquisarem conteúdo musical usando consultas estruturadas musicalmente.

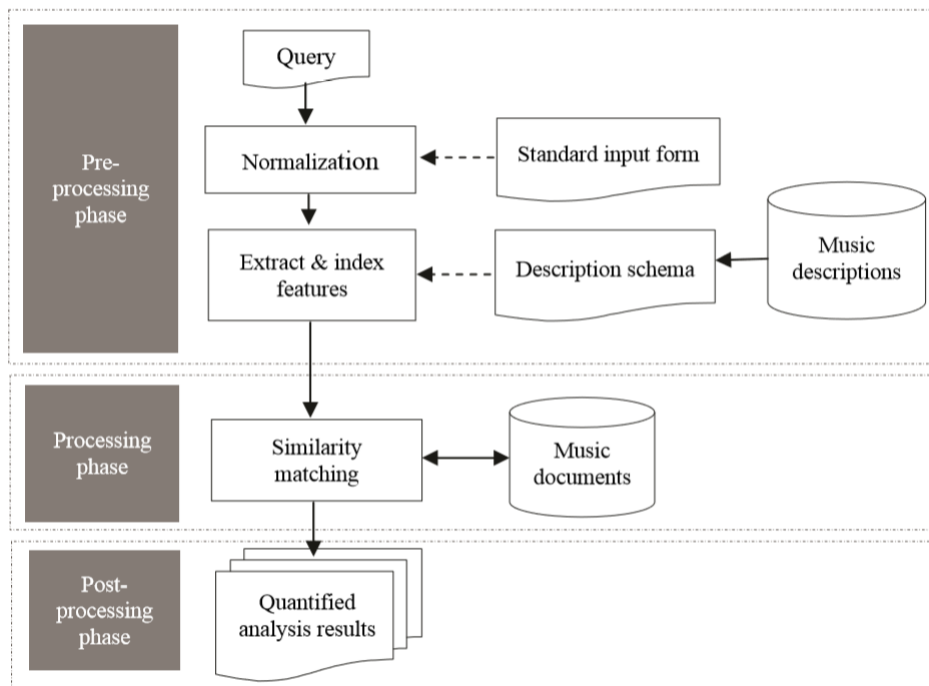
Tais consultas podem ser estruturadas musicalmente de diversas maneiras, onde comumente em MIR vêm sido identificadas como consultas "Query-By", onde podem incluir técnicas como consulta por canto, consulta por exemplo (por exemplo, enviar um MP3 conhecido para encontrar peças semelhantes) e / ou consulta por notação (por exemplo, colocar notas em um quadro musical para formar a consulta), etc.

Este documento descreve a pesquisa realizada no intuito de fazer com que buscas Query-By-Playing, possam realizar estas consultas por similaridade, utilizando dados simbólicos musicais, de maneira mais precisa melhorando os resultados destas buscas, tornando-as mais Musicais do que apenas utilizar a similaridade musical calculada através da análise tonal da música (Notas Musicais).

1.1 Fundamentação Teórica

De uma maneira geral um sistema MIR funciona em etapas distintas responsáveis pelo processamento do som e detalhadas no artigo (Kuldeep Gurjar, 2018)(Figura 1).

Figura 1 – MIR System: Estrutura de Processamento



Fonte: (Kuldeep Gurjar, 2018)

1.1.1 A etapa de Pré-Processamento

A grande maioria dos dados musicais utilizados, sendo arquivos de áudio, ou arquivos com dados musicais simbólicos são armazenados procurando seguir a mesma estrutura musical original, onde uma música é considerada como o encadeamento sequencial de notas musicais, que de maneira cadenciada forma a sua estrutura fundamental.

Na etapa de pré-processamento de uma *Query*, a informação recebida para a busca, é extraída, convertida e estruturada, para que possa ser utilizada na etapa posterior de processamento.

De maneira geral, esta conversão ocorre de acordo com o tipo e os propósitos da consulta sendo realizada.

Como identificado por (MARTINIANO; N. Silla, 2017), a grande maioria dos sistemas de consultas musicais, acaba convertendo dados de áudio, para dados simbólicos como exemplificado na Figura 2.

Figura 2 – Pré-Processamento: Áudio para Dados Simbólicos

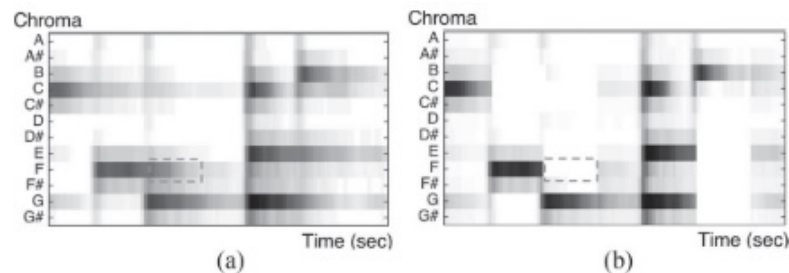


Fonte: Autoria Própria

Outros tipos de conversão podem ser realizados nesta etapa, como por exemplo em Em 2013 no artigo ([PARK; LEE, 2013](#)), Min Woo Park & Eui Chul Lee , através de uma abordagem de processamento de dados simbólico converteram faixas MIDI em um vetor unidimensional numerico contendo apenas algumas características musicais. A base desta pesquisa foi a proposta de um algoritmo de conversao das notas musicais simbolicas MIDI (pitch e tempo), em um vetor unidimensional de notas numericas.

Nesta etapa os dados podem ser estruturados de forma diferente ao convencional onde por exemplo, Em 2016, no artigo ([LI; DUAN, 2016](#)), é apresentado uma forma de processar o calculo de score de similaridade convertendo dados de Áudio em um Cromagrama para possibilitar o mapeamento de variações tonais causadas por alavancas tonais presentes nos teclados musicais.

Figura 3 – Pré-Processamento: Áudio para Chromagrama

Fonte: ([LI; DUAN, 2016](#))

Em 2017 no artigo ([MARTINIANO; N. Silla, 2017](#)) "BIRITS: A Music Information Retrieval System Using Query-by-Playing Techniques", Martimiano & Silla apresentam um metodo de Query-by-Playing em dois estágios utilizando dados simbólicos da musica. Em seu pré-processamento é realizado a conversão de músicas simbólicas MIDI em cadeias de strings , para o processamento de calculo de similaridade, conforme a Figura 4.

A conversão de dados MIDI em strings possibilitou a redução das notas para o calculo de similaridade a 12 notas pré-definidas presentes na escala diatônica ocidental, convertidas dentro da escala de Dó(Escala 5 MIDI).

Figura 4 – Pré-Processamento: MIDI para Strings



Fonte: ([MARTINIANO; N. Silla, 2017](#))

1.1.2 A etapa de Processamento

A etapa de processamento caracteriza o núcleo da funcionalidade de busca (Query) e em um sistema MIR abrange primordialmente o cálculo de pontuação de similaridade entre as músicas envolvidas na busca.

Da mesma maneira que no Pré-Processamento as técnicas e algoritmos utilizados nesta etapa variam de acordo com o tipo de busca sendo realizada, que precisa estar em concordância com os dados recebidos pelo pré-processamento.

Várias técnicas para avaliar músicas simbólicas similares foram introduzidas nos últimos anos, desde algoritmos geométricos, que consideram representações geométricas de melodias e calculam a distância entre estas músicas, até utilização de aprendizagem de máquina e métodos probabilísticos como redes de Markov para determinar esta similaridade.

Em 2013, Ching-Hua Chuan no artigo ([CHUAN, 2013](#)) "A multimodal approach to song-level style identification in pop/rock using similarity metrics" é utilizado a geração de n-grams a partir de sequências de acordes realizados por músicos para a identificação do seu estilo individual (Signature Style). Neste artigo foram gerados três (3) n-grams gerados a partir de cálculos de Similaridade Melódica, Perfil de Harmonia e Características acústicas, e que alimentam um Classificador Binário em SVM (Support Vector Machines) para classificar o estilo de cada música.

No mesmo ano, Alessio Degani, Marco Dalai, Riccardo Leonardi e Pierangelo Miglioni no artigo ([DEGANI et al., 2013](#)) "A Heuristic for Distance Fusion in Cover Song Identification" detalham uma pesquisa de fusão de duas features musicais que otimizam os resultados de cálculo de similaridade entre músicas. Para os experimentos são utilizadas a combinação de duas features (Salience Feature e Harmonic Pitch Class Profile - HPCP) sendo utilizadas em dois algoritmos de cálculo de distância (Dinamic Time Warping e Qmax Cross Recurrence Plot).

Em 2017 no artigo ([MARTINIANO; N. Silla, 2017](#)) "BIRITS: A Music Information Retrieval System Using Query-by-Playing Techniques", Martiniano & Silla utilizam um cálculo de similaridade por strings de notas e propõem uma Matriz de Pesos contendo parâmetros de Acertos/Erros baseada na distância em semi-tons entre as notas musicais para a otimização dos resultados do alinhamento local chamada MUSSUM (Music Substitution Matrix).

É importante notar que apesar destes métodos e algoritmos se aplicarem ao cálculo da medida de similaridade, os únicos algoritmos que levam em consideração as características principais de uma música (Encadeamento sequencial de notas no tempo), são os algoritmos de alinhamento de sequências Global e Local.

1.1.2.1 O algoritmo Needleman & Wunsch (NW)

Como descrito em (FERRARO; HANNA, 2007) , em 1970 Needleman & Wunsch (NEEDLEMAN; WUNSCH, 1970) e, em seguida, Wagner & Fisher (1974) propuseram algoritmos que computam uma medida de similaridade entre duas cadeias de Símbolos como a sequência mínima de pontuação das operações elementares necessárias para transformar uma das cadeias na outra.

Dadas duas cadeias de símbolos $S1$ e $S2$ com seus respectivos comprimentos $|S1|$ e $|S2|$, ou seja, o número de notas em melodias, um conjunto de operadores elementares em strings, chamados de operações de edição e uma pontuação associada a cada operação de edição, uma pontuação entre essas duas strings é definida como a pontuação da sequência de operações de edição transformando $S1$ em $S2$ com uma pontuação mínima. Esta medida de similaridade faz uso do princípio de programação dinâmica para alcançar um algoritmo com complexidade quadrática, ou seja, em $O(|S1| \times |S2|)$.

A idéia principal do algoritmo de alinhamento é mapear o número de operações de Substituição, Alteração ou Exclusão necessárias para transformar $S1$ em $S2$, onde o score de similaridade seria provindo do número de operações realizadas para cada item da string, armazenados durante o andamento do algoritmo em uma matriz de distancias (número de operações) ou score.

Mesmo utilizando-se do algoritmo e uma matriz de distâncias, executando os algoritmos anteriores obtemos apenas o score de similaridade. Caso seja necessário obter também a string resultante deste alinhamento, é necessário armazenar todas as operações realizadas para que posteriormente possamos reconstruir o caminho percorrido na matriz.

Esta reconstrução do caminho é comumente chamada de Backtracing, onde traçamos o caminho percorrido das operações e determinamos o seu resultado. Este resultado final determina o seu nome de algoritmo, pois através do alinhamento das distancias e com o backtracing das operações resultando em uma string alinhada em sua totalidade, o nome alinhamento global foi adotado para este tipo de algoritmo.

Algoritmo 1: O Algoritmo Needleman & Wunsch**Result:** Matriz de Scores Calculada(Bottom-Up).

```

1 Inicialização;
2  $D(i,0) = i$ ;
3  $D(0,j) = j$ ;
4 Calculo de Scores;
5 for  $i \leftarrow 1$  to  $M$  do
6   for  $j \leftarrow 1$  to  $N$  do
7      $D(i,j) = \text{Min}(i,j)$ ;
8   end
9 end

```

Algoritmo 2: A Funcao Min**input** : A posicao atual i,j **output** : Distancia Minima**Result:** O resultado sera o score minimo de distancia.

```

10  $embaixo = M[i-1,j] + 1$ ;
11  $lado = M[i,j-1] + 1$ ;
12  $diagonal = M[i-1,j-1] + \text{Score de Erro} : \text{Acerto}$ ;
13  $retorno = \text{Minimo}(embaixo, lado, diagonal)$ 

```

Figura 5 – Processamento NW: a) Scores e b) Backtracing

N	9	8	9	10	11	12	11	10	9	8
O	8	7	8	9	10	11	10	9	8	9
I	7	6	7	8	9	10	9	8	9	10
T	6	5	6	7	8	9	8	9	10	11
N	5	4	5	6	7	8	9	10	11	10
E	4	3	4	5	6	7	8	9	10	9
T	3	4	5	6	7	8	7	8	9	8
N	2	3	4	5	6	7	8	7	8	7
I	1	2	3	4	5	6	7	6	7	8
#	0	1	2	3	4	5	6	7	8	9
#	E	X	E	C	U	T	I	O	N	

n	9	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖
o	8	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖
i	7	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖
t	6	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖
n	5	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖
e	4	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖
t	3	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖
n	2	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖
i	1	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖
#	0	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖	↖
#	e	x	e	c	u	t	i	o	n										

Fonte: Autoria própria

1.1.2.2 O algoritmo Dynamic Time Warping (DTW)

O algoritmo DTW é um algoritmo de alinhamento global comumente utilizado no processamento de músicas polifônicas por possuir um cálculo geométrico de distância (Distância Euclideana) na construção da matriz de pesos.

Tecnicamente a única diferença entre os algoritmos NW e o DTW consiste na soma do peso atribuído ao score mínimo em cada iteração do algoritmo, onde no algoritmo NW

esta soma ocorre apenas quando um `match(diagonal)` ocorre.

Algoritmo 3: O Algoritmo DTW

Result: Matriz de Scores Calculada(Bottom-Up).

```

14 initialization;
15 D(i,0) = infinito;
16 D(0,j) = infinito;
17 D(0,0) = 0;
18 for i ← 1 to M do
19   for j ← 1 to N do
20     custo = EuclideanDistance(i,j);
21     D(i,j) = custo + Min(i,j);
22   end
23 end
  
```

Algoritmo 4: A Funcao Min

input : A posicao atual i,j

output : Distancia Minima

Result: O resultado sera o score minimo de distancia.

```

24 embaixo = M[i-1,j ];
25 lado = M[i ,j-1];
26 diagonal = M[i-1,j-1];
27 retorno = Min(embaixo,lado,diagonal)
  
```

1.1.2.3 O algoritmo Smith Waterman (SW)

Como descrito também em (FERRARO; HANNA, 2007), Em muitas aplicações, duas cadeias podem não ser muito semelhantes na sua totalidade, mas podem conter regiões altamente semelhantes (substrings). Nesse caso, a tarefa é localizar e extrair um par de regiões, uma a uma das duas cadeias onde existem altas similaridades. Isto é chamado de alinhamento local ou problema de similaridade local definido por Smith & Waterman em 1981 (SMITH; WATERMAN, 1981) e é definido como: dada duas sequências S1 e S2, encontrar as substrings p1 e p2 de S1 e S2 respectivamente, cuja similaridade é máxima sobre todos os pares de substrings de S1 e S2.

Como no algoritmo de alinhamento global, o algoritmo de alinhamento local, também precisa realizar um backtracing na sua matriz (Chamada de Substitution Matrix), e que resulta em diferenças significativas.

A principal diferença está nas repetições de score máximo , onde para cada score encontrado, uma substring local é gerada.

Outra diferença significativa no algoritmo relativo a seu antecessor é a possibilidade de processamento dos Gaps entre estas substrings, permitindo que uma tolerância a erros possa ser mapeada.

Algoritmo 5: O Algoritmo Smith Waterman

Result: Matriz de Scores Calculada(Top-Down).

```

28 Inicialização;
29  $D(i,j) = 0$ ;
30 Calculo de Scores;
31 for  $i \leftarrow 1$  to  $M$  do
32   for  $j \leftarrow 1$  to  $N$  do
33      $D(i,j) = \text{Max}(i,j)$ ;
34   end
35 end
```

Algoritmo 6: A Funcao Max

input : A posicao atual i,j

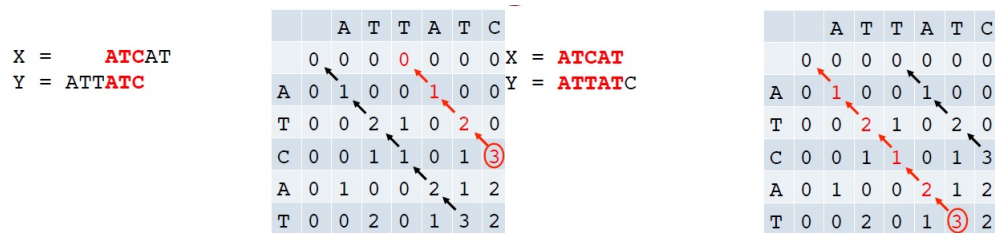
output : Distancia Maxima

Result: O resultado sera o score maximo de distancia.

```

36  $lado = M[i-1,j] + gap$ ;
37  $topo = M[i,j-1] + gap$ ;
38  $diagonal = M[i-1,j-1] + ScoredeErro : Acerto$ ;
39  $retorno = \text{Maximo}(topo, lado, diagonal)$ 
```

Figura 6 – Processamento SW: Backtracing



Fonte: Autoria própria

1.1.3 A etapa de Pós-Processamento

De maneira geral na etapa de pós-processamento os resultados do processamento de similaridade musical são coletados.

É na etapa de pós-processamento que dados estatísticos como acuracia e ranqueamento de similaridade são coletados para uma posterior análise quantitativa das informações geradas em cada busca MIR.

Em 2007 no artigo ([FERRARO; HANNA, 2007](#)) "Optimizations of Local Edition for Evaluating Similarity Between Monophonic Musical Sequences", Ferraro & Hanna realizam comparações entre os algoritmos de similaridade utilizados até então. Nestas comparações são propostos novos parâmetros de configuração dos algoritmos baseado em regras musicais para a melhoria da acurácia destes algoritmos. Para a avaliação dos resultados foi criado pós-processamento onde a técnica Average Dynamic Recall (ADR), é utilizada para avaliar todos os experimentos.

No artigo ([WANG; TSAI; SU, 2014](#)) "Note-based alignment using score-driven non-negative matrix factorisation for audio recordings" onde Tien-Ming Wang, Pei-Yin Tsai & Alvin W. Y. Su, apresentam a divisão de faixas de áudio em notas simbólicas MIDI e colocadas em uma estrutura de conversão chamada Piano-Roll feature. Todo o Pós-Processamento das comparações levaram em conta a acurácia resultante como medida de comparação.

Conforme exposto no artigo ([DOWNIE et al., 2010](#)) "The Music Information Retrieval Evaluation eXchange: Some Observations and Insights", Stephen Downie coloca "a seleção ou criação de métricas de avaliação apropriadas é crucial para a avaliação científica apropriada do desempenho do sistema MIR."

A seleção das métricas de avaliação também possui um forte componente emocional, à medida que os participantes se esforçam para mostrar o sucesso de seus algoritmos e sistemas da melhor maneira possível.

Ainda no mesmo artigo é apresentado uma lista de métricas para a avaliação de sistemas MIR, estas métricas são utilizadas na competição de algoritmos MIR, chamada MIREX - Music Information Retrieval Evaluation eXchange, e seguem abaixo elencadas as principais métricas utilizadas para avaliação de algoritmos MIR.

1. Precision, Recall e F-Measure
2. Mean Reciprocal Rank
3. Average Accuracy

1.2 Questão de Pesquisa e Hipótese

Para este trabalho temos a seguinte questão de pesquisa:

1. É Possível obter resultados mais "Musicais", que se aproxime da forma como o ser humano identifica Músicas?

A hipótese desta pesquisa é:

1. Existe uma forma de obter scores de similaridade mais musicais através de parametrização de alinhamento de sequência local entre músicas.

1.3 Objetivo desta Pesquisa

O Cálculo de Pontuação de Similaridade, em todos os trabalhos de MIR, desempenha papel central, onde através dele são elencadas as músicas sejam para análise de resultados, posterior classificação, ou até mesmo para a preparação de dados para futura utilização.

Portanto o objetivo deste trabalho é otimizar o cálculo de pontuação de similaridade para buscas de músicas simbólicas (SMS) Query-By-Playing, a fim de obtermos maior precisão nos resultados de um modo geral.

2 Análise do Problema

Em todas as pesquisas analisadas para este trabalho, quando o cálculo de similaridade leva em conta a estrutura de uma música, a mesma realiza este calculo de pontuação levando-se em conta apenas sua estrutura tonal(notas musicais), onde outros aspectos musicais que podem auxiliar no aumento da precisão destes resultados não é utilizada e quando utilizada, a mesma é feita após ou antes do calculo de pontuação, onde não foi encontrado nenhuma, em que outras estruturas musicais fossem utilizadas durante o processamento de calculo.

Isto se deve em alguns casos, por não haver a necessidade de uma precisão alta para a seleção das músicas, em outros casos foi identificado que devido aos algoritmos de alinhamento de sequência garantirem o melhor alinhamento, o maior score sempre é resultante da música correta não havendo necessidade de precisão maior.

Em Query-By-Playing além de possuímos apenas fragmentos de músicas para alinhamento, a precisão no retorno de tais músicas deve ser refletida em toda a lista de músicas e não somente na primeira música mais similar, onde a ocorrência de um número alto de repetições de pontuação descaracterizam o resultado final da busca.

Por Exemplo, no artigo ([MARTINIANO; N. Silla, 2017](#))"BIRITS: A Music Information Retrieval System Using Query-by-Playing Techniques", Martimiano & Silla apresentam um metodo de Query-by-Playing em dois estágios utilizando dados simbólicos da musica.Em seu trabalho, no segundo estágio de processamento é aplicado um filtro SVM , para a diminuição de repetições de pontuação de similaridade,onde mesmo com a aplicação de filtro SVM com 87% de acurácia, a efetividade na filtragem dos resultados foi em torno de 1% dos resultados.

No trabalho BIRITS , foi obtido resultados significativos relativos a forma de utilização do algoritmo de alinhamento e consequentemente no calculo de pontuação de similaridade, onde fora criada a matriz MUSSUM - Music Substitution Matrix.A idéia principal desta matriz foi a utilização da distância em semi-tons das notas da escala diatônica para a atribuição das penalidades (Gaps) no algoritmo de alinhamento local Smith-Waterman (SW).

Isto abriu o precedente principal da hipótese deste trabalho, onde viu-se que, seria possível otimizar os resultados destes alinhamentos e obter pontuação de similaridade mais precisa e com um menor indice de repetições.

Aliado a isto, devido a entrada da consulta musical ser realizada por músicos, a precisão do retorno deve dar a impressão ao ser humano, que não apenas o acerto de notas

está sendo levada em consideração, onde a utilização de estruturas musicais independentes da característica tonal da música precisa ser levada em conta também.

2.1 Contorno Melódico , Harmônico e Rítmico

Um contorno de maneira geral pode ser descrito como o desenho do relevo formado ao longo de uma ou mais medidas distribuídas através de uma sequência.

Para a definição de uma música, existem três componentes fundamentais, a melodia , a Harmonia e o Ritmo.

Uma melodia pode ser considerada como o encadeamento sequencial de notas musicais, que de maneira cadenciada forma a estrutura fundamental da Música(Figura 7).

Figura 7 – Contorno Melódico de uma Música.



Fonte: Autoria Própria

Cada nota emitida por um instrumento musical possui uma frequência específica de acordo com a escala tonal do instrumento musical, e as mesmas são cadenciadas na melodia através do tempo de vibração de cada nota formando o Ritmo Musical.

Cada uma das notas musicais da escala diatônica (Dó, Ré , Mi , Fá , Sol, La , Si) distanciam-se umas das outras através de uma medida de Tons e Semi-Tons, onde a menor dela é o Semi-Tom.

Portanto um Contorno Melódico pode ser definido como o numero de semi-tons entre uma nota e sua nota subsequente na melodia. Conforme citado por (DUREY; CLEMENTS, 2003), o contorno melódico é encontrado quando calculamos o numero de semi-tons de uma nota para a próxima.

Extendendo esta teoria para as demais estruturas musicais , o Contorno Harmônico é o contorno desenhado pelas frequências harmônicas de cada uma das notas da cadência, e o Contorno Rítmico é a Duração de cada nota da cadência e que define o ritmo da música.

Foi percebido em (MARVIN, 1988) que a percepção do contorno para o ser humano é mais imediata do que a do tom e Parece provável que a percepção do contorno, em algum nível, tenha um papel na maneira como os ouvintes podem diferenciar melodias, associar relações temáticas, determinar formalidades esquemas e assim por diante, pois atendem a uma composição não-tonal.

Aplicações analíticas da teoria dos contornos, portanto, modelam alguns aspectos da percepção musical do ouvinte que a análise de pitch sozinha não aborda.

2.2 A atitude do Músico

Outro aspecto importante no intuito de simular a percepção humana na detecção de similaridades musicais é relativo ao modo de reconhecimento de padrões rítmicos e harmônicos provindos da atitude do músico na composição musical.

De acordo com o estudo de (MARVIN, 1988):

“Finalmente, a generalização da teoria do contorno em outros domínios permite ao analista comparar diversas facetas da estrutura musical em uma única escala seqüencial. Os exemplos analíticos mostraram maneiras pelas quais a análise de sucessões de duração como contornos rítmicos esclarece alguns aspectos da linguagem musical de um compositor.”

Mesmo com excelentes resultados em seu trabalho, boa parte do estudo de MIR busca reproduzir a reação humana em identificar a similaridade do ouvinte, porém quase nenhum estudo foi encontrado na tentativa de tentar reproduzir como o músico tenta expressar seus sentimentos e expressões através de sua música.

Portanto chega-se a conclusão que, não basta apenas a verificação por similaridade Tonal somada ao Contorno Melódico, mas também as características de Atitude do Músico na composição musical também possuem um grau significativo da percepção musical do ser humano, já que acabam expressando aspectos culturais e sociais do mesmo.

2.3 Convergências Conceituais

Apesar de parecerem bastante promissores os resultados avaliados por (MARVIN, 1988), em 2007 no artigo (FERRARO; HANNA, 2007) "Optimizations of Local Edition for Evaluating Similarity Between Monophonic Musical Sequences", Ferraro & Hanna realizam comparações entre os algoritmos de similaridade utilizados até então e identificaram que além de haver diferenças entre resultados provindos de alinhamentos locais e Globais, a utilização de contorno melódico foi avaliada da seguinte forma:

“Os resultados mostram claramente que a representação do contorno leva aos piores resultados. A diferença de acurácia é muito significativa: 38%, enquanto os demais resultados são maiores que 50%. Isso se justifica principalmente pela falta de informações contidas nessa representação. É óbvio que várias melodias podem ser representadas pelo mesmo contorno melódico. Assim, o vocabulário proposto parece ser muito limitado para esta aplicação musical. As representações relativas absolutas e principais levam aproximadamente às mesmas medidas médias de RAM. No entanto, as principais representações

relativas e de intervalo apresentam a grande vantagem de propor um escore de similaridade invariável na transposição.”

Através de uma análise do trabalho de (CHUAN, 2013), foi possível verificar também o mapeamento dos sentimentos intuitivos de similaridade na perspectiva do ouvinte, concentrando-se na criação de características inseridas em um n-gram de valores numéricos, para treinamento de um Classificador SVM , identificando o estilo musical de um Músico.

Como resultado Ching-Hua Chuan mostra que a abordagem alcança o melhor resultado na taxa correta de 85% usando apenas sete características de similaridade.

Portanto é possível concluir que as soluções utilizando a estrutura tonal da música possuem a limitação da dependência da escala tonal da música para transposição, e que as soluções que propoe a utilização de estruturas independentes de transposição como o contorno melódico, apesar de possuírem esta vantagem , possuem baixo índice de acertos na pontuação de similaridade, bem como , além disto a atitude do músico em suas composições oferecem promissoras informações para a identificação de similaridade sem que a verificação nota por nota precise ser feita.

2.4 Conclusão da Análise

Levando em consideração os trabalhos realizados anteriormente e os conceitos musicais levantados para a análise do problema , fica aparente que, é necessário encontrar uma forma de cálculo de pontuação de similaridade que seja capaz de unir a estrutura tonal da música , com a estrutura invariante de contornos musicais, bem como o mapeamento da atitude do músico, para atingirmos o resultado esperado por este trabalho.

3 A Matriz MUSSIX (Mu6)

A matriz MUSSIX , também apelidada de Mu6 , surgiu no intuito de se obter uma matriz de parametrização do cálculo de similaridade capaz de retornar scores de músicas próximas da forma que o ser humano identifica a similaridade entre as músicas, ou seja , resultados mais "Musicais".

Inspirada na matriz MUSSUM ([MARTINIANO; N. Silla, 2017](#)) , a Mu6 procura expandir o cálculo de Match/Mismatch e Gaps baseados apenas nas notas musicais (realizado pela MUSSUM), e passa a utilizar outros aspectos da percepção musical, como Melodia , Harmonia e Ritmo.

Para isto, expandiu-se o conceito de notas musicais, para notas musicais Multi-Dimensionais, onde cada nota da música possui um espaço específico dentro do espaço musical contido em uma música.

Isto possibilitou aplicar os conceitos de contornos musicais durante o processamento de calculo do score de similaridade(Figura 8).

Figura 8 – A Matriz MUSSIX (Mu6).

Notas da Música						
	C	D	E	F	G	A
Pitch	60	62	64	65	67	69
Octave	5	5	5	5	5	5
Melodic Contour	0	2	2	1	2	2
Harmonic Contour	261.63	293.67	329.63	349.23	392.00	440.00
Rhythm Contour	0.66	0.33	0.023	0.23	0.54	0.20

Diagram labels: **Espaço da Música** (bracketed on the left), **Espaço da Nota** (bracketed on the bottom), **Notas da Música** (bracketed on the top).

Mussix = Music Space Matrix (Matriz de Espaço da Música)

Fonte: Autoria Própria

3.1 Cálculos da Mu6

O valor final de pontuação de similaridade é proveniente do alinhamento local realizado pelo algoritmo, e seu valor é diretamente influenciado pelos valores de Match/Mismatch e Penalidades (Gaps) aplicados durante o processamento.

Diferente do algoritmo original Smith-Waterman (SMITH; WATERMAN, 1981) com valores de Match/Mismatch e Gaps Fixos, bem como do calculo realizado pela MUSSUM (MARTINIANO; N. Silla, 2017) que utiliza a distancia em semi-tons entre as notas para a definição das penalidades, a matriz Mu6 utiliza o espaço determinado pelos valores de Pitch, Contorno Melódico e a Oitava da Nota para a definição do valor de Match/Mismatch, e utiliza os valores de contorno Harmônico e Rítmico para determinar a penalidade a ser atribuída.

3.1.1 Definição de Match/Mismatch da Mu6

Durante o processamento convolutivo do algoritmo de alinhamento local , é realizada uma comparação de Match/Mismatch para determinar a pontuação e direção a qual este valor foi originado (Esquerda , Diagonal ou Topo).

É neste momento em que é aplicado o calculo de match/mismatch da matriz Mu6, utilizando um cálculo de distância entre os componentes Pitch,Oitava e Contorno melódico da nota do Sinal A , com o Sinal B.

O cálculo de distância sendo aplicado atualmente é a distância Manhattan , onde considerando A e B como as notas multi-dimensionais sendo comparadas, x sendo considerada como o valor de pitch da nota, y sendo o valor da oitava da nota e z como sendo o contorno melódico da nota, temos especificado a equação 3.1.

$$d(A, B) = |Ax - Bx| + |Ay - By| + |Az - Bz| \quad (3.1)$$

Após calculado a distância entre as notas, é comparado se a distância é maior que zeros, caso a distância for zero temos um Match, caso contrário temos um Mismatch. O valor final atribuído ao resultado do Match e Mismatch é a Oitava da nota para um Match e o Valor da Oitava negativado em caso de um Mismatch.

3.1.2 Definição de Penalidades da Mu6

Ainda durante o processamento convolutivo do algoritmo de alinhamento local , nos casos em que ocorre um Mismatch um cálculo de penalidade é aplicado e atribuído aos parâmetros de funcionamento.

O cálculo de penalidade foi inspirado na reação que um músico tem ao errar uma nota em seu instrumento musical, onde ao errar uma nota, por reflexo tende-se a interromper a execução da nota.

Portanto, dificilmente um músico que cometa um erro de execução irá manter a nota sendo tocada por um longo período, ou seja, a penalidade precisa ser inversamente proporcional ao tempo decorrido da nota errada.

Outro detalhe importante, é a característica dos instrumentos musicais em adaptar-se às limitações fisiológicas do ser humano, referente a localização das oitavas, onde o erro ocorre frequentemente na mesma escala onde o contorno harmônico está sendo desenhado, e dificilmente um músico será capaz de errar uma nota que não esteja dentro da oitava atual.

Conceitualmente o cálculo de penalidades da Mu6 pode ser considerado como o valor normalizado do contorno Harmônico (dentro da escala atual), versus o tempo em que o erro ocorre. Como cada escala possui um intervalo fixo de frequências, foi possível calcular os valores mínimos e máximos da fórmula de normalização no mesmo momento em que a duração da nota ocorre.

Considerando $NormHz$ como a Frequência Normalizada, Hz como a frequência da nota executada, $MinHz$ a Frequência Mínima da Escala e $MaxHz$ a frequência máxima da escala temos a equação de Normalização 3.2.

$$NormHz = (Hz - MinHz) / (MaxHz - MinHz) \quad (3.2)$$

Figura 9 – Cálculo de Penalidades (Mu6).



Fonte: Autoria Própria

Consecutivamente considerando P como a Penalidade a ser aplicada, $NormHz$ como a Frequência previamente normalizada e Td como a duração em milisegundos da nota executada, temos a equação do cálculo de penalidades da Mu6 3.3.

$$P = NormHz/Td \quad (3.3)$$

3.2 O Algoritmo Mu6

Para o processamento do cálculo de pontuação de similaridade foram feitas adaptações no algoritmo de alinhamento Local (SW) para o processamento correto da Mu6.

Estas adaptações consistem na conversão de utilização de strings de notas musicais, para a utilização da Mu6 em cada trecho musical, e a inclusão do cálculo de distância e cálculo de penalidades.

Para isto a estrutura principal do algoritmo SW foi preservada, convertendo-se a estrutura em strings, para a matriz Mussix possibilitando que as notas multi-dimensionais sejam acessadas a cada iteração do algoritmo.

Algoritmo 7: O Algoritmo Mu6

Result: Matriz de Scores Calculada(Top-Down).

```

40 Inicialização;
41 D(i,j) = 0;
42 Calculo de Scores;
43 for i ← 1 to M do
44   for j ← 1 to N do
45     D(i,j) = Max(i,j);
46   end
47 end
```

A cada iteração as notas multi-dimensionais A e B são comparadas através do cálculo de distância Manhattan, utilizando os componentes Pitch, Oitava da Nota e Contorno Melódico.

Algoritmo 8: A Funcao Distancia

input : Notas MultiDimensionais A , B **output** : Distancia entre as Notas Multi-Dimensionais**Result:** O resultado sera a distancia Manhattan entre as notas.

```

48 tDistances[0] = A.Pitch - B.Pitch;
49 tDistances[1] = A.Contour - B.Contour;
50 tDistances[2] = A.Octave - B.Octave;
51 for  $i \leftarrow 1$  to 3 do
52   |  $tResultDistance += tDistances[i]$ ;
53 end
54 retorno = tResultDistance

```

Após o calculo de Similaridade entre as Notas, é aplicado então os valores de Match/Mismatch ou Penalidades de acordo com o definido pela análise conceitual de funcionamento da Matriz Mu6.

Algoritmo 9: A Funcao DistanciaSimilaridade

input : A posicao atual i, j **output** : Distancia de Similaridade**Result:** O resultado sera a distancia de similaridade entre as notas.

```

55  $A = Mu6A[i-1]$ ;
56  $B = Mu6B[j-1]$ ;
57 if  $Distancia(A, B) = 0$  then
58   |  $Acerto = A.Octave$ ;
59 else
60   |  $ScoredeErro = -A.Octave$ ;
61   |  $NormHz = (A.Frequency - MinHz) / (MaxHz - MinHz)$ ;
62   |  $gap = NormHz / A.Duration$ ;
63 end
64 retorno = Maximo( $topo, lado, diagonal$ )

```

3.3 O Attitude-Gram

Inspirado conceitualmente em identificar como a atitude de um musico pode ser traduzida em uma serie de características e tecnicamente pelo trabalho proposto por (CHUAN, 2013) , o Attitude-Gram é um vetor de características que é extraído da Mu6, no intuito de mapear a atitude do músico durante a performance musical e compará-la com a musica de comparação antes de realizar o alinhamento e calculo de pontuação de similaridade.

A idéia principal do Attitude-Gram é atribuir uma pontuação zerada, para toda música que estiver distante do Attitude-Gram, ou seja, somente músicas com a mesma atitude mapeada na musica sendo tocada será alvo de calculo de similaridade.

Com isto é possível otimizar e filtrar a lista final de músicas a serem retornadas pelo sistema ao Músico/Usuário.

No total foram extraídas oito (8) características da Mu6 com o objetivo principal de identificar similaridades entre a atitude empregada pelo músico no momento da composição musical.

Como Exemplo, podemos citar uma dupla de músicos que realizam uma performance com 2Cielos utilizando duas músicas vindas de músicos aparentemente sem relação nenhuma (5º Sinfonia de Bethoven e a música Whole Lotta Love da Banda de Rock Led Zeppelin) ilustrado na Figura 10.

Figura 10 – Attitude-Gram - Bethoven vs Led Zeppelin.



Fonte: <<https://www.youtube.com/watch?v=x8yymm3DtVA>>

No vídeo, podemos identificar diversas características que fazem com que as duas musicas sejam similares baseado principalmente na repetição de notas, ritmos e contorno harmônico, e foi baseado neste vídeo que serviu de inspiração para a definição das 8 características do Attitude-Gram.

3.3.1 As Características Attitude-Gram

No Total foram calculadas oito(8) características extraídas da Mu6 listadas na tabela 1.

Tabela 1 – Características do Attitude-Gram

Campo	Descrição
Oitavas	Número de Oitavas utilizada durante a performance do músico.
Bemois	Número de Bemóis utilizado durante a performance.
Média de Duração Notas	A média de duração de todas as notas da música.
Nota Mais Utilizada	A nota da Escala diatônica mais utilizada durante a performance.
Notas Únicas	Quantidade das notas da escala diatônica utilizadas durante a performance.
Repetições Consecutivas	A maior quantidade de repetições consecutivas de uma nota durante a performance.
BPM Médio	A quantidade de Batidas por Minuto Média da performance.
Atitude do Músico	Valor Composto das características anteriores representando a Atitude do músico durante a performance.

Fonte: Autoria Própria

O objetivo principal do Attitude-Gram é mapear a atitude do músico durante a performance musical ou sua composição, e para tal realiza uma análise rítmica da obra mapeando principalmente repetições de notas, a média de velocidade na execução da obra e notas mais utilizadas.

O valor final de atitude do músico é calculado através da divisão entre a duração média das notas e o número de repetições consecutivas da nota mais utilizada. Considerando Tm como a duração média das notas, Rt como a maior quantidade de repetições da nota base e Att como sendo a atitude do músico temos a seguinte equação 3.4

$$Att = Tm/Rt \quad (3.4)$$

Podem ocorrer casos em que não há repetições de notas mais usadas onde para se mapear a atitude do músico foi utilizado a quantidade de notas únicas da escala diatônica, portanto considerando Nu como a quantidade de notas únicas utilizadas na composição, Tm como a duração média das notas e Att como sendo a atitude do músico temos a seguinte equação 3.5

$$Att = Tm/Nu \quad (3.5)$$

Outra característica importante para a definição do Attitude-Gram é o cálculo de Batidas por Minuto Médio, onde é possível estabelecer uma área comum entre as frequentes transições no andamento da performance musical, que em conjunto com a atitude do músico

conseguem determinar a distancia de atitude entre uma música e outra. Considerando Tm como a duração média das notas e $MedBPM$ como sendo as batidas por minuto médias da performance chegou-se a equação 3.6.

$$MedBPM = 60000/Tm \quad (3.6)$$

Para se determinar se os músicos de cada composição, a da performance de entrada e da música pré-existente sendo consultada, um cálculo de distância Manhattan é realizado entre as oito características extraídas conforme o algoritmo de Cálculo de Distância Attitude-Gram.

Algoritmo 10: A Funcao Distancia Attitude-Gram

input : Attitude-Gram A , B

output : Distancia entre dois Attitude-Grams

Result: O resultado sera a distancia Manhattan entre as atitudes.

```

65 tDistances[0] = A.Octave - B.Octave;
66 tDistances[1] = A.Bemois - B.Bemois;
67 tDistances[2] = A.DurMed - B.DurMed;
68 tDistances[3] = A.RNote - B.RNote;
69 tDistances[4] = A.UNote - B.UNote;
70 tDistances[5] = A.CNote - B.CNote;
71 tDistances[6] = A.MedBPM - B.MedBPM;
72 tDistances[7] = A.Attitude - B.Attitude;
73 for  $i \leftarrow 0$  to 7 do
74   |  $tResultDistance \ += tDistances[i]$ ;
75 end
76 retorno = tResultDistance
```

4 Protocolo Experimental

Para execução do protocolo experimental foi desenvolvido um software chamado QBPLab, contendo um motor de busca que executa as 3 etapas de MIR, Rotinas automatizadas para a execução de buscas parametrizadas utilizando dados de controle, bem como foram implementados 3 algoritmos de alinhamento para o cálculo de pontuação de similaridade (SW , MUSSUM , Mussix(Mu6) e Attitude-Gram).

O objetivo principal deste protocolo experimental é possibilitar que as diversas técnicas de calculo de similaridade possuam o mesmo ambiente computacional para a coleta das métricas de avaliação.

4.1 O Sistema de Busca

Para orquestrar a execução da consulta ao dataset musical e realizar a etapa de processamento é necessário um sistema de busca para que estes itens computacionais gerem as metrcas de avaliação.

Neste trabalho foi desenvolvido um sistema de busca customizado para a execução do protocolo experimental e pode ser visualizado na figura 11.

A implementação da etapa de processamento foi realizada de forma que diversos algoritmos de processamento possam ser escolhidos sem que a estrutura principal de busca seja alterada.

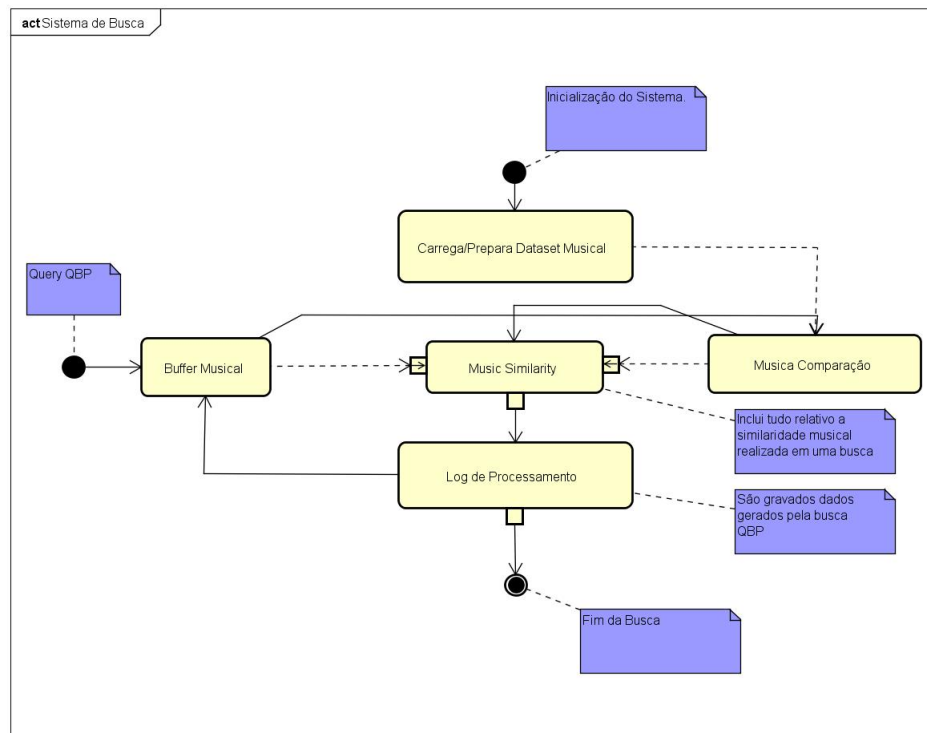
O sistema de busca implementado comporta as etapas de pré-Processamento, processamento e pós-processamento, gerando todas as informações necessárias para a análise de resultados desta pesquisa.

4.1.1 Pré-Processamento QBPLab

Para a realização deste trabalho está sendo utilizado uma lista de 10.198 músicas tradicionais irlandesas em formato MIDI utilizadas no artigo ([MARTINIANO; N. Silla, 2017](#)).

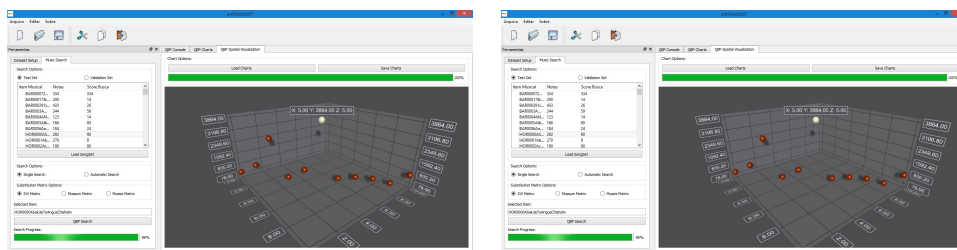
Todas as músicas apresentam-se separadas em folders de gênero musical , onde cada música apresenta uma média de 350 notas por música e uma média de 1000 músicas por gênero(Imagem 12). Esta lista de músicas é importada ao QBPLab , e na etapa de Pré-Processamento é criado um dataset musical customizado contendo a conversão das faixas MIDI e cálculo das notas musicais e matrizes utilizadas (MUSSUM e Mu6).

Figura 11 – Sistema de Busca.



Fonte: Autoria Própria

Figura 12 – Músicas por Gênero Musical: a) 3D e b) Folders



Fonte: Autoria própria

Nesta etapa também são geradas de acordo com a necessidade, listas randômicas de músicas utilizando parâmetros de controle como tamanho da amostragem musical.

4.1.2 Processamento QBPLab

Para o Processamento das Buscas, o QBPLab foi criado para possibilitar o processamento de buscas utilizando o algoritmo SW padrão, utilizando a matriz MUSSUM, utilizar a nova matriz Mu6 proposta por este trabalho e adicionalmente possui o processamento da Mu6 em conjunto com a filtragem Attitude-Gram.

Além de poder escolher o algoritmo a ser utilizado no experimento, foram criados alguns campos de controle de processamento para podermos avaliar a execução do modelo

e possuir melhores informações estatísticas de processamento, ou seja, estes campos são utilizados para determinar a classificação de cada musica durante a busca, e a quantidade de musicas que serão retornadas a cada busca.

4.1.2.1 O Balanço de Precisão

4.1.2.2 O Balanço de Tamanho de Resultado

4.1.3 Pós-Processamento QBPLab

O Pós-Processamento QBPLab é responsável em armazenar e persistir o histórico de execuções dos experimentos de pesquisa.

5 Avaliação dos Resultados

5.1 Experimentos

5.1.1 Experimento 1

5.1.2 Experimento 2

5.1.3 Experimento 3

5.1.4 Experimento 4

5.2 Resultados Obtidos

5.3 Análise dos Resultados

6 Considerações Finais

6.1 Relevância do Estudo

6.2 Contribuições da Pesquisa

6.3 Limitações da Pesquisa

6.4 Trabalhos Futuros

Nesta pesquisa exploratória foi possível explorar todas as três etapas de uma pesquisa em MIR , estudar as técnicas atuais de processamento de similaridade musical e foram geradas contribuições científicas não somente no nível de teoria, mas também um extenso trabalho que contribui na metodologia empregada ao estudo de computação musical e contribuições significativas em nível de aplicação.

Os resultados obtidos nos experimentos de pesquisa foram bastante expressivos e serão utilizados na realização de trabalhos futuros, e consecutivamente gerando novas contribuições científicas.

Estima-se que proveniente deste trabalho surjam em torno de cinco novos estudos de MIR e processamento de áudio/música envolvendo os tópicos listados abaixo:

- Utilização de programação heterogênea para aceleração em GPU.
- Separação de instrumentos musicais em faixas polifônicas.
- Mixagem e Posicionamento Espacial de Áudio.
- Melhorias no Buffer Musical(Inclusão de Outras Estruturas como chromagrama e histograma).
- sincronização de múltiplos instrumentos em faixas polifônicas simbolicas.

Sem sombra de dúvidas este trabalho de pesquisa abriu precedente para a extensão de estudos mais longos e aprofundados onde espera-se obter também resultados expressivos para a comunidade científica de um modo geral.

Há muito trabalho de pesquisa ainda a ser realizado e com certeza as contribuições científicas geradas nesta pesquisa servirão como uma base sólida para a realização destas novas descobertas.

Referências

CHUAN, C. H. A multimodal approach to song-level style identification in pop/rock using similarity metrics. *2013 12th International Conference on Machine Learning and Applications*, v. 2, p. 321–324, 2013. Disponível em: <http://ieeexplore.ieee.org/document/6786128/>. Citado 3 vezes nas páginas 6, 16 e 21.

DEGANI, A. et al. A Heuristic for Distance Fusion in Cover Song Identification. *2013 (WIAMIS)*, 2013. Citado na página 6.

DOWNIE, J. S. et al. The music information retrieval evaluation eXchange: Some observations and insights. *Studies in Computational Intelligence*, v. 274, p. 93–115, 2010. ISSN 1860949X. Citado na página 11.

DUREY, A.; CLEMENTS, M. Direct estimation of musical pitch contour from audio data. *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03)*, v. 5, p. V–561–4, 2003. ISSN 1520-6149. Disponível em: <http://ieeexplore.ieee.org/document/1200031/>. Citado na página 14.

FERRARO, P.; HANNA, P. Optimizations of local edition for evaluating similarity between monophonic musical sequences. *Large Scale Semantic Access to Content (Text, Image, Video, and Sound)*, n. 1970, p. 64–69, 2007. Citado 4 vezes nas páginas 7, 9, 11 e 15.

Kuldeep Gurjar, Y.-S. M. A comparative analysis of music similarity measures in music information retrieval systems. *Journal of Information Processing Systems · January 2018*, n. January, 2018. ISSN 2092805X. Citado na página 4.

LI, B.; DUAN, Z. An Approach to Score Following for Piano Performances with the Sustained Effect. *IEEE/ACM Transactions on Audio Speech and Language Processing*, v. 24, n. 12, p. 2425–2438, 2016. ISSN 23299290. Citado na página 5.

MARTINIANO, L.; SILLA, C. BIRITS: A Music Information Retrieval System Using Query-by-Playing Techniques. In: *2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI)*. IEEE, 2017. p. 535–542. ISBN 978-1-5386-3876-7. Disponível em: <https://ieeexplore.ieee.org/document/8371990/>. Citado 7 vezes nas páginas 4, 5, 6, 13, 17, 18 e 25.

MARVIN, E. W. A generalized theory of musical contour: its application to melodic and rhythmic analysis of non-tonal music and its perceptual and pedagogical implications. 1988. Disponível em: <http://www.mendeley.com/research/a-generalized-theory-of-musical-contour-its-application-to-melodic-and-rhythmic-analysis-of-nontonal> >. Citado 2 vezes nas páginas 14 e 15.

NEEDLEMAN, S. B.; WUNSCH, C. D. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, v. 48, n. 3, p. 443–453, 1970. ISSN 00222836. Citado na página 7.

PARK, M. W.; LEE, E. C. Similarity measurement method between two songs by using the conditional euclidean distance. *WSEAS Transactions on Information Science and Applications*, v. 10, n. 12, p. 381–388, 2013. ISSN 17900832. Citado na página 5.

SMITH, T.; WATERMAN, M. Identification of common molecular subsequences. *Molecular Biology*, v. 147, p. 195–197, 1981. ISSN 0022-2836. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/7265238>. Citado 2 vezes nas páginas 9 e 18.

WANG, T.-M.; TSAI, P.-Y.; SU, A. W.-Y. Note-based alignment using score-driven non-negative matrix factorisation for audio recordings. *IET Signal Processing*, v. 8, n. 1, p. 1–9, 2014. ISSN 17519675. Disponível em: <http://digital-library.theiet.org/content/journals/10.1049/iet-spr.2012.0157>. Citado na página 11.