

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
INSTITUTO DE COMPUTAÇÃO
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

RODRIGO CÔRTEZ NOGUEIRA DA ROCHA PITA

IDENTIFICAÇÃO DE ACORDES VIA MODELOS DE MARKOV OCULTOS
MUSICALMENTE INFORMADOS

RIO DE JANEIRO
2024

RODRIGO CÔRTEZ NOGUEIRA DA ROCHA PITA

IDENTIFICAÇÃO DE ACORDES VIA MODELOS DE MARKOV OCULTOS
MUSICALMENTE INFORMADOS

Trabalho de conclusão de curso de graduação apresentado ao Departamento de Ciência da Computação da Universidade Federal do Rio de Janeiro como parte dos requisitos para obtenção do grau de Bacharel em Ciência da Computação.

Orientador: Prof. João Antônio Recio da Paixão
Co-orientador: Prof. Hugo Tremonte de Carvalho

RIO DE JANEIRO

2024

CIP - Catalogação na Publicação

P681i Pita, Rodrigo Côrtes Nogueira da Rocha
Identificação de acordes via modelos de Markov
ocultos musicalmente informados / Rodrigo Côrtes
Nogueira da Rocha Pita. -- Rio de Janeiro, 2024.
97 f.

Orientador: João Antônio Recio da Paixão.
Coorientador: Hugo Tremonte de Carvalho.
Trabalho de conclusão de curso (graduação) -
Universidade Federal do Rio de Janeiro, Instituto
de Computação, Bacharel em Ciência da Computação,
2024.

1. Modelos de Markov. 2. Música. 3. Identificação
automática de acordes. I. Paixão, João Antônio Recio
da, orient. II. Carvalho, Hugo Tremonte de,
coorient. III. Título.


RODRIGO CÔRTEZ NOGUEIRA DA ROCHA PITA

IDENTIFICAÇÃO DE ACORDES VIA MODELOS DE MARKOV OCULTOS
MUSICALMENTE INFORMADOS

Trabalho de conclusão de curso de graduação
apresentado ao Departamento de Ciência da
Computação da Universidade Federal do Rio
de Janeiro como parte dos requisitos para ob-
tenção do grau de Bacharel em Ciência da
Computação.

Aprovado em 12 de março de 2024

BANCA EXAMINADORA:

Documento assinado digitalmente
 JOAO ANTONIO RECIO DA PAIXAO
Data: 05/04/2024 13:51:39-0300
Verifique em <https://validar.iti.gov.br>

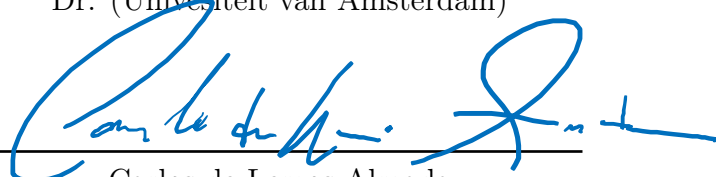
João Antônio Recio da Paixão
Dr. (PUC-RJ)



Hugo Tremonte de Carvalho
Dr. (UFRJ)



Hugo de Holanda Cunha Nobrega
Dr. (Univesiteit van Amsterdam)



Carlos de Lemos Almada
Dr. (UNIRIO)

Dedico este projeto à minha família, em especial minha mãe Mariana Nogueira, meu pai Sergio Pita e irmã Carolina Pita que sempre estiveram presentes me apoiando e incentivando em todas as etapas da minha vida. Gostaria também de aproveitar para agradecer a todos os meus amigos que me ajudaram, seja academicamente ou apenas estando por perto, principalmente os grupos *Prog* e *Winx*.

AGRADECIMENTOS

Em primeiro lugar gostaria de agradecer aos meus orientadores João Paixão e Hugo Carvalho por acreditar no meu potencial desde o início e me incentivar a seguir com este projeto que começou como um sonho distante. Além disso, preciso agradecer ao Carlos Almada, que junto com Hugo Carvalho coordena o Projeto MPB que contribuiu fortemente no desenvolvimento do projeto.

Quero também agradecer ao Roberto Stelling e Adriana Vivacqua por serem meus orientadores na minha Iniciação Científica, de onde nasceu a ideia para este projeto.

Por fim, preciso agradecer ao Eduardo Gonçalves, meu atual gestor na Hashdex, por ter me dado todo o suporte e flexibilidade para eu conseguir balancear o trabalho e o desenvolvimento deste projeto.

*"I realized by using the high notes of the chords
as a melodic line, and by the right harmonic progression,
I could play what I heard inside me."*

Charlie Parker

RESUMO

Este projeto visa abordar o problema da identificação de acordes através de modelos de Markov ocultos (HMM) musicalmente informados. Para que seja um projeto autocon-tido, serão explicados três grandes tópicos. Primeiramente, as noções básicas de teoria musical, como intervalos e progressão harmônica. Em segundo lugar, o processamento de sinais de áudio, abordando transformadas de Fourier para em seguida poder apresentar a transformada de Fourier de tempo curto (STFT) e também cromagramas, que empenham um papel fundamental para o funcionamento do modelo HMM. Por fim, em terceiro, as cadeias de Markov, que são a base para o entendimento de HMMs. Para embutir a infor-mação musical no modelo, será utilizada uma matriz de transição gerada a partir de uma base de músicas da MPB. Os experimentos realizados serão feitos inicialmente num con-texto geral, com músicas de gêneros diversos com o intuito de comparar os modelos HMM e *template-based* assim como os diferentes tipos de cromagrama. Depois serão realizados somente no contexto da MPB, onde os resultados provarão a eficácia da proposta.

Palavras-chave: cadeias de Markov; HMM; processamento de sinais musicais; MPB.

ABSTRACT

This project aims to tackle the problem of identifying chords using musically informed hidden Markov models (HMM). To make it a self-contained project, three main topics will be explained. Firstly, the basics of music theory, such as intervals and harmonic progression. Secondly, audio signal processing, covering Fourier transforms and then the short-time Fourier transform (STFT), as well as chromagrams, which play a fundamental role in the functioning of the HMM model. Thirdly, Markov chains, which are the basis for understanding HMMs. To incorporate musical information into the model, a transition matrix generated from a database of MPB (Brazilian Popular Music, acronym in Portuguese) songs will be used. The experiments will initially be carried out in a general context, with songs from different genres in order to compare the HMM and *template-based* models as well as the different types of chromagram. They will then be restricted to the MPB context, where the results will prove the effectiveness of the proposal.

Keywords: Markov chains; HMM; music signal processing; MPB.

LISTA DE ILUSTRAÇÕES

Figura 1 – (a) Círculo Cromático, com a escala diatônica representada pelas setas vermelhas. (b) Hélice Tonal (SHEPARD, 1964).	16
Figura 2 – Teclado de piano com oitavas numeradas e uma em destaque.	16
Figura 3 – Tríades básicas. (a) Dó Maior. (b) Dó Aumentado. (c) Dó Menor. (d) Dó Diminuto.	17
Figura 4 – Tétrades. (a) Dó com Sétima Maior. (b) Dó com Sétima. (c) Dó Menor com Sétima. (d) Dó Menor com Sétima Maior. (e) Dó Diminuto com Sétima. (f) Dó Meio Diminuto. (g) Dó com Sétima e Quinta Bemol. (h) Dó com Sétima e Quinta Aumentada.	18
Figura 5 – Esquema de correlação das forças tonais. T - tônica, D - dominante, SD - subdominante.	19
Figura 6 – Módulo da Transformada de Fourier de um trecho da música <i>Teresinha</i> , de Chico Buarque.	21
Figura 7 – STFT de um trecho da música <i>Teresinha</i> de Chico Buarque.	23
Figura 8 – Cromagrama de um trecho da música <i>Teresinha</i> de Chico Buarque.	25
Figura 9 – Cromagramas gerados para um violino tocando a Escala de Lá maior.	26
Figura 10 – Ilustração de uma cadeia de Markov com três estados. (a) Grafo da cadeia de Markov. (b) Matriz de transição da cadeia de Markov.	28
Figura 11 – Modelo Oculto de Markov no contexto de reconhecimento de acordes.	29
Figura 12 – Etapas do processo de reconhecimento baseado em modelos.	32
Figura 13 – Vetores de modelos binários para as 24 tríades maiores e menores. Células em branco têm valor zero, e células em vermelho têm valor um. Cada coluna corresponde ao acorde ilustrado na partitura acima.	33
Figura 14 – Resultado do reconhecimento baseado em modelos para um trecho da música " <i>Teresinha</i> " de Chico Buarque.	34
Figura 15 – Etapas do processo de reconhecimento baseado em HMM.	35
Figura 16 – Matriz de transição invariante à transposição. A matriz foi computada usando um conjunto de dados dos Beatles.	36
Figura 17 – Matriz uniforme de transição com probabilidade de transição ao mesmo estado $p = 0.15$	37
Figura 18 – Matriz de triádica alterada. (a) Matriz de transição. (b) Matriz de transição com log probabilidade.	38
Figura 19 – Matriz de emissão por IIR para um trecho da música <i>Teresinha</i> de Chico Buarque.	40
Figura 20 – Cromagramas gerados para a música <i>When I was your man</i>	42

Figura 21 – Resultados do algoritmo <i>template-based</i> para a música <i>When I was your man</i>	42
Figura 22 – Resultados do HMM para a música <i>When I was your man</i>	43
Figura 23 – Resultados para a música <i>Até segunda-feira</i>	45
Figura 24 – Cromagramas gerados para a música de Carol Pita: <i>Um cafuné e um violão</i>	46
Figura 25 – Resultados do reconhecimento de acordes com a matriz uniforme para a música de Carol Pita: <i>Um cafuné e um violão</i>	47
Figura 26 – Resultados do reconhecimento de acordes com a matriz MPB 24×24 para a música de Carol Pita: <i>Um cafuné e um violão</i>	48
Figura 27 – Cromagramas gerados para a música <i>everything i wanted</i>	52
Figura 28 – Cromagramas gerados para a música <i>Attention</i>	52
Figura 29 – Cromagramas gerados para a música <i>Teresinha</i>	52
Figura 30 – Cromagramas gerados para a música <i>The scientist</i>	53
Figura 31 – Cromagramas gerados para a música <i>Best part</i>	53
Figura 32 – Cromagramas gerados para a música <i>Your song</i>	53
Figura 33 – Cromagramas gerados para a música <i>Take me out</i>	53
Figura 34 – Cromagramas gerados para a música <i>Somewhere over the rainbow</i>	54
Figura 35 – Cromagramas gerados para a música <i>Maybe I'm amazed</i>	54
Figura 36 – Cromagramas gerados para a música <i>Stitches</i>	54
Figura 37 – Cromagramas gerados para a música <i>Pride and joy</i>	54
Figura 38 – Cromagramas gerados para a música <i>House of the rising sun</i>	55
Figura 39 – Resultados do Template-based para a música <i>everything i wanted</i>	56
Figura 40 – Resultados do HMM para a música <i>everything i wanted</i>	57
Figura 41 – Resultados do Template-based para a música <i>Attention</i>	58
Figura 42 – Resultados do HMM para a música <i>Attention</i>	59
Figura 43 – Resultados do Template-based para a música <i>Teresinha</i>	60
Figura 44 – Resultados do HMM para a música <i>Teresinha</i>	61
Figura 45 – Resultados do Template-based para a música <i>The scientist</i>	62
Figura 46 – Resultados do HMM para a música <i>The scientist</i>	63
Figura 47 – Resultados do Template-based para a música <i>Best part</i>	64
Figura 48 – Resultados do HMM para a música <i>Best part</i>	65
Figura 49 – Resultados do Template-based para a música <i>Your song</i>	66
Figura 50 – Resultados do HMM para a música <i>Your song</i>	67
Figura 51 – Resultados do Template-based para a música <i>Take me out</i>	68
Figura 52 – Resultados do HMM para a música <i>Take me out</i>	69
Figura 53 – Resultados do Template-based para a música <i>Somewhere over the rainbow</i>	70
Figura 54 – Resultados do HMM para a música <i>Somewhere over the rainbow</i>	71

Figura 55 – Resultados do Template-based para a música <i>Maybe I'm amazed</i>	72
Figura 56 – Resultados do HMM para a música <i>Maybe I'm amazed</i>	73
Figura 57 – Resultados do Template-based para a música <i>Stitches</i>	74
Figura 58 – Resultados do HMM para a música <i>Stitches</i>	75
Figura 59 – Resultados do Template-based para a música <i>Pride and joy</i>	76
Figura 60 – Resultados do HMM para a música <i>Pride and joy</i>	77
Figura 61 – Resultados do Template-based para a música <i>House of the rising sun</i> . .	78
Figura 62 – Resultados do HMM para a música <i>House of the rising sun</i>	79
Figura 63 – Cromagramas gerados para as músicas <i>Comeu</i> e <i>Deixa sangrar</i>	80
Figura 64 – Cromagramas gerados para as músicas <i>Até segunda-feira</i> e <i>Juca</i>	80
Figura 65 – Cromagramas gerados para as músicas <i>Atrás Poeira</i> e <i>Começar de novo</i> . .	81
Figura 66 – Cromagramas gerados para as músicas <i>As Praias Desertas</i> e <i>Aula de matemática</i>	81
Figura 67 – Resultados com a Matriz Uniforme para a música <i>Comeu</i>	82
Figura 68 – Resultados com a Matriz Triádica Alterada para a música <i>Comeu</i>	83
Figura 69 – Resultados com a Matriz Uniforme para a música <i>Deixa sangrar</i>	84
Figura 70 – Resultados com a Matriz Triádica Alterada para a música <i>Deixa sangrar</i> . .	85
Figura 71 – Resultados com a Matriz Uniforme para a música <i>Até segunda-feira</i> . . .	86
Figura 72 – Resultados com a Matriz Triádica Alterada para a música <i>Até segunda-feira</i>	87
Figura 73 – Resultados com a Matriz Uniforme para a música <i>Juca</i>	88
Figura 74 – Resultados com a Matriz Triádica Alterada para a música <i>Juca</i>	89
Figura 75 – Resultados com a Matriz Uniforme para a música <i>Atrás poeira</i>	90
Figura 76 – Resultados com a Matriz Triádica Alterada para a música <i>Atrás poeira</i> . .	91
Figura 77 – Resultados com a Matriz Uniforme para a música <i>Começar de novo</i> . . .	92
Figura 78 – Resultados com a Matriz Triádica Alterada para a música <i>Começar de novo</i>	93
Figura 79 – Resultados com a Matriz Uniforme para a música <i>As praias desertas</i> . .	94
Figura 80 – Resultados com a Matriz Triádica Alterada para a música <i>As praias desertas</i>	95
Figura 81 – Resultados com a Matriz Uniforme para a música <i>Aula de matemática</i> . .	96
Figura 82 – Resultados com a Matriz Triádica Alterada para a música <i>Aula de matemática</i>	97

SUMÁRIO

1	INTRODUÇÃO	13
2	TEORIA MUSICAL	14
2.1	NOTAS E INTERVALOS	14
2.2	ESCALAS DIATÔNICAS E CROMÁTICA	15
2.3	FUNDAMENTOS DE HARMONIA E TEORIA DOS ACORDES	16
2.3.1	Tríades	17
2.3.2	Inversões	17
2.3.3	Tétrades	18
2.4	PROGRESSÃO DE ACORDES	18
2.4.1	Progressões harmônicas comuns	19
3	PROCESSAMENTO DE SINAIS	21
3.1	TRANSFORMADA DE FOURIER	21
3.2	TRANSFORMADA DISCRETA DE FOURIER (DFT)	22
3.3	TRANSFORMADA DE FOURIER DE TEMPO CURTO (STFT)	23
3.4	CROMAGRAMA	24
3.4.1	Funcionamento	25
4	MODELOS DE MARKOV	27
4.1	CADEIA DE MARKOV	27
4.2	MODELO OCULTO DE MARKOV (HMM)	28
4.2.1	Problema da Descoberta	29
5	APLICAÇÃO A RECONHECIMENTO DE ACORDES	32
5.1	BASEADO EM MODELOS (<i>TEMPLATE-BASED</i>)	32
5.1.1	Procedimento	32
5.1.2	Avaliação dos resultados	34
5.2	BASEADO EM HMM (<i>HMM-BASED</i>)	35
6	REFINAMENTO DAS MATRIZES DO HMM	37
6.1	MATRIZ DE TRANSIÇÃO	37
6.1.1	Matriz Triádica Uniforme	37
6.1.2	Matriz Triádica Alterada	37
6.1.3	Matriz Tetrádica	38
6.2	MATRIZ DE EMISSÃO	39

7	APLICAÇÕES À MPB	41
7.1	EXPERIMENTOS PRELIMINARES	41
7.2	EXPERIMENTOS NO CONTEXTO DA MPB	44
7.3	EXEMPLO-SUMÁRIO	45
8	CONCLUSÃO E PROJETOS FUTUROS	49
	REFERÊNCIAS	51
	APÊNDICE A – CROMAGRAMAS GERADOS PARA OS EX- PERIMENTOS PRELIMINARES.	52
	APÊNDICE B – RESULTADOS DOS EXPERIMENTOS PRELI- MINARES.	56
	APÊNDICE C – CROMAGRAMAS GERADOS PARA OS EX- PERIMENTOS NO CONTEXTO DA MPB.	80
	APÊNDICE D – RESULTADOS DOS EXPERIMENTOS NO CON- TEXTO DA MPB.	82

1 INTRODUÇÃO

Este projeto tem por objetivo abordar o problema da identificação automática de acordes. Para ilustrar melhor a motivação por trás desse projeto, um exemplo claro é o de uma pessoa querendo aprender uma certa música, mas a cifra da música não se encontra em nenhum dos veículos principais, obrigando a pessoa a usar o próprio ouvido para descobrir os acordes presentes. A proposta do projeto é solucionar o problema por meio de Modelos de Markov Ocultos (HMM - Seção 4.2) musicalmente informados¹, como sugerido em (MÜLLER, 2015, Cap. 5), a partir de uma base de dados formada pelo Projeto MPB (ALMADA; CARVALHO, 2024).

Para melhor entender o projeto, será necessário passar por uma breve contextualização de teoria musical no Capítulo 2. Em seguida, no Capítulo 3 serão apresentadas as noções de processamento de sinais, como a de transformadas de Fourier, espectrogramas e cromagramas, que servirão de ferramentas para a construção de um HMM. Finalmente no Capítulo 4 será introduzido o conceito de Cadeias de Markov e matrizes de transição, conseqüentemente de HMMs também.

Após esse embasamento teórico, com o Capítulo 5 será formalizada a ideia de aplicação dos conceitos apresentados num contexto de reconhecimento de acordes. Esse capítulo irá explorar duas possíveis abordagens para o problema: *template-based* e por HMM. Além disso informará como é feita a avaliação dos resultados. No Capítulo 6 será explorado o refinamento das matrizes do HMM como forma de especificar a estética harmônica desejada na identificação. O refinamento será responsável por informar musicalmente o modelo através de dados de progressões musicais da MPB.

No Capítulo 7, num primeiro momento, serão apresentados os resultados dos experimentos preliminares, com músicas de diversos gêneros e contendo diversas progressões harmônicas. Em seguida, serão apresentados os resultados de experimentos de identificação no contexto da MPB, utilizando as matrizes refinadas (Capítulo 6). Por fim, será feito um último experimento que servirá para sumarizar os resultados.

Códigos utilizados nos experimentos foram tirados de (MÜLLER, 2015) e aperfeiçoados para as necessidades do projeto. Todos se encontram em (PITA, 2024).

¹ Leia-se musicalmente informados os modelos em que para gerar a matriz de transição, utilizou-se dados reais de transições entre acordes presentes em uma base musical selecionada, como forma de contextualizar o modelo em um estilo musical específico.

2 TEORIA MUSICAL

Este capítulo é dedicado a apresentar brevemente os conceitos básicos da teoria musical, essenciais para entender o projeto em sua totalidade. Como este trabalho está intimamente ligado à música, termos específicos da teoria musical serão frequentemente utilizados nos capítulos seguintes. Estes termos serão explicados brevemente nas seções subsequentes. Contudo, esse capítulo não tem a intenção de ser extensivo a ponto de ser um texto introdutório. Para uma referência nessa direção, veja (MED, 1996; ALMADA, 2009).

2.1 NOTAS E INTERVALOS

Para fins de convenção, este capítulo será baseado no sistema *de temperamento igual*, sistema musical predominante no Ocidente. Nele, um **tom**¹ é caracterizado por três qualidades sonoras: altura, volume e timbre. Quando combinado com ataque e duração, o resultado é uma **nota**. Ou seja, uma nota é um tom em um contexto temporal.

Frequência é a medida física de vibrações por segundo de um determinado corpo. **Altura**, por sua vez é um conceito mais abstrato, dando-se pela percepção auditiva de uma frequência. Na afinação de instrumentos musicais utiliza-se o *diapasão*, um instrumento metálico que emite um tom em determinada frequência. O mais comum é encontrar um diapasão que emita a frequência de 440 Hz (LOY, 2006, Cap. 2). Ao usá-lo, a altura que escutamos é convencionalmente chamada de A440, ou seja, a nota Lá₃ (como será explicado abaixo) de frequência 440 Hz.

O **semitom** é a menor distância (**intervalo**) entre duas notas musicais no sistema de temperamento igual (BRAGA, 2020). Existem outros sistemas que exploram outras distâncias (dito *microtonalidade*), porém isso foge do escopo deste projeto. Existem dois tipos de intervalos melódicos²: *ascendentes* e *descendentes*. O primeiro mede a distância a partir da nota mais grave para a mais aguda, enquanto o segundo faz o oposto. Para simplificar, a presente introdução focará apenas em intervalos ascendentes.

O intervalo mais básico na música é a **oitava**, definida como a distância entre uma altura e outra com a metade ou o dobro de sua frequência fundamental. O A880, por exemplo, é a oitava acima do A440. Uma oitava é composta por 12 semitons, cujas frequências são separadas por múltiplos inteiros de $\sqrt[12]{2}$. Os principais intervalos ascendentes podem ser encontrados na Tabela 1, partindo da classe de altura Dó. Existem intervalos maiores que a oitava, entretanto como será explicado na Seção 2.2, qualquer intervalo maior do que uma oitava, intervalo composto, pode ser mapeado para um intervalo simples. Vale

¹ Neste projeto, tom procura traduzir a palavra "tone" do inglês.

² Além dos intervalos melódicos, onde uma nota é tocada após a outra, existem também os harmônicos, onde as notas são tocadas simultaneamente.

notar que alguns intervalos, apesar de apresentarem a mesma distância em semitons, têm notações distintas. Foi assumida a **equivalência enarmônica** (WRIGHT, 2009) como simplificação necessária para a efetivação do estudo das bases consideradas ao longo do projeto. Enarmonia diz respeito à alturas diferentes que apresentam o mesmo som, como o caso de $Dó\sharp$ e $Réb$.

Tabela 1 – Tabela dos intervalos.

Fonte: Elaboração própria.

Semitons	Intervalos	Nota relativa a $Dó$
1	Segunda Menor	$Dó\sharp$ ou $Réb$
2	Segunda Maior	$Ré$
3	Terça Menor	$Ré\sharp$ ou Mib
4	Terça Maior	Mi
5	Quarta Justa	$Fá$
6	Trítono	$Fá\sharp$ ou $Solb$
7	Quinta Justa	Sol
8	Quinta Aumentada	$Sol\sharp$ ou $Láb$
9	Sexta Maior	$Lá$
10	Sétima Menor	$Lá\sharp$ ou Sib
11	Sétima Maior	Si
12	Oitava	$Dó$

2.2 ESCALAS DIATÔNICAS E CROMÁTICA

O protótipo da maioria das escalas no sistema de temperamento igual é a **escala diatônica maior** (LOY, 2006, Cap. 3). Exemplificando com a escala diatônica de $Dó$, ela apresenta sete classes de alturas por oitava, nomeadas com as sete letras C, D, E, F, G, A e B correspondentes aos sete graus da escala ($Dó$, $Ré$, Mi , $Fá$, Sol , $Lá$ e Si respectivamente, notas das teclas brancas do piano).

Como a escala diatônica apresentada acima é a escala maior de $Dó$, esta é apenas uma das 12 variações possíveis: desde que os intervalos entre os graus da escala sejam respeitados, basta girar o polígono formado pelas setas vermelhas na Figura 1a, alterando a nota inicial, para encontrar as outras escalas diatônicas maiores.

Além de representar o intervalo mais básico, o termo “oitava” também pode ser compreendido como o grupo das 12 classes de alturas presentes na **escala cromática**, como demonstrado pela Figura 2. Uma possível analogia matemática para a oitava é a do grupo dos inteiros módulo 12, onde cada classe de equivalência é mapeada na respectiva nota da oitava, sendo $C = 0$, $C\sharp = 1$, \dots , $B = 11$, de forma que ao chegar no 12, a sequência se repete, uma oitava acima. As Figuras 1a e 1b ilustram bem esta ideia (SHEPARD, 1964). Com este sistema em mente, não é mais necessário usar a frequência para diferenciar duas alturas diferentes de mesmo nome, no lugar podendo ser usada a posição da oitava em

relação ao piano, como destacado na Figura 2 com as notas C precedidas do índice de sua oitava.

Tal qual no grupo dos inteiros módulo 12, podemos considerar classes de equivalência de notas, entidades chamadas de **classes de alturas**: de certo ponto de vista, todas as notas, digamos, Dó são equivalentes pois todas têm frequência que são múltiplas de uma frequência de referência, sendo o fator multiplicativo 2^n , para algum $n \in \mathbb{Z}$. Em determinados contextos musicais faz sentido considerar apenas a classe de altura, enquanto que em outros, é necessário levar em consideração a altura específica.

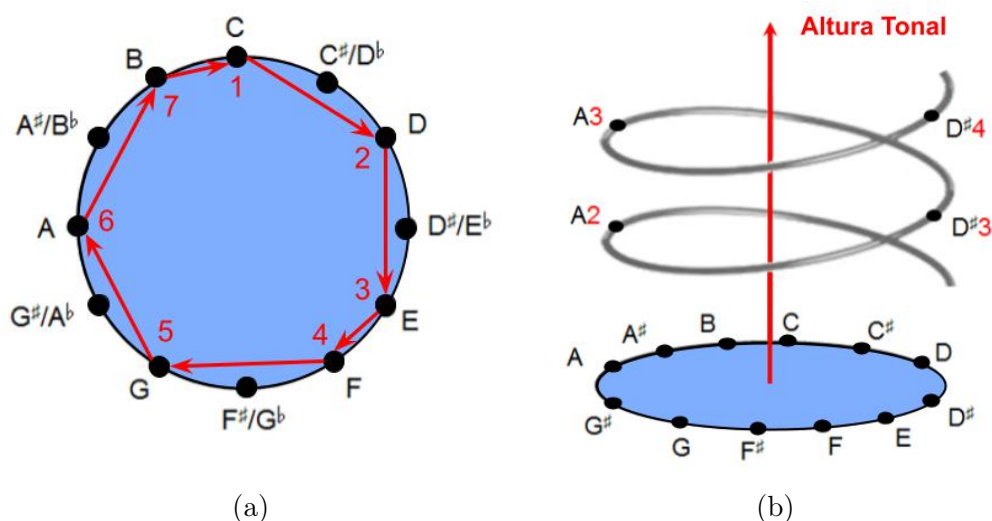


Figura 1 – (a) Círculo Cromático, com a escala diatônica representada pelas setas vermelhas. (b) Hélice Tonal (SHEPARD, 1964).

Fonte: Adaptado da Figura 1.3 de (MÜLLER, 2015).



Figura 2 – Teclado de piano com oitavas numeradas e uma em destaque.

Fonte: Elaboração própria a partir do teclado de piano do Musecore 4.

2.3 FUNDAMENTOS DE HARMONIA E TEORIA DOS ACORDES

A teoria da harmonia é um pilar fundamental da música, particularmente no contexto da tradição ocidental. A **harmonia**³ se refere à combinação de notas diferentes que, quando tocadas ou percebidas simultaneamente, formam uma entidade coesa na mente do ouvinte. Os componentes centrais da harmonia são os **acordes**, que são determinados conjuntos de notas tocadas ou percebidas conjuntamente e são essenciais para a construção musical, sua progressão e interação.

³ Para um contexto historiográfico veja (RAMEAU, 1722), o primeiro livro sobre teoria da harmonia.

Os acordes são definidos como grupos de várias notas que soam simultaneamente. Geralmente, um acorde é composto por pelo menos três notas, sendo um acorde de exatamente três notas dito uma **tríade**. Porém há situações em que acordes apresentam apenas duas notas (**díades**), quatro notas (**tétrades**), ou mais. A análise harmônica muitas vezes considera classes de alturas em vez de notas distintas, devido ao fato de que notas em oitavas diferentes são percebidas como tendo a mesma "qualidade de som".

2.3.1 Tríades

As tríades podem ser divididas em quatro categorias, **maiores**, **menores**, **diminutas** e **aumentadas**. Independente da categoria, uma tríade é composta de três notas: uma **tônica**, uma **terça** e uma **quinta**. A tônica especificará o nome do acorde, enquanto a terça e quinta definem sua qualidade. Para as tríades perfeitas (maiores e menores), utiliza-se a quinta justa, enquanto a terça definirá se o acorde será maior, com uma terça maior, ou menor, com uma terça menor (consultar a Tabela 1).

A quinta, porém, pode aparecer de outras duas formas: diminuta (trítono), que quando acompanhada de uma tônica e uma terça menor forma a tríade diminuta, ou aumentada, que acompanhada de uma tônica e uma terça maior forma a tríade aumentada. A Figura 3 ilustra no piano as tríades apresentadas.

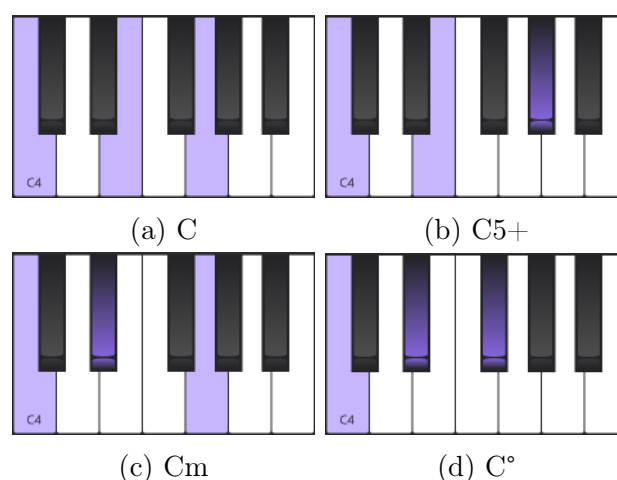


Figura 3 – Tríades básicas. (a) Dó Maior. (b) Dó Aumentado. (c) Dó Menor. (d) Dó Diminuto.

Fonte: Elaboração própria a partir do teclado de piano do Musecore 4.

2.3.2 Inversões

Inversão é a ação de alterar a ordem em que as notas de um acorde são dispostas. Para tríades existem três estados possíveis para um acorde: o estado fundamental, primeira inversão e segunda inversão.

No estado fundamental, as notas aparecem na ordem trivial, como explicitado acima (Seção 2.3.1). Já na primeira inversão, a terça se encontra na posição de baixo, ou seja, a mais grave. O mesmo ocorre para a segunda inversão, dessa vez, com a quinta.

2.3.3 Tétrades

Assim como mencionado na introdução desta seção, as **tétrades** são formadas por quatro notas. As três primeiras notas vêm das tríades, e a quarta nota é uma terça adicionada à tríade, ou seja, um intervalo de terça da quinta, o que resulta numa sétima em relação à tônica (consultar a Tabela 1).

Pelo fato de as tétrades herdarem a estrutura das tríades, as mesmas regras de qualidade se aplicam. Por conta disto, as tétrades também podem ser maiores, menores, diminutas ou aumentadas. Onde para as maiores existem acordes com sétima maior e sétima menor (esse último caso dito *acordes dominantes*); enquanto para menores, existem acordes menores com sétima menor ou maior, acordes diminutos com sétima menor e acordes diminutos com sétima diminuta (três semitons abaixo da oitava, enarmônica à sexta maior).

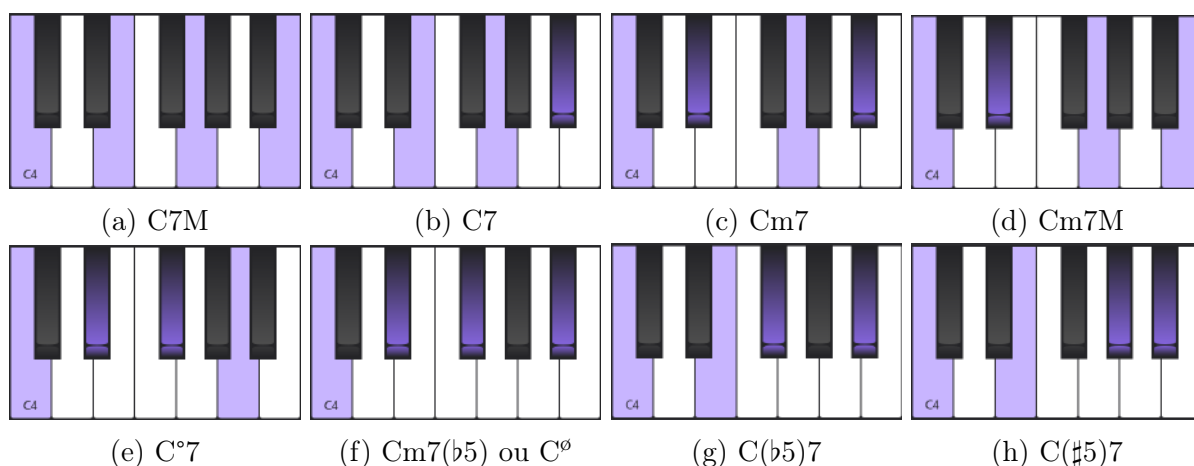


Figura 4 – Tétrades. (a) Dó com Sétima Maior. (b) Dó com Sétima. (c) Dó Menor com Sétima. (d) Dó Menor com Sétima Maior. (e) Dó Diminuto com Sétima. (f) Dó Meio Diminuto. (g) Dó com Sétima e Quinta Bemol. (h) Dó com Sétima e Quinta Aumentada.

Fonte: Elaboração própria a partir do teclado de piano do Muscore 4.

2.4 PROGRESSÃO DE ACORDES

Antes de falar sobre progressão de acordes é necessário apresentar o conceito de **funcionalidade**⁴. No escala maior existem 7 graus de acordes, representados da seguinte

⁴ Conceito abordado com detalhes em (ALMADA, 2009).

forma: I - ii - iii - IV - V - vi - vii°. Onde os graus em maiúsculo são maiores, em minúsculo, menores e, o símbolo '°' representa o diminuto. Os três graus maiores, I, IV e V, podem ser vistos como as principais funções harmônicas existentes, respectivamente *tônica*, *subdominante* e *dominante*.

A função tônica é a mais importante de todas, pois no contexto de sua escala, ela sugere repouso, relaxamento. Todas as restantes estão hierarquicamente subordinadas à ela. A função dominante, por sua vez, pode ser entendida como o reverso da tônica, remete instabilidade, tensão. Por isso, um acorde da área dominante procura de qualquer forma dirigir-se para um da área tônica, se estabilizar, *resolver*. Já a função subdominante, pode ser percebida auditivamente como uma espécie de *afastamento* do poder central da tônica. Essa característica da função subdominante cria um contraste em relação às outras duas, podendo ser esquematizado pela Figura 5.

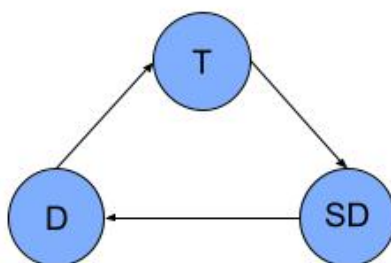


Figura 5 – Esquema de correlação das forças tonais. T - tônica, D - dominante, SD - subdominante.

Fonte: Exemplo de (ALMADA, 2009).

A partir do repouso tônico, o contraste subdominante encaminha-se de volta à estabilidade através da dominante. Este esquema (Figura 5) é o alicerce harmônico para a estrutura da maioria das músicas produzidas no Ocidente desde o advento do Tonalismo⁵. Os demais graus⁶ não trazem nada de novo em termos de funções, simplesmente se enquadram como acordes secundários em uma das três funções já existentes. Para a tônica: I, iii e vi; subdominante: IV e ii; e dominante: V e vii°. A progressão de acordes, então, é uma sequência de acordes que leva em conta essas funções, criando uma narrativa harmônica.

2.4.1 Progressões harmônicas comuns

As progressões de acordes são essenciais para definir o caráter e o estilo de diferentes gêneros musicais. Dentre as inúmeras progressões possíveis, algumas se destacam, marcadas por um gênero musical específico:

⁵ Para mais informações sobre tonalismo, olhar (ALMADA, 2009) ou (HURON, 2006).

⁶ Também conhecidos como *graus modais*, em oposição aos *graus tonais*, I, IV e V.

- No blues, uma das progressões mais icônicas é a que segue o padrão I - IV - V. Mapeando ela para a tonalidade de Dó maior, a progressão seria C7 - F7 - G7. O uso de acordes dominantes, apresentados na Seção 2.3.3, adiciona uma tensão característica que é resolvida parcialmente ao longo da progressão.
- Na música pop, as progressões tendem a ser cativantes e memoráveis com padrões repetitivos. Uma progressão extremamente popular é a I - V - vi - IV. Em Dó maior seria C - G - Am - F. Uma grande parte das músicas pop de sucesso fazem uso deste padrão.
- O jazz, por sua vez, é conhecido por suas progressões harmônicas complexas e uso extensivo de acordes com tensões. Uma das progressões mais famosas é o ii - V - I. Que na verdade se faz presente na grande maioria dos gêneros musicais, porém no jazz é usual encontrar acordes mais rebuscados. Em Dó maior seria Dm9 - G9(13) - C9 ou em Dó menor, Dm7(b5) - G7alt - Cm.
- Por fim, a MPB incorpora uma ampla variedade de influências, incluindo gêneros como jazz e samba. Por conta de todas essas influências, os padrões de progressão são muito variados.

3 PROCESSAMENTO DE SINAIS

3.1 TRANSFORMADA DE FOURIER

A Transformada de Fourier, uma técnica matemática fundamental no processamento de sinais, tem suas raízes nas pesquisas do matemático francês Jean-Baptiste Joseph Fourier. Sua contribuição essencial foi a ideia de que qualquer função periódica pode ser representada como uma soma de funções senoidais e cossenoidais. Esta descoberta, de significativa importância histórica, revolucionou o modo como a análise de sinais é feita em diversos campos, incluindo a acústica e a música. Apresentaremos aqui alguns conceitos básicos referentes à Transformada de Fourier, e para mais detalhes, veja (MÜLLER, 2015, Cap. 2).

A essência da Transformada de Fourier reside na análise de sinais no domínio do tempo e sua conversão para o domínio da frequência. Isto implica decompor um sinal, contendo por exemplo uma nota musical ou um acorde, em seus componentes de frequência. Em termos práticos, isso significa que um sinal musical complexo pode ser entendido como uma combinação de várias frequências diferentes, como ilustrado na Figura 6.

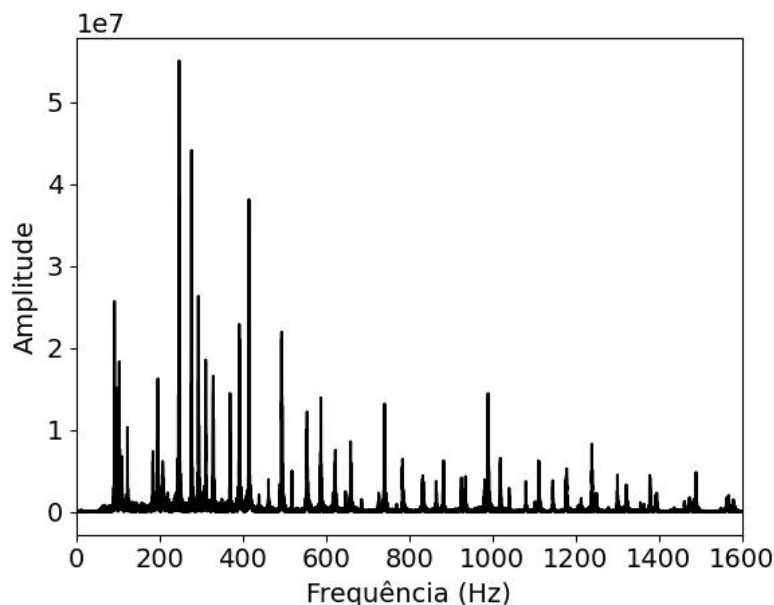


Figura 6 – Módulo da Transformada de Fourier de um trecho da música *Teresinha*, de Chico Buarque.

Fonte: Elaboração própria. Gráfico feito em Python com a biblioteca Matplotlib.

A Transformada de Fourier é matematicamente representada por uma fórmula que transforma uma função de tempo em uma função de frequência. Tal função de partida pode ser de diversas naturezas, levando a diferentes formulações: funções em domínio contínuo (periódicas ou não) e funções em domínio discreto (periódicas ou não). Em

resumo, a Transformada de Fourier fornece uma representação frequencial de sinais¹ de diferentes naturezas, o que é crucial para a análise espectral.

A fim de exemplificação, seja $f : \mathbb{R} \rightarrow \mathbb{R}$ uma função. A sua Transformada de Fourier é dada pela função $\hat{f} : \mathbb{R} \rightarrow \mathbb{C}$ definida conforme abaixo:

$$\hat{f}(\omega) = \int_{-\infty}^{+\infty} f(t) \exp(-2\pi i \omega t) dt. \quad (3.1)$$

Dessa forma, o valor $\hat{f}(\omega)$ pode ser interpretado como a "projeção" da função f na função periódica $\exp(-2\pi i \omega t) = \cos(2\pi \omega t) - i \sin(2\pi \omega t)$, ou seja, o quanto que tal função "contribui" para "formar" a função f . Estão fora do escopo deste trabalho uma discussão aprofundada das condições em f para termos \hat{f} bem definida, bem como porque a intuição aqui apresentada faz sentido. Para mais detalhes veja (MÜLLER, 2015, Cap. 2).

No processamento de sinais, a Transformada de Fourier (e seus objetos matemáticos derivados) são amplamente utilizados para *análise espectral*, ou seja, para entender as componentes de frequência de determinado sinal. Na música, sua aplicação auxilia a desvendar, por exemplo, a estrutura harmônica e as características de timbre de uma peça musical, a partir de uma gravação. Por exemplo, ao analisar um sinal de áudio musical, a Transformada de Fourier pode ser usada para identificar e separar as diferentes notas e acordes presentes.

Porém, um dos principais desafios ao usar a Transformada de Fourier no reconhecimento de acordes é a sua natureza estática, que não captura as variações de frequência ao longo do tempo. Além disso, até o momento foi definida somente a Transformada de Fourier para uma função $f : \mathbb{R} \rightarrow \mathbb{R}$, e essa não é a forma como dispõe-se de sinais de áudio armazenados digitalmente para análise. Nas próximas seções, a Transformada de Fourier é "transportada" para outros domínios a fim de contornar essas dificuldades.

3.2 TRANSFORMADA DISCRETA DE FOURIER (DFT)

A DFT adapta os princípios da Transformada de Fourier original para o domínio digital, ou seja, para funções $x : \mathbb{Z} \rightarrow \mathbb{R}$, semelhante a como dispomos de sinais de áudio armazenados digitalmente. Essa evolução foi crucial para o processamento de sinais digitais de comprimento finito, como os encontrados em gravações de música. Ela surgiu como uma resposta às necessidades da era digital, mantendo a essência das ideias de Fourier enquanto introduzia novos conceitos para lidar com dados discretizados.

Para um sinal $x : \{0, \dots, N-1\} \subset \mathbb{Z} \rightarrow \mathbb{R}$, a DFT é representada pela fórmula

$$X(k) = \sum_{n=0}^{N-1} x(n) \exp(-2\pi i kn/N), \quad (3.2)$$

¹ Em Processamento de Sinais, os termos "sinal" e "função" são usado de modos intercambiáveis, apesar do primeiro termo abranger uma categoria mais geral de objetos que o segundo.

que pode ser interpretada como uma "discretização" da Transformada de Fourier, apresentada na Equação 3.1 (MÜLLER, 2015, Cap. 2). A fórmula na Equação 3.2 também pode ser interpretada como uma mudança de base em \mathbb{C}^N .

A DFT é particularmente vantajosa para sinais digitais, proporcionando uma análise precisa de sinais de comprimento finito, típicos em gravações digitais de música. Contudo, enfrenta desafios como o *vazamento espectral*, decorrente da limitação na resolução devido ao número finito de amostras. Este fenômeno pode afetar a precisão na identificação de frequências específicas, um aspecto crítico no reconhecimento de acordes. O principal problema é a incapacidade de localizar no tempo onde cada frequência acontece, como ilustra a Figura 6.

3.3 TRANSFORMADA DE FOURIER DE TEMPO CURTO (STFT)

A inovação de Fourier na análise de sinais estabeleceu as bases para técnicas mais avançadas como a STFT, introduzida por Dennis Gabor em 1946, que surgiu para superar algumas das limitações da Transformada de Fourier original no tratamento de sinais não estacionários, ou seja, cujas características estatísticas variam ao longo do tempo.

Diferentemente da Transformada de Fourier, que fornece uma representação frequencial de um sinal para todo o seu período, a STFT permite uma análise frequencial que varia ao longo do tempo, como observado na Figura 7. A chave para esta técnica é a utilização de uma "janela" de análise, que se move ao longo do sinal, permitindo a observação de como as frequências variam no tempo. Esta janela de análise é essencial para capturar as dinâmicas temporais dos sinais musicais, especialmente importantes no reconhecimento de acordes.

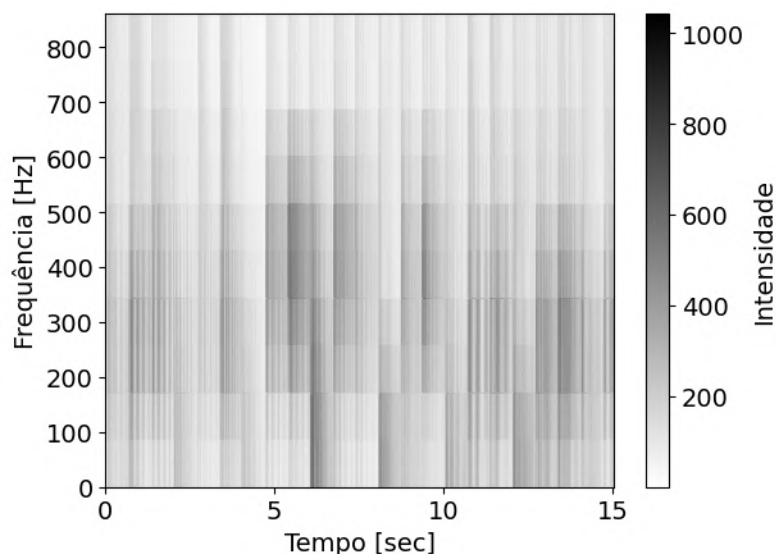


Figura 7 – STFT de um trecho da música *Teresinha* de Chico Buarque.

Fonte: Elaboração própria. Gráfico feito em Python com a biblioteca Matplotlib.

Seja $x : \{0, \dots, N - 1\} \subset \mathbb{Z} \rightarrow \mathbb{R}$ um sinal a tempo discreto, e $w : \mathbb{Z} \rightarrow \mathbb{R}$ uma *função-janela*.² A STFT é definida então como

$$\mathcal{X}(m, k) := \sum_{n=0}^{N-1} x(n + mH)w(n) \exp(-2\pi i kn/N), \quad (3.3)$$

onde o parâmetro H representa o *tamanho do passo da janela*. Esta formulação permite a análise do sinal em segmentos temporais, mantendo a informação frequencial. A representação gráfica da função \mathcal{X} ilustrada na Figura 7 é dita um *espectrograma*.

A STFT proporciona, portanto, uma análise mais detalhada do sinal de áudio, permitindo a identificação de variações de frequência ao longo do tempo. Esta característica é crucial para analisar composições musicais com variações de timbres e campos tonais, onde a estrutura temporal e frequencial do som é fundamental. A STFT é amplamente utilizada em várias aplicações na música e na engenharia de áudio, como na equalização, na redução de ruído, na compressão de dados, e no reconhecimento automático de música. A habilidade de analisar e manipular sinais tanto no domínio da frequência quanto do tempo é essencial para o aprimoramento da qualidade do áudio e para o desenvolvimento de novas tecnologias na indústria musical e campos relacionados.

A principal vantagem da STFT sobre a DFT é sua capacidade de fornecer uma análise de frequência que varia no tempo, crucial para lidar com sinais não estacionários, comuns em música. Contudo, há um desafio inerente na escolha da janela de análise, pois janelas maiores proporcionam melhor resolução de frequência, mas menor resolução temporal, e vice-versa. Esse compromisso entre resolução temporal e frequencial é um aspecto crucial a ser considerado na análise musical, especialmente quando comparado aos desafios enfrentados pela Transformada de Fourier, que fornece uma visão estática do espectro de frequência.

3.4 CROMAGRAMA

A evolução no processamento de áudio musical trouxe à tona técnicas avançadas como os *cromagramas*, essenciais na análise harmônica de música e som. O início dessa jornada analítica se dá com os espectrogramas, cruciais para entender a estrutura musical.

A Figura 1b traz uma boa analogia para a forma como um espectrograma é mapeado num cromagrama, sendo o espectrograma análogo à hélice tonal e o cromagrama análogo ao círculo cromático projetado pela hélice. Porém diferentemente da figura, os cromagramas levam em conta a variação ao longo do tempo, assim como o espectrograma.

Existem vários tipos de espectrograma, porém, por uma questão de simplicidade, este projeto foca em apenas três desses tipos, com o STFT sendo o primeiro, já explicado na

² A janela escolhida impacta diretamente nos espectros obtidos na STFT. Tipicamente, as janelas são escolhidas dependendo da aplicação. A escolha de janelas é um problema muito complexo e foge do escopo deste trabalho. Para mais detalhes, veja (MÜLLER, 2015, Cap. 2).

Seção 3.3. O segundo espectrograma é o CQT (Constant Q Transform), onde Q se refere ao fator de qualidade dos filtros usados na transformada. O CQT é uma adaptação do DFT (COSTA, 2020, Cap. 2), apresentado na Seção 3.2. Por fim, o IIR (Infinite Impulse Response), apresentado em (MÜLLER, 2007), também aplica pré- e pós-filtragens no procedimento da DFT. Em resumo, tanto o CQT quanto o IIR tentam replicar de forma mais específica o funcionamento do ouvido humano, e sua explicação mais detalhada está fora do escopo deste trabalho.

3.4.1 Funcionamento

Na análise de áudio, os cromagramas surgem como uma evolução dos espectrogramas. Eles condensam as informações de frequência em 12 categorias distintas, alinhadas com as notas da escala cromática, apresentada na Seção 2.2. Essa simplificação facilita o manejo do conteúdo harmônico do áudio para análises mais profundas.

Os cromagramas se destacam na análise de áudio pela sua habilidade em dividir o tom em duas componentes: a altura, que se relaciona com a oitava, e o croma, identificando a classe de altura específica (como C, C \sharp , D). Esta técnica agrupa tons em oitavas distintas que compartilham a mesma classe de altura, revelando padrões harmônicos.

Essencialmente, os cromagramas mapeiam a estrutura harmônica de uma composição musical, "enrolando" as notas em torno de um eixo cromático, como evidenciado na Figura 8. Isso ressalta a equivalência de oitavas na música, onde notas em diferentes oitavas, ou seja, alturas, mas com a mesma designação (por exemplo, C3, C4), são representadas no mesmo ponto do cromagrama. Isso ilustra como uma única classe de altura pode energizar várias bandas cromáticas, refletindo a sua riqueza harmônica. Comparando este cromagrama com o espectrograma da Figura 7, percebe-se claramente uma facilidade maior de classificar quais classes de alturas foram emitidas em um dado intervalo. Essa vantagem é evidente no contexto de reconhecimento de acordes, que será abordado no Capítulo 5.

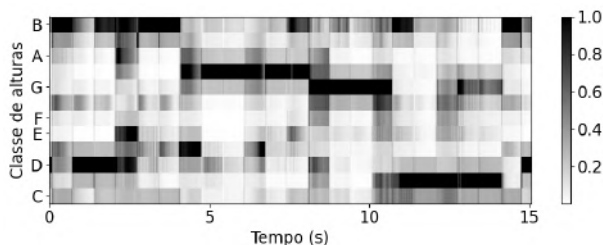


Figura 8 – Cromagrama de um trecho da música *Teresinha* de Chico Buarque.

Fonte: Elaboração própria. Gráfico feito em Python com a biblioteca Matplotlib.

Matematicamente, um cromagrama é obtido através da projeção do espectrograma em um espaço de 12 dimensões, onde cada dimensão representa uma das 12 notas. Esta transformação é feita utilizando-se uma série de filtros ou uma transformação de Fourier,

que concentra a energia de frequências em suas respectivas classes de alturas na escala cromática. Para mais informações sobre como um cromagrama é calculado veja (MÜLLER, 2015, Cap. 2).

A Figura 9 ilustra a diferença de cromagramas gerados para cada um dos tipos (STFT, CQT e IIR), através da escala diatônica de Lá maior, composta pelas classes de alturas Lá, Si, Do♯, Ré, Mi, Fá♯ e Sol♯. Nela já é possível notar a diferença entre as abordagens, e que o STFT acaba compreendendo mais ruído do que os outros dois. Além disso pode ser observada a taxa de 21.5 Hz no título das figuras: esta taxa é um hiper-parâmetro da função que gera os cromagramas, que estabelece que na janela de 1 segundo serão observados 21.5 *vetores cromas*.³

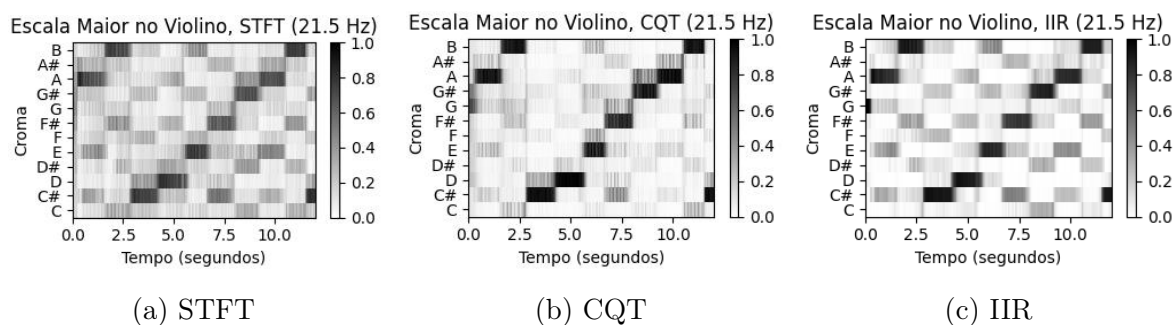


Figura 9 – Cromagramas gerados para um violino tocando a Escala de Lá maior.

Fonte: Testes Iniciais do (PITA, 2024).

³ Croma diz respeito à classe de altura apresentada na Seção 2.2.

4 MODELOS DE MARKOV

4.1 CADEIA DE MARKOV

As Cadeias de Markov (ROCHA, 2020), introduzidas por Andrey Markov no início do século XX, oferecem um modelo para analisar sequências de eventos onde a probabilidade de cada evento depende apenas de um evento imediatamente anterior. Esta característica é conhecida como a "propriedade de Markov", matematicamente expressa pela Equação 4.1:

$$\mathbb{P}(X_{n+1} = x | X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = \mathbb{P}(X_{n+1} = x | X_n = x_n) \quad (4.1)$$

Em uma definição geral, uma cadeia de Markov permite que a probabilidade de transição entre estados dependa do momento n em que ocorre. Isso significa que a probabilidade de mudar de um estado para outro pode variar conforme o tempo avança. Formalmente, para quaisquer dois estados i e j em um espaço de estados S e para qualquer instante n , a probabilidade de transição do estado i para o estado j é dada por $A_{ij}(n) = \mathbb{P}(X_{n+1} = j | X_n = i)$. O espaço de estados S em uma Cadeia de Markov é o conjunto de todos os possíveis estados que o sistema pode ocupar. Vale mencionar a existências de cadeias de Markov definidas em espaços de estados contínuos e/ou a tempo contínuo, porém estes modelos estão além do escopo do projeto: apenas transições em tempos discretos e espaços de estados finitos serão utilizadas.

As cadeias homogêneas são uma particularização para o caso em que a probabilidade de transição é constante ao longo do tempo. Neste caso, a probabilidade de transição de um estado para outro não depende de n , ou seja $A_{ij}(n) = A_{ij}$ para todos os n . Isso simplifica significativamente a análise da cadeia, pois a dinâmica de transição é invariável no tempo. As Cadeias de Markov homogêneas são descritas por uma matriz estocástica A , como a da Figura 10b, onde cada elemento a_{ij} representa a probabilidade de transição do estado i para o estado j . Note que o estado de "partida" é indicado nas linhas da matriz, e o estado de "chegada" está nas colunas. As propriedades de uma matriz estocástica incluem o fato de que todas as probabilidades de transição são não-negativas e que a soma de cada linha é igual a um (ou seja, as probabilidades de sair de qualquer estado somam um). A figura 10a ilustra uma Cadeia de Markov como um grafo direcionado: os vértices representam os estados e as arestas (com seus respectivos pesos) representam as probabilidades de transição.

Para fins de padronização e simplicidade, o termo "Cadeias de Markov" será utilizado para referir-se especificamente a cadeias de Markov homogêneas a tempo discreto com espaço de estados finito.

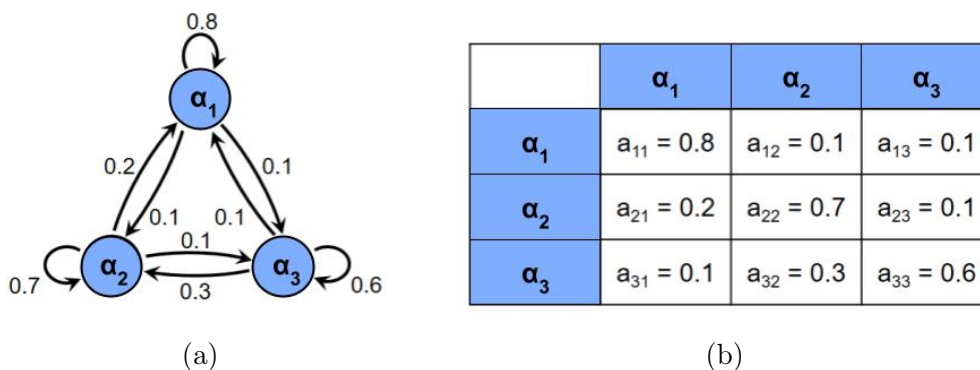


Figura 10 – Ilustração de uma cadeia de Markov com três estados. (a) Grafo da cadeia de Markov. (b) Matriz de transição da cadeia de Markov.

Fonte: Figura 5.24 de (MÜLLER, 2015).

4.2 MODELO OCULTO DE MARKOV (HMM)

Os Modelos Ocultos de Markov (HMM – *Hidden Markov Models*) são uma extensão das Cadeias de Markov, que são fundamentais no campo do processamento estatístico de sinais. Enquanto uma Cadeia de Markov tradicional é usada para modelar a probabilidade de transições entre estados visíveis, num HMM a propriedade de Markov é levada para uma camada oculta, que não é observada diretamente.

No contexto do reconhecimento de acordes musicais, um HMM pode ser utilizado para vincular sequências de vetores de cromas (retirado de um cromagrama), observáveis, a uma sequência de acordes subjacentes, que são os estados ocultos do modelo. Tomando como base a Figura 10a, suponha que os estados α_1 , α_2 e α_3 sejam atribuídos aos acordes C (Dó maior), G (Sol maior) e F (Fá maior) respectivamente. A Figura 11 ilustra os pontos principais de um HMM no contexto de acordes a partir desta configuração.

Diferentemente das cadeias de Markov, o HMM apresenta uma segunda matriz importante, a *matriz de emissão*. Seguindo no exemplo da Figura 10a e partindo do que foi explicado na Seção 2.3.1, pode-se concluir que o acorde de Dó maior é composto pelas notas Dó-Mi-Sol, Fá maior por Fá-Lá-Dó e Sol maior por Sol-Si-Ré. Ao comparar os três acordes é evidente que os acordes F e G compartilham notas em comum com o acorde de C. Num contexto de reconhecimento de acordes, essas notas em comum podem gerar uma certa confusão, dependendo da acústica do instrumento em questão.

A matriz de emissão modela essa ambiguidade, associando, por exemplo, as probabilidades do estado oculto de C e G serem observados como iguais. A matriz de emissão do HMM ilustrado na Figura 11 pode ser encontrada na Tabela 2, onde os vetores de possíveis observações são representados por o_1 , o_2 , o_3 respectivamente da esquerda para a direita. Por acaso, esta é uma matriz quadrada, porém havendo mais estados observáveis, a matriz aumentaria de 3×3 para $3 \times m$, sendo m o número de estados observáveis. A Seção 5.2 explica o procedimento para a geração dos vetores de estados observáveis.

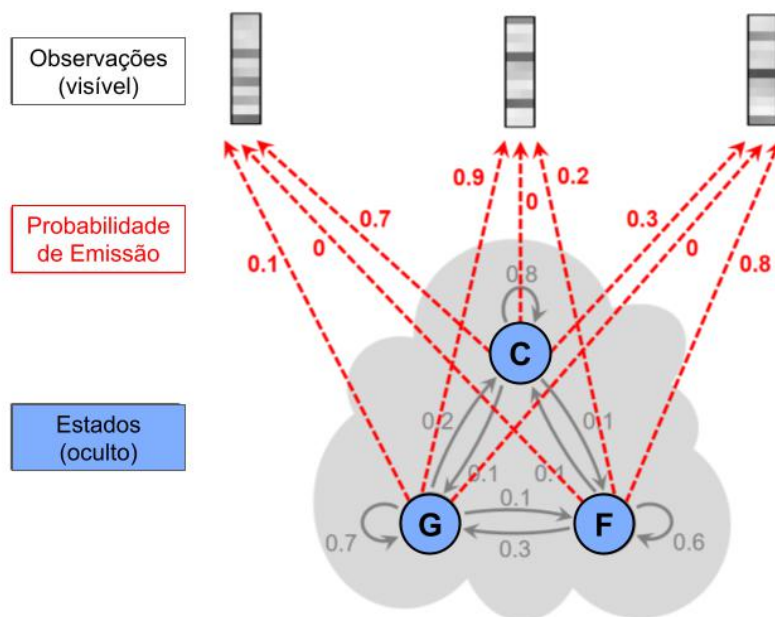


Figura 11 – Modelo Oculto de Markov no contexto de reconhecimento de acordes.

Fonte: Figura 5.25 do (MÜLLER, 2015).

Tabela 2 – Matriz de Probabilidades de Emissão.

	o_1	o_2	o_3
C	0.7	0	0.3
G	0.1	0.9	0
F	0	0.2	0.8

Um HMM é caracterizado pela tupla $\Theta = (\mathcal{A}, A, C, \mathcal{B}, B)$, onde \mathcal{A} é o conjunto de estados ocultos, A a matriz de probabilidades de transição de estado, C as probabilidades do estado inicial, \mathcal{B} um conjunto de símbolos de observação possíveis, e B as probabilidades de emissão. Essas probabilidades de emissão ligam cada estado oculto a uma probabilidade de emitir um certo símbolo ou observação.

4.2.1 Problema da Descoberta

O problema da descoberta, também conhecido como problema de decodificação, é central na aplicação de HMMs, e refere-se à tarefa de determinar a sequência de estados ocultos mais provável que tenha gerado uma sequência observada de eventos. No contexto do reconhecimento de acordes, isso se traduz em identificar a progressão de acordes com base numa sequência de vetores de corama extraídos de uma gravação de áudio, podendo ser pensado como uma automatização do processo de "tirar uma música de ouvido".

Para encontrar essa sequência ótima de estados de maneira eficiente existe o algoritmo de Viterbi (Tabela 3 e Código 1), baseado em programação dinâmica. Como entrada, ele recebe a tupla $\Theta = (\mathcal{A}, A, C, \mathcal{B}, B)$ e uma sequência de observações $O = (o_1, o_2, \dots, o_N)$,

onde cada o_n corresponde a uma observação no tempo n . A saída é, por sua vez, a sequência ótima de estados $S^* = (s_1^*, s_2^*, \dots, s_N^*)$ que tem a maior probabilidade de ter resultado na sequência de observações dada, com base no modelo HMM fornecido. Para mais detalhes no algoritmo de Viterbi, veja (MÜLLER, 2015, Cap. 5).

Tabela 3 – Pseudocódigo do Algoritmo de Viterbi.

Fonte: Adaptado da Tabela 5.2 de (MÜLLER, 2015).

Algoritmo:	VITERBI
Entrada:	HMM especificado por $\Theta = (\mathcal{A}, A, C, \mathcal{B}, B)$ Sequência de Observação $O = (o_1, o_2, \dots, o_N)$
Saída:	Sequência ótima de estados $S^* = (s_1^*, s_2^*, \dots, s_N^*)$
Procedimento:	Inicializar a matriz $\mathbf{D}_{I \times N}$ como $\mathbf{D}(i, l) = c_i b_{ik_l}$ para $i \in [1 : l]$. Então computar num segundo <i>loop</i> para $n = 2, \dots, N$ e $i = 1, \dots, l$:
	$\mathbf{D}(i, n) = \max_{j \in [1:l]} (a_{ij} \cdot \mathbf{D}(j, n-1)) \cdot b_{ik_n}$ $\mathbf{E}(i, n-1) = \arg \max_{j \in [1:l]} (a_{ij} \cdot \mathbf{D}(j, n-1))$
	Atribuir $i_N = \arg \max_{j \in [1:l]} \mathbf{D}(j, N)$ e computar para $n = N-1, \dots, 1$ os índices de maximização
	$i_n = \arg \max_{j \in [1:l]} (a_{ij_{n+1}} \cdot \mathbf{D}(j, n)) = \mathbf{E}(i_{n+1}, n).$
	A sequência ótima de estados $S^* = (s_1^*, \dots, s_N^*)$ é definida por $s_n^* = \alpha_{i_n}$ para $n \in [1 : N]$.

Código 1 – Algoritmo de Viterbi.
 Fonte: Implementação de (MÜLLER, 2015).

```

import numpy as np

def viterbi(A, C, B, O):
    """Algoritmo de Viterbi para resolver o problema da descoberta

    Notebook: C5/C5S3_Viterbi.ipynb

    Args:
        A (np.ndarray): Matriz de probabilidade de transicao de estado
            de dimensao I x I
        C (np.ndarray): Distribuicao inicial de estado de dimensao I
        B (np.ndarray): Matriz de probabilidade de saida de dimensao I x
            K
        O (np.ndarray): Sequencia de observacao de tamanho N

    Retorna:
        S_opt (np.ndarray): Sequencia otima de estado de tamanho N
        D (np.ndarray): Matriz de probabilidade acumulada
        E (np.ndarray): Matriz de backtracking
    """
    I = A.shape[0]      # Numero de estados
    N = len(O)         # Tamanho da sequencia de observacao

    # Inicializa as matrizes D e E
    D = np.zeros((I, N))
    E = np.zeros((I, N-1)).astype(np.int32)
    D[:, 0] = np.multiply(C, B[:, 0[0]])

    # Computar D e E em um loop aninhado
    for n in range(1, N):
        for i in range(I):
            temp_product = np.multiply(A[:, i], D[:, n-1])
            D[i, n] = np.max(temp_product) * B[i, O[n]]
            E[i, n-1] = np.argmax(temp_product)

    # Backtracking
    S_opt = np.zeros(N).astype(np.int32)
    S_opt[-1] = np.argmax(D[:, -1])
    for n in range(N-2, -1, -1):
        S_opt[n] = E[int(S_opt[n+1]), n]

    return S_opt, D, E

```

5 APLICAÇÃO A RECONHECIMENTO DE ACORDES

5.1 BASEADO EM MODELOS (*TEMPLATE-BASED*)

O reconhecimento de acordes baseado em modelos (*template-based*) é um processo básico utilizado na análise de gravações musicais para identificar os acordes tocados ao longo do tempo. Esta tarefa é geralmente dividida em duas etapas principais, como pode ser observado na Figura 12, com as etapas envolvidas pela linha tracejada. Além disso, este processo servirá de base para o reconhecimento de acordes baseado em HMM, explicado na Seção 5.2.

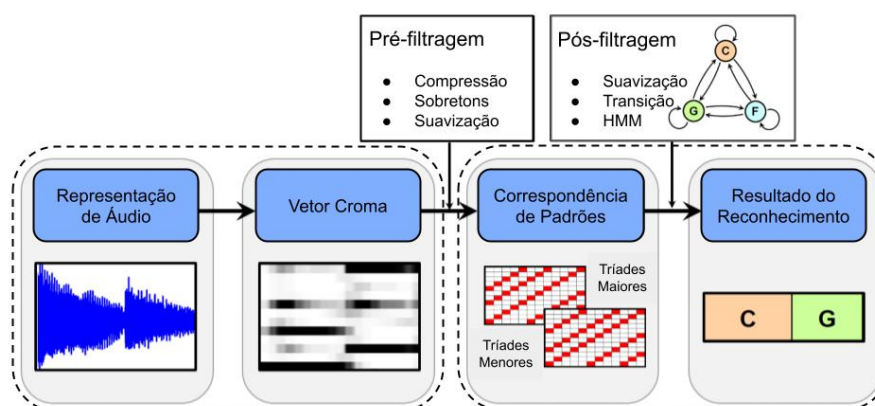


Figura 12 – Etapas do processo de reconhecimento baseado em modelos.

Fonte: Figura 5.13 do (MÜLLER, 2015).

5.1.1 Procedimento

A primeira etapa transforma a gravação de áudio em uma sequência de vetores de características, que frequentemente se baseiam em atributos cromáticos (vetores de cromas) devido à sua capacidade de capturar o conteúdo harmônico de curta duração presente no sinal de áudio.

A segunda etapa do processo envolve técnicas de correspondência de padrões, que mapeiam cada vetor observado de características a um modelo de acorde pré-definido (como na Figura 13). Este conjunto de modelos de acordes geralmente inclui as 24 tríades básicas, maiores e menores, explicadas no início da Seção 2.3.1. O melhor ajuste entre o vetor de características e os modelos de acordes é utilizado para atribuir um rótulo ao vetor observado correspondente. Para aprimorar os resultados do reconhecimento de acordes, técnicas adicionais de processamento de sinais podem ser aplicadas antes ou depois da etapa de correspondência de padrões, conhecidas respectivamente como *pré-filtragem* e *pós-filtragem*.

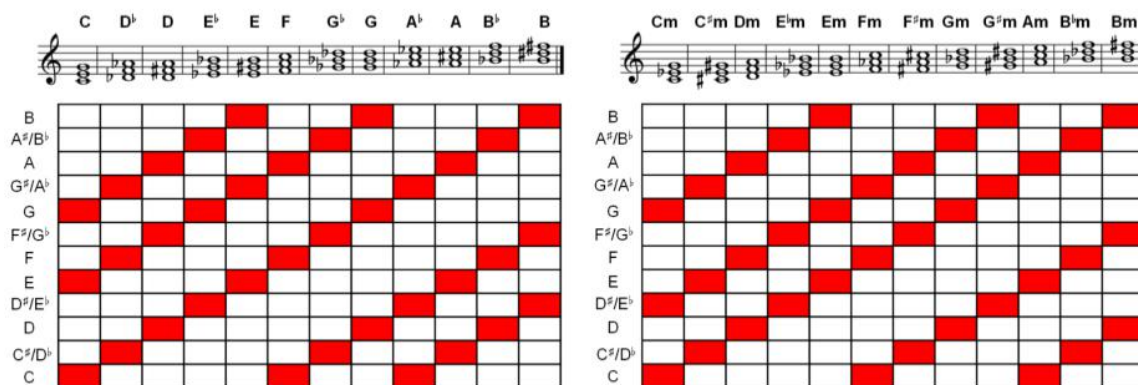


Figura 13 – Vetores de modelos binários para as 24 tríades maiores e menores. Células em branco têm valor zero, e células em vermelho têm valor um. Cada coluna corresponde ao acorde ilustrado na partitura acima.

Fonte: Figura 5.6 do (MÜLLER, 2015)

A tarefa de reconhecimento é, portanto, reduzida a encontrar o rótulo de acorde $\alpha_n \in \mathcal{A} = \{C, C\sharp, \dots, B, Cm, C\sharp m, \dots, Bm\}$ que maximiza a similaridade entre o modelo correspondente e o vetor de características observado. Esta similaridade pode ser medida através do produto interno entre versões normalizadas dos vetores binários (Figura 13) com o vetor observado (também normalizado), uma técnica que simplifica a comparação mantendo a robustez contra variações de intensidade. A Equação 5.1 sistematiza essa comparação, onde, por exemplo, x pode ser um vetor de croma observado e y pode ser um dos 24 modelos binários apresentados na Figura 13. Fazendo um paralelo com o HMM (Seção 4.2), olhamos somente para a camada observável do modelo, desconsiderando probabilidades de transição entre os acordes.

$$s(x, y) = \frac{\langle x, y \rangle}{\|x\|_2 \cdot \|y\|_2}. \quad (5.1)$$

Durante o processo de reconhecimento de acordes, várias escolhas de projeto influenciam significativamente o desempenho do sistema. Estas escolhas incluem quais acordes considerar no conjunto de modelos, como medir a similaridade entre os vetores de características e os modelos de acordes e quais técnicas de pré- e pós-processamento aplicar aos vetores de características. A limitação a apenas os 24 acordes maiores e menores pode ser problemática e não captura todas as nuances harmônicas possíveis, mas é frequentemente uma restrição aceita na literatura de reconhecimento de acordes.

O procedimento de reconhecimento de acordes geralmente termina com a representação visual da sequência de acordes ao longo do tempo, proporcionando uma quantificação musicalmente informada da representação cromática original. Essa representação pode ser comparada com anotações de especialistas para avaliar a precisão do reconhecimento de acordes, e ajustes no modelo ou no algoritmo podem ser feitos para melhorar a correspondência com as expectativas harmônicas.

Finalmente, pode haver trechos de silêncio que não se encaixam com nenhum dos rótulos presentes no conjunto \mathcal{A} . Para resolver isso é adicionado um rótulo \mathbf{N} para indicar a ausência de acorde.

5.1.2 Avaliação dos resultados

Para avaliar os resultados é feita uma comparação dos rótulos obtidos no procedimento com os rótulos de referência, gerados manualmente através de uma análise por um especialista, ou a partir da partitura da música. Essa comparação resulta em quatro grupos: os **verdadeiros positivos** (TP, *true positive*), quando o rótulo estimado é igual ao rótulo de referência; os **falsos positivos** (FP, *false positives*), quando um rótulo de referência que não corresponde ao acorde é incorretamente estimado e a referência não for \mathbf{N} ; os **falsos negativos** (FN, *false negatives*), quando o rótulo de referência correto não é devidamente detectado; por fim os **verdadeiros negativos** (TN, *true negatives*), que são rótulos de referência que não correspondem ao acorde e que de fato não são estimados. A partir desses grupos, são geradas três métricas (MÜLLER, 2015), **precisão** (P), **revocação** (R) e **estatística F** (F):

$$P = \frac{\#TP}{\#TP + \#FP}, \quad R = \frac{\#TP}{\#TP + \#FN}, \quad F = \frac{2 \cdot P \cdot R}{P + R} \quad (5.2)$$

Os resultados são avaliados de acordo com a estatística F, que varia no intervalo $[0, 1]$. A Figura 14 ilustra um exemplo de resultado do reconhecimento baseado em modelos, já com as métricas calculadas.

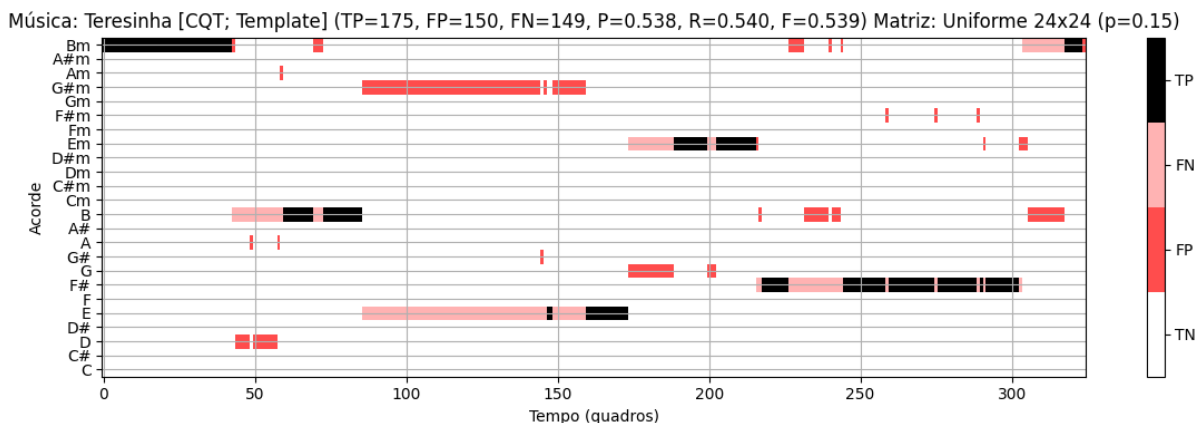


Figura 14 – Resultado do reconhecimento baseado em modelos para um trecho da música "Teresinha" de Chico Buarque.

Fonte: Testes iniciais do (PITA, 2024).

O formato dos resultados de identificação, como exemplificado na Figura 14, seguem um padrão de acordes que só apresenta acordes sustentados (\sharp). Isto se dá como forma de simplificar a leitura, já que $C\sharp$ e $D\flat$, por exemplo, são enarmônicos (Seção 2.1), ou

seja, têm o mesmo conteúdo acústico, porém são funcionalmente distintos em uma peça musical.

5.2 BASEADO EM HMM (*HMM-BASED*)

O processo de reconhecimento baseado em HMM, ocorre com a parametrização do modelo no contexto de acordes, brevemente explicada na Seção 4.2. Assim como o processo baseado em modelos (Seção 5.1), o reconhecimento baseado por HMM pode também ser dividido em duas etapas: a de treinamento do modelo para ajustar as matrizes de transição e emissão, seguida pela de avaliação do modelo, testando seus resultados com novas amostras de observação. As etapas do processo estão ilustradas na Figura 15. O projeto será focado na segunda etapa do processo, pois a matriz de transição será fornecida pelo Projeto MPB (ALMADA; CARVALHO, 2024). A partir desta matriz, será gerada a de emissão não necessitando ter uma etapa de treinamento. Uma outra abordagem de atacar o problema de reconhecimento de acordes foi explorada em (LI, 2021) que combina o HMM com uma Rede Neural Convolutacional.

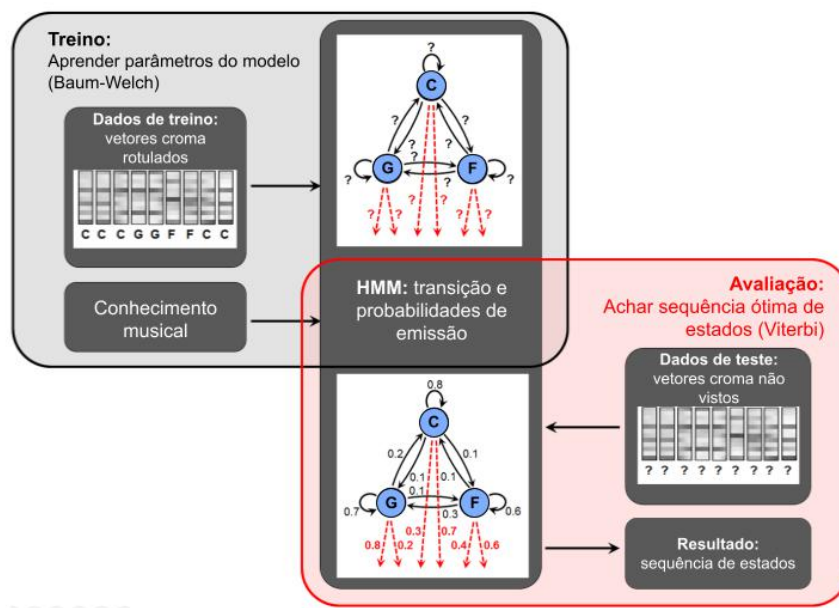


Figura 15 – Etapas do processo de reconhecimento baseado em HMM.

Fonte: Figura 5.33 do (MÜLLER, 2015).

A sequência de observações é obtida de um cromagrama (Seção 3.4) gerado a partir de um áudio. As observações são constituídas de vetores retirados de trechos temporais do cromagrama, como pode ser visto nas Figuras 11 e 15. Os estados ocultos, por sua vez, são exatamente o propósito do projeto, os acordes a serem descobertos.

A matriz de transição constituirá nas probabilidades de um dado acorde transicionar para ele mesmo ou outro qualquer. Dependendo do gênero musical, e de acordo com a

teoria de progressão de acordes, introduzida na Seção 2.4, existem acordes que apresentam uma grande chance de transicionar para outros, como um acorde dominante para um tônico. A matriz de transição serve para evidenciar isto. Por esse motivo, não existe uma única matriz ideal, mas sim matrizes específicas seja para uma banda ou cantor-solo, como é o caso na Figura 16, ou para um determinado gênero musical, como será apresentada na Seção 6.1.2.

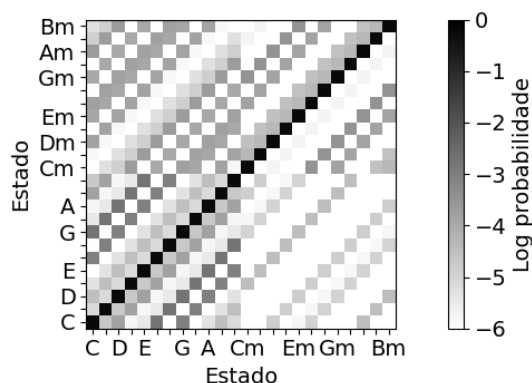


Figura 16 – Matriz de transição invariante à transposição. A matriz foi computada usando um conjunto de dados dos Beatles.

Fonte: Figura 5.30 do (MÜLLER, 2015).

As probabilidades de emissão representam a chance que um acorde α tem de ser mapeado numa observação o . Em outras palavras, o quão provável é de um acorde apresentar a assinatura cromática de uma dada observação. Assim como mencionado na Seção 4.2, a matriz de probabilidades de emissão, num contexto de reconhecimento de acordes, tem a função de combater ambiguidades harmônicas encontradas nas observações. Na Seção 6.2 será explicado o processo usado para gerar a matriz de emissão a partir do banco do Projeto MPB (ALMADA; CARVALHO, 2024).

6 REFINAMENTO DAS MATRIZES DO HMM

6.1 MATRIZ DE TRANSIÇÃO

Como mencionado na Seção 5.2, as matrizes de transição no contexto de reconhecimento de acordes estabelecem a probabilidade de transição entre acordes. (MÜLLER, 2015) usa uma matriz triádica uniforme como ponto de partida; mais à frente no livro, também é usada uma matriz gerada a partir da coletânea de músicas dos Beatles, porém esta não será abordada, visto que o alvo do projeto é reconhecer acordes em contextos harmonicamente complexos.

A primeira etapa da Figura 15 destaca um possível procedimento de gerar a matriz de transição. Este se dá através do treinamento do modelo a partir de dados já rotulados. Porém, como dito da Seção 5.2, este não será o procedimento aqui utilizado.

6.1.1 Matriz Triádica Uniforme

A matriz triádica uniforme se dá por uma matriz 24×24 uniforme, onde a probabilidade de transição para o mesmo estado é parametrizada e seu complemento é dividido igualmente entre os outros estados, como evidenciado na Figura 17.

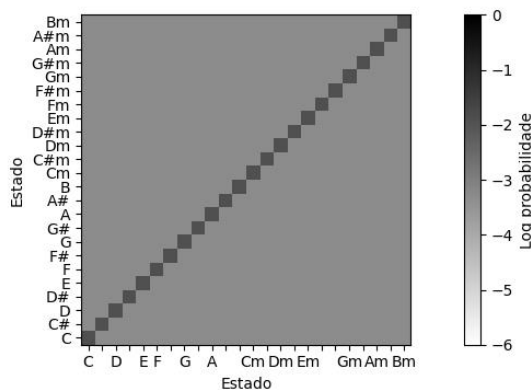


Figura 17 – Matriz uniforme de transição com probabilidade de transição ao mesmo estado $p = 0.15$.

Fonte: Seção 5.3.4 de (MÜLLER, 2015).

6.1.2 Matriz Triádica Alterada

A matriz triádica alterada ainda é uma matriz 24×24 , porém ela foi obtida através do Projeto MPB (ALMADA; CARVALHO, 2024). Para gerar a matriz alterada foi utilizado um banco de 550 músicas dos grupos, Samba, Jazz, Choro, Jobim, Ivan Lins, Edu Lobo, Chico Buarque, Caetano Veloso, Djavan, Milton Nascimento, João Bosco e Gilberto Gil, sendo 50 músicas de cada.

As anotações foram feitas manualmente por um único especialista, Carlos Almada, professor e compositor da Escola de Música da UFRJ, para garantir consistência, tomando como referência os *songBooks* oficiais dos compositores. Estas anotações dizem respeito, dentre outras coisas, à sequência de acordes de cada uma das 550 músicas. A partir desses dados, foi gerada uma matriz 24×24 com suas células sendo o somatório de ocorrências de transições de um acorde α_i para um acorde α_j . A partir dessa matriz (dita *matriz de contagem*) foi possível normalizar os valores para que as linhas somem 1, gerando uma matriz de transição musicalmente informada (Figura 18).

Pelo fato de o procedimento não considerar transições de mesmo estado, foi necessário uma segunda normalização para estabelecer um valor fixo à diagonal da matriz, mantendo a proporção das linhas.

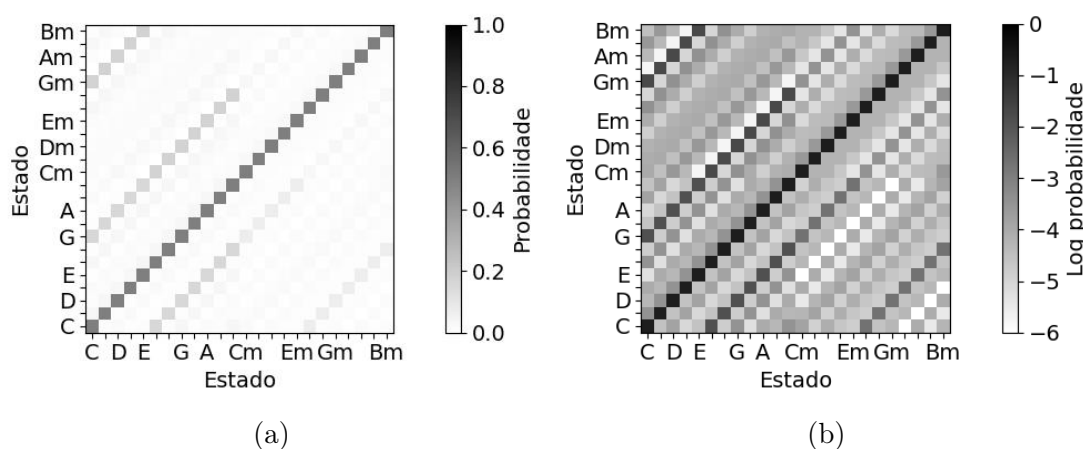


Figura 18 – Matriz de triádica alterada. (a) Matriz de transição. (b) Matriz de transição com log probabilidade.

Fonte: Elaboração de (ALMADA; CARVALHO, 2024).

6.1.3 Matriz Tetrádica

Assim como a matriz anterior, esta também foi gerada pelo Projeto MPB (ALMADA; CARVALHO, 2024) seguindo o mesmo procedimento, entretanto por ser tetrádica, como visto na Seção 2.3.3, ela possui uma complexidade superior. Além de contemplar os 24 acordes triádicos que compõe as matrizes anteriores, ela também compreende as qualidades $X7M$, $X7$, $Xm7$, $Xm7M$, $X^\circ7$, X° , $X(b5)7$, $X(\#5)7$ (respectivas às subfiguras da Figura 4), onde X representa uma das 12 classes de alturas (Seção 2.2), resultando numa matriz de dimensão 120×120 .

A maior complexidade da matriz, apesar de poder aumentar tempos de execução, possibilita potencialmente uma precisão maior ao reconhecimento de acordes. Ao criar os rótulos de referência (Seção 5.1.2) para os testes com as matrizes triádicas, certos acordes tiveram de ser alterados para acomodar o espaço de estados. Ou seja, em músicas da MPB, que apresentam uma grande diversidade de subtipos acordais acabaram sendo mapeados

nas 24 tríades básicas. Agora com uma matriz tetrádica, as alterações necessárias virão em menor frequência, possibilitando um reconhecimento mais específico, porém ainda distante de cobrir as situações reais. No artigo (LEE; SLANEY, 2006, Cap. 4) é proposta uma matriz de transição 36×36 ao adicionar as tríades diminutas na matriz básica 24×24 . A proposta da matriz tetrádica é seguir nesse aumento de complexidade para obter resultados mais precisos.

6.2 MATRIZ DE EMISSÃO

A matriz de emissão (introduzida na Seção 4.2) pode ser calculada através da similaridade entre um vetor de modelo binário (Figura 13) e um vetor de observação (*vetor cromático*). Tal similaridade pode ser calculada através do produto interno normalizado destes vetores, recapitulado na Equação 6.1.

$$s(x, y) = \frac{\langle x, y \rangle}{\|x\|_2 \cdot \|y\|_2} \quad (6.1)$$

A matriz de emissão para o modelo *template-based* pode ser vista como uma matriz onde cada modelo binário tem probabilidade 1 de emitir o acorde que ele representa. Enquanto isso, a matriz de emissão para o HMM realmente exerce o papel de tentar eliminar ambiguidades harmônicas provenientes das observações. A Figura 19 evidencia que a matriz de emissão pode ser vista analogamente como um grande cromograma para acordes, onde cada célula mostra a probabilidade (de acordo com a energia presente) da observação no instante dado (eixo horizontal) ser um acorde α (eixo vertical).

Explica-se, brevemente, como é calculada a matriz de emissão. Seja $O = (o_1, \dots, o_n)$ uma sequência observada de vetores de cromático e denote por t_i os vetores de modelo ($i = 1, \dots, 24$ no caso de modelos triádicos – Figura 13 – e $i = 1, \dots, 120$ no caso de modelos tetrádicos). A entrada na linha i e coluna n da matriz de emissão é dada pela fórmula na Equação 6.2:

$$\frac{s(o_n, t_i)}{\sum_j s(o_n, t_j)}. \quad (6.2)$$

Dessa forma, a coluna n da matriz de emissão é um vetor cujas entradas modelam a probabilidade da observação o_n estar associada a cada um dos acordes em consideração. A matriz de emissão tem, portanto, a mesma quantidade de linhas que a quantidade de acordes considerados, e a mesma quantidade de colunas que observações de vetores de cromático.



Figura 19 – Matriz de emissão por IIR para um trecho da música *Teresinha* de Chico Buarque.

Fonte: Testes iniciais do (PITA, 2024).

7 APLICAÇÕES À MPB

7.1 EXPERIMENTOS PRELIMINARES

Os experimentos preliminares tiveram por objetivo entender como o algoritmo HMM do (MÜLLER, 2015, Cap. 5) performa em comparação com o algoritmos baseado em modelos (*template-based*) e explorar potenciais locais de aprimoramento. A matriz de transição utilizada foi a matriz uniforme com probabilidade de transição ao mesmo estado de $p = 0.15$. Além disso, tais experimentos também tiveram o intuito de verificar a diferença entre os tipos de cromagrama (STFT, CQT e IIR - Seção 3.4).

A base de dados (Tabela 4) para os testes dessa etapa preliminar compreendeu músicas de diferentes estilos, timbres e qualidade sonora em questões de ruído; isso se deu para observar os limites do algoritmo. Para fins de otimização do tempo, todos os experimentos foram feitos em trechos de 15 segundos de cada uma das músicas presentes na base. Decidiu-se que esta duração é suficiente para entender a estrutura harmônica de uma dada música e observar a performance dos algoritmos de reconhecimento. O limite temporal estipulado para os trechos foi escolhido para agilizar o processo manual de rotulação.

Tabela 4 – Base de dados para os experimentos preliminares.

Banda/Artista	Música
Billie Eilish	<i>everything i wanted</i>
Bruno Mars	<i>When I was your man</i>
Charlie Puth	<i>Attention</i>
Chico Buarque	<i>Teresinha</i>
Coldplay	<i>The scientist</i>
Daniel Caesar & H.E.R.	<i>Best part</i>
Elton John	<i>Your song</i>
Franz Ferdinand	<i>Take me out</i>
Israel Kamakawiwo'ole	<i>Somewhere over the rainbow</i>
Paul McCartney	<i>Maybe I'm amazed</i>
Shawn Mendes	<i>Stitches</i>
Stevie Ray Vaughan	<i>Pride and joy</i>
The Animals	<i>House of the rising sun</i>

A Figura 20 ilustra, a título de exemplo, os três tipos de cromagramas gerados para a música *When I was your man*. Ao analisá-los, pode-se concluir que o STFT acaba gerando mais ruído que as outras duas opções. Isso faz sentido, pois como discutido na Seção 3.4, tanto o CQT quanto o IIR tentam replicar como o ouvido humano funciona, resultando em cromagramas mais claros. Devido à alta incidência de artefatos nos cromagramas gerados através da STFT, o reconhecimento de acordes usando-a não se provou tão eficaz. Por este motivo, daqui em diante, serão mostrados apenas os resultados dos reconhecimentos

usando CQT e IIR. Para a base de dados desse experimento, todos os três tipos de cromagramas são mostrados no Apêndice A; veja também (PITA, 2024).

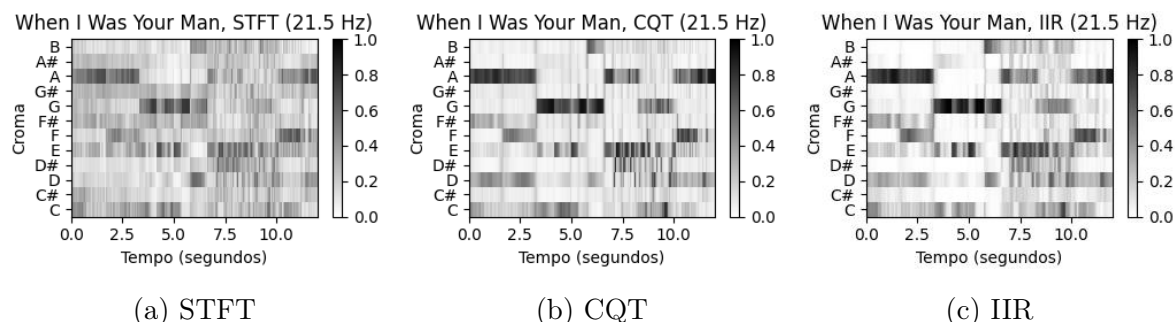
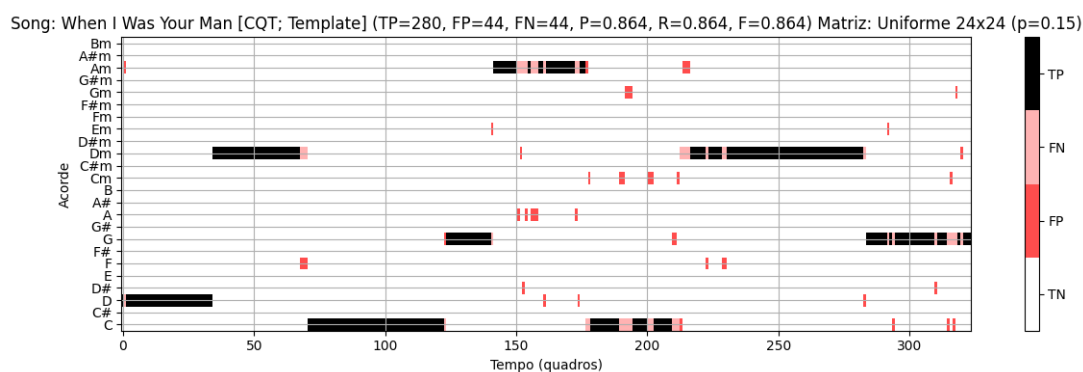


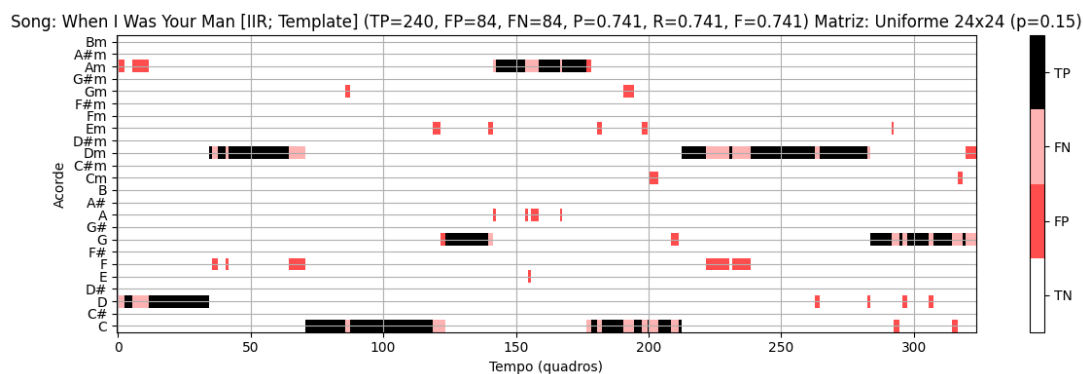
Figura 20 – Cromagramas gerados para a música *When I was your man*.

Fonte: Testes Iniciais do (PITA, 2024).

Por uma questão de brevidade, serão discutidos os resultados da análise somente na música *When I was your man*. Os resultados para as outras músicas da base de dados desse experimento encontram-se no Apêndice B e não serão extensamente discutidos, visto que são semelhantes ao que é apresentado.



(a) Template-based + CQT



(b) Template-based + IIR

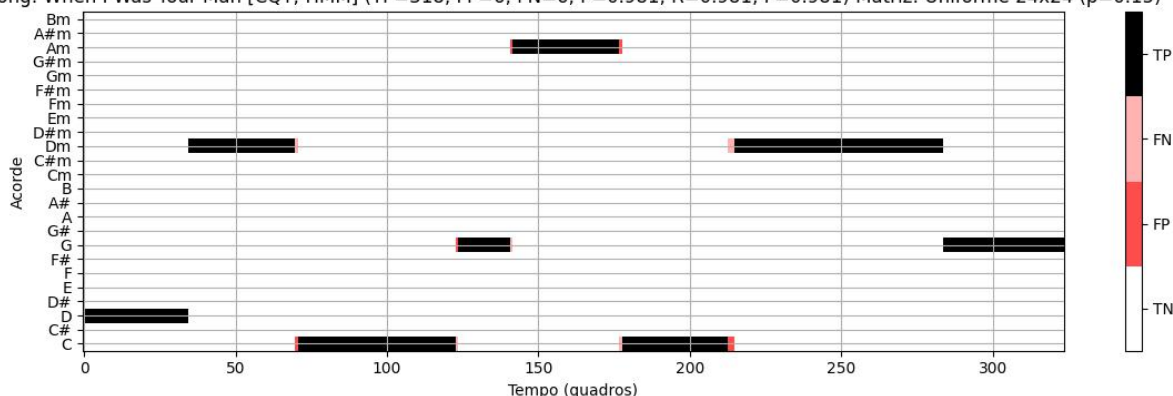
Figura 21 – Resultados do algoritmo *template-based* para a música *When I was your man*.

Fonte: Testes Iniciais do (PITA, 2024).

Comparando o algoritmo baseado em modelos (*template-based* - Seção 5.1) e o baseado em HMM (Seção 5.2), ilustrados nas Figuras 21 e 22, respectivamente, pode-se notar uma melhora do HMM nas três métricas (apresentadas na Seção 5.1.2). Apesar de a matriz de transição usada não ser musicalmente informada, o fato de ela ser uma matriz uniforme com $p > \frac{1}{24}$ faz com os elementos da diagonal tenham um valor maior do que o resto. Isto acarreta uma menor probabilidade de mudança de acordes, ou seja, o modelo opta por mudar de acorde quando de fato há evidência de que isso ocorre. Em contrapartida, o modelo *template-based* age de forma independente a cada instante de tempo, não impondo uma coerência com os instantes anteriores, o que acaba causando breves identificações equivocadas, independentemente do cromagrama ser originado de CQT ou IIR.

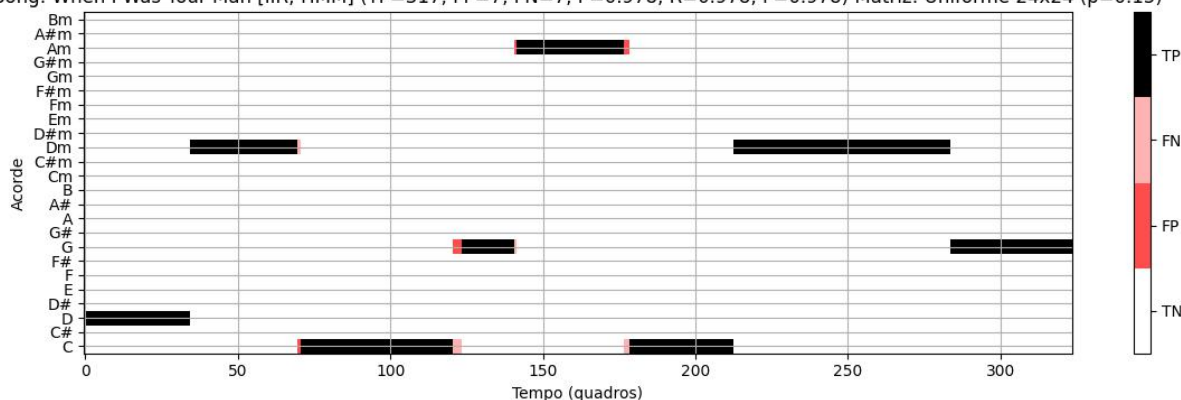
Dito isso, mesmo que o modelo HMM tenha se mostrado mais promissor, vale notar que o *template-based*, apesar da hipótese de independência temporal, apresentou resultados acima das expectativas.

Song: When I Was Your Man [CQT; HMM] (TP=318, FP=6, FN=6, P=0.981, R=0.981, F=0.981) Matriz: Uniforme 24x24 ($p=0.15$)



(a) HMM + CQT

Song: When I Was Your Man [IIR; HMM] (TP=317, FP=7, FN=7, P=0.978, R=0.978, F=0.978) Matriz: Uniforme 24x24 ($p=0.15$)



(b) HMM + IIR

Figura 22 – Resultados do HMM para a música *When I was your man*.

Fonte: Testes Iniciais do (PITA, 2024).

7.2 EXPERIMENTOS NO CONTEXTO DA MPB

Os resultados dos experimentos preliminares permitiram a conclusão de que o STFT é inferior aos outros dois espectrogramas de log-frequência (CQT e IIR) para a geração do cromagrama. Além disso, também foi possível perceber uma vantagem do HMM comparado ao *template-based*. Agora os experimentos serão focados em comparar a performance da matriz uniforme com a de uma matriz musicalmente informada por uma base de dados da MPB (Seção 6.1.2). Um trabalho que não segue a linha MPB, porém também preza por melhorar o reconhecimento de acordes "raros" pode ser encontrado em (BORTOLOZZO; SCHRAMM; JUNG, 2020).

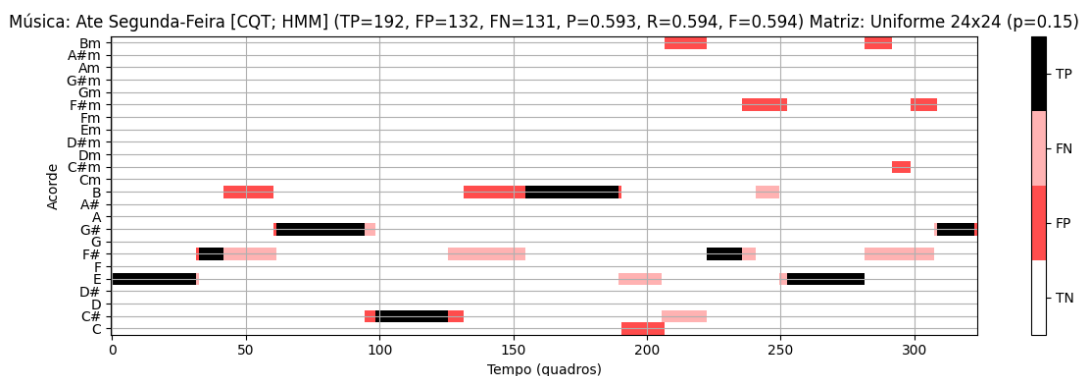
Para melhor avaliar os resultados foi escolhida uma nova base de músicas, em que todas são da MPB. Houve um esforço para garantir que nenhuma fez parte da confecção da matriz de transição informada, como forma de evitar o sobreajuste (*overfitting*) nos resultados. Como evidenciado na Tabela 5, a base conta com duas músicas para cada um dos quatro compositores: Caetano Veloso, Chico Buarque, Ivan Lins e Tom Jobim. Com o intuito de manter o padrão estabelecido nos experimentos preliminares, no contexto da MPB também foi utilizado um limite de 15 segundos para os trechos das músicas da base nova. Para a matriz uniforme foi mantida a probabilidade de transição ao próprio estado $p = 0.15$, porém para a nova matriz, exatamente por não ser uniforme e apresentar uma assinatura única das transições para outros estados, foi optado por começar os testes com $p = 0.5$.

Tabela 5 – Base de dados para os Experimentos no Contexto da MPB.

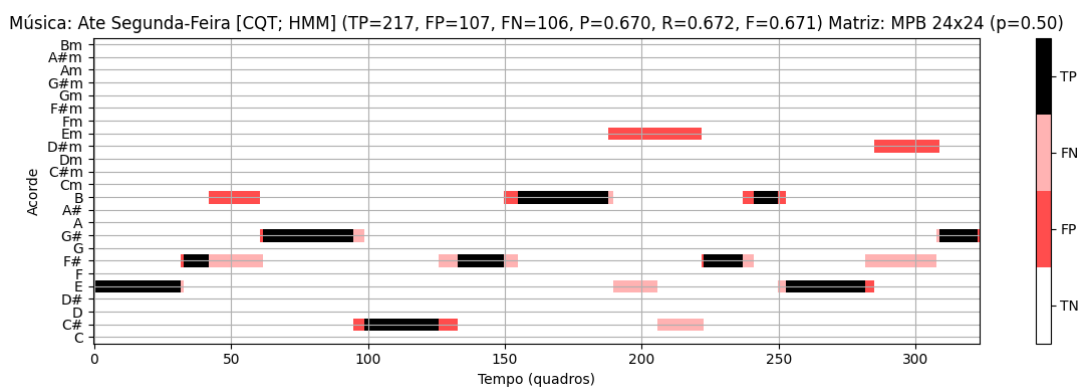
Banda/Artista	Música
Caetano Veloso	<i>Comeu</i>
	<i>Deixa sangrar</i>
Chico Buarque	<i>Até segunda-feira</i>
	<i>Juca</i>
Ivan Lins	<i>Atrás poeira</i>
	<i>Começar de novo</i>
Tom Jobim	<i>As Praias Desertas</i>
	<i>Aula de matemática</i>

Ao comparar a Figura 23a com a Figura 23b, pode-se notar a diferença que a matriz de transição musicalmente informada faz em questão de precisão das decisões de reconhecimento: mesmo nos acordes falsos positivos existe uma estabilidade maior. Isto ocorre pois a matriz MPB foi gerada a partir de músicas com a mesma estética da música analisada, ou seja, essa matriz informa ao modelo quais transições são mais prováveis em um contexto da MPB, resultando em identificações mais precisas. Logo alguns erros cometidos pelo experimento com a matriz uniforme não serão replicados, pois de certa forma contradizem o conteúdo musical da matriz MPB. Um segundo ponto positivo observável

da matriz MPB é que, mesmo nos erros, demonstra mais confiança, gerando um resultado de reconhecimento mais estável. Isto é benéfico, pois pode facilitar enormemente o trabalho de alguém que eventualmente busque corrigir essa anotação automática. Essa característica também é observada nos outros resultados do experimento no contexto da MPB (veja mais detalhes no Apêndice D).



(a) Matriz Uniforme + CQT



(b) Matriz MPB + CQT

Figura 23 – Resultados para a música *Até segunda-feira*.

Fonte: Testes MPB do (PITA, 2024).

7.3 EXEMPLO-SUMÁRIO

Para finalizar com um exemplo-sumário, foi escolhida a música *Um cafuné e um violão*, de Carol Pita, devido à sua progressão harmônica não trivial. Para este último experimento, decidiu-se usar a música completa em vez de apenas um trecho de 15 segundos.

Assim como já observado nos experimentos preliminares e no contexto MPB (Apêndices A e C), o cromagrama gerado a partir do STFT na Figura 24a, comporta uma quantidade maior de ruído em relação aos outros dois (Figuras 24b e 24c).

Comparando os experimentos feitos usando a matriz uniforme na Figura 25 com os experimentos da matriz MPB na Figura 26, é interessante observar que, mesmo com todos resultados estando satisfatórios quanto à estatística F, a matriz musicalmente informada

acaba atrapalhando certos instantes do reconhecimento, como os momentos em que o acorde $D\sharp m$ é tocado. Isso mesmo depois fazer testes para a matriz MPB com p variando de $[0, 1]$ com passo 0.02 para encontrar o melhor p .

Uma possível explicação para a matriz uniforme ter se saído melhor nestes experimentos é que a progressão harmônica da música apresenta uma estética diferente do que seria encontrado nas músicas MPB em geral, com a estrofe: $B\flat m - F - B\flat m - F - B\flat m - G\flat - C7 - F$; e o refrão: $B\flat m - B\flat m - G\flat - F - E\flat m - E\flat m/D\flat - C7 - F$. Isso faz com que a matriz MPB acabe não esperando certas transições, enquanto a uniforme, por definição, não discrimina as transições com exceção da diagonal de transições para o mesmo estado.

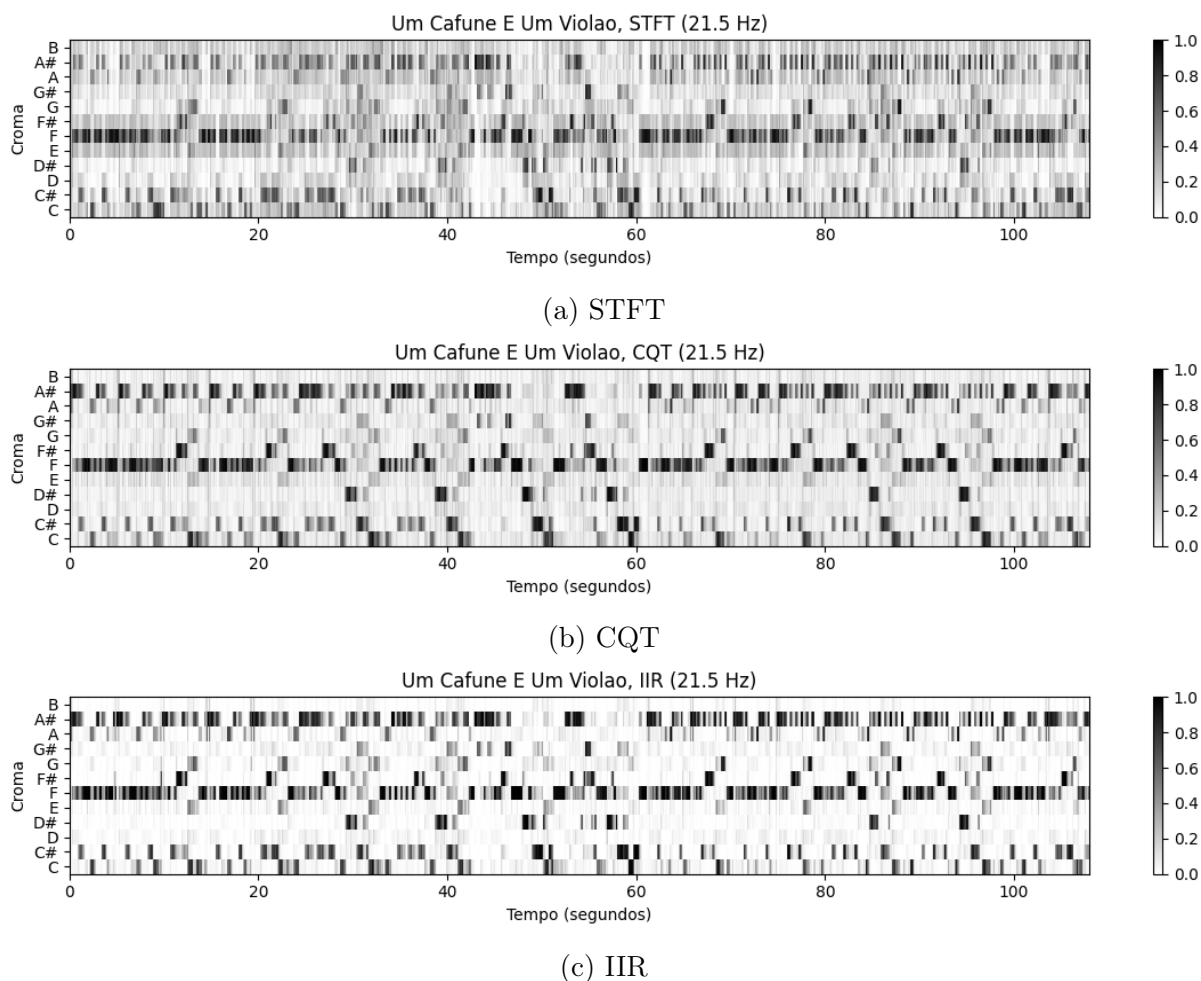
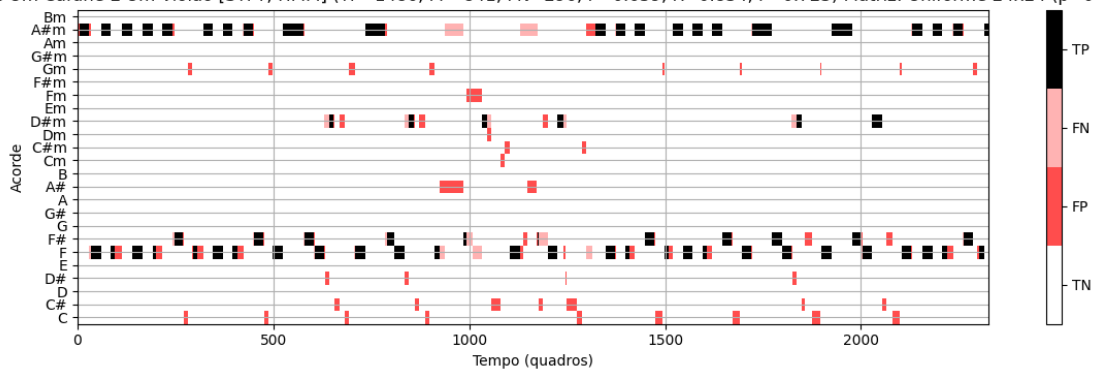


Figura 24 – Cromagramas gerados para a música de Carol Pita: *Um cafuné e um violão*.

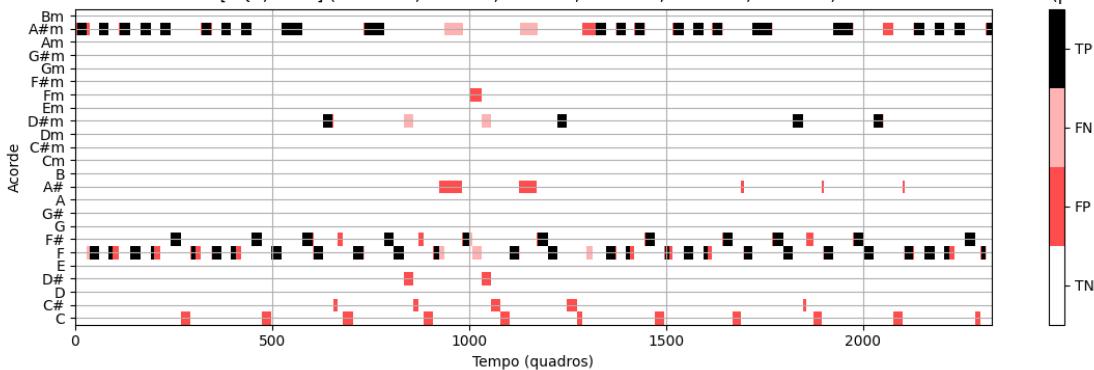
Fonte: Testes finais do (PITA, 2024).

Música: Um Cafune E Um Violao [STFT; HMM] (TP=1486, FP=841, FN=296, P=0.639, R=0.834, F=0.723) Matriz: Uniforme 24x24 ($p=0.15$)



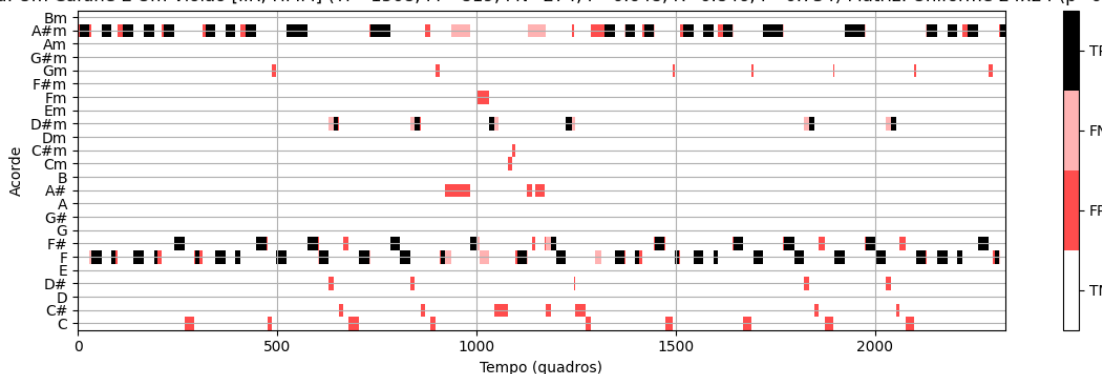
(a) STFT

Música: Um Cafune E Um Violao [CQT; HMM] (TP=1535, FP=792, FN=247, P=0.660, R=0.861, F=0.747) Matriz: Uniforme 24x24 ($p=0.15$)



(b) CQT

Música: Um Cafune E Um Violao [IIR; HMM] (TP=1508, FP=819, FN=274, P=0.648, R=0.846, F=0.734) Matriz: Uniforme 24x24 ($p=0.15$)

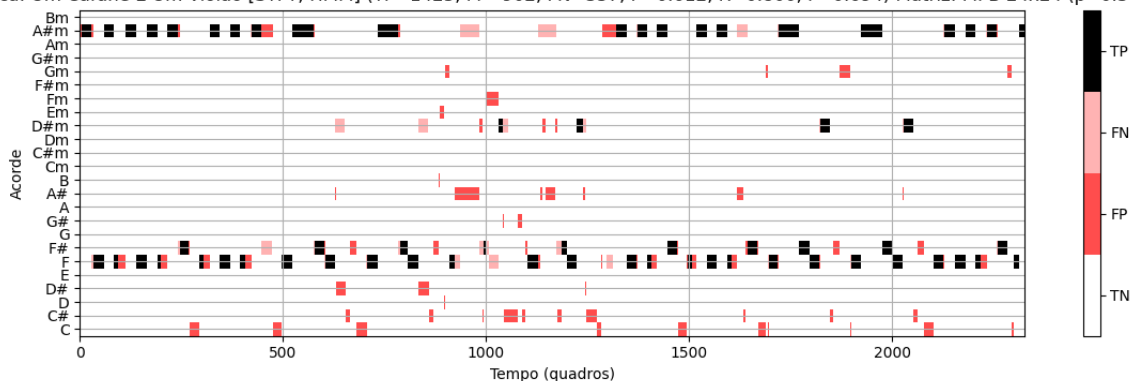


(c) IIR

Figura 25 – Resultados do reconhecimento de acordes com a matriz uniforme para a música de Carol Pita: *Um cafuné e um violão*.

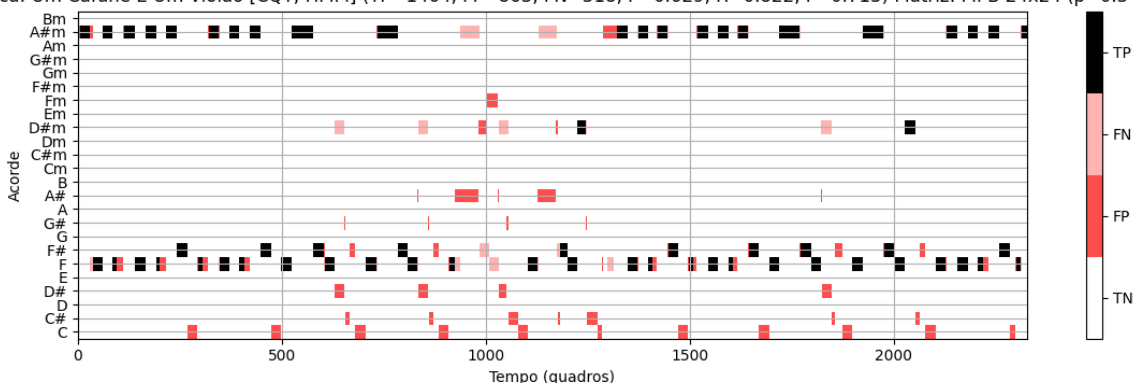
Fonte: Testes finais do (PITA, 2024).

Música: Um Cafune E Um Violao [STFT; HMM] (TP=1425, FP=902, FN=357, P=0.612, R=0.800, F=0.694) Matriz: MPB 24x24 (p=0.34)



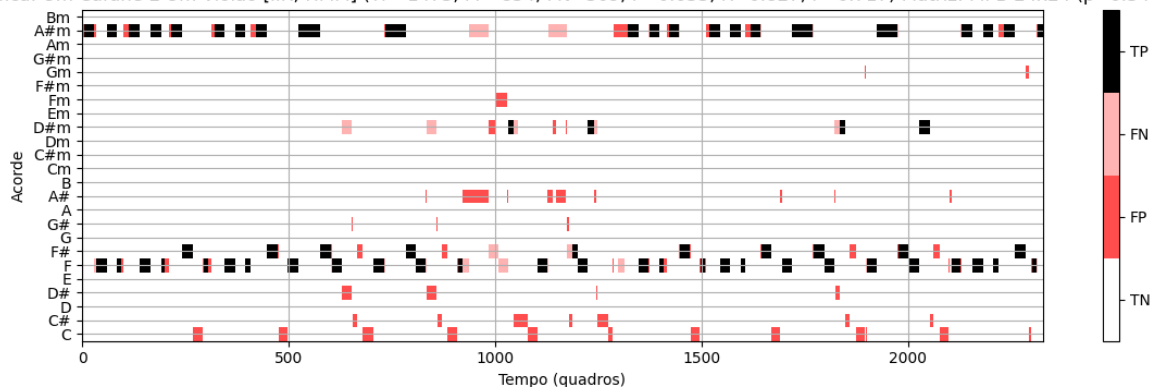
(a) STFT

Música: Um Cafune E Um Violao [CQT; HMM] (TP=1464, FP=863, FN=318, P=0.629, R=0.822, F=0.713) Matriz: MPB 24x24 (p=0.34)



(b) CQT

Música: Um Cafune E Um Violao [IIR; HMM] (TP=1473, FP=854, FN=309, P=0.633, R=0.827, F=0.717) Matriz: MPB 24x24 (p=0.34)



(c) IIR

Figura 26 – Resultados do reconhecimento de acordes com a matriz MPB 24×24 para a música de Carol Pita: *Um cafuné e um violão*.

Fonte: Testes finais do (PITA, 2024).

8 CONCLUSÃO E PROJETOS FUTUROS

Este projeto apresentou experimentos de identificação de acordes através de duas abordagens principais, *template-based* e HMM. Para conseguir falar destes modelos e avaliar os resultados foi crucial entender os conceitos de construção de acorde e estrutura harmônica passados no Capítulo 2. Além disso, a base para o projeto foi estabelecida no Capítulo 3 com a introdução de transformadas de Fourier e conseqüentemente as transformadas de Fourier de tempo curto (STFT), que por sua vez tornou possível a apresentação do cromagrama, ferramenta fundamental para o objetivo do projeto.

Chegando nos Capítulos 4 e 5, foram evidenciadas as ideias de cadeias de Markov, e mais importante, modelos ocultos de Markov (HMM), o coração do projeto. Não apenas essas ideias, mas como aplicar o HMM ao problema da identificação de acordes. O Capítulo 6 então direcionou a proposta do projeto através do refinamento das matrizes do HMM, fazendo o modelo ser musicalmente informado. Além disso, foi proposta uma expansão da matriz triádica para uma que contemple as tétrades também, aumentando a complexidade das possíveis transições.

O Capítulo 7 trouxe três grandes etapas de experimentos: preliminares, no contexto da MPB e um exemplo-sumário. Como mencionado no capítulo, os experimentos preliminares serviram para entender os funcionamentos do modelo HMM disponibilizado por (MÜLLER, 2015) e como ele se compara em relação ao modelo *template-based*. Além disso, estes experimentos ajudaram na percepção de performance dos tipos de cromagrama (STFT, CQT e IIR). Essa primeira etapa contou com músicas de gêneros variados e a partir dela conclui-se que o modelo HMM de fato performa melhor que o *template-based*, apesar do segundo também apresentar resultados acima das expectativas. Outro dado tirado desta leva de experimentos foi que o STFT se provou menos eficaz, dada sua alta incidência de ruídos em relação aos outros dois tipos de cromagrama. Chegando ao segundo tópico do capítulo, foram feitos os experimentos no contexto da MPB com músicas de Caetano Veloso, Chico Buarque, Ivan Lins e Tom Jobim. Os experimentos focaram em comparar a performance da identificação usando a matriz uniforme e usando a matriz musicalmente informada pela base do Projeto MPB (ALMADA; CARVALHO, 2024). Os resultados mostraram que o modelo musicalmente informado foi superior em relação ao uniforme e que mesmo nos momentos em que o modelo errou, demonstrou mais confiança no erro, gerando uma estabilidade maior no reconhecimento. Para finalizar, foi feito um último experimento com a música *Um cafuné e um violão* de Carol Pita, para sumarizar os resultados. O resultado do teste vai contra a expectativa, pelo modelo uniforme ter performado melhor do que o modelo MPB. Porém uma possível explicação vem da progressão não trivial da música, que apresenta transições de acordes inesperadas no contexto da MPB, fazendo com que a informação musical do modelo acabe prejudicando

o reconhecimento.

Tendo em vista tudo o que foi analisado, o projeto teve sucesso em implementar um identificador de acordes por meio de um HMM musicalmente informado, porém ainda há margem para melhora. Foram feitos alguns testes preliminares com a matriz tetrádica, todavia não houve tempo de colocar os resultados nesta versão do projeto. Um ponto intermediário de melhora seria o aumentar a complexidade do espaço de estados em menor escala, abrangendo acordes diminutos e aumentados, contemplando todas as tríades. Outro ponto que pode ser implementado para observar uma melhora no reconhecimento é no pré-processamento das músicas, removendo por meio de filtros a percussão presente, a voz e outros ruídos, como forma de deixar o áudio mais limpo e claro. Por fim, uma alteração possível de trazer resultados mais consistentes seria a de estabelecer um limite temporal dinâmico para o trecho de música analisado. Dinâmico no sentido de estar atrelado ao BPM (batidas por minuto) da música em questão. Partindo da suposição de que a transição de acordes ocorre comumente a cada compasso, quatro compassos seria o suficiente para descobrir a progressão da maioria das músicas pop, por exemplo. Caso a música esteja em 80 BPM, então quatro segundos já seria o suficiente para o limite temporal, poupando tempo de rotulação.

REFERÊNCIAS

- ALMADA, C. **Harmonia funcional**. Campinas: Editora da Unicamp, 2009.
- ALMADA, C.; CARVALHO, H. **Projeto MPB**. 2024. <https://projetompb.com.br>.
- BORTOLOZZO, M.; SCHRAMM, R.; JUNG, C. Improving the classification of rare chords with unlabeled data. 12 2020.
- BRAGA, J. **Uma aplicação de Teoria de Grupos em Teoria Musical**. 2020. TRABALHO DE CONCLUSÃO DE CURSO.
- COSTA, M. da. **Novel Time-frequency Representations For Music Information Retrieval**. Tese (Doutorado) — UFRJ, COPPE, 2020.
- HURON, D. **Sweet Anticipation: Music and the Psychology of Expectation**. Cambridge: The MIT Press, 2006. v. 1.
- LEE, K.; SLANEY, M. Automatic Chord Recognition from Audio Using a HMM with Supervised Learning. p. 133–137, 01 2006.
- LI, T. Study on a cnn-hmm approach for audio-based musical chord recognition. **Journal of Physics: Conference Series**, v. 1802, p. 032033, 03 2021.
- LOY, G. **Musimathics: The Mathematical Foundations of Music, Volume 1**. Cambridge: The MIT Press, 2006.
- MED, B. **Teoria da música**. Brasília: MusiMed, 1996.
- MÜLLER, M. **Information Retrieval for Music and Motion**. New York City: Springer, 2007.
- MÜLLER, M. **Fundamentals of Music Processing**. New York City: Springer, 2015.
- PITA, R. **Repositório de códigos: Identificação de Acordes via Modelos de Markov Ocultos Musicalmente Informados**. 2024. <https://github.com/RodrigoPita/TCC>.
- RAMEAU, J.-P. **Traité de l’harmonie réduite à ses principes naturels**. [S.l.]: J. B. C. Ballard, 1722.
- ROCHA, N. **ESTATÍSTICA E PROBABILIDADE**. Rio de Janeiro, Instituto de Matemática - UFRJ: Distribuído pelo próprio autor, 2020.
- SHEPARD, R. N. Circularity in judgments of relative pitch. **Journal of the Acoustical Society of America**, v. 36, p. 2346–2353, 1964. Disponível em: <https://api.semanticscholar.org/CorpusID:121024452>.
- WRIGHT, D. **Mathematics and music**. Providence, Rhode Island: American Mathematical Soc., 2009. v. 28.

APÊNDICE A – CROMAGRAMAS GERADOS PARA OS EXPERIMENTOS PRELIMINARES.

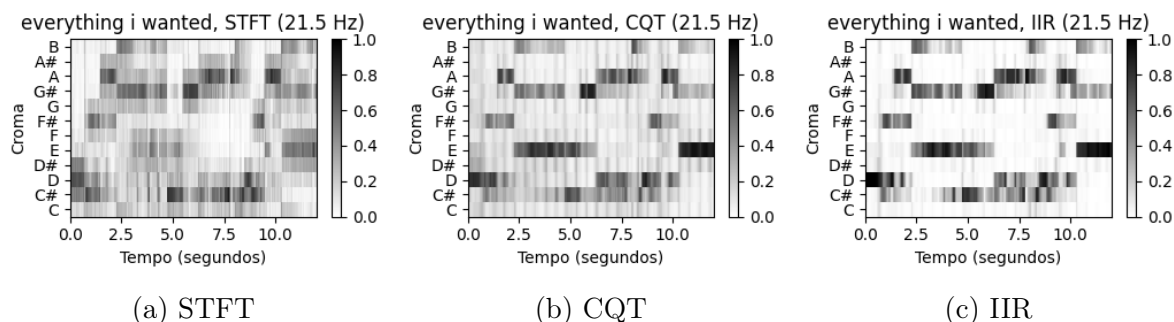


Figura 27 – Cromagramas gerados para a música *everything i wanted*.

Fonte: Testes Iniciais do (PITA, 2024).

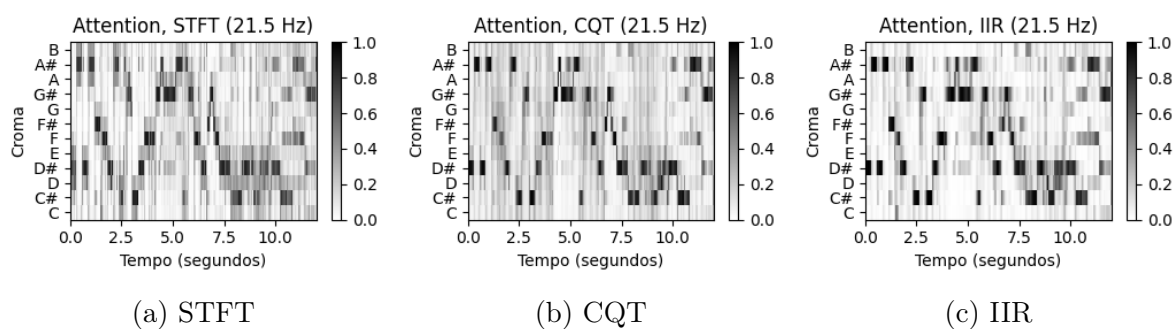


Figura 28 – Cromagramas gerados para a música *Attention*.

Fonte: Testes Iniciais do (PITA, 2024).

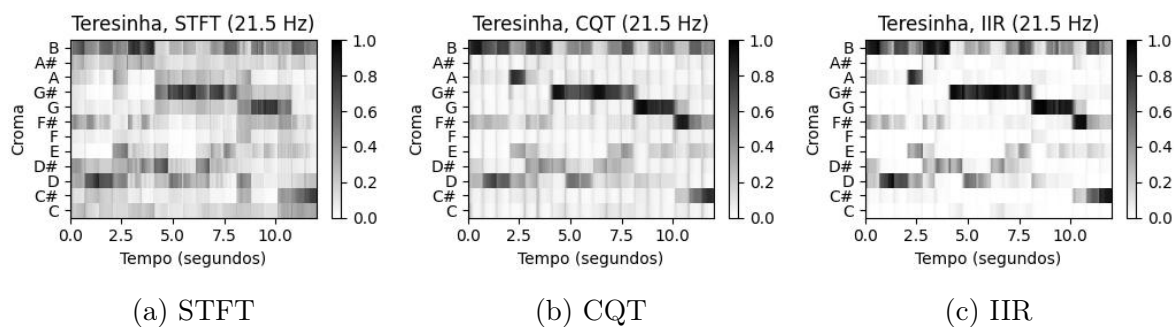


Figura 29 – Cromagramas gerados para a música *Teresinha*.

Fonte: Testes Iniciais do (PITA, 2024).

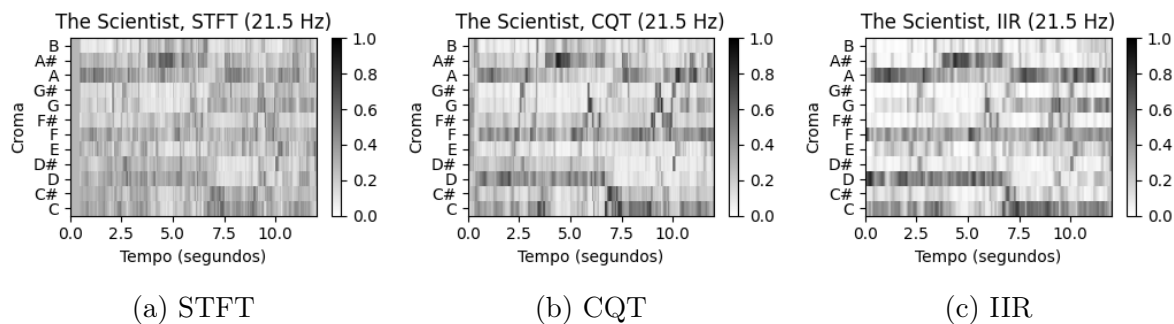


Figura 30 – Cromagramas gerados para a música *The scientist*.

Fonte: Testes Iniciais do (PITA, 2024).

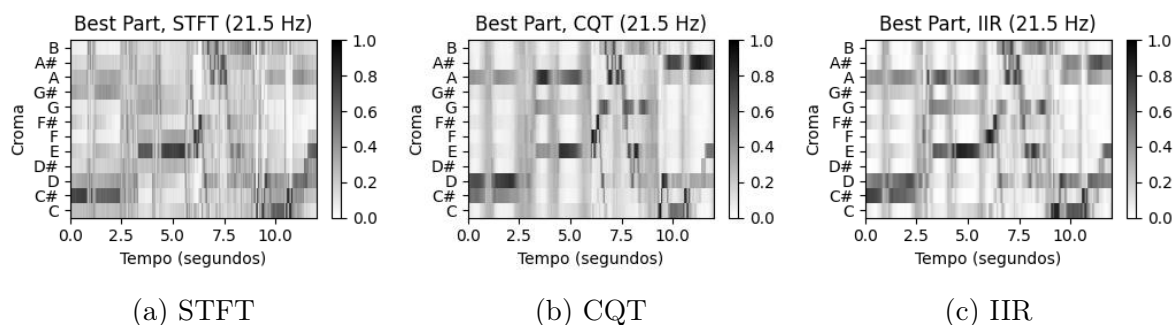


Figura 31 – Cromagramas gerados para a música *Best part*.

Fonte: Testes Iniciais do (PITA, 2024).

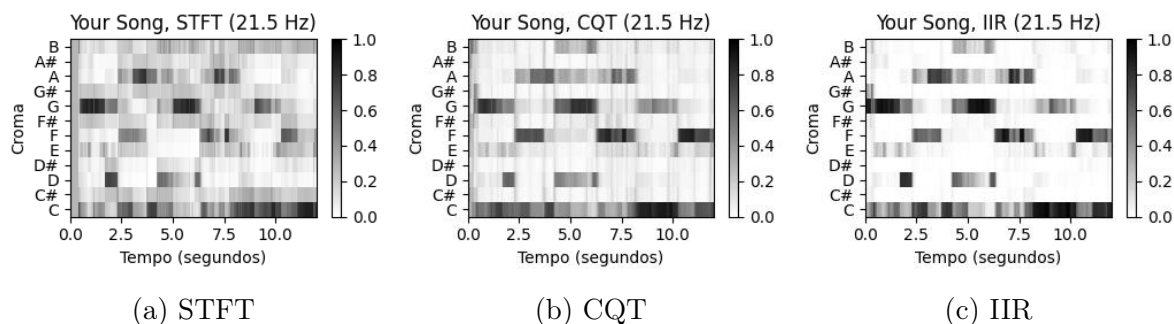


Figura 32 – Cromagramas gerados para a música *Your song*.

Fonte: Testes Iniciais do (PITA, 2024).

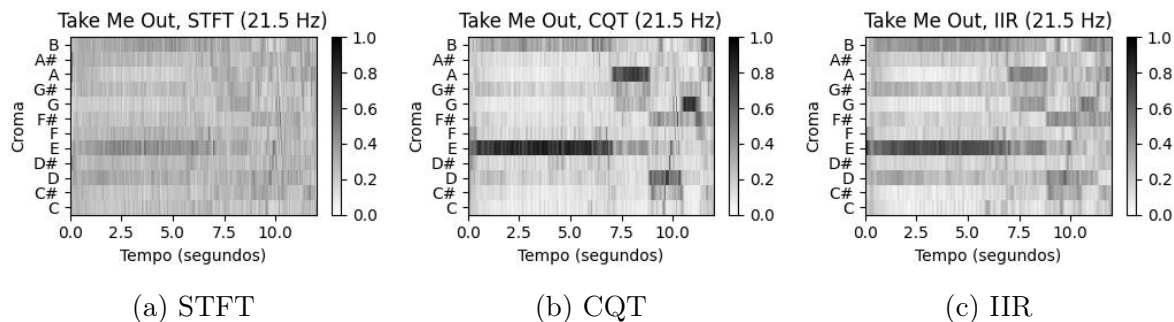


Figura 33 – Cromagramas gerados para a música *Take me out*.

Fonte: Testes Iniciais do (PITA, 2024).

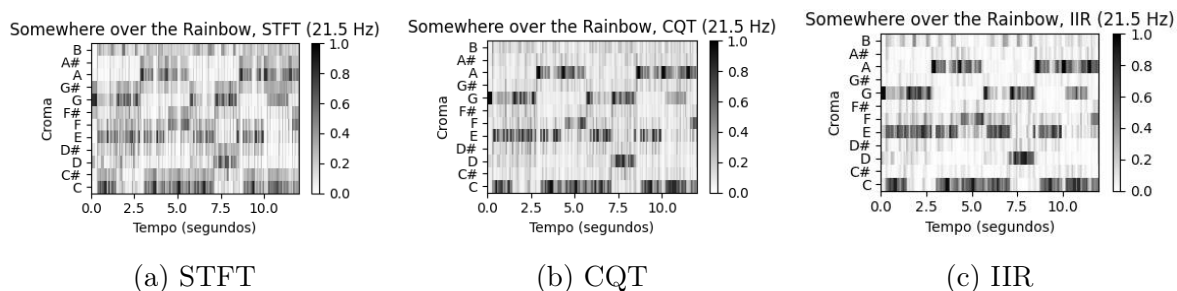


Figura 34 – Cromagramas gerados para a música *Somewhere over the rainbow*.

Fonte: Testes Iniciais do (PITA, 2024).

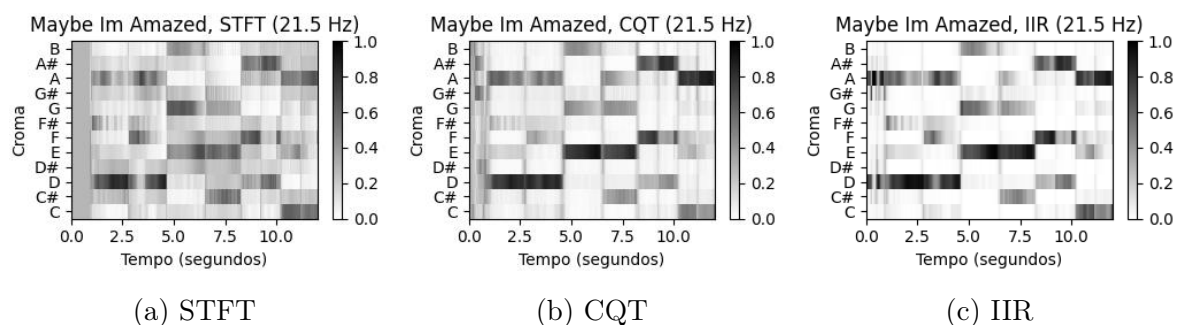


Figura 35 – Cromagramas gerados para a música *Maybe I'm amazed*.

Fonte: Testes Iniciais do (PITA, 2024).

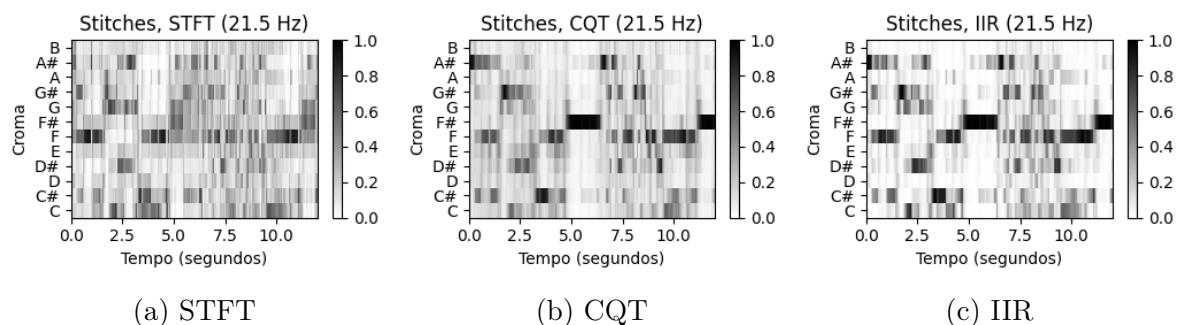


Figura 36 – Cromagramas gerados para a música *Stitches*.

Fonte: Testes Iniciais do (PITA, 2024).

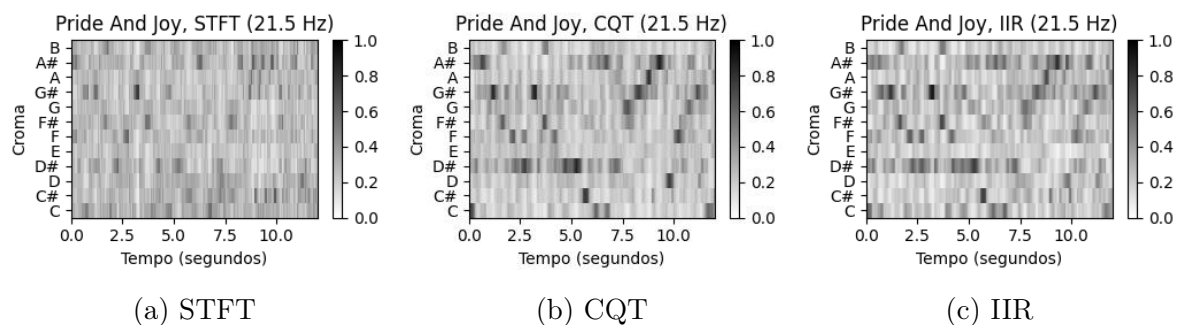


Figura 37 – Cromagramas gerados para a música *Pride and joy*.

Fonte: Testes Iniciais do (PITA, 2024).

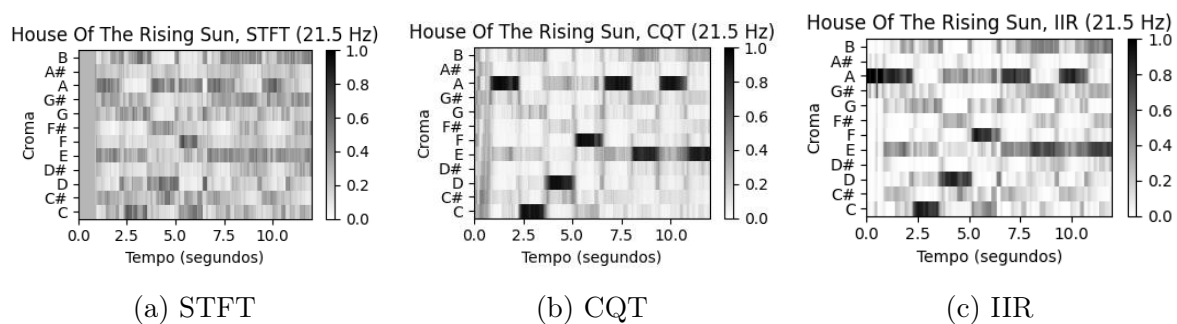
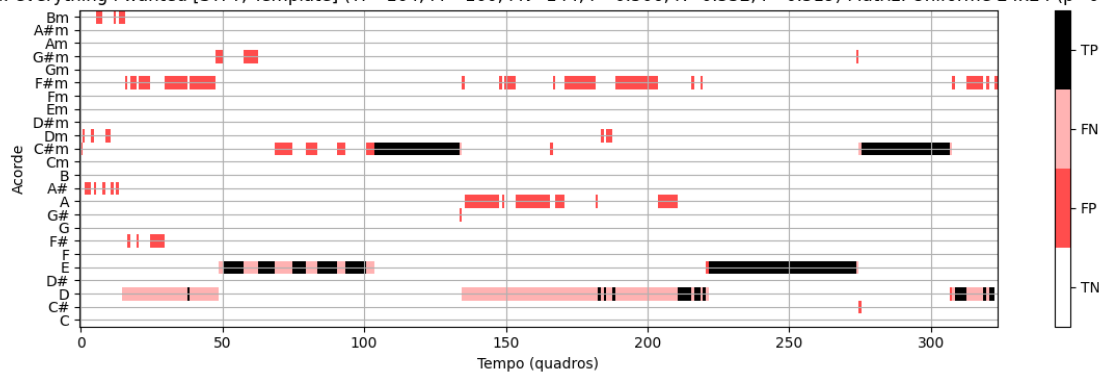


Figura 38 – Cromagramas gerados para a música *House of the rising sun*.

Fonte: Testes Iniciais do (PITA, 2024).

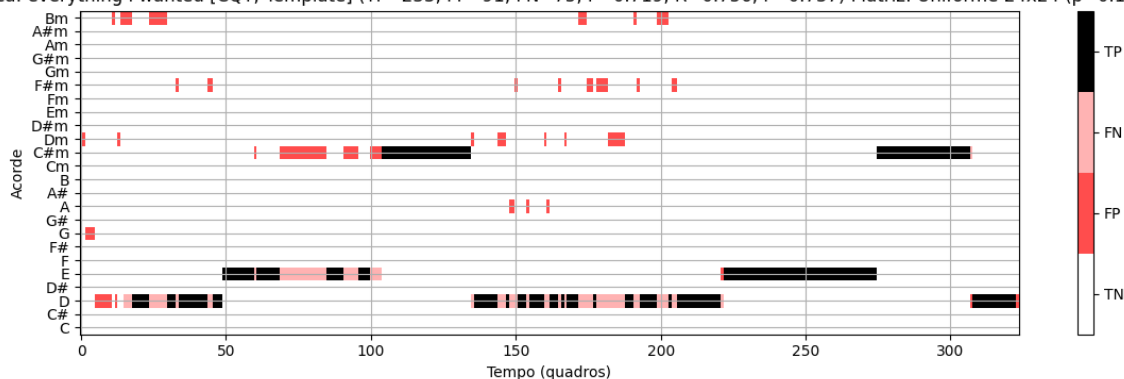
APÊNDICE B – RESULTADOS DOS EXPERIMENTOS PRELIMINARES.

Música: everything i wanted [STFT; Template] (TP=164, FP=160, FN=144, P=0.506, R=0.532, F=0.519) Matriz: Uniforme 24x24 ($p=0.15$)



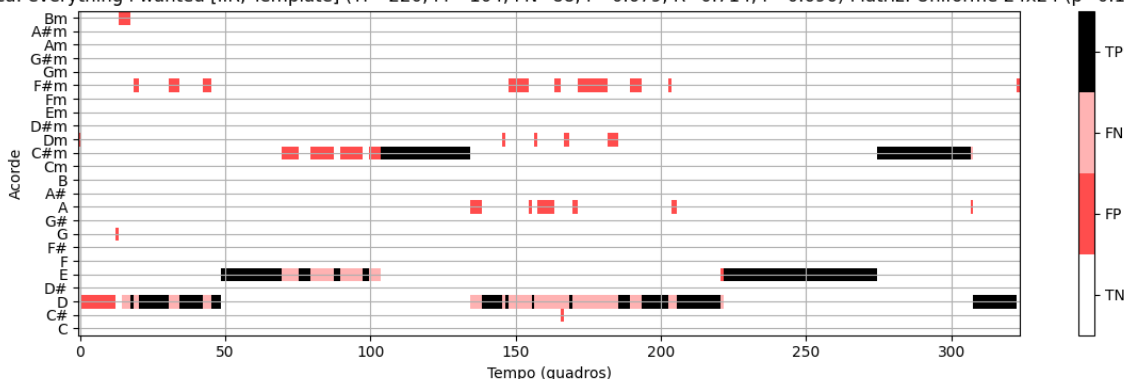
(a) Template-based + STFT

Música: everything i wanted [CQT; Template] (TP=233, FP=91, FN=75, P=0.719, R=0.756, F=0.737) Matriz: Uniforme 24x24 ($p=0.15$)



(b) Template-based + CQT

Música: everything i wanted [IIR; Template] (TP=220, FP=104, FN=88, P=0.679, R=0.714, F=0.696) Matriz: Uniforme 24x24 ($p=0.15$)

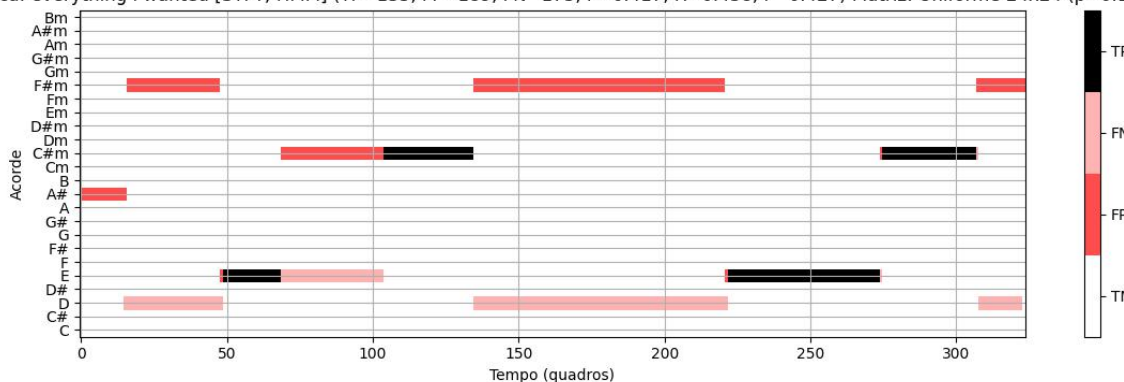


(c) Template-based + IIR

Figura 39 – Resultados do Template-based para a música *everything i wanted*.

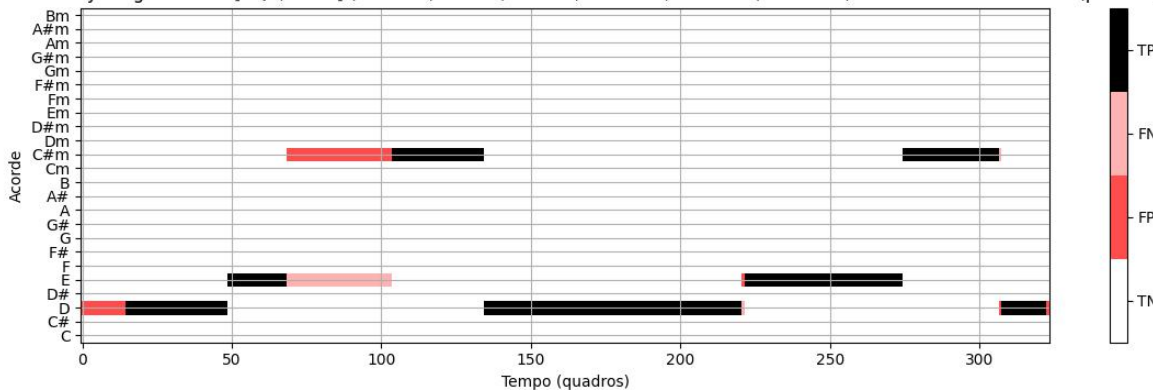
Fonte: Testes Iniciais do (PITA, 2024).

Música: everything i wanted [STFT; HMM] (TP=135, FP=189, FN=173, P=0.417, R=0.438, F=0.427) Matriz: Uniforme 24x24 ($p=0.15$)



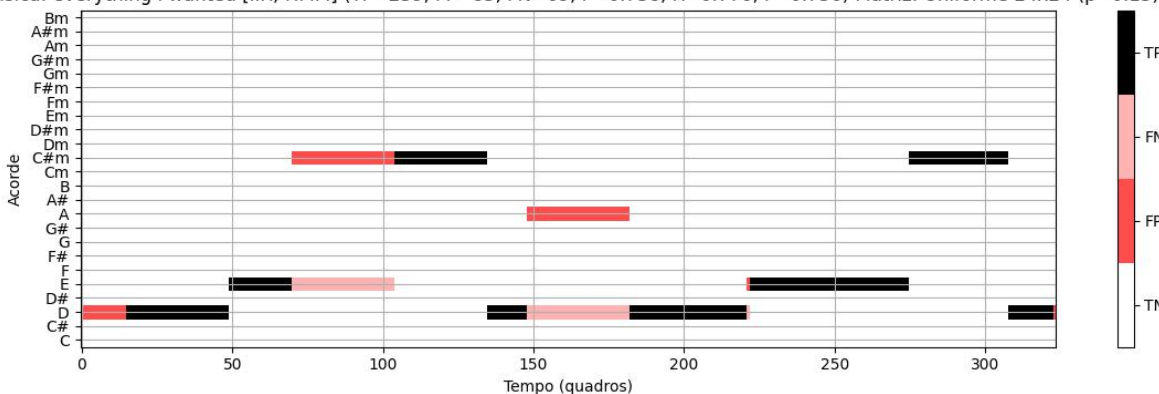
(a) HMM + STFT

Música: everything i wanted [CQT; HMM] (TP=271, FP=53, FN=37, P=0.836, R=0.880, F=0.858) Matriz: Uniforme 24x24 ($p=0.15$)



(b) HMM + CQT

Música: everything i wanted [IIR; HMM] (TP=239, FP=85, FN=69, P=0.738, R=0.776, F=0.756) Matriz: Uniforme 24x24 ($p=0.15$)

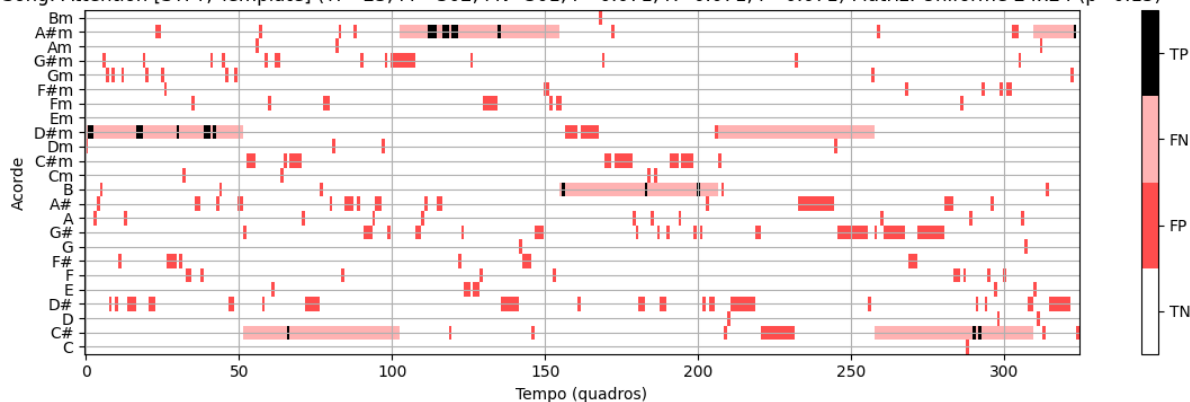


(c) HMM + IIR

Figura 40 – Resultados do HMM para a música *everything i wanted*.

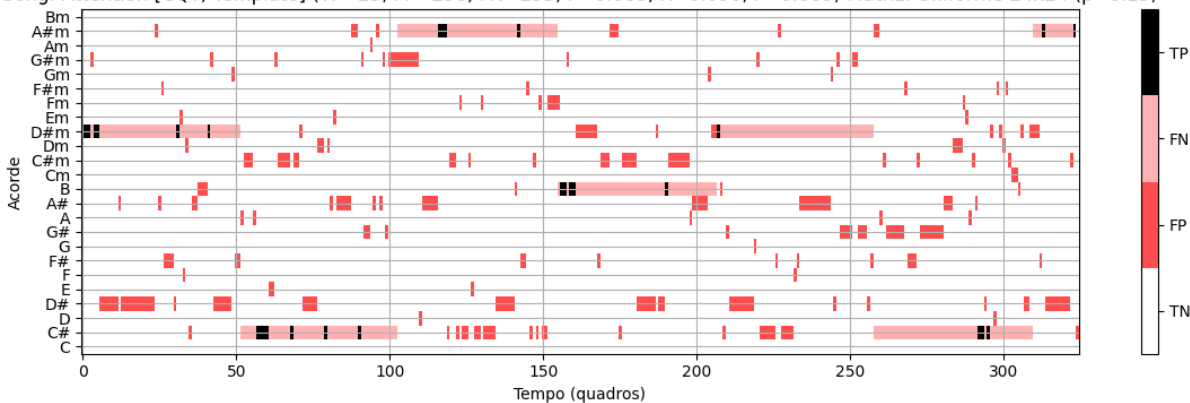
Fonte: Testes Iniciais do (PITA, 2024).

Song: Attention [STFT; Template] (TP=23, FP=302, FN=301, P=0.071, R=0.071, F=0.071) Matriz: Uniforme 24x24 ($p=0.15$)



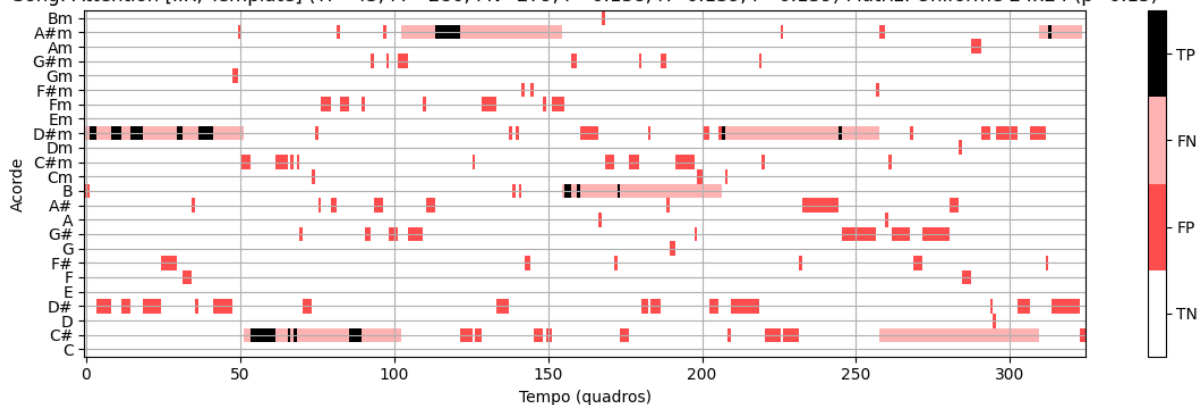
(a) Template-based + STFT

Song: Attention [CQT; Template] (TP=29, FP=296, FN=295, P=0.089, R=0.090, F=0.089) Matriz: Uniforme 24x24 ($p=0.15$)



(b) Template-based + CQT

Song: Attention [IIR; Template] (TP=45, FP=280, FN=279, P=0.138, R=0.139, F=0.139) Matriz: Uniforme 24x24 ($p=0.15$)

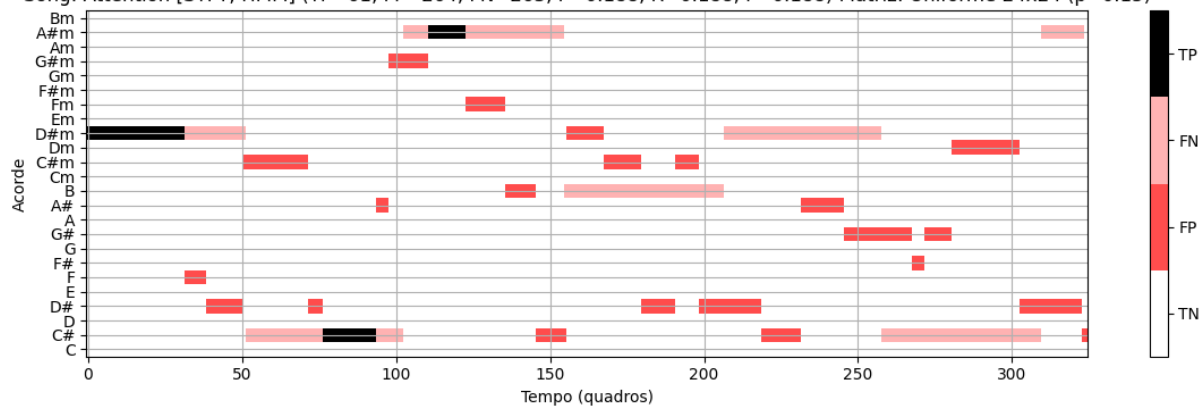


(c) Template-based + IIR

Figura 41 – Resultados do Template-based para a música *Attention*.

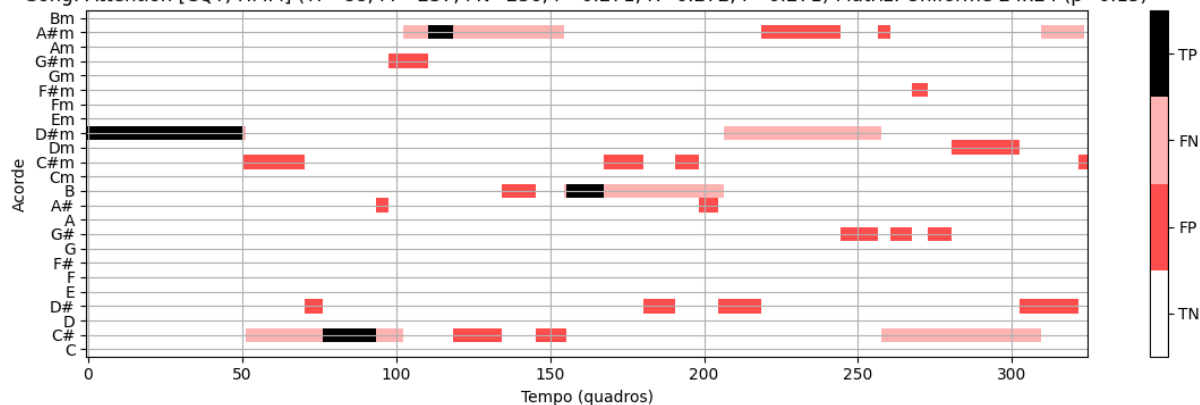
Fonte: Testes Iniciais do (PITA, 2024).

Song: Attention [STFT; HMM] (TP=61, FP=264, FN=263, P=0.188, R=0.188, F=0.188) Matriz: Uniforme 24x24 ($p=0.15$)



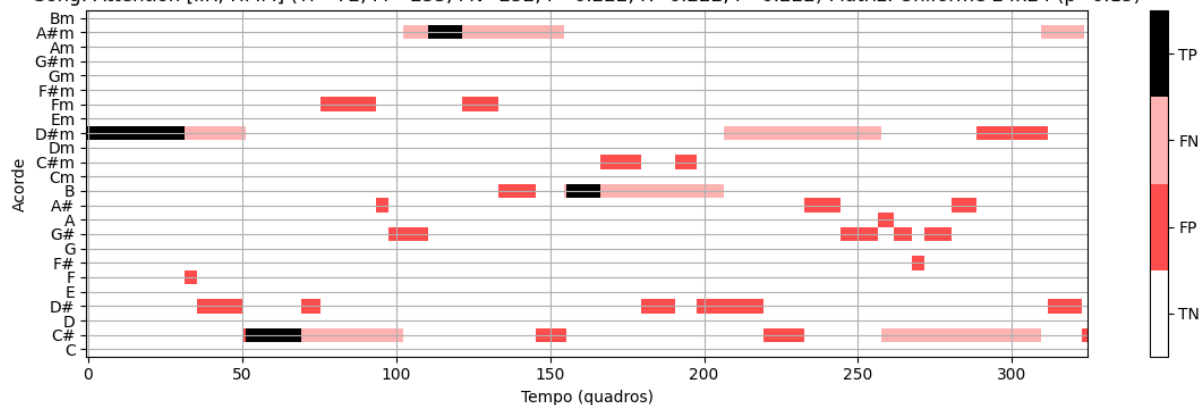
(a) HMM + STFT

Song: Attention [CQT; HMM] (TP=88, FP=237, FN=236, P=0.271, R=0.272, F=0.271) Matriz: Uniforme 24x24 ($p=0.15$)



(b) HMM + CQT

Song: Attention [IIR; HMM] (TP=72, FP=253, FN=252, P=0.222, R=0.222, F=0.222) Matriz: Uniforme 24x24 ($p=0.15$)

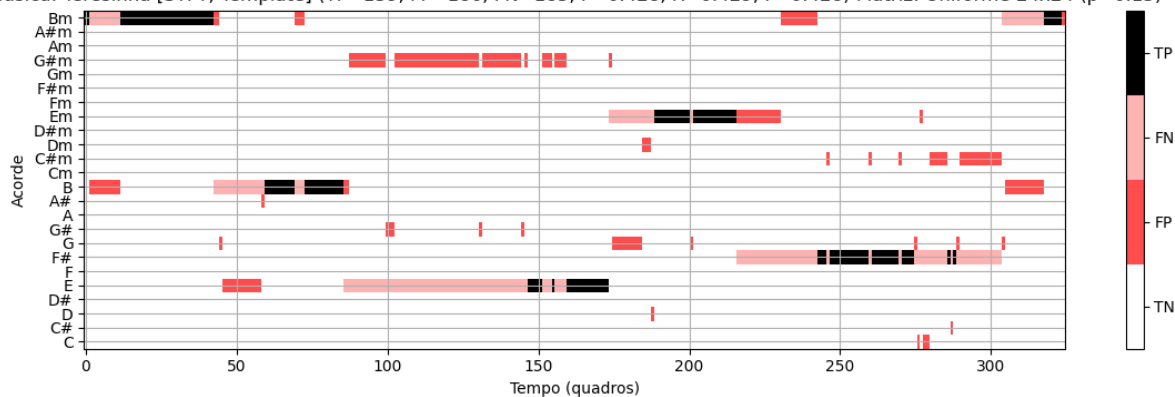


(c) HMM + IIR

Figura 42 – Resultados do HMM para a música *Attention*.

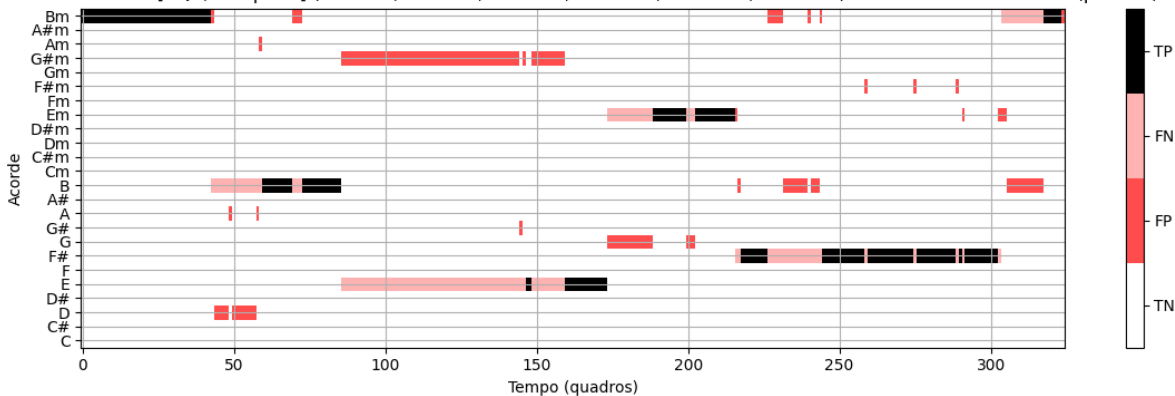
Fonte: Testes Iniciais do (PITA, 2024).

Música: Teresinha [STFT; Template] (TP=139, FP=186, FN=185, P=0.428, R=0.429, F=0.428) Matriz: Uniforme 24x24 ($p=0.15$)



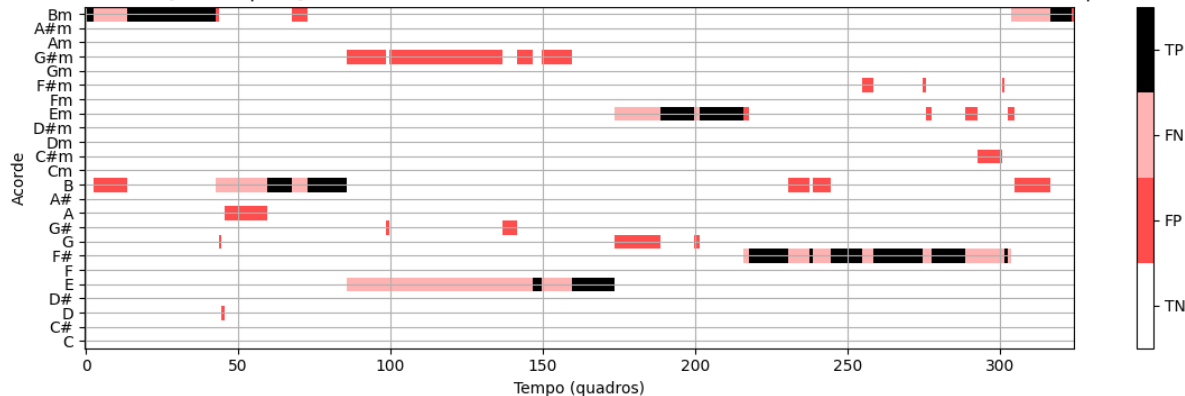
(a) Template-based + STFT

Música: Teresinha [CQT; Template] (TP=175, FP=150, FN=149, P=0.538, R=0.540, F=0.539) Matriz: Uniforme 24x24 ($p=0.15$)



(b) Template-based + CQT

Música: Teresinha [IIR; Template] (TP=154, FP=171, FN=170, P=0.474, R=0.475, F=0.475) Matriz: Uniforme 24x24 ($p=0.15$)

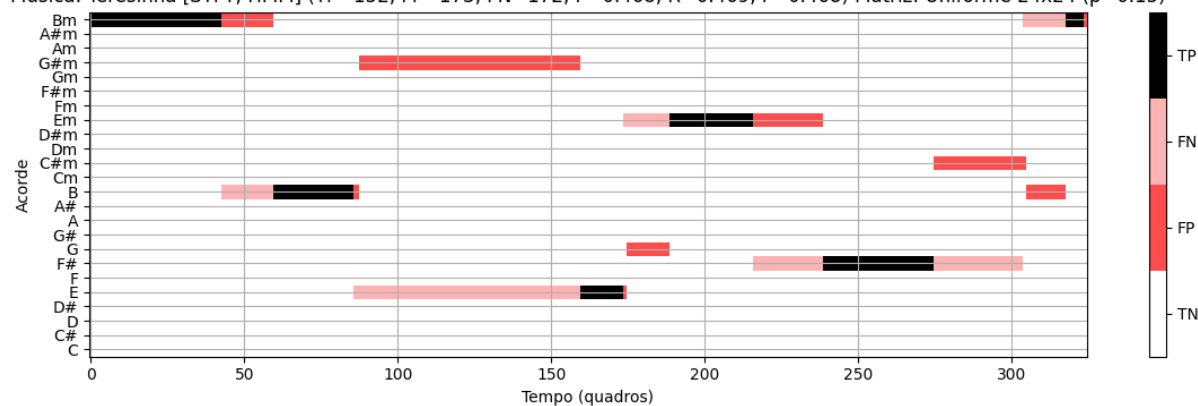


(c) Template-based + IIR

Figura 43 – Resultados do Template-based para a música *Teresinha*.

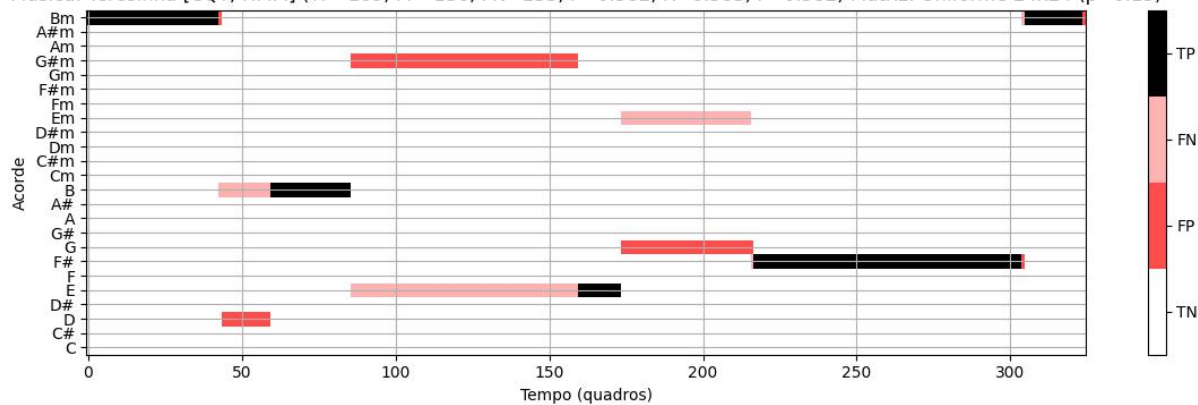
Fonte: Testes Iniciais do (PITA, 2024).

Música: Teresinha [STFT; HMM] (TP=152, FP=173, FN=172, P=0.468, R=0.469, F=0.468) Matriz: Uniforme 24x24 ($p=0.15$)



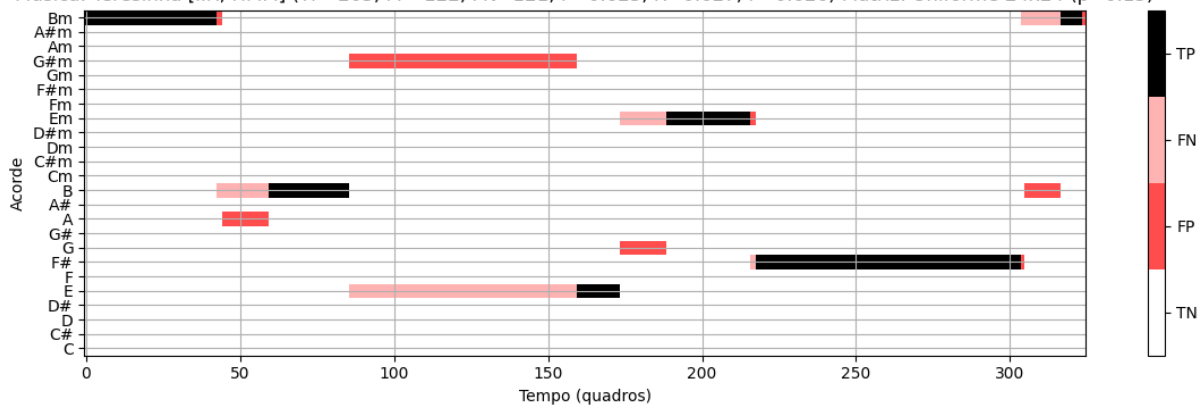
(a) HMM + STFT

Música: Teresinha [CQT; HMM] (TP=189, FP=136, FN=135, P=0.582, R=0.583, F=0.582) Matriz: Uniforme 24x24 ($p=0.15$)



(b) HMM + CQT

Música: Teresinha [IIR; HMM] (TP=203, FP=122, FN=121, P=0.625, R=0.627, F=0.626) Matriz: Uniforme 24x24 ($p=0.15$)

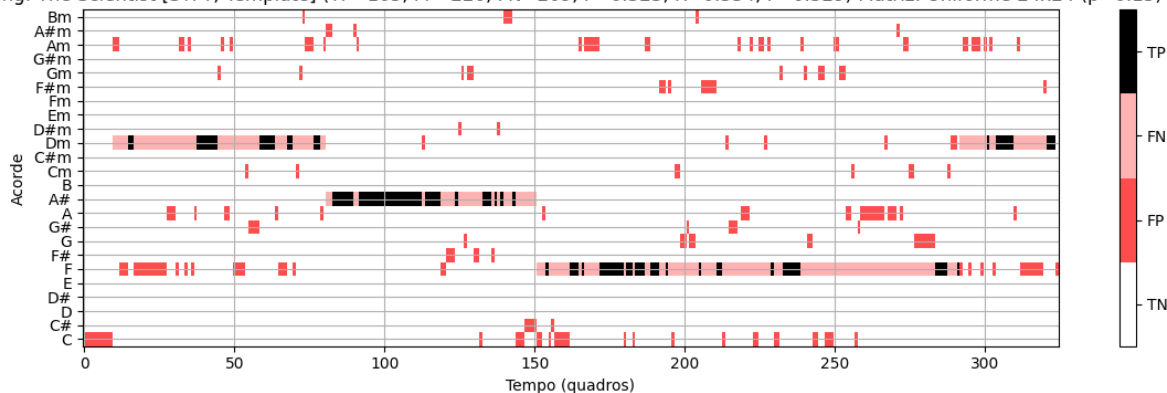


(c) HMM + IIR

Figura 44 – Resultados do HMM para a música *Teresinha*.

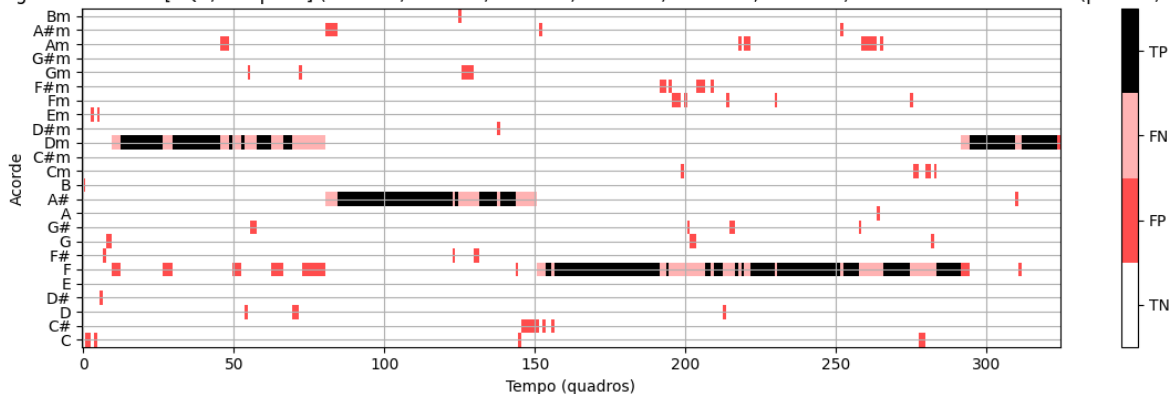
Fonte: Testes Iniciais do (PITA, 2024).

Song: The Scientist [STFT; Template] (TP=105, FP=220, FN=209, P=0.323, R=0.334, F=0.329) Matriz: Uniforme 24x24 ($p=0.15$)



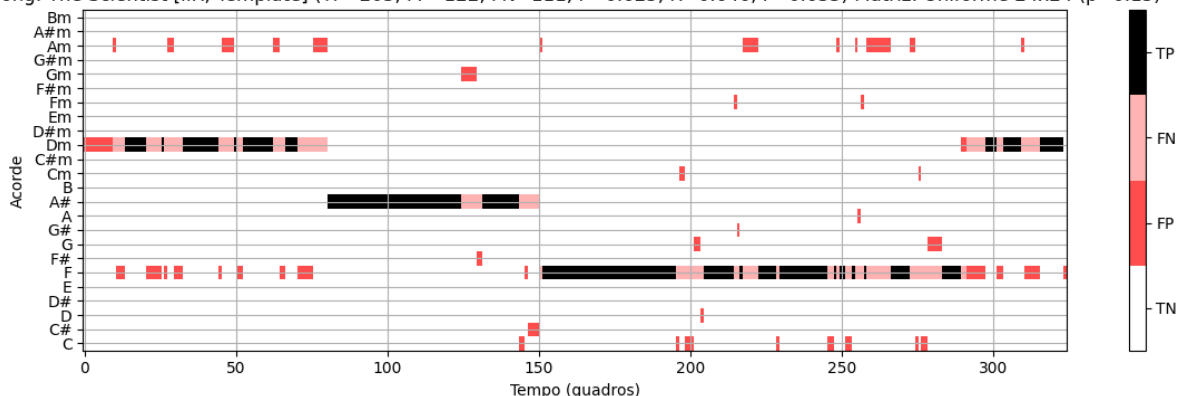
(a) Template-based + STFT

Song: The Scientist [CQT; Template] (TP=213, FP=112, FN=101, P=0.655, R=0.678, F=0.667) Matriz: Uniforme 24x24 ($p=0.15$)



(b) Template-based + CQT

Song: The Scientist [IIR; Template] (TP=203, FP=122, FN=111, P=0.625, R=0.646, F=0.635) Matriz: Uniforme 24x24 ($p=0.15$)

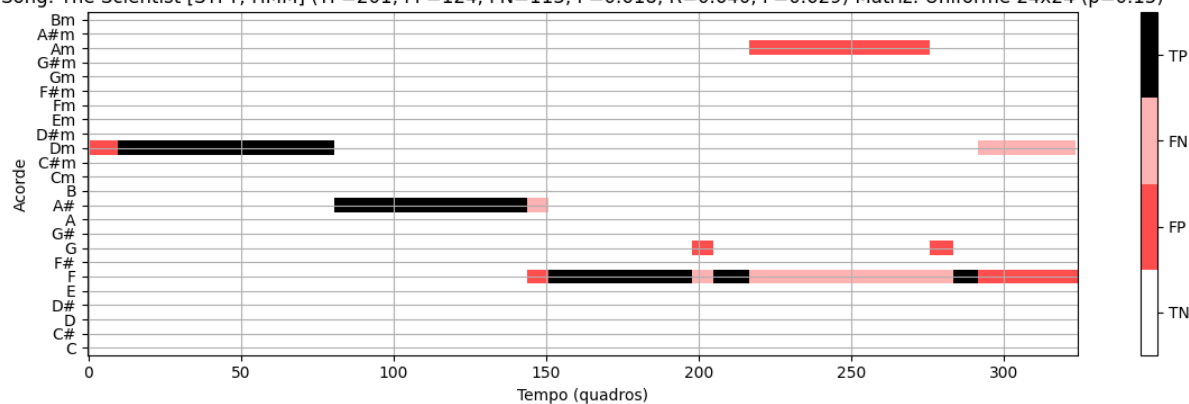


(c) Template-based + IIR

Figura 45 – Resultados do Template-based para a música *The scientist*.

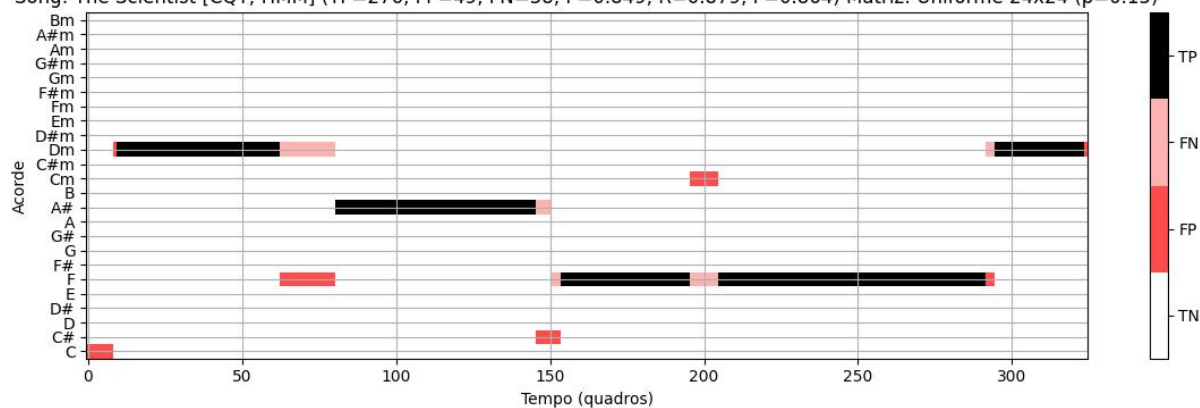
Fonte: Testes Iniciais do (PITA, 2024).

Song: The Scientist [STFT; HMM] (TP=201, FP=124, FN=113, P=0.618, R=0.640, F=0.629) Matriz: Uniforme 24x24 ($p=0.15$)



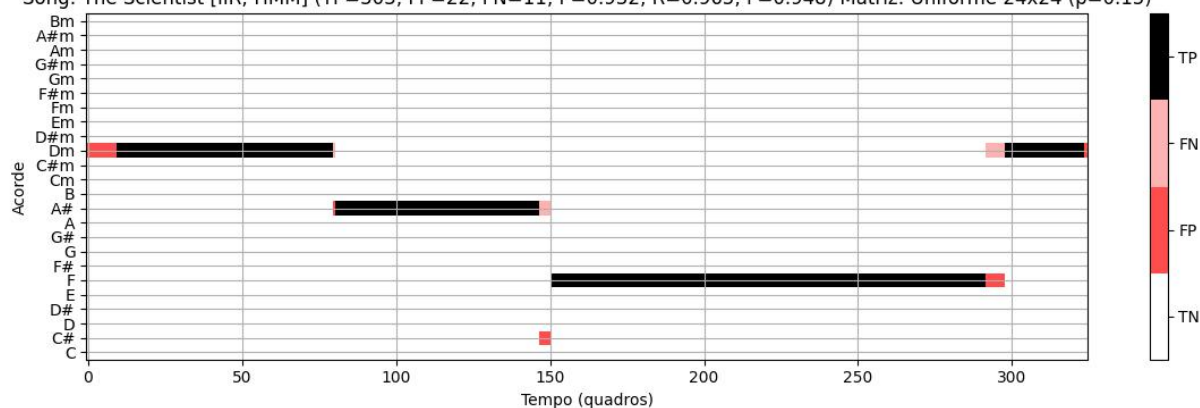
(a) HMM + STFT

Song: The Scientist [CQT; HMM] (TP=276, FP=49, FN=38, P=0.849, R=0.879, F=0.864) Matriz: Uniforme 24x24 ($p=0.15$)



(b) HMM + CQT

Song: The Scientist [IIR; HMM] (TP=303, FP=22, FN=11, P=0.932, R=0.965, F=0.948) Matriz: Uniforme 24x24 ($p=0.15$)

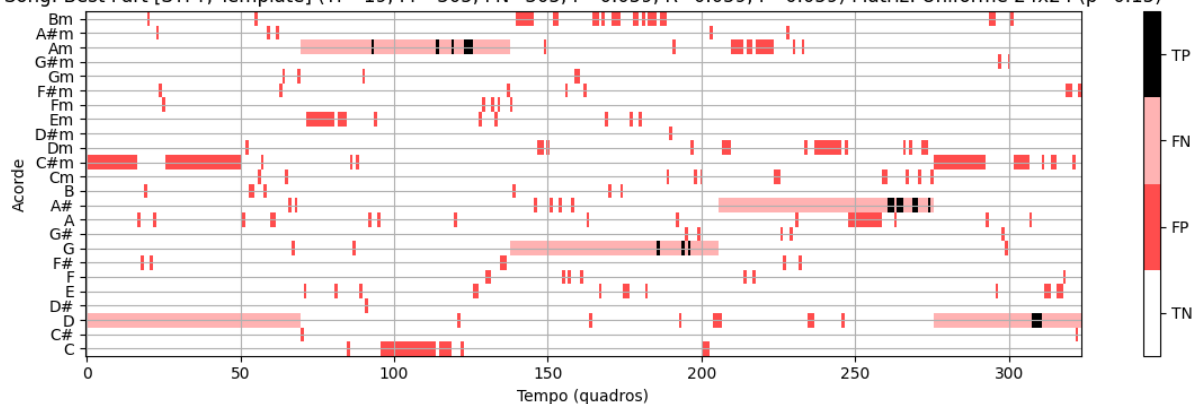


(c) HMM + IIR

Figura 46 – Resultados do HMM para a música *The scientist*.

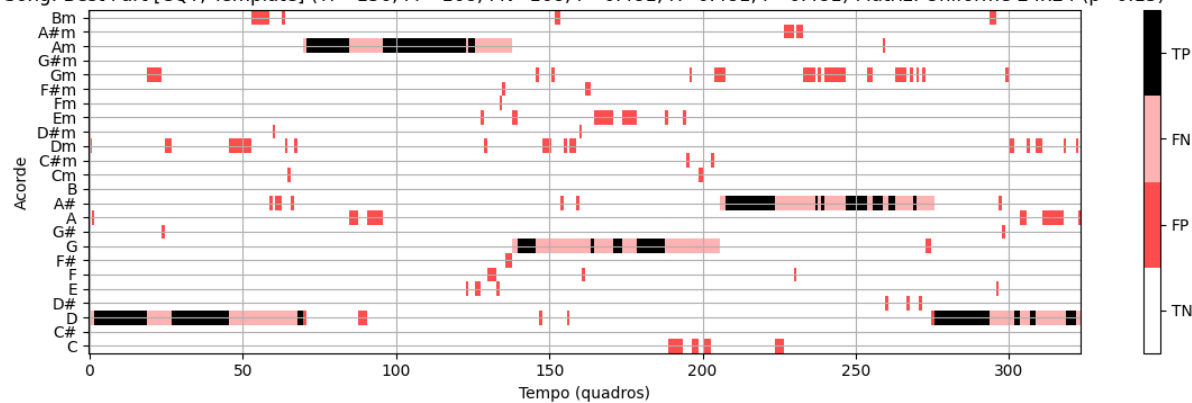
Fonte: Testes Iniciais do (PITA, 2024).

Song: Best Part [STFT; Template] (TP=19, FP=305, FN=305, P=0.059, R=0.059, F=0.059) Matriz: Uniforme 24x24 ($p=0.15$)



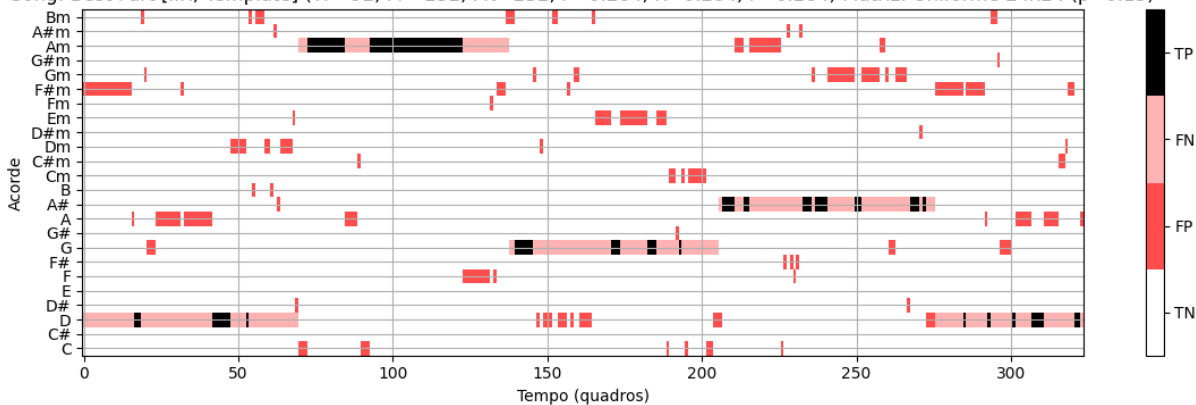
(a) Template-based + STFT

Song: Best Part [CQT; Template] (TP=156, FP=168, FN=168, P=0.481, R=0.481, F=0.481) Matriz: Uniforme 24x24 ($p=0.15$)



(b) Template-based + CQT

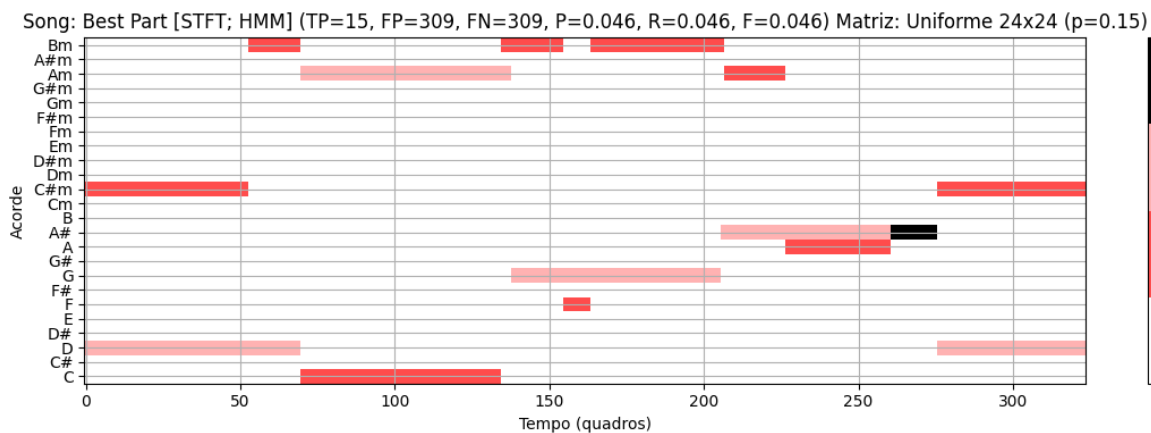
Song: Best Part [IIR; Template] (TP=92, FP=232, FN=232, P=0.284, R=0.284, F=0.284) Matriz: Uniforme 24x24 ($p=0.15$)



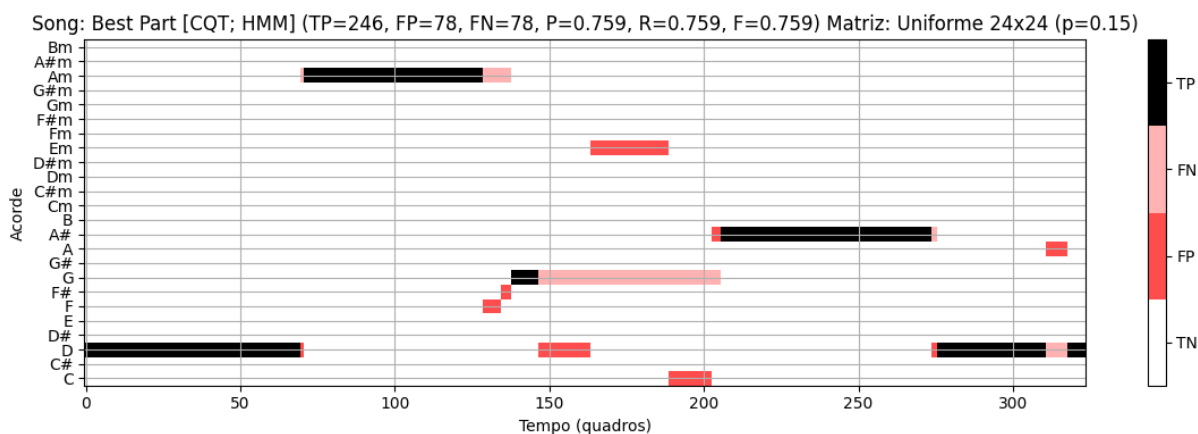
(c) Template-based + IIR

Figura 47 – Resultados do Template-based para a música *Best part*.

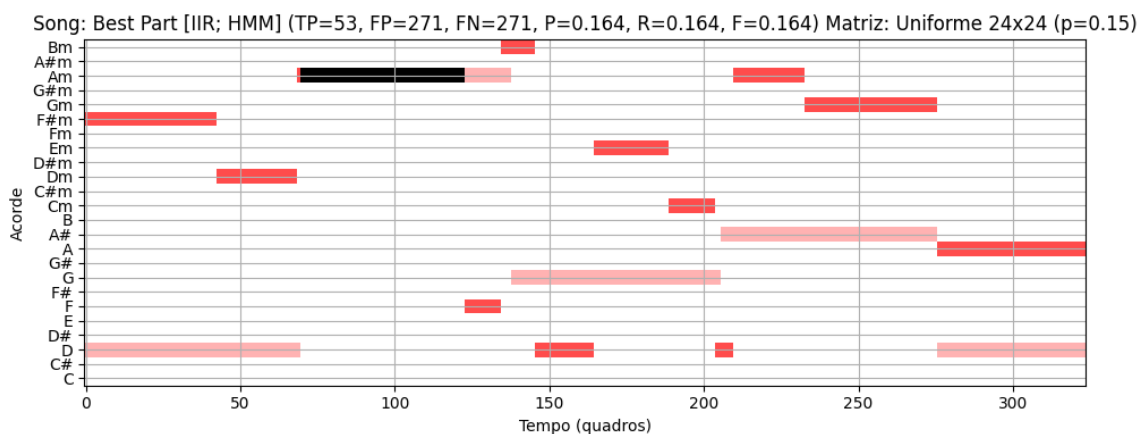
Fonte: Testes Iniciais do (PITA, 2024).



(a) HMM + STFT



(b) HMM + CQT

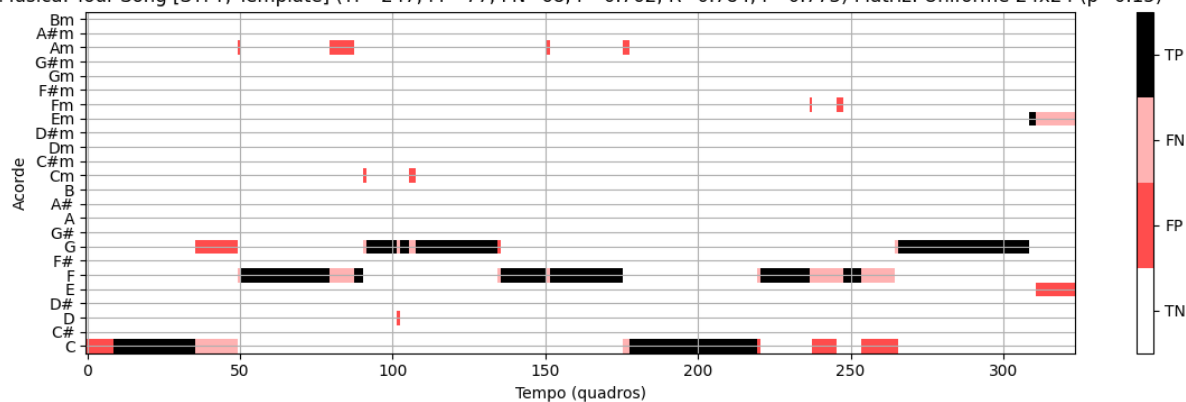


(c) HMM + IIR

Figura 48 – Resultados do HMM para a música *Best part*.

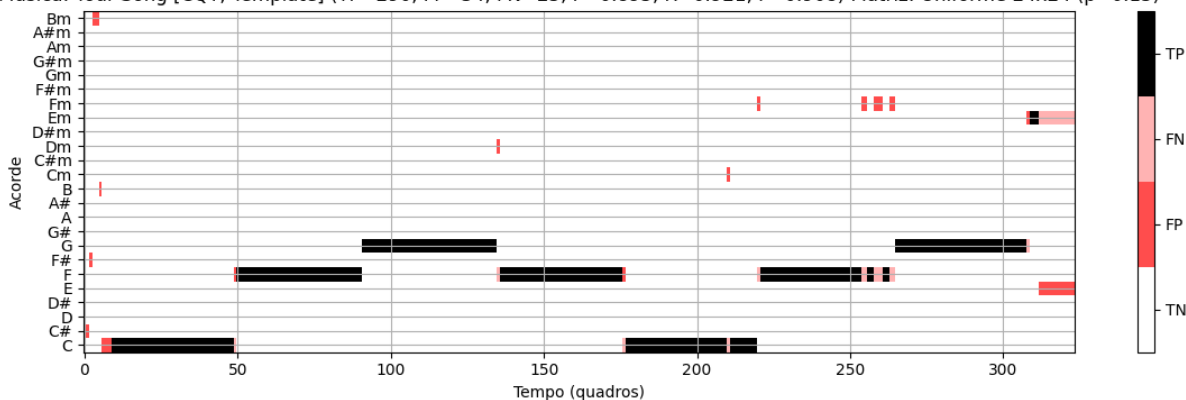
Fonte: Testes Iniciais do (PITA, 2024).

Música: Your Song [STFT; Template] (TP=247, FP=77, FN=68, P=0.762, R=0.784, F=0.773) Matriz: Uniforme 24x24 ($p=0.15$)



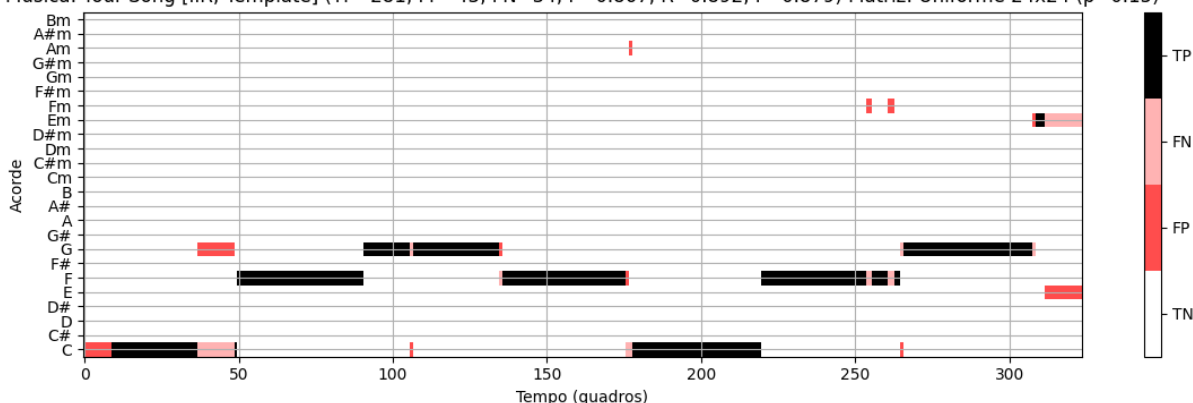
(a) Template-based + STFT

Música: Your Song [CQT; Template] (TP=290, FP=34, FN=25, P=0.895, R=0.921, F=0.908) Matriz: Uniforme 24x24 ($p=0.15$)



(b) Template-based + CQT

Música: Your Song [IIR; Template] (TP=281, FP=43, FN=34, P=0.867, R=0.892, F=0.879) Matriz: Uniforme 24x24 ($p=0.15$)

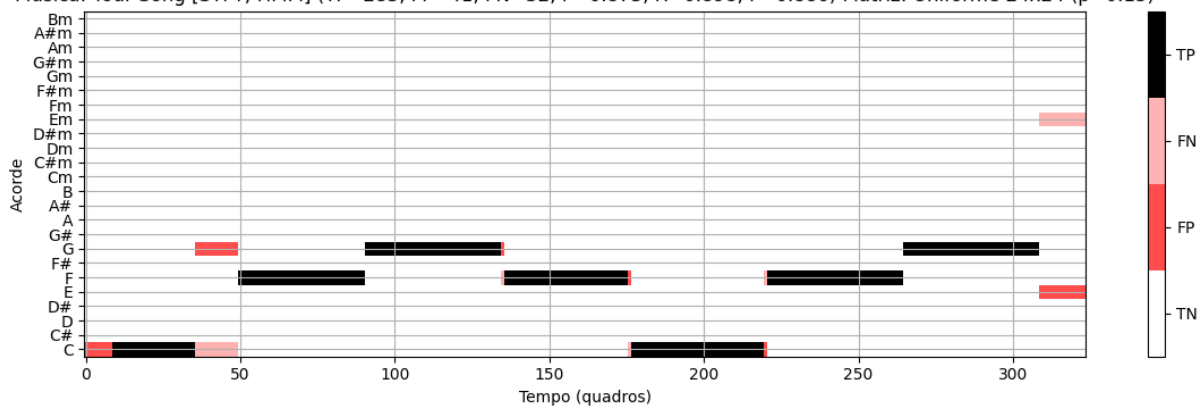


(c) Template-based + IIR

Figura 49 – Resultados do Template-based para a música *Your song*.

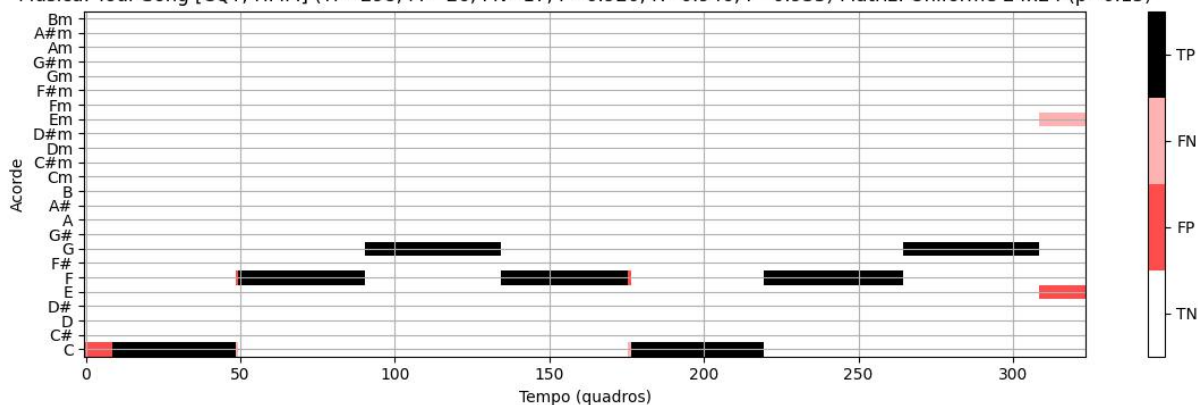
Fonte: Testes Iniciais do (PITA, 2024).

Música: Your Song [STFT; HMM] (TP=283, FP=41, FN=32, P=0.873, R=0.898, F=0.886) Matriz: Uniforme 24x24 ($p=0.15$)



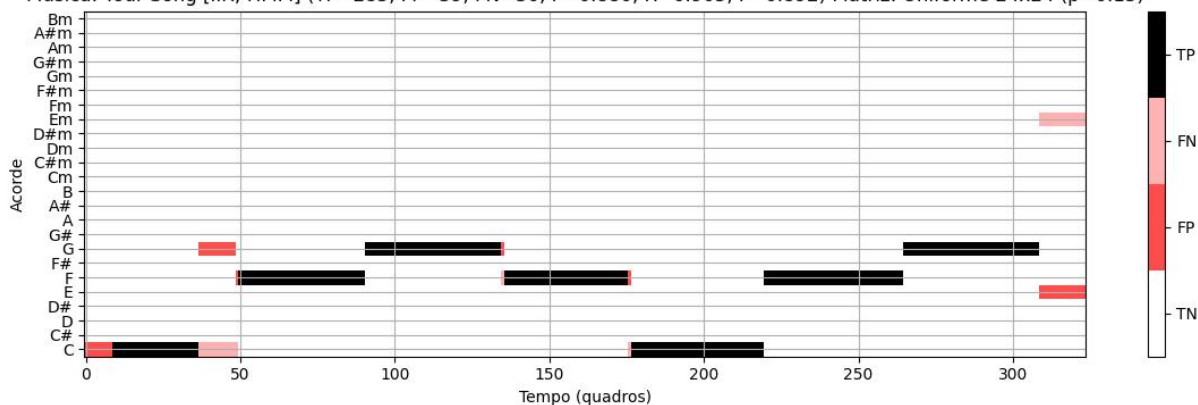
(a) HMM + STFT

Música: Your Song [CQT; HMM] (TP=298, FP=26, FN=17, P=0.920, R=0.946, F=0.933) Matriz: Uniforme 24x24 ($p=0.15$)



(b) HMM + CQT

Música: Your Song [IIR; HMM] (TP=285, FP=39, FN=30, P=0.880, R=0.905, F=0.892) Matriz: Uniforme 24x24 ($p=0.15$)

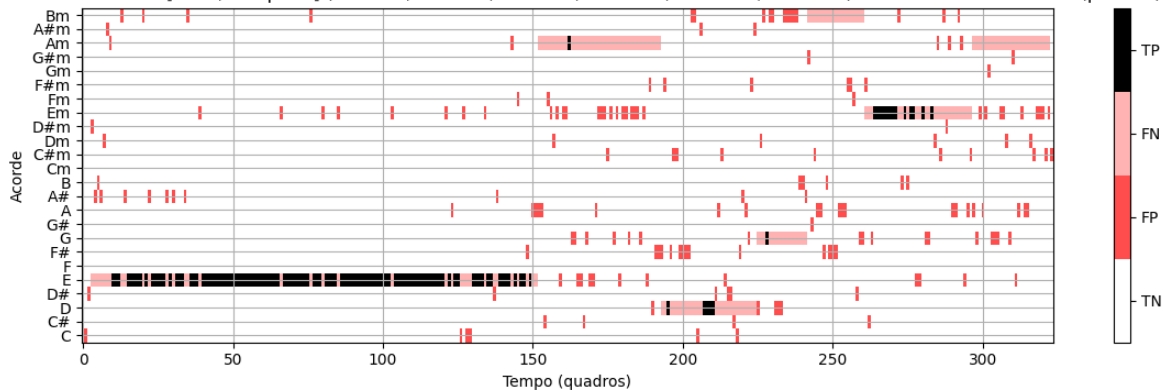


(c) HMM + IIR

Figura 50 – Resultados do HMM para a música *Your song*.

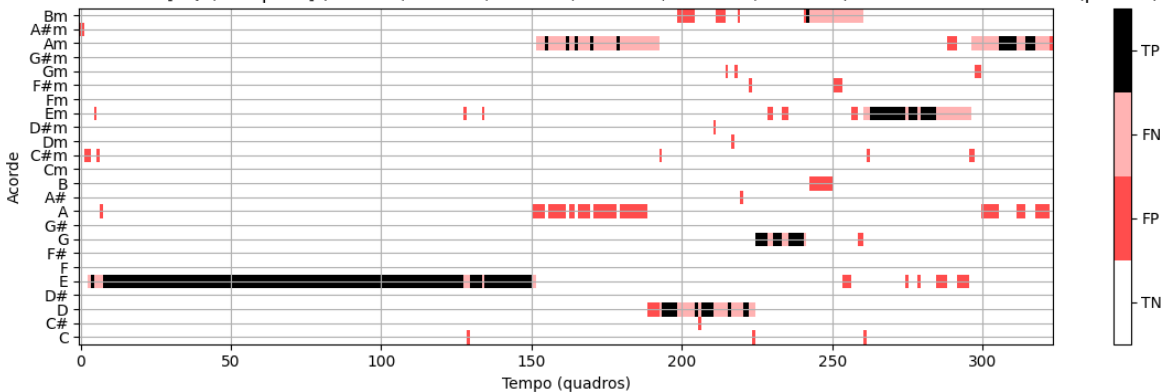
Fonte: Testes Iniciais do (PITA, 2024).

Música: Take Me Out [STFT; Template] (TP=134, FP=190, FN=186, P=0.414, R=0.419, F=0.416) Matriz: Uniforme 24x24 ($p=0.15$)



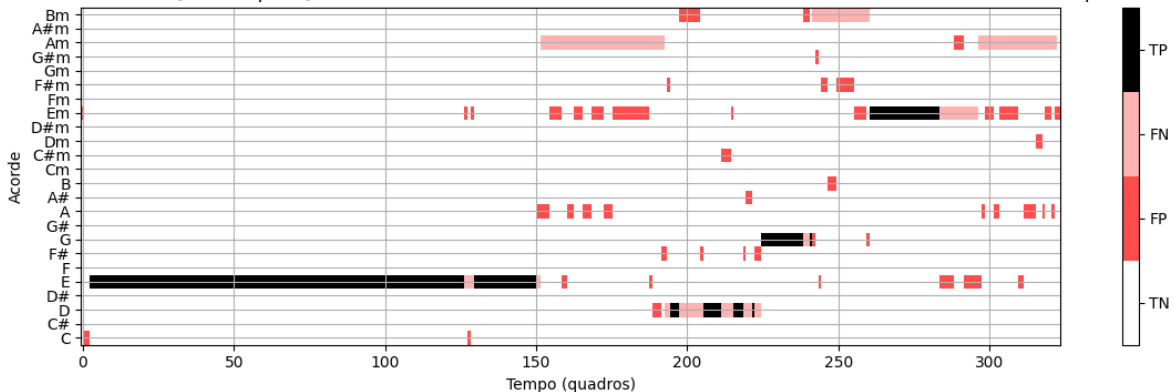
(a) Template-based + STFT

Música: Take Me Out [CQT; Template] (TP=201, FP=123, FN=119, P=0.620, R=0.628, F=0.624) Matriz: Uniforme 24x24 ($p=0.15$)



(b) Template-based + CQT

Música: Take Me Out [IIR; Template] (TP=196, FP=128, FN=124, P=0.605, R=0.613, F=0.609) Matriz: Uniforme 24x24 ($p=0.15$)

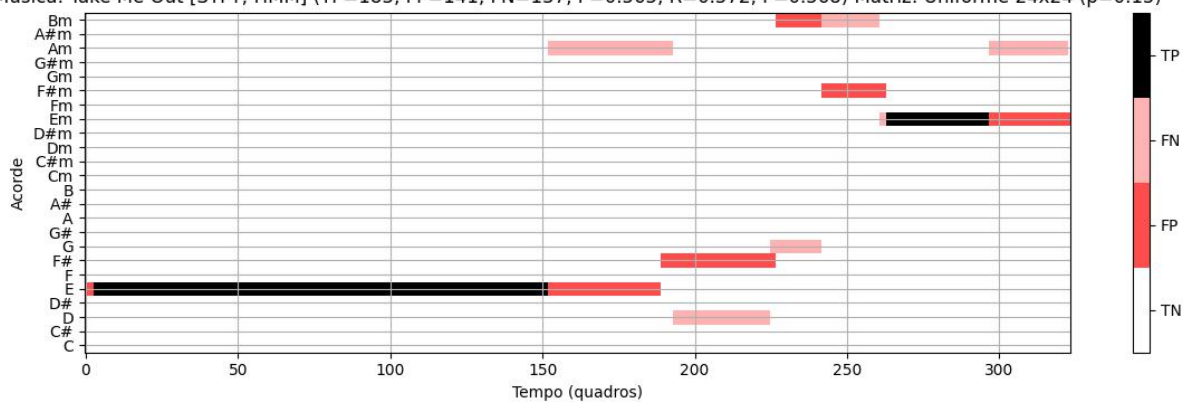


(c) Template-based + IIR

Figura 51 – Resultados do Template-based para a música *Take me out*.

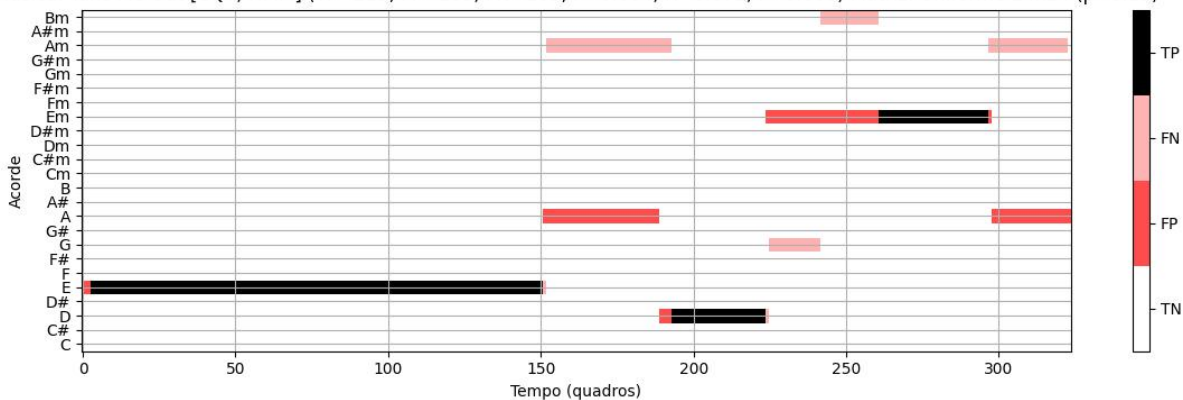
Fonte: Testes Iniciais do (PITA, 2024).

Música: Take Me Out [STFT; HMM] (TP=183, FP=141, FN=137, P=0.565, R=0.572, F=0.568) Matriz: Uniforme 24x24 ($p=0.15$)



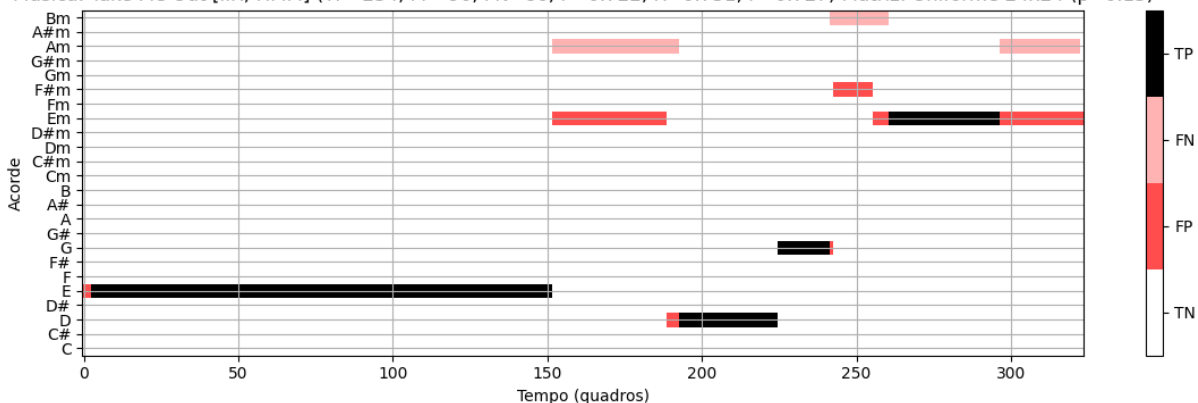
(a) HMM + STFT

Música: Take Me Out [CQT; HMM] (TP=215, FP=109, FN=105, P=0.664, R=0.672, F=0.668) Matriz: Uniforme 24x24 ($p=0.15$)



(b) HMM + CQT

Música: Take Me Out [IIR; HMM] (TP=234, FP=90, FN=86, P=0.722, R=0.731, F=0.727) Matriz: Uniforme 24x24 ($p=0.15$)

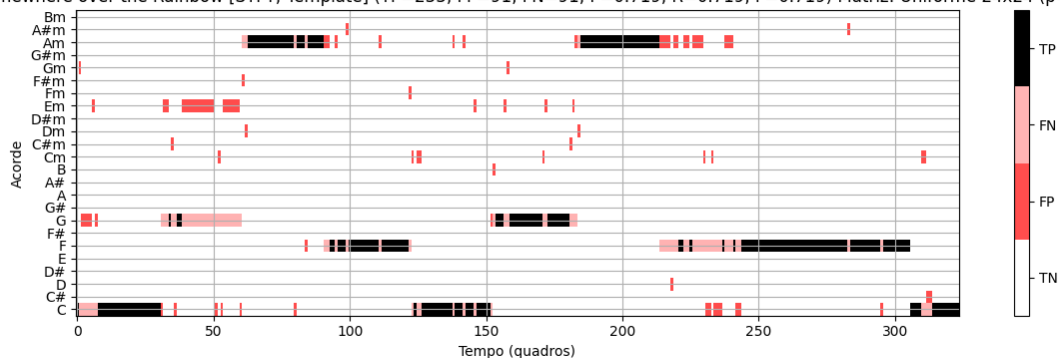


(c) HMM + IIR

Figura 52 – Resultados do HMM para a música *Take me out*.

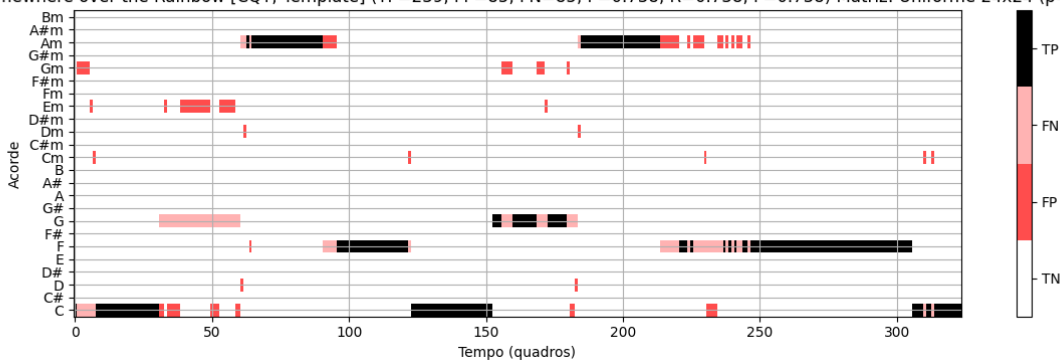
Fonte: Testes Iniciais do (PITA, 2024).

Song: Somewhere over the Rainbow [STFT; Template] (TP=233, FP=91, FN=91, P=0.719, R=0.719, F=0.719) Matriz: Uniforme 24x24 ($p=0.15$)



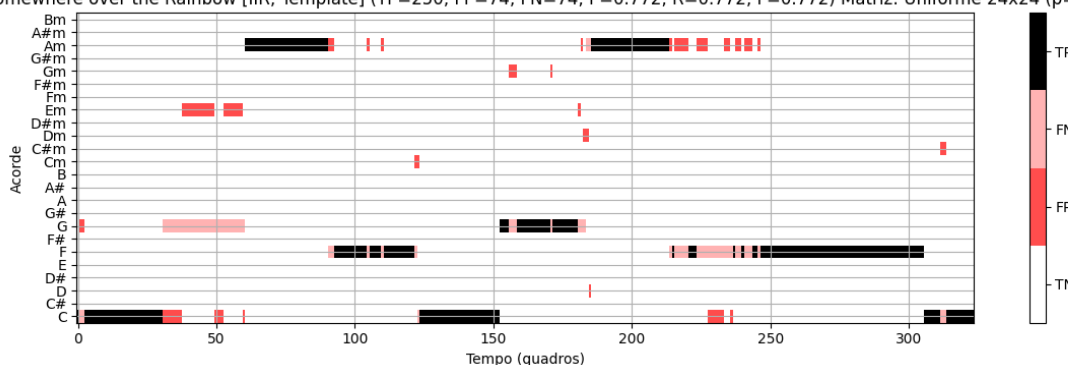
(a) Template-based + STFT

Song: Somewhere over the Rainbow [CQT; Template] (TP=239, FP=85, FN=85, P=0.738, R=0.738, F=0.738) Matriz: Uniforme 24x24 ($p=0.15$)



(b) Template-based + CQT

Song: Somewhere over the Rainbow [IIR; Template] (TP=250, FP=74, FN=74, P=0.772, R=0.772, F=0.772) Matriz: Uniforme 24x24 ($p=0.15$)

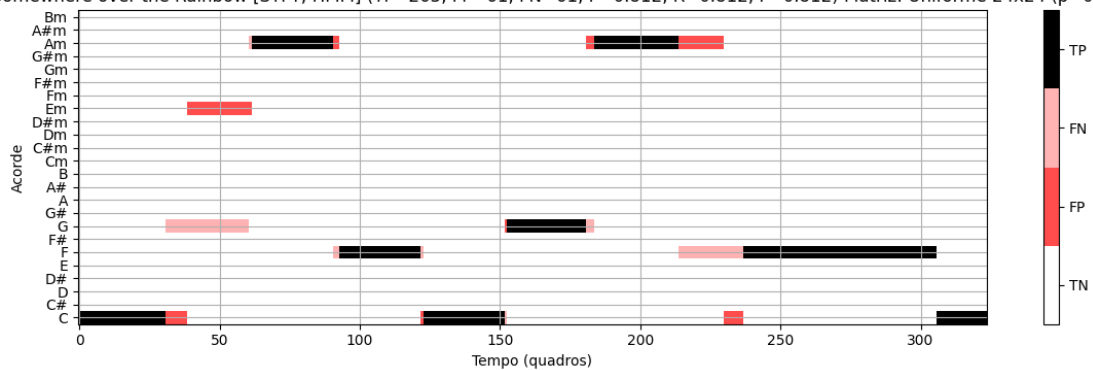


(c) Template-based + IIR

Figura 53 – Resultados do Template-based para a música *Somewhere over the rainbow*.

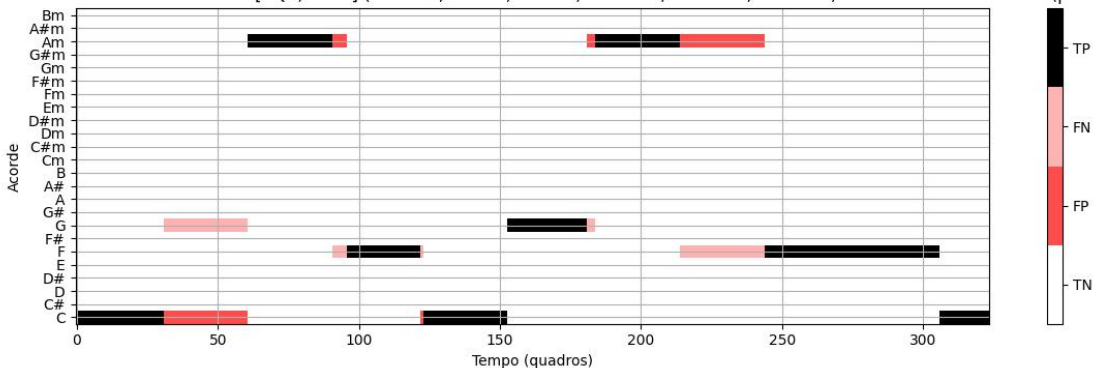
Fonte: Testes Iniciais do (PITA, 2024).

Song: Somewhere over the Rainbow [STFT; HMM] (TP=263, FP=61, FN=61, P=0.812, R=0.812, F=0.812) Matriz: Uniforme 24x24 ($p=0.15$)



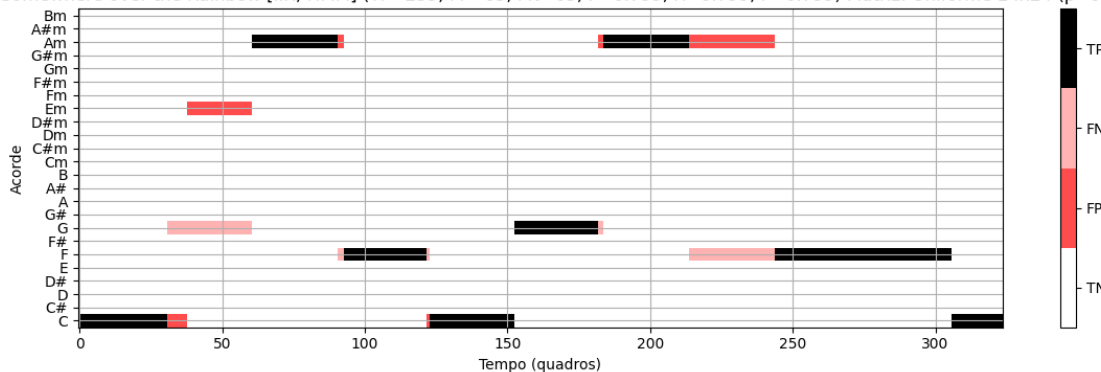
(a) HMM + STFT

Song: Somewhere over the Rainbow [CQT; HMM] (TP=255, FP=69, FN=69, P=0.787, R=0.787, F=0.787) Matriz: Uniforme 24x24 ($p=0.15$)



(b) HMM + CQT

Song: Somewhere over the Rainbow [IIR; HMM] (TP=259, FP=65, FN=65, P=0.799, R=0.799, F=0.799) Matriz: Uniforme 24x24 ($p=0.15$)

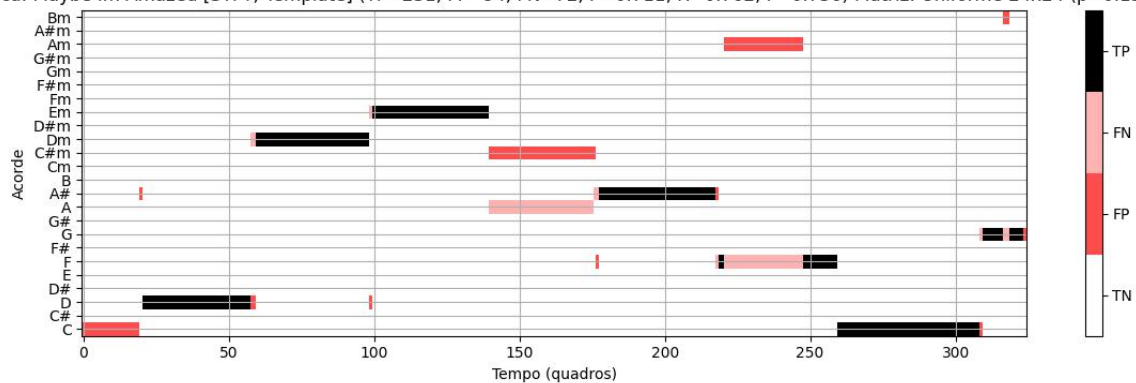


(c) HMM + IIR

Figura 54 – Resultados do HMM para a música *Somewhere over the rainbow*.

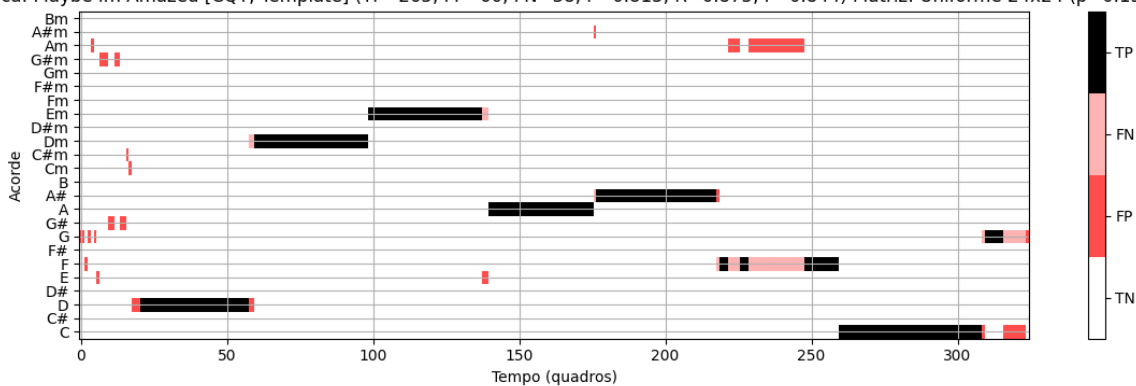
Fonte: Testes Iniciais do (PITA, 2024).

Música: Maybe Im Amazed [STFT; Template] (TP=231, FP=94, FN=72, P=0.711, R=0.762, F=0.736) Matriz: Uniforme 24x24 ($p=0.15$)



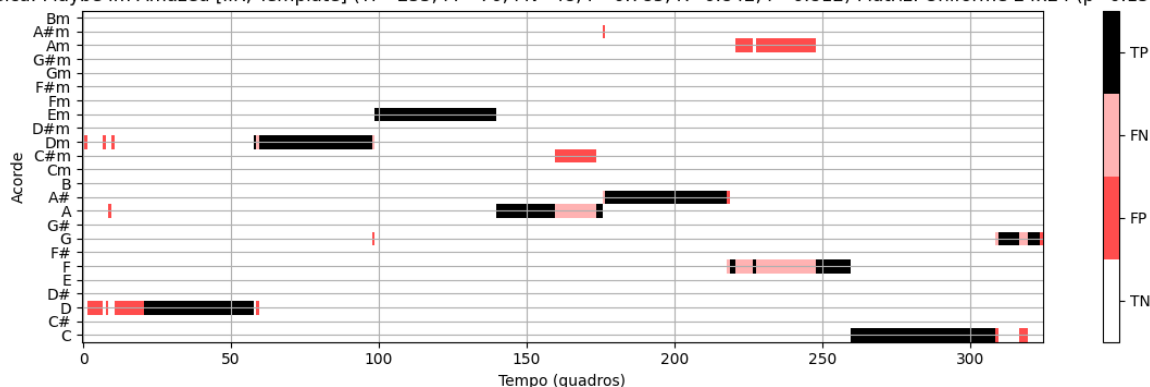
(a) Template-based + STFT

Música: Maybe Im Amazed [CQT; Template] (TP=265, FP=60, FN=38, P=0.815, R=0.875, F=0.844) Matriz: Uniforme 24x24 ($p=0.15$)



(b) Template-based + CQT

Música: Maybe Im Amazed [IIR; Template] (TP=255, FP=70, FN=48, P=0.785, R=0.842, F=0.812) Matriz: Uniforme 24x24 ($p=0.15$)

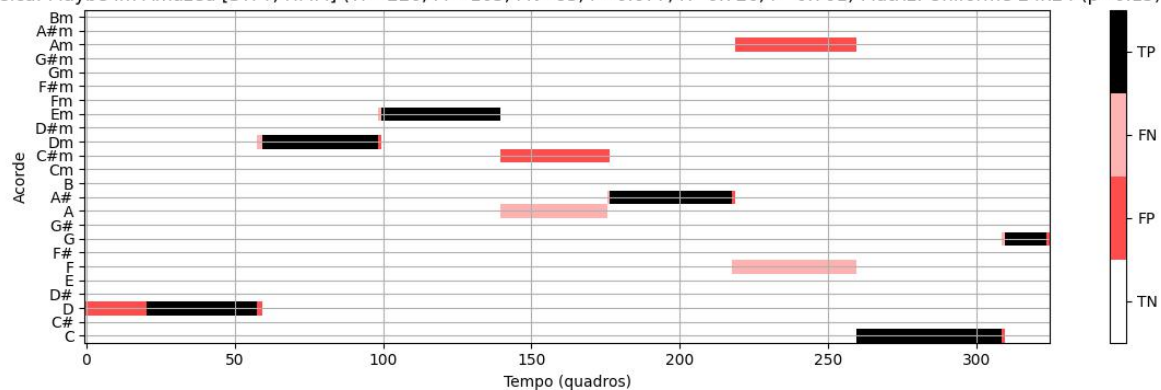


(c) Template-based + IIR

Figura 55 – Resultados do Template-based para a música *Maybe I'm amazed*.

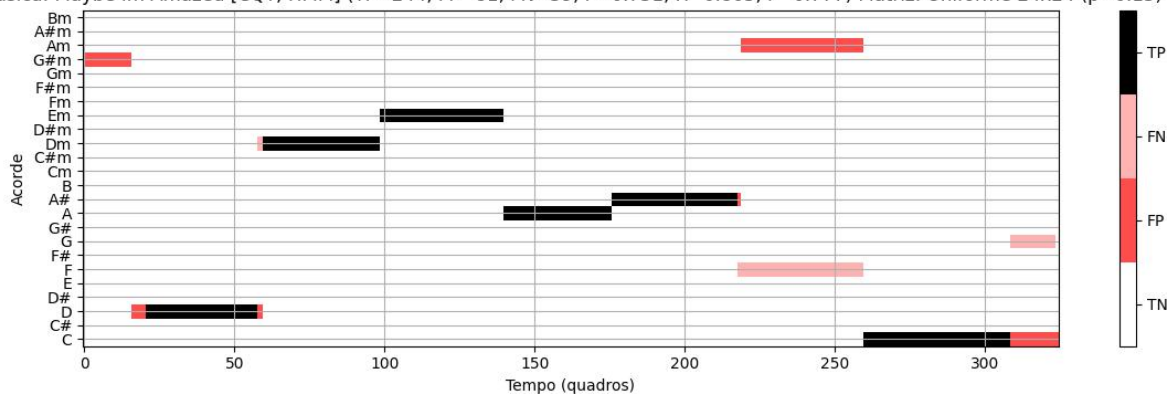
Fonte: Testes Iniciais do (PITA, 2024).

Música: Maybe Im Amazed [STFT; HMM] (TP=220, FP=105, FN=83, P=0.677, R=0.726, F=0.701) Matriz: Uniforme 24x24 ($p=0.15$)



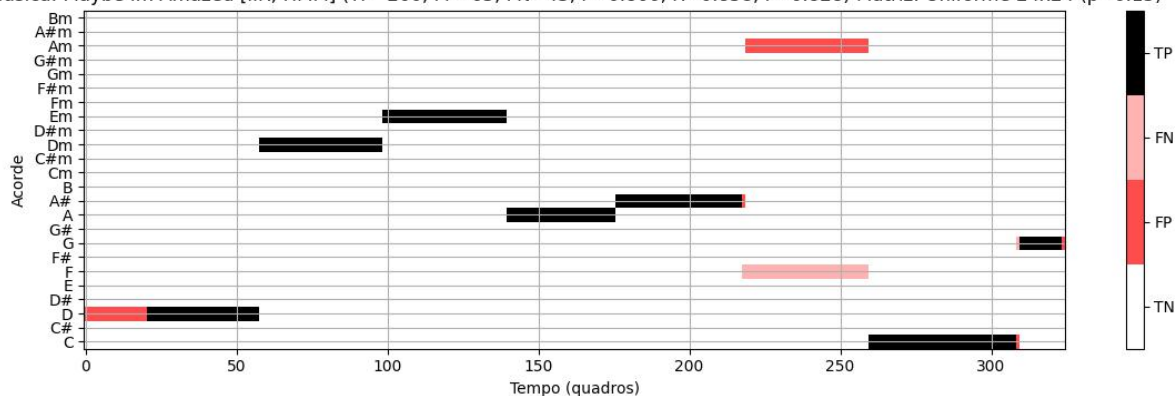
(a) HMM + STFT

Música: Maybe Im Amazed [CQT; HMM] (TP=244, FP=81, FN=59, P=0.751, R=0.805, F=0.777) Matriz: Uniforme 24x24 ($p=0.15$)



(b) HMM + CQT

Música: Maybe Im Amazed [IIR; HMM] (TP=260, FP=65, FN=43, P=0.800, R=0.858, F=0.828) Matriz: Uniforme 24x24 ($p=0.15$)

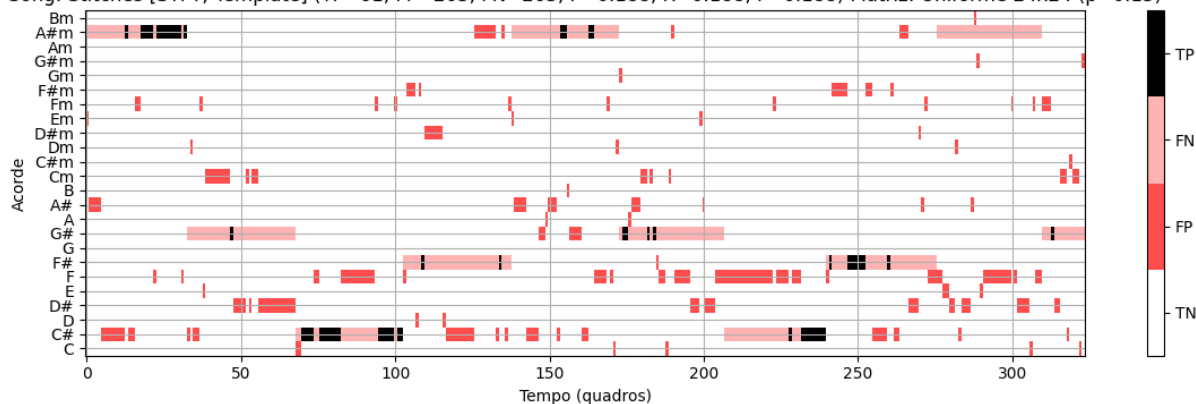


(c) HMM + IIR

Figura 56 – Resultados do HMM para a música *Maybe I'm amazed*.

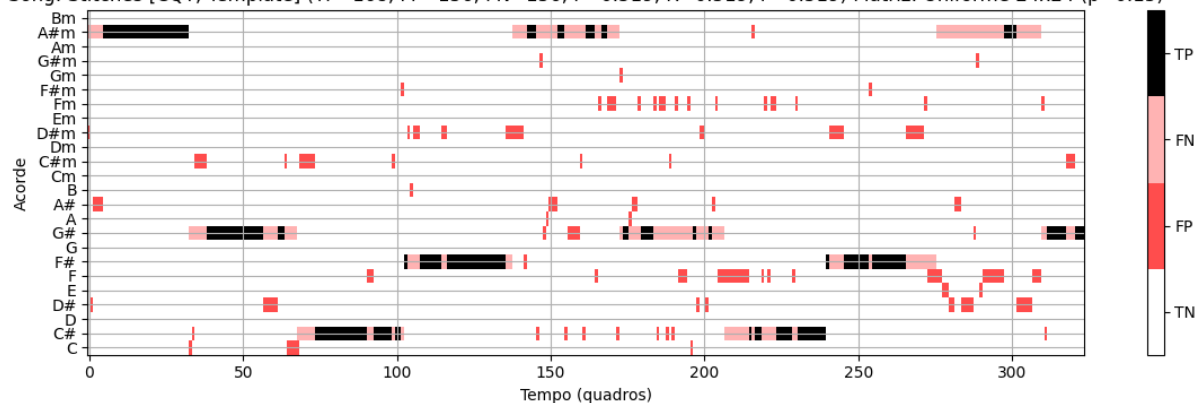
Fonte: Testes Iniciais do (PITA, 2024).

Song: Stitches [STFT; Template] (TP=61, FP=263, FN=263, P=0.188, R=0.188, F=0.188) Matriz: Uniforme 24x24 ($p=0.15$)



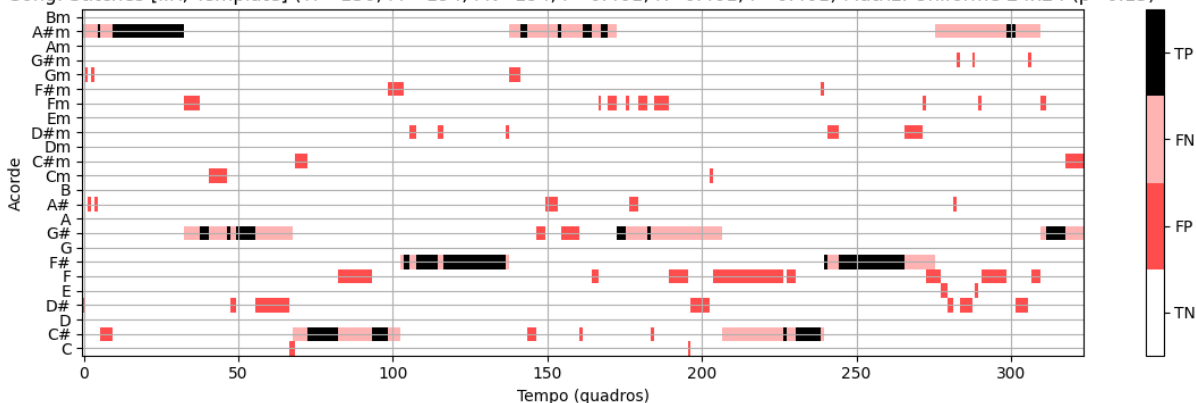
(a) Template-based + STFT

Song: Stitches [CQT; Template] (TP=168, FP=156, FN=156, P=0.519, R=0.519, F=0.519) Matriz: Uniforme 24x24 ($p=0.15$)



(b) Template-based + CQT

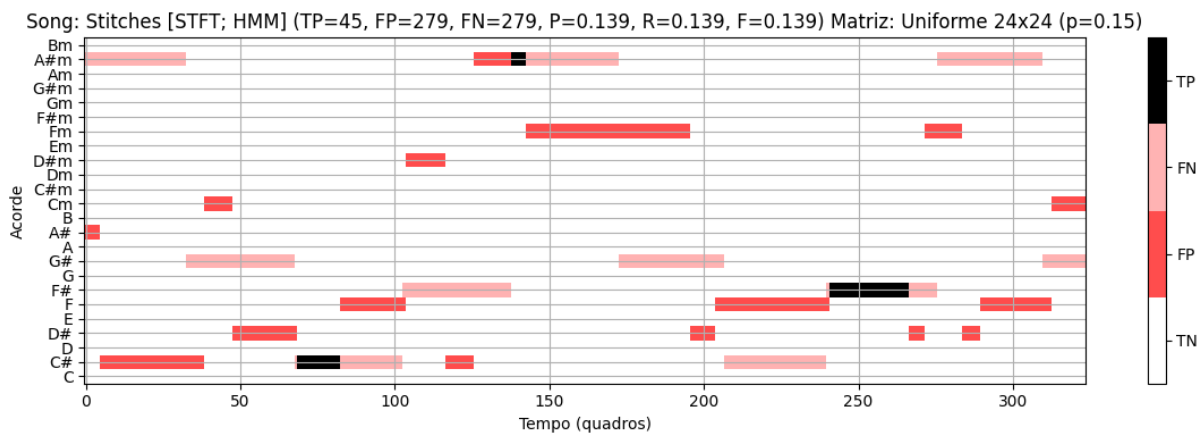
Song: Stitches [IIR; Template] (TP=130, FP=194, FN=194, P=0.401, R=0.401, F=0.401) Matriz: Uniforme 24x24 ($p=0.15$)



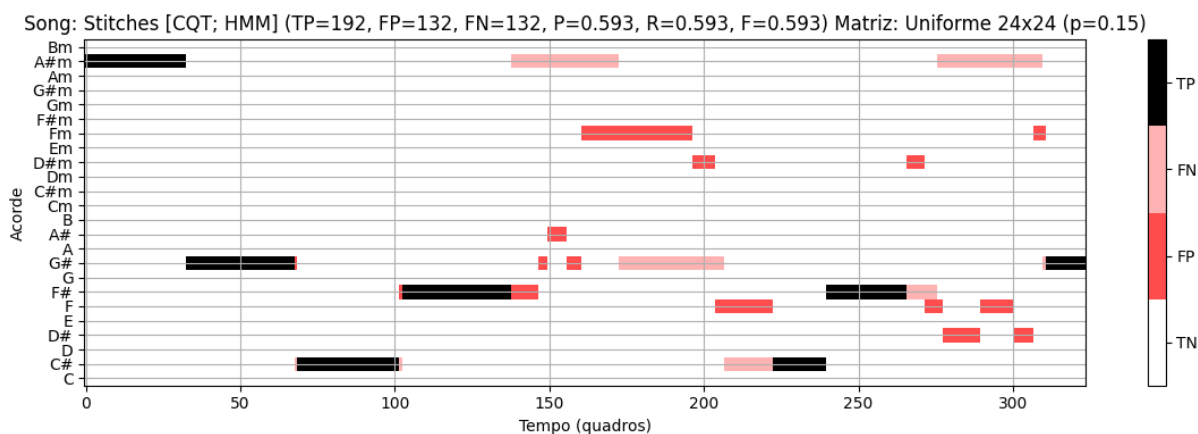
(c) Template-based + IIR

Figura 57 – Resultados do Template-based para a música *Stitches*.

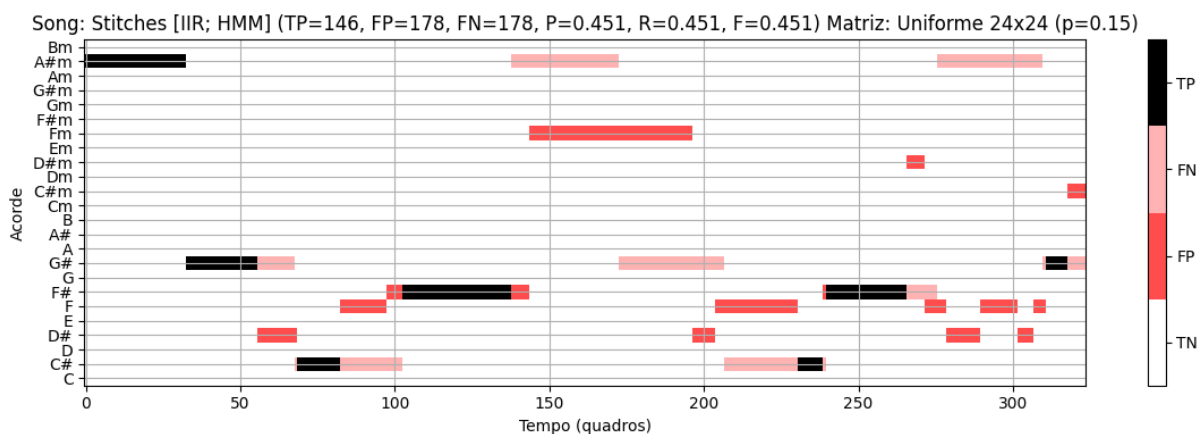
Fonte: Testes Iniciais do (PITA, 2024).



(a) HMM + STFT



(b) HMM + CQT

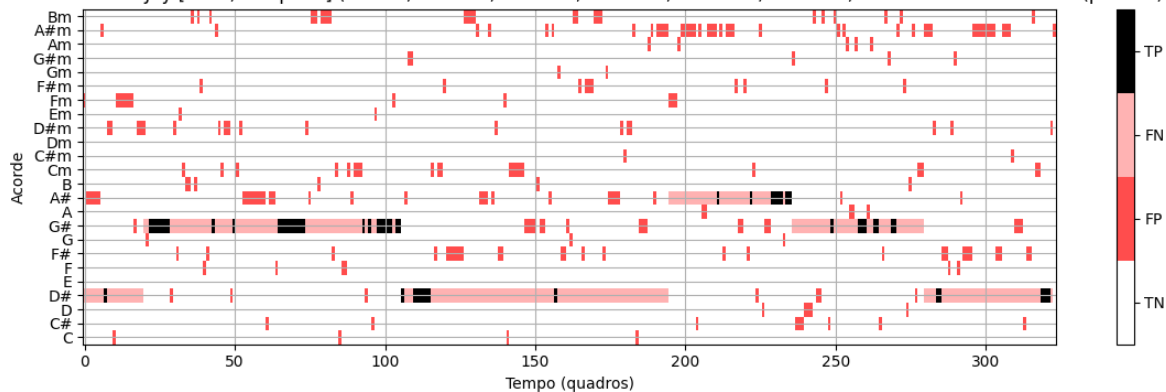


(c) HMM + IIR

Figura 58 – Resultados do HMM para a música *Stitches*.

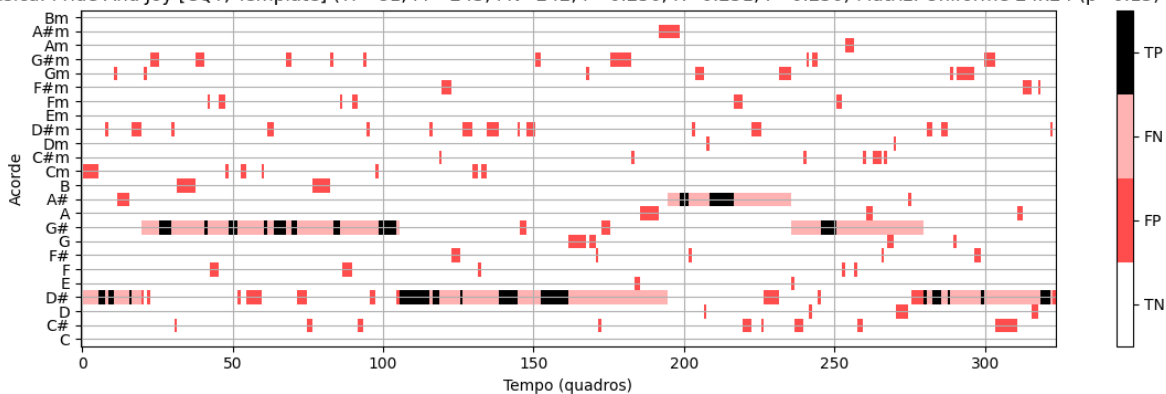
Fonte: Testes Iniciais do (PITA, 2024).

Música: Pride And Joy [STFT; Template] (TP=57, FP=267, FN=266, P=0.176, R=0.176, F=0.176) Matriz: Uniforme 24x24 ($p=0.15$)



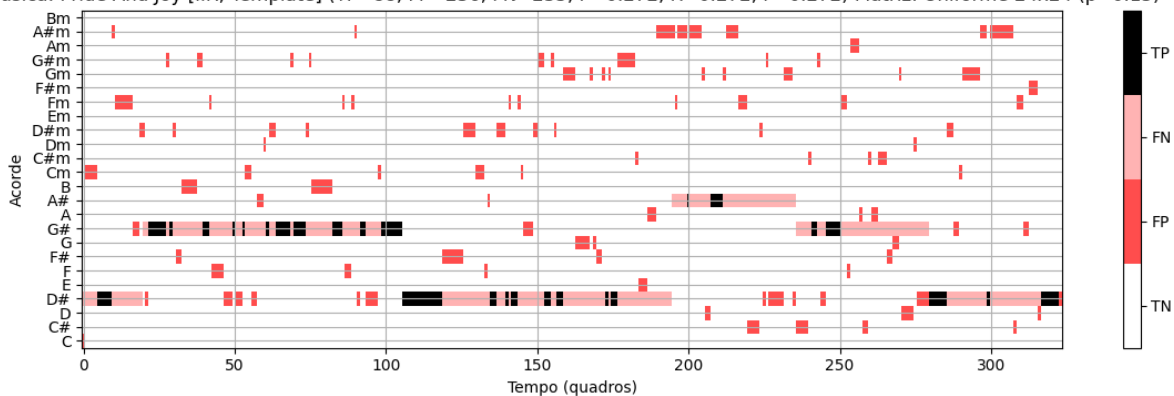
(a) Template-based + STFT

Música: Pride And Joy [CQT; Template] (TP=81, FP=243, FN=242, P=0.250, R=0.251, F=0.250) Matriz: Uniforme 24x24 ($p=0.15$)



(b) Template-based + CQT

Música: Pride And Joy [IIR; Template] (TP=88, FP=236, FN=235, P=0.272, R=0.272, F=0.272) Matriz: Uniforme 24x24 ($p=0.15$)

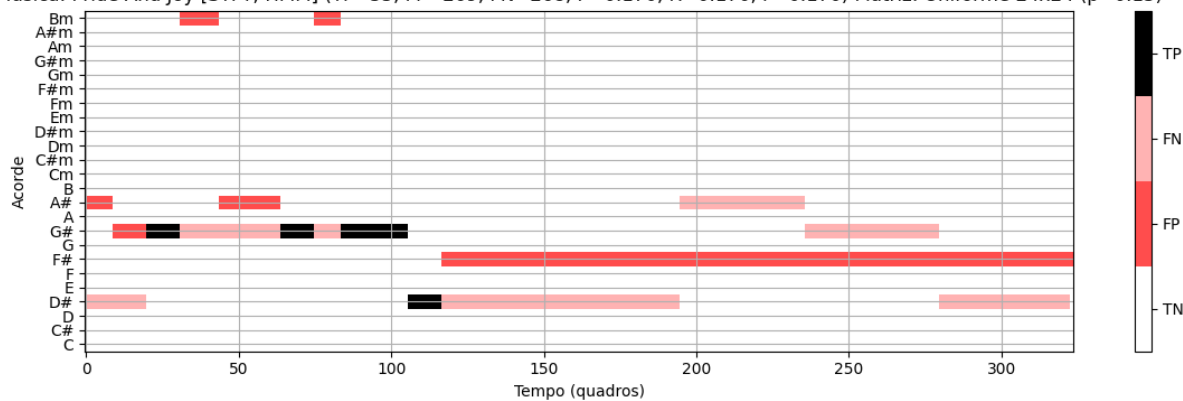


(c) Template-based + IIR

Figura 59 – Resultados do Template-based para a música *Pride and joy*.

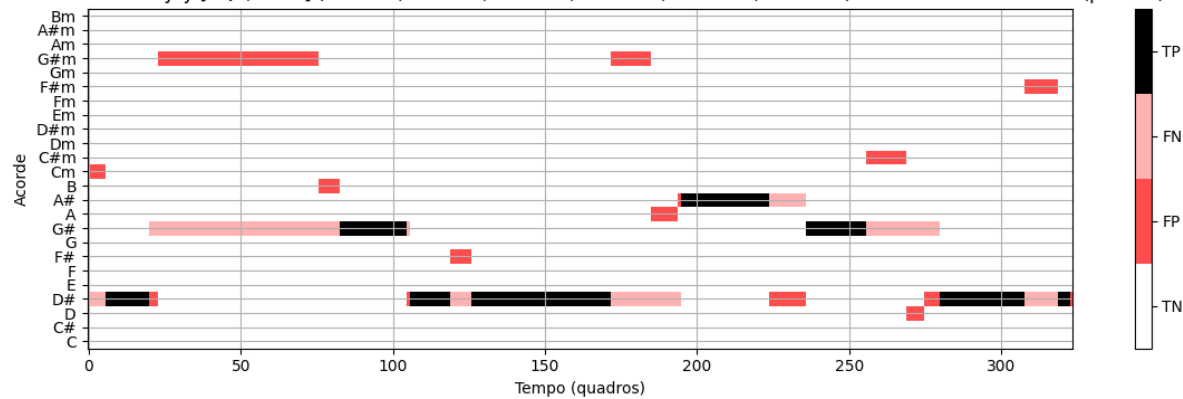
Fonte: Testes Iniciais do (PITA, 2024).

Música: Pride And Joy [STFT; HMM] (TP=55, FP=269, FN=268, P=0.170, R=0.170, F=0.170) Matriz: Uniforme 24x24 ($p=0.15$)



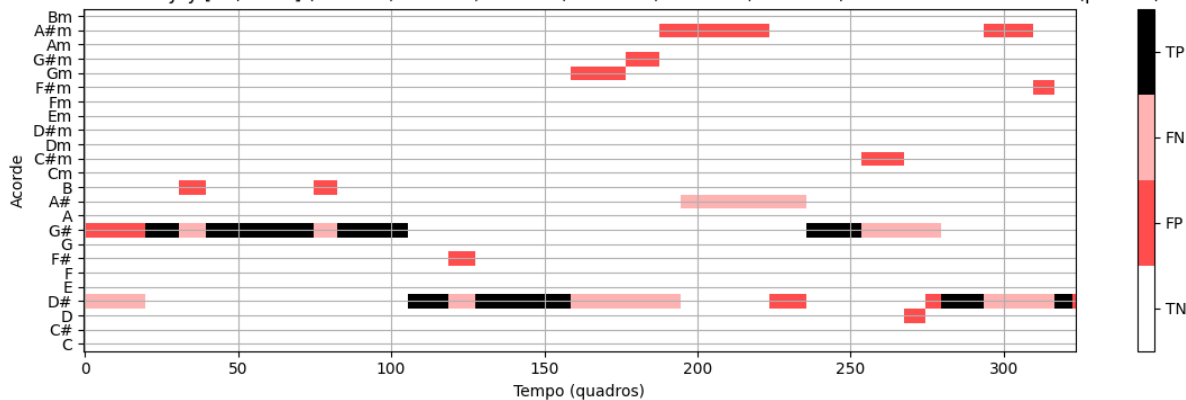
(a) HMM + STFT

Música: Pride And Joy [CQT; HMM] (TP=176, FP=148, FN=147, P=0.543, R=0.545, F=0.544) Matriz: Uniforme 24x24 ($p=0.15$)



(b) HMM + CQT

Música: Pride And Joy [IIR; HMM] (TP=151, FP=173, FN=172, P=0.466, R=0.467, F=0.467) Matriz: Uniforme 24x24 ($p=0.15$)

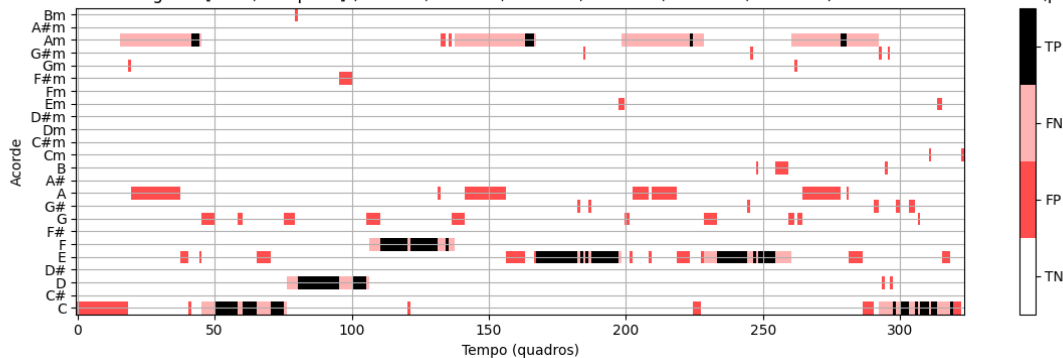


(c) HMM + IIR

Figura 60 – Resultados do HMM para a música *Pride and joy*.

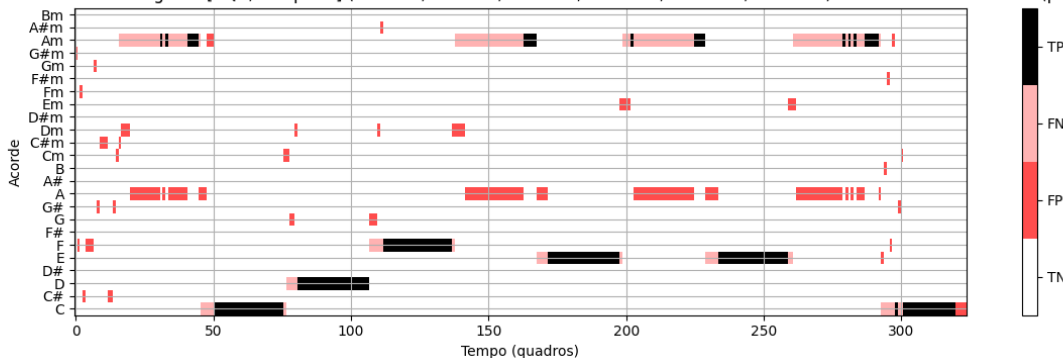
Fonte: Testes Iniciais do (PITA, 2024).

Música: House Of The Rising Sun [STFT; Template] (TP=124, FP=200, FN=180, P=0.383, R=0.408, F=0.395) Matriz: Uniforme 24x24 ($p=0.15$)



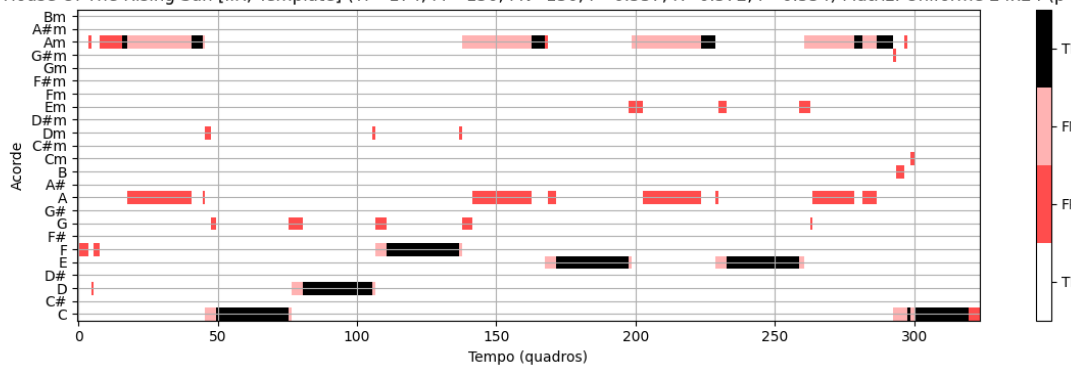
(a) Template-based + STFT

Música: House Of The Rising Sun [CQT; Template] (TP=171, FP=153, FN=133, P=0.528, R=0.562, F=0.545) Matriz: Uniforme 24x24 ($p=0.15$)



(b) Template-based + CQT

Música: House Of The Rising Sun [IIR; Template] (TP=174, FP=150, FN=130, P=0.537, R=0.572, F=0.554) Matriz: Uniforme 24x24 ($p=0.15$)

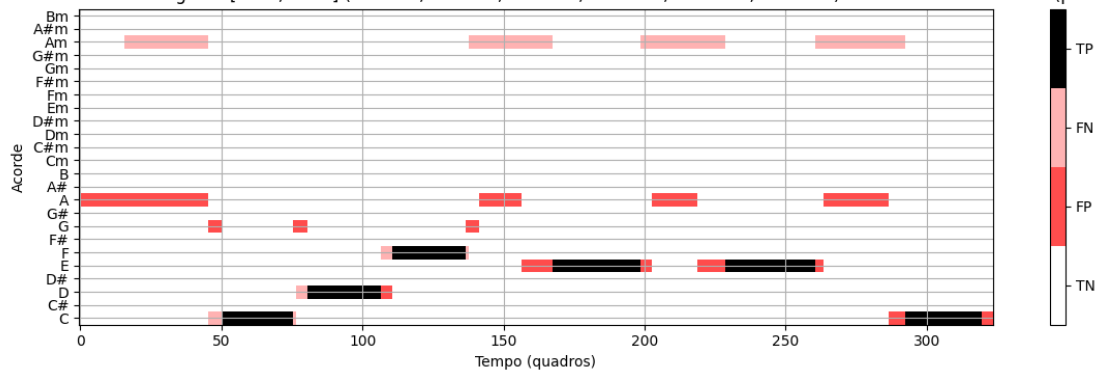


(c) Template-based + IIR

Figura 61 – Resultados do Template-based para a música *House of the rising sun*.

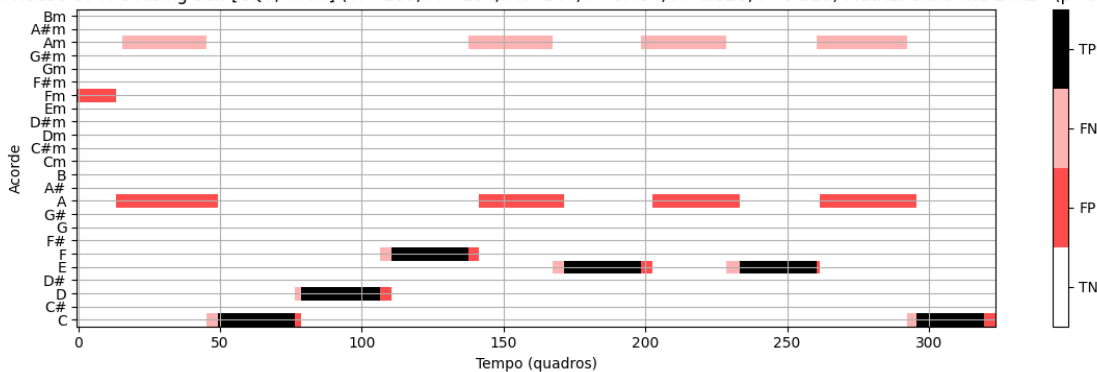
Fonte: Testes Iniciais do (PITA, 2024).

Música: House Of The Rising Sun [STFT; HMM] (TP=167, FP=157, FN=137, P=0.515, R=0.549, F=0.532) Matriz: Uniforme 24x24 ($p=0.15$)



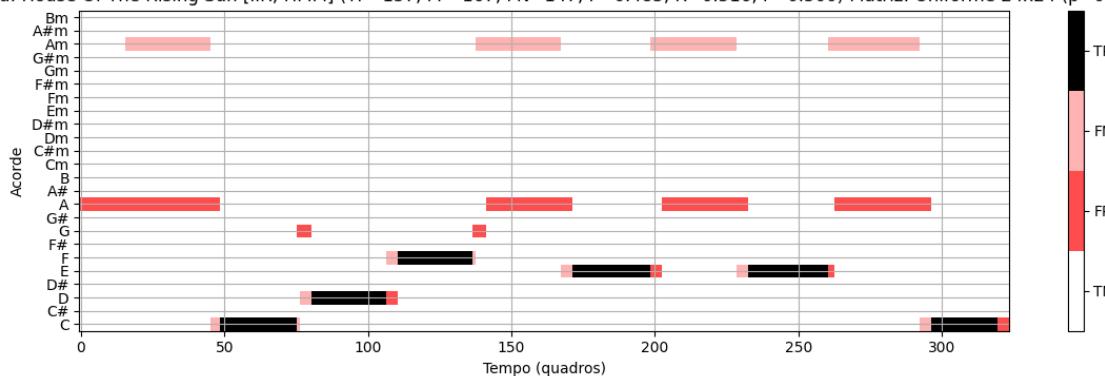
(a) HMM + STFT

Música: House Of The Rising Sun [CQT; HMM] (TP=160, FP=164, FN=144, P=0.494, R=0.526, F=0.510) Matriz: Uniforme 24x24 ($p=0.15$)



(b) HMM + CQT

Música: House Of The Rising Sun [IIR; HMM] (TP=157, FP=167, FN=147, P=0.485, R=0.516, F=0.500) Matriz: Uniforme 24x24 ($p=0.15$)



(c) HMM + IIR

Figura 62 – Resultados do HMM para a música *House of the rising sun*.

Fonte: Testes Iniciais do (PITA, 2024).

APÊNDICE C – CROMAGRAMAS GERADOS PARA OS EXPERIMENTOS NO CONTEXTO DA MPB.

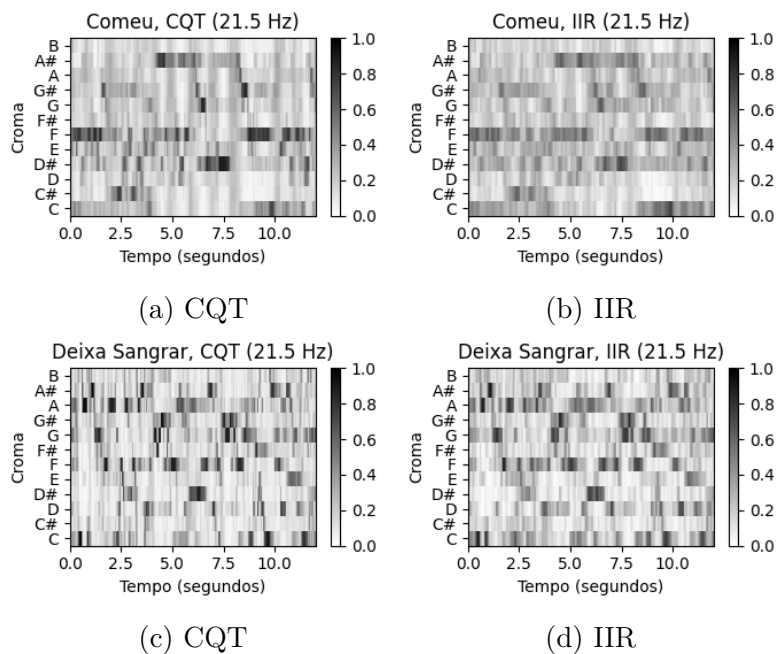


Figura 63 – Cromagramas gerados para as músicas *Comeu* e *Deixa sangrar*.

Fonte: Testes MPB do (PITA, 2024).

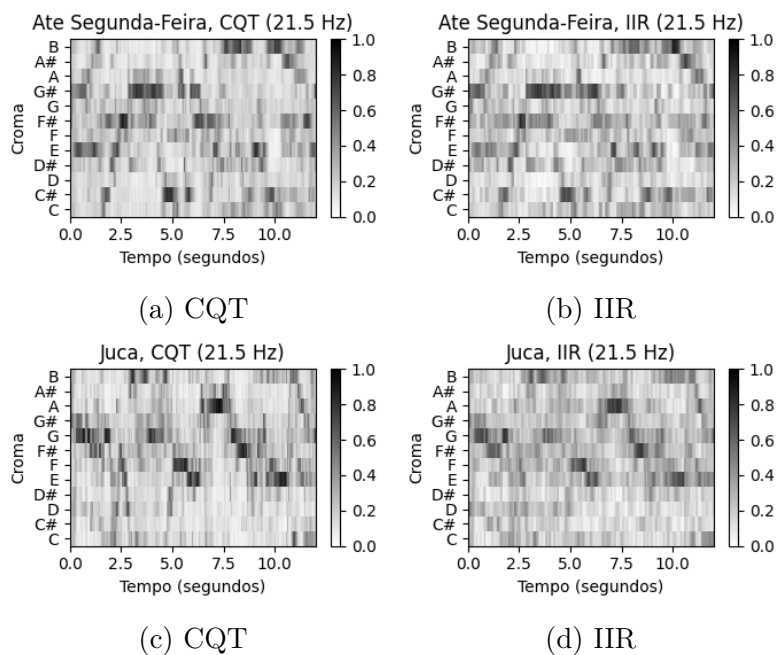


Figura 64 – Cromagramas gerados para as músicas *Até segunda-feira* e *Juca*.

Fonte: Testes MPB do (PITA, 2024).

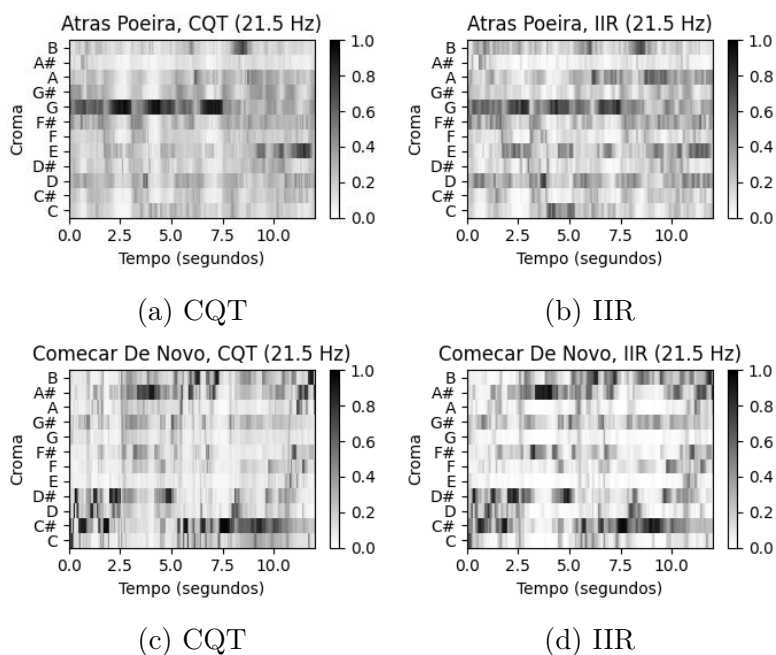


Figura 65 – Cromagramas gerados para as músicas *Atras Poeira* e *Começar de novo*.

Fonte: Testes MPB do (PITA, 2024).

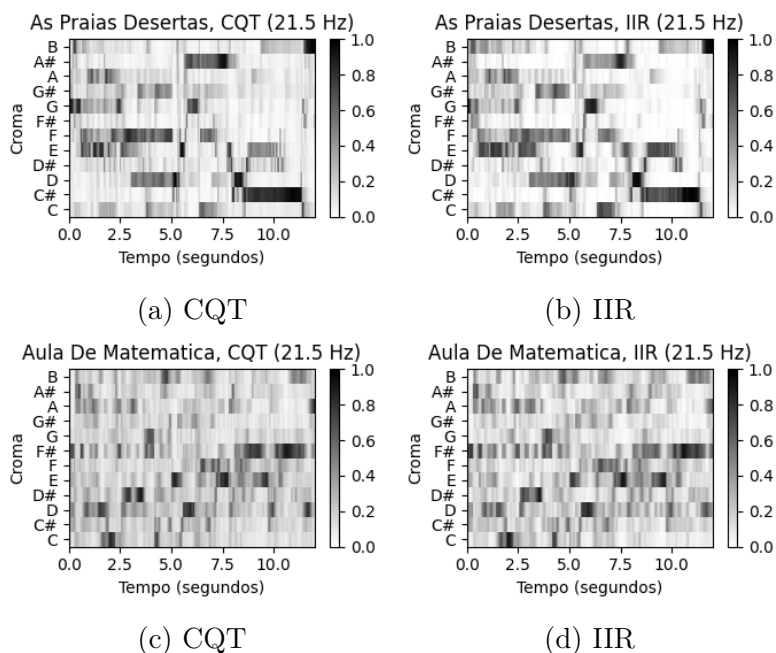
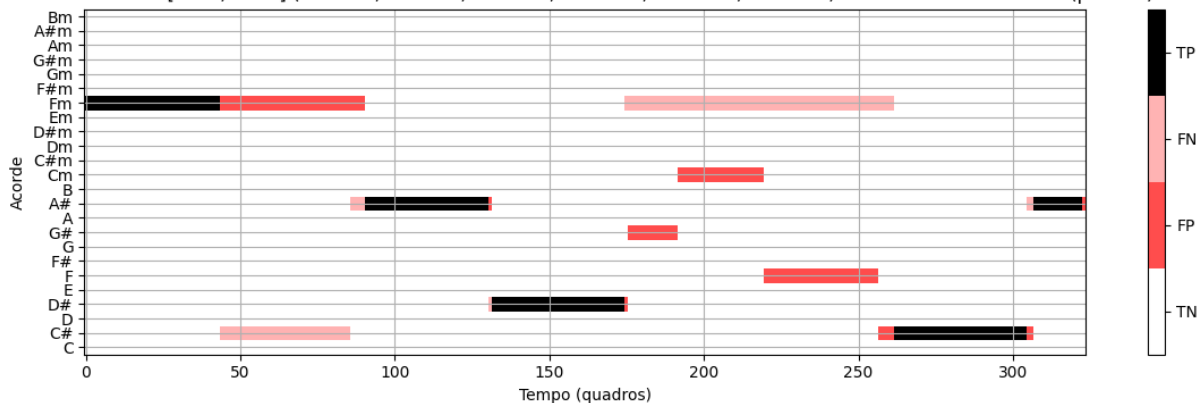


Figura 66 – Cromagramas gerados para as músicas *As Praias Desertas* e *Aula de matemática*.

Fonte: Testes MPB do (PITA, 2024).

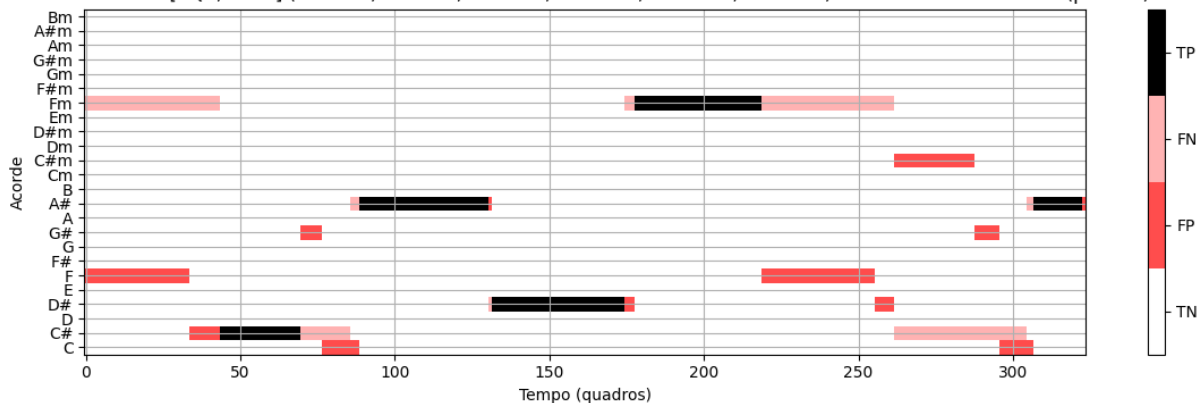
APÊNDICE D – RESULTADOS DOS EXPERIMENTOS NO CONTEXTO DA MPB.

Música: Comeu [STFT; HMM] (TP=186, FP=138, FN=137, P=0.574, R=0.576, F=0.575) Matriz: Uniforme 24x24 ($p=0.15$)



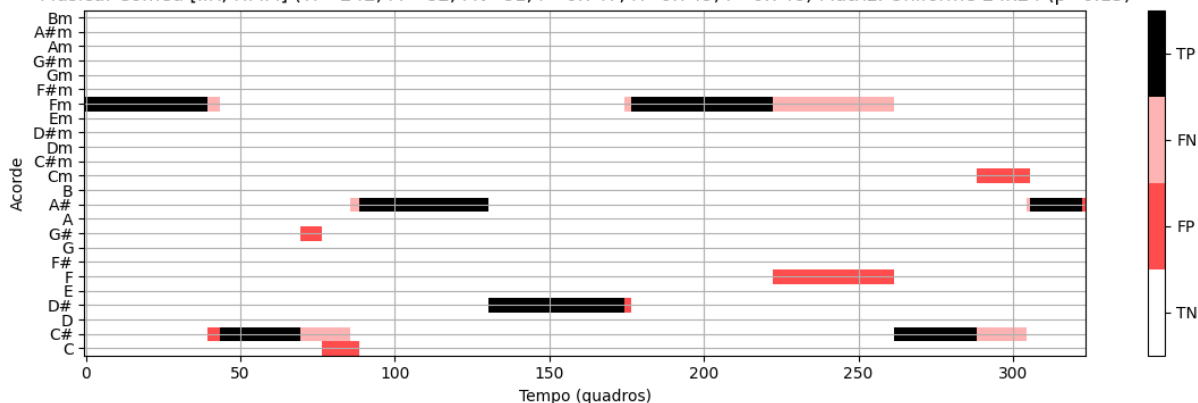
(a) Matriz Uniforme + STFT

Música: Comeu [CQT; HMM] (TP=168, FP=156, FN=155, P=0.519, R=0.520, F=0.519) Matriz: Uniforme 24x24 ($p=0.15$)



(b) Matriz Uniforme + CQT

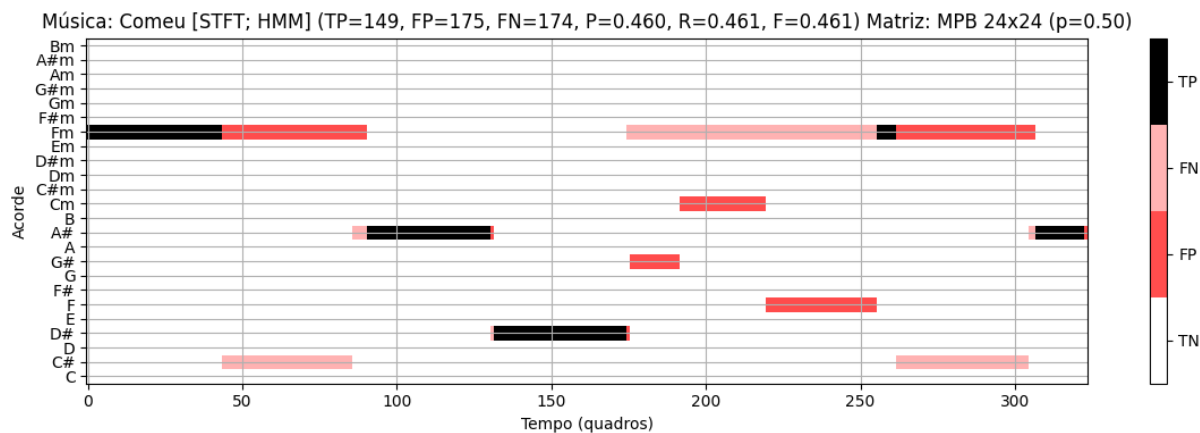
Música: Comeu [IIR; HMM] (TP=242, FP=82, FN=81, P=0.747, R=0.749, F=0.748) Matriz: Uniforme 24x24 ($p=0.15$)



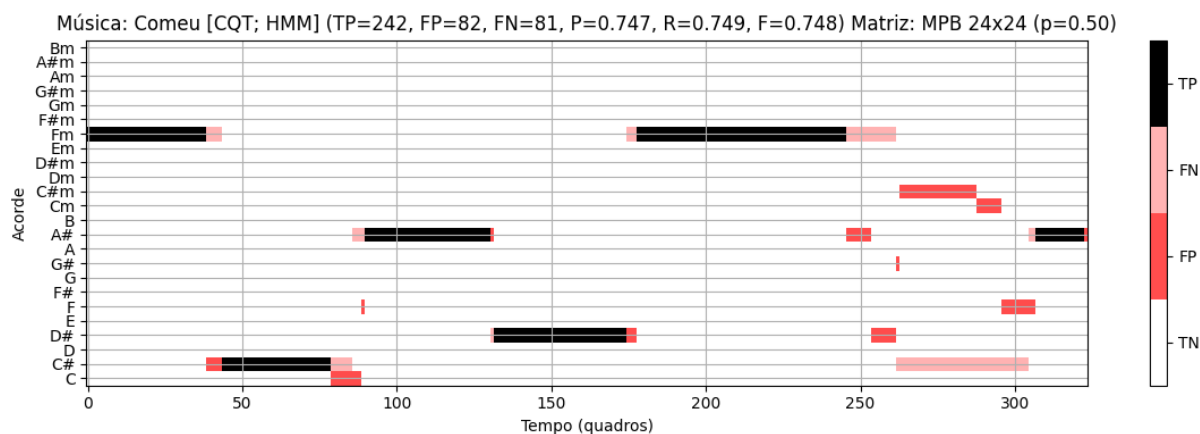
(c) Matriz Uniforme + IIR

Figura 67 – Resultados com a Matriz Uniforme para a música *Comeu*.

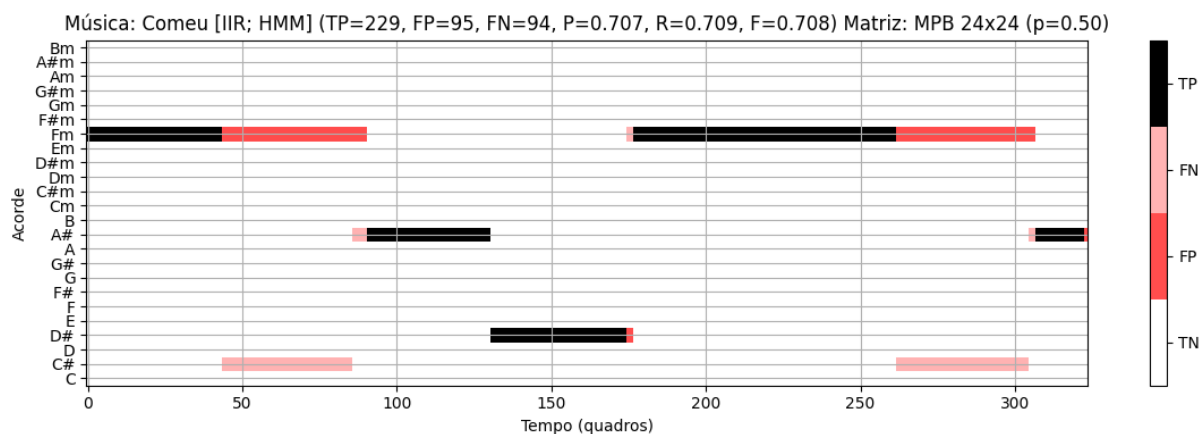
Fonte: Testes MPB do (PITA, 2024).



(a) Matriz MPB + STFT



(b) Matriz MPB + CQT

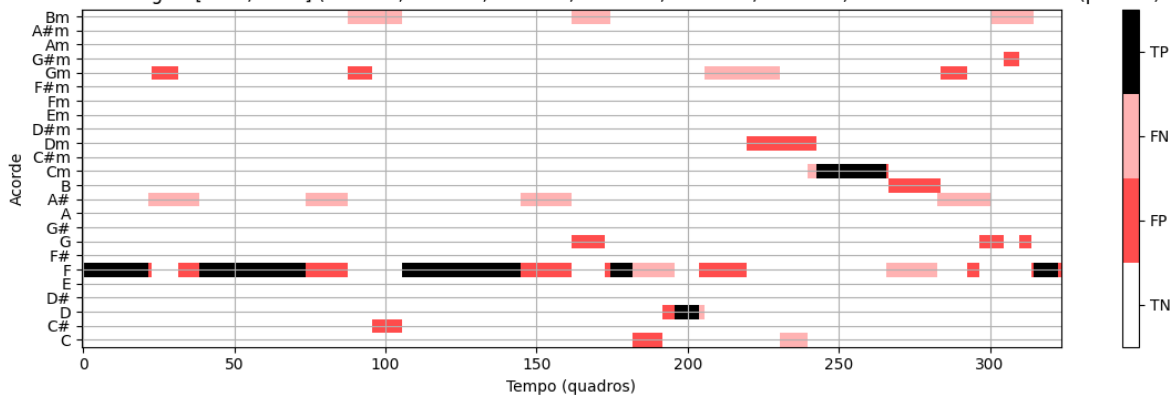


(c) Matriz MPB + IIR

Figura 68 – Resultados com a Matriz Triádica Alterada para a música *Comeu*.

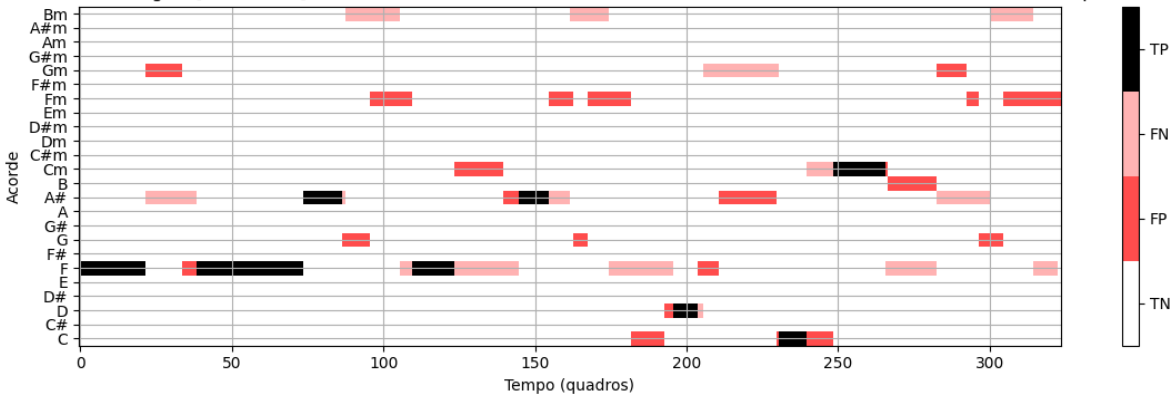
Fonte: Testes MPB do (PITA, 2024).

Música: Deixa Sangrar [STFT; HMM] (TP=142, FP=182, FN=181, P=0.438, R=0.440, F=0.439) Matriz: Uniforme 24x24 (p=0.15)



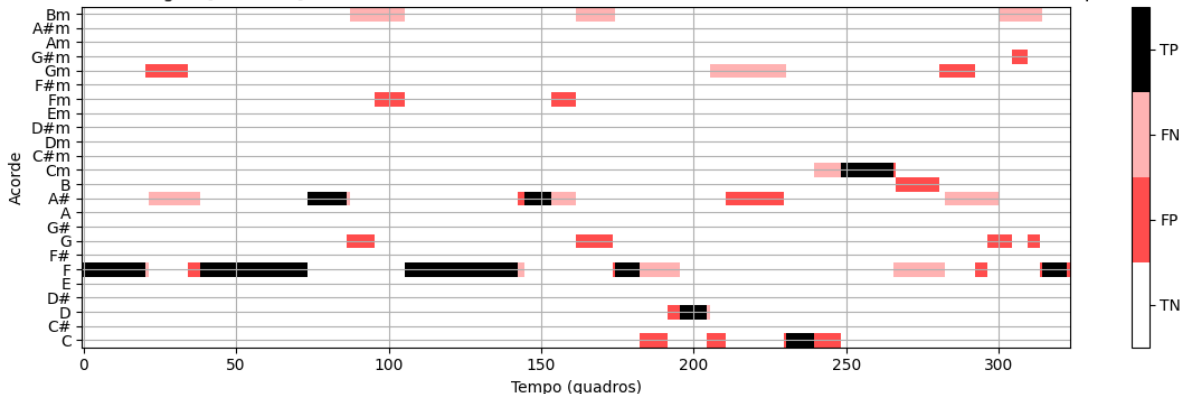
(a) Matriz Uniforme + STFT

Música: Deixa Sangrar [CQT; HMM] (TP=128, FP=196, FN=195, P=0.395, R=0.396, F=0.396) Matriz: Uniforme 24x24 (p=0.15)



(b) Matriz Uniforme + CQT

Música: Deixa Sangrar [IIR; HMM] (TP=166, FP=158, FN=157, P=0.512, R=0.514, F=0.513) Matriz: Uniforme 24x24 (p=0.15)

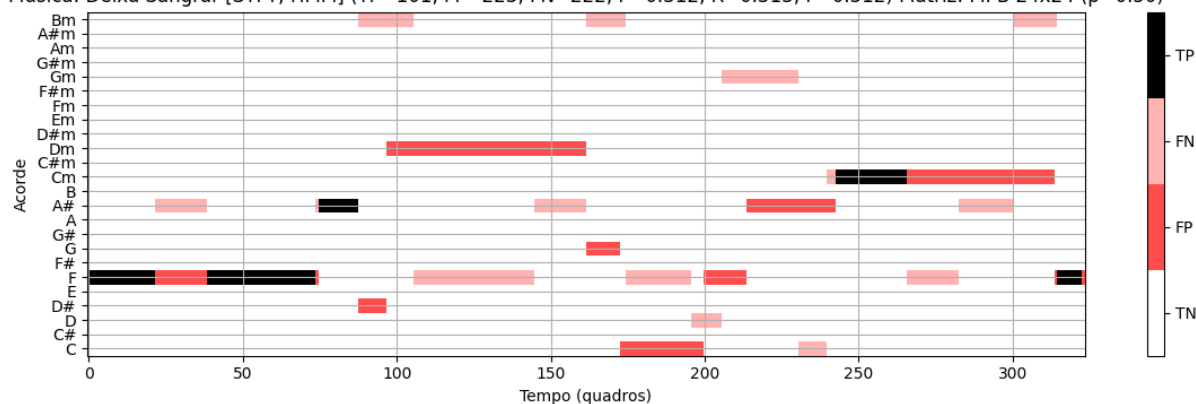


(c) Matriz Uniforme + IIR

Figura 69 – Resultados com a Matriz Uniforme para a música *Deixa sangrar*.

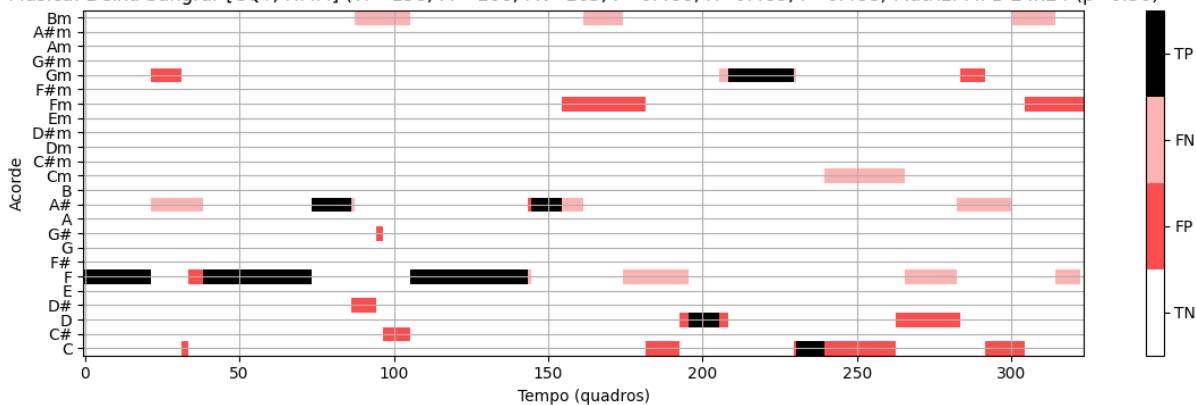
Fonte: Testes MPB do (PITA, 2024).

Música: Deixa Sangrar [STFT; HMM] (TP=101, FP=223, FN=222, P=0.312, R=0.313, F=0.312) Matriz: MPB 24x24 ($p=0.50$)



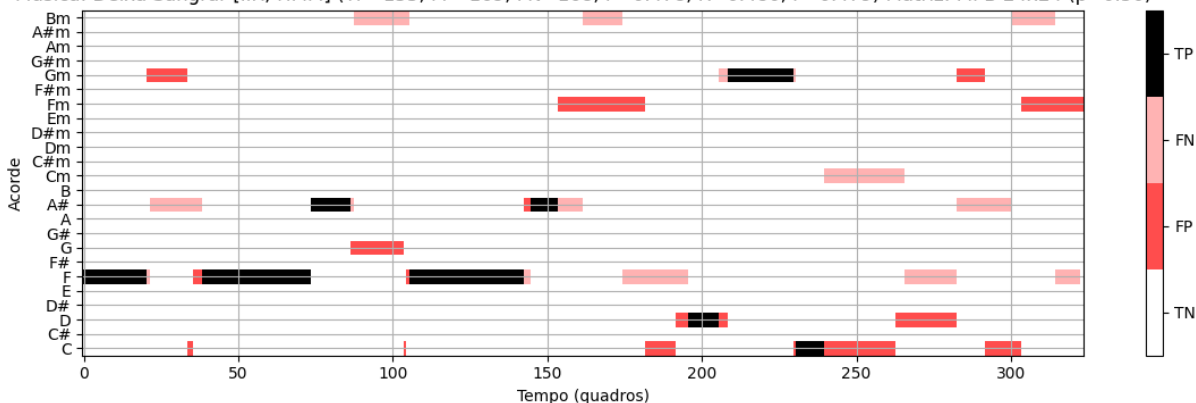
(a) Matriz MPB + STFT

Música: Deixa Sangrar [CQT; HMM] (TP=158, FP=166, FN=165, P=0.488, R=0.489, F=0.488) Matriz: MPB 24x24 ($p=0.50$)



(b) Matriz MPB + CQT

Música: Deixa Sangrar [IIR; HMM] (TP=155, FP=169, FN=168, P=0.478, R=0.480, F=0.479) Matriz: MPB 24x24 ($p=0.50$)

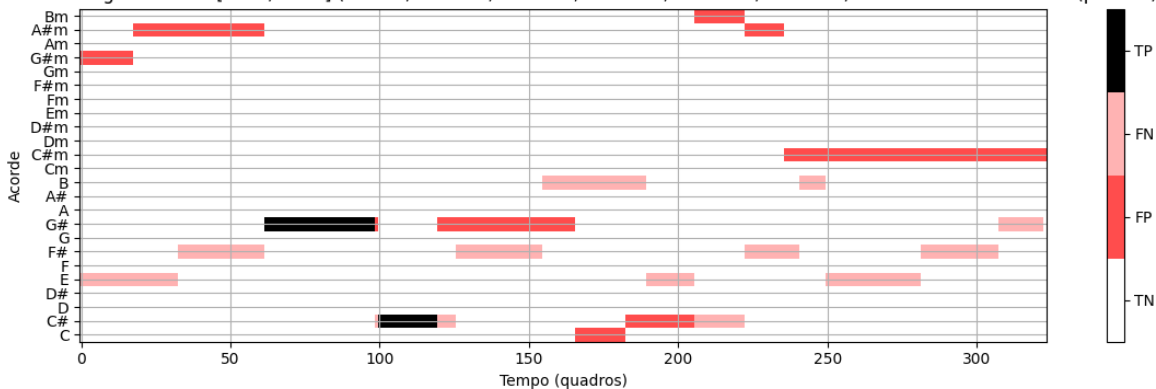


(c) Matriz MPB + IIR

Figura 70 – Resultados com a Matriz Triádica Alterada para a música *Deixa sangrar*.

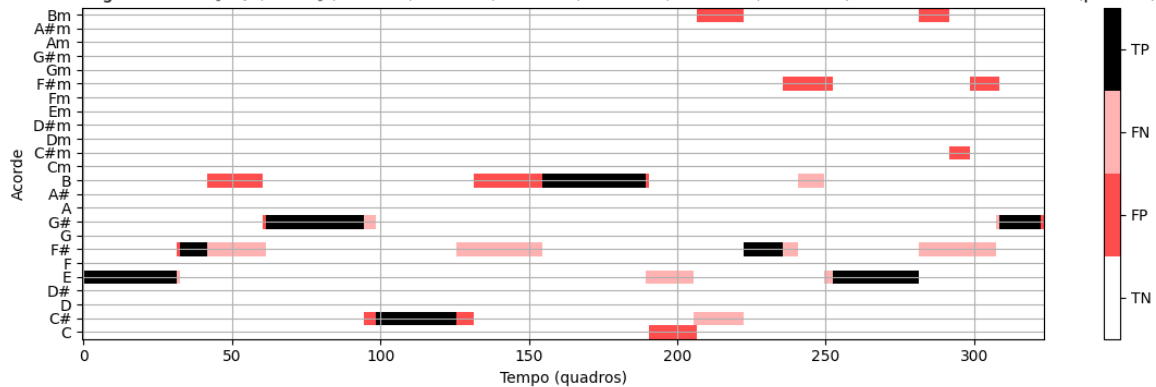
Fonte: Testes MPB do (PITA, 2024).

Música: Até Segunda-Feira [STFT; HMM] (TP=57, FP=267, FN=266, P=0.176, R=0.176, F=0.176) Matriz: Uniforme 24x24 (p=0.15)



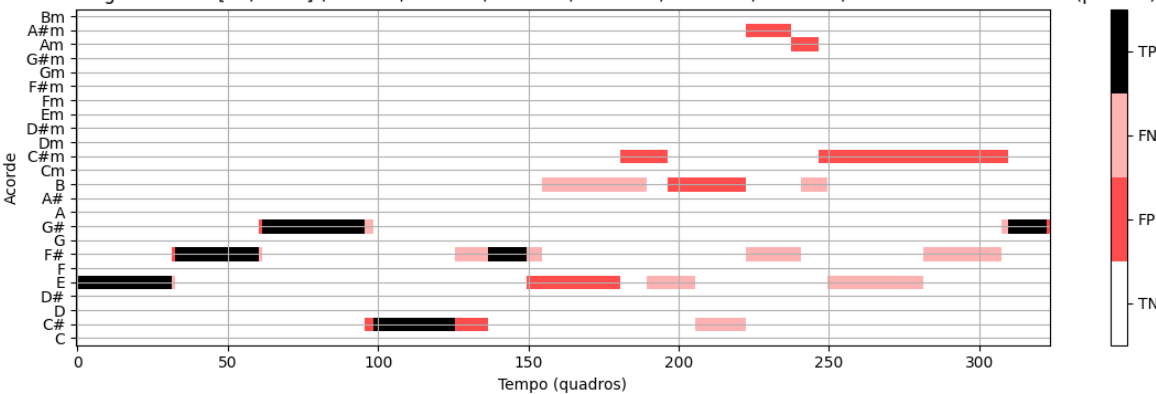
(a) Matriz Uniforme + STFT

Música: Até Segunda-Feira [CQT; HMM] (TP=192, FP=132, FN=131, P=0.593, R=0.594, F=0.594) Matriz: Uniforme 24x24 (p=0.15)



(b) Matriz Uniforme + CQT

Música: Até Segunda-Feira [IIR; HMM] (TP=147, FP=177, FN=176, P=0.454, R=0.455, F=0.454) Matriz: Uniforme 24x24 (p=0.15)

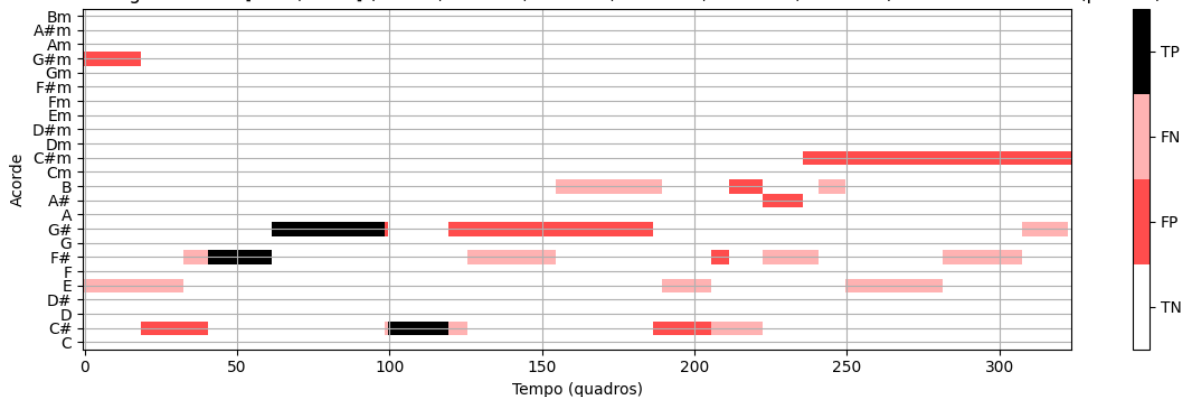


(c) Matriz Uniforme + IIR

Figura 71 – Resultados com a Matriz Uniforme para a música *Até segunda-feira*.

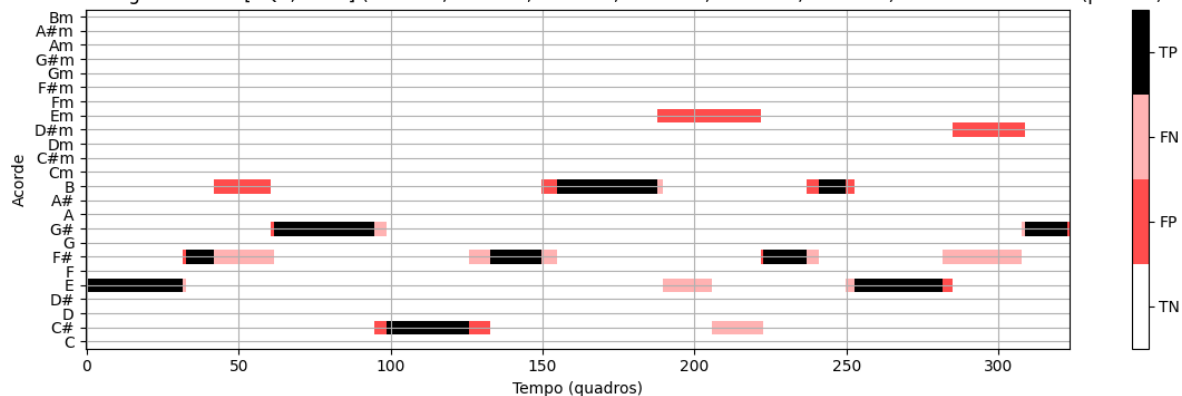
Fonte: Testes MPB do (PITA, 2024).

Música: Ate Segunda-Feira [STFT; HMM] (TP=78, FP=246, FN=245, P=0.241, R=0.241, F=0.241) Matriz: MPB 24x24 (p=0.50)



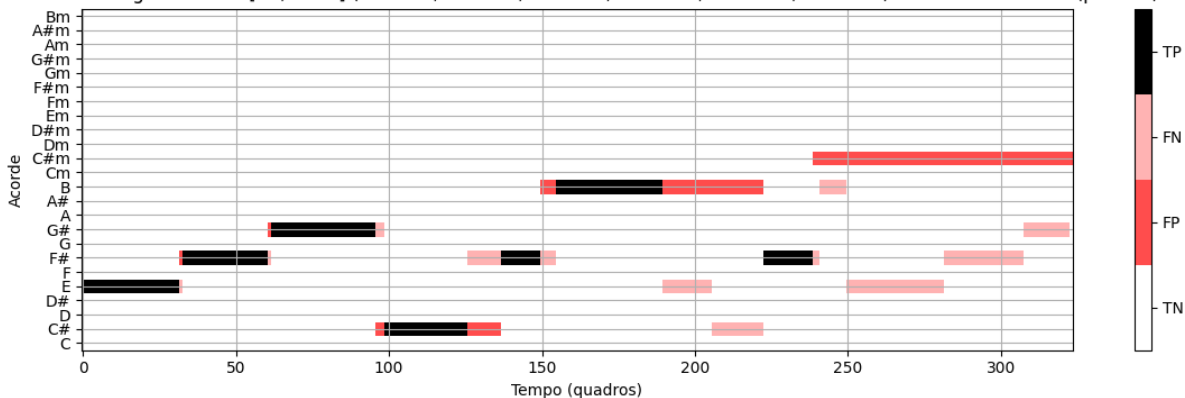
(a) Matriz MPB + STFT

Música: Ate Segunda-Feira [CQT; HMM] (TP=217, FP=107, FN=106, P=0.670, R=0.672, F=0.671) Matriz: MPB 24x24 (p=0.50)



(b) Matriz MPB + CQT

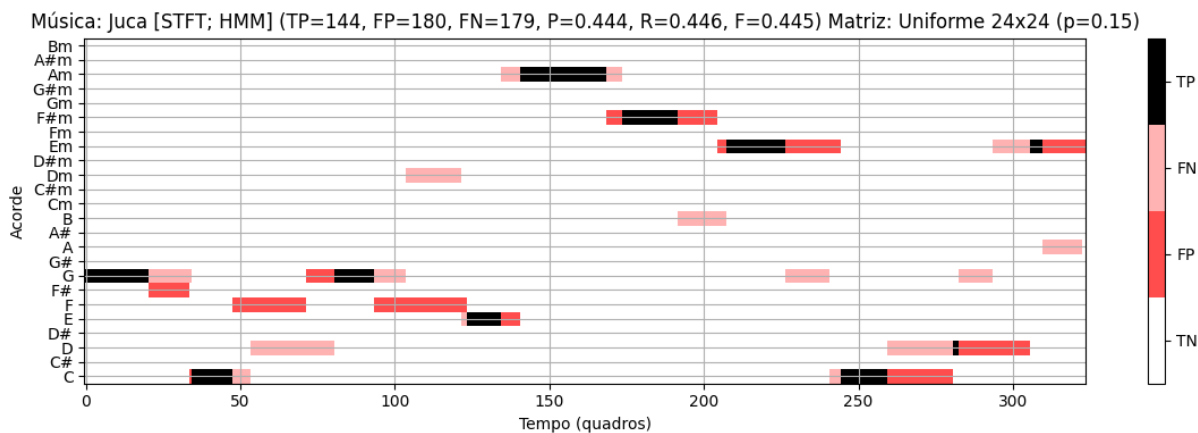
Música: Ate Segunda-Feira [IIR; HMM] (TP=185, FP=139, FN=138, P=0.571, R=0.573, F=0.572) Matriz: MPB 24x24 (p=0.50)



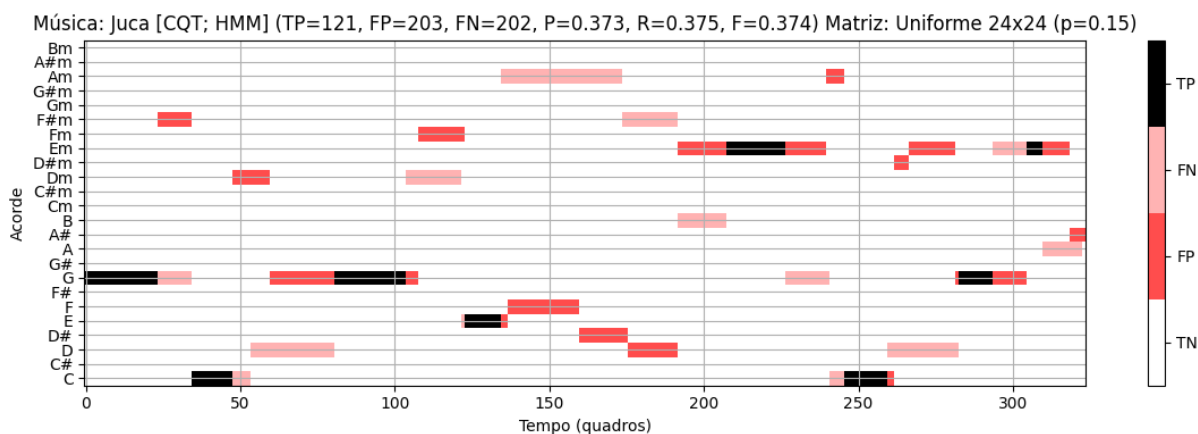
(c) Matriz MPB + IIR

Figura 72 – Resultados com a Matriz Triádica Alterada para a música *Até segunda-feira*.

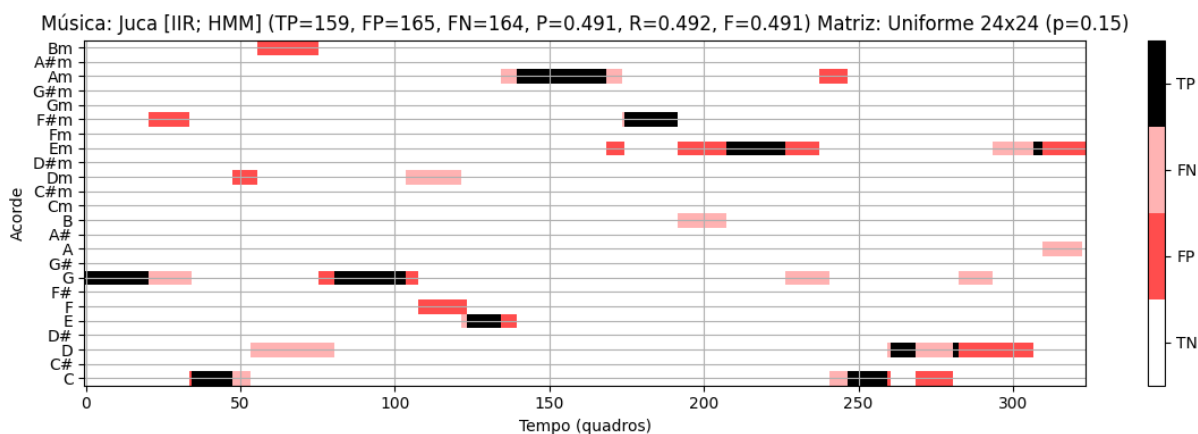
Fonte: Testes MPB do (PITA, 2024).



(a) Matriz Uniforme + STFT



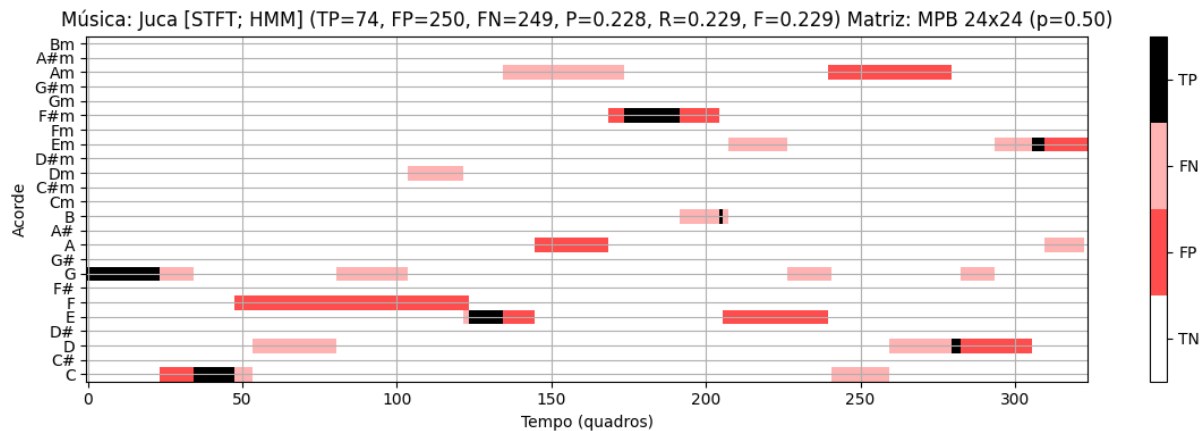
(b) Matriz Uniforme + CQT



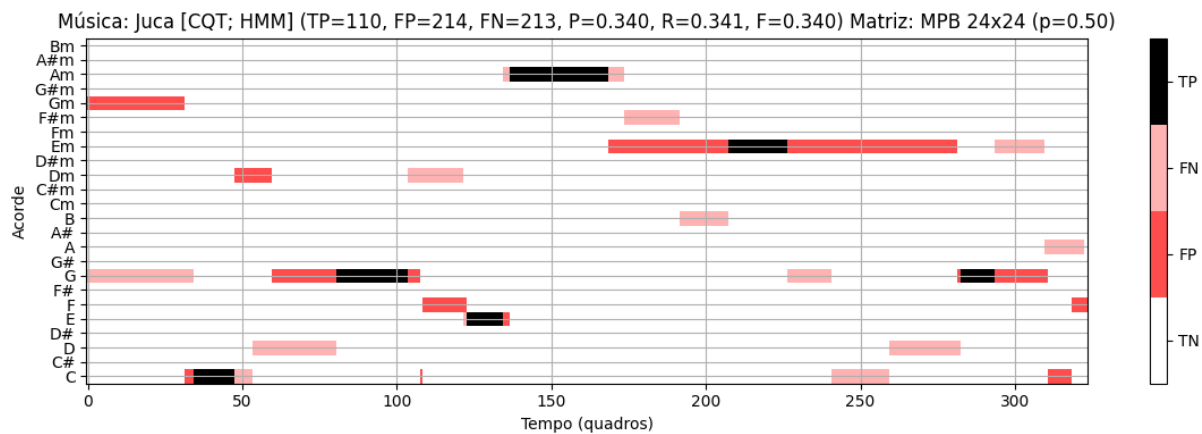
(c) Matriz Uniforme + IIR

Figura 73 – Resultados com a Matriz Uniforme para a música *Juca*.

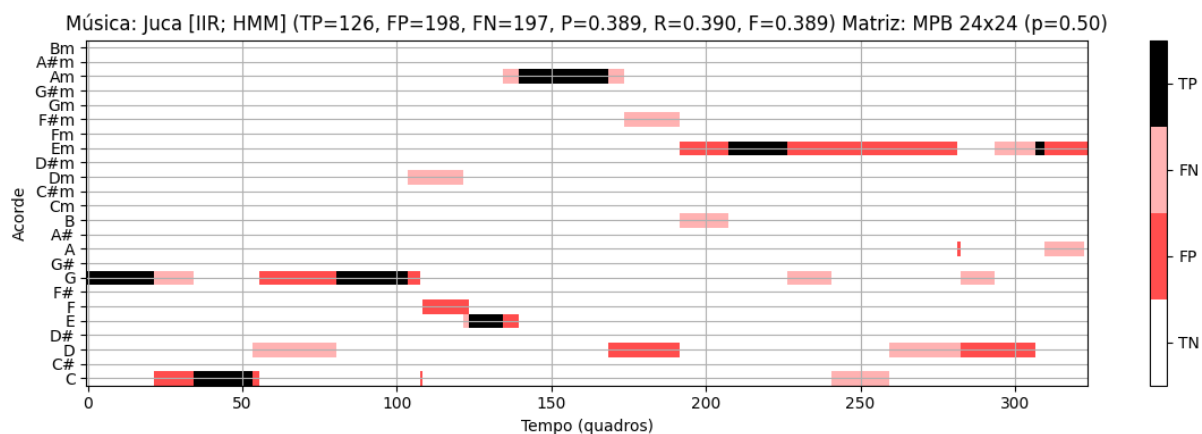
Fonte: Testes MPB do (PITA, 2024).



(a) Matriz MPB + STFT



(b) Matriz MPB + CQT

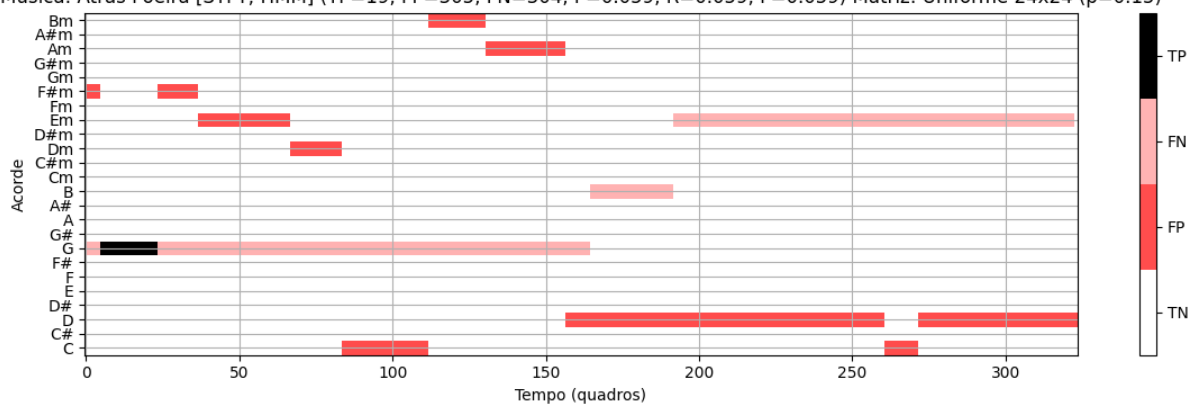


(c) Matriz MPB + IIR

Figura 74 – Resultados com a Matriz Triádica Alterada para a música *Juca*.

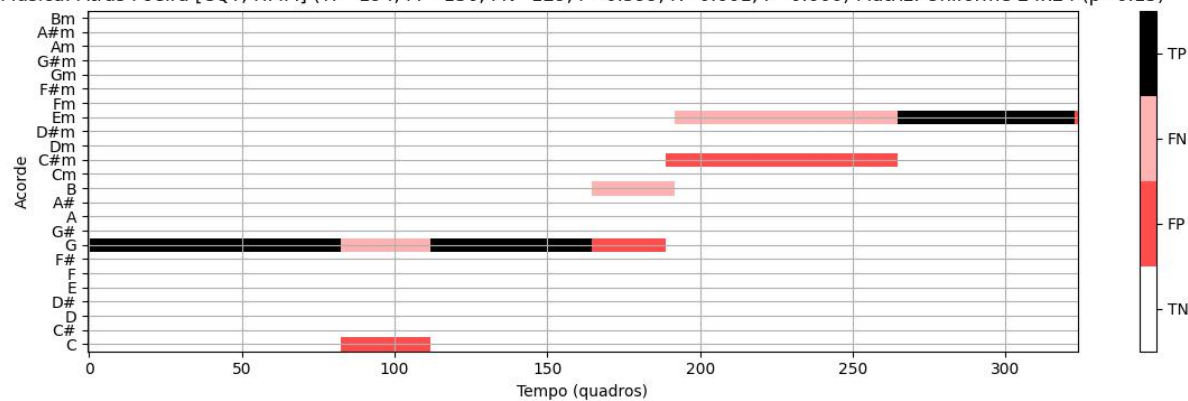
Fonte: Testes MPB do (PITA, 2024).

Música: Atras Poeira [STFT; HMM] (TP=19, FP=305, FN=304, P=0.059, R=0.059, F=0.059) Matriz: Uniforme 24x24 ($p=0.15$)



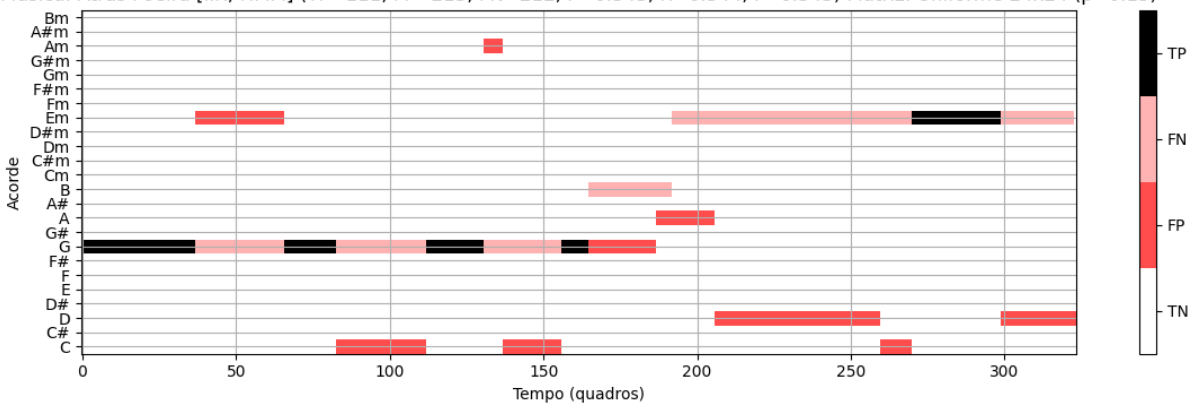
(a) Matriz Uniforme + STFT

Música: Atras Poeira [CQT; HMM] (TP=194, FP=130, FN=129, P=0.599, R=0.601, F=0.600) Matriz: Uniforme 24x24 ($p=0.15$)



(b) Matriz Uniforme + CQT

Música: Atras Poeira [IIR; HMM] (TP=111, FP=213, FN=212, P=0.343, R=0.344, F=0.343) Matriz: Uniforme 24x24 ($p=0.15$)

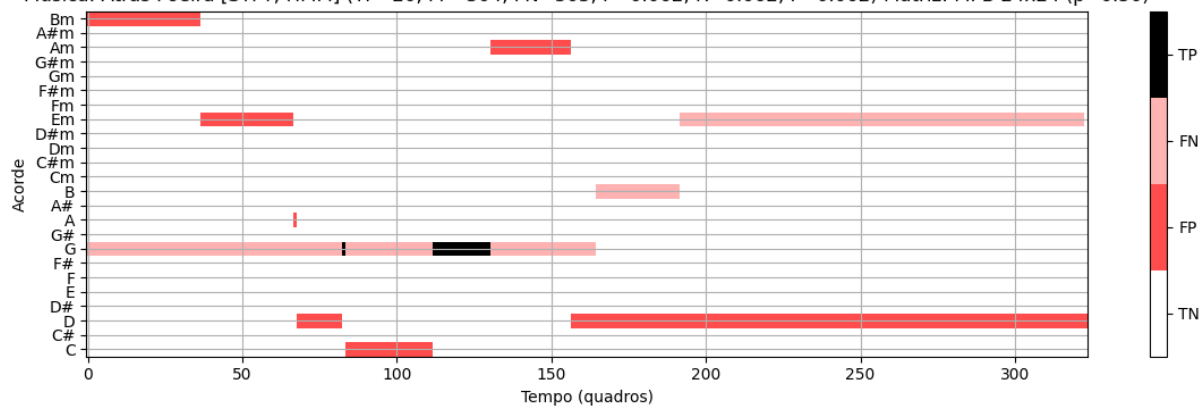


(c) Matriz Uniforme + IIR

Figura 75 – Resultados com a Matriz Uniforme para a música *Atras poeira*.

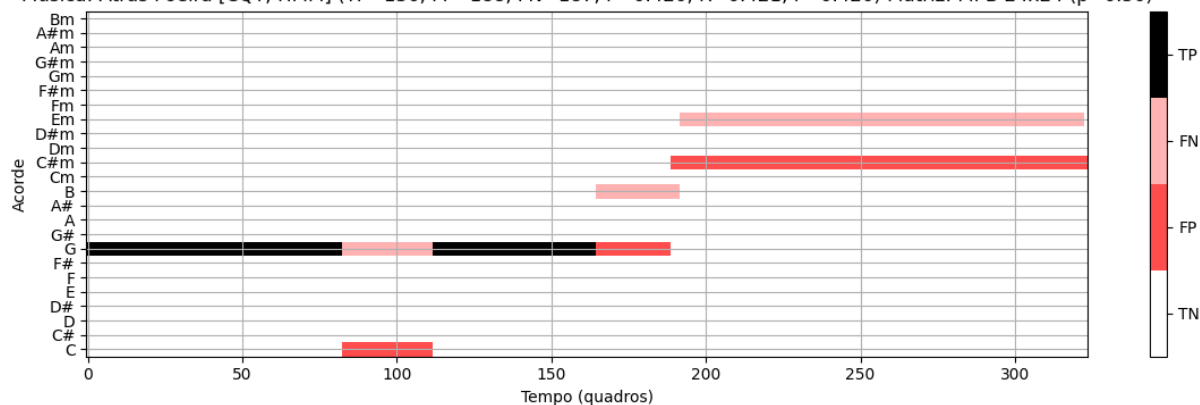
Fonte: Testes MPB do (PITA, 2024).

Música: Atras Poeira [STFT; HMM] (TP=20, FP=304, FN=303, P=0.062, R=0.062, F=0.062) Matriz: MPB 24x24 ($p=0.50$)



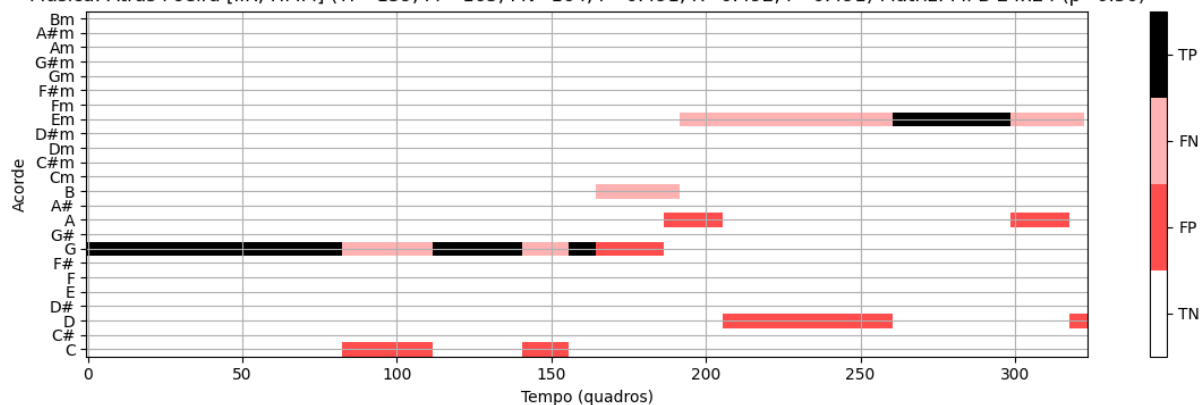
(a) Matriz MPB + STFT

Música: Atras Poeira [CQT; HMM] (TP=136, FP=188, FN=187, P=0.420, R=0.421, F=0.420) Matriz: MPB 24x24 ($p=0.50$)



(b) Matriz MPB + CQT

Música: Atras Poeira [IIR; HMM] (TP=159, FP=165, FN=164, P=0.491, R=0.492, F=0.491) Matriz: MPB 24x24 ($p=0.50$)

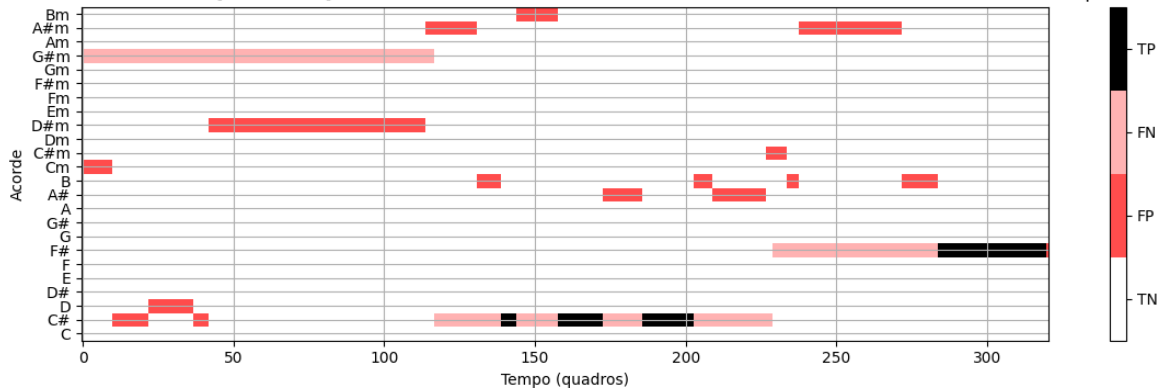


(c) Matriz MPB + IIR

Figura 76 – Resultados com a Matriz Triádica Alterada para a música *Atrás poeira*.

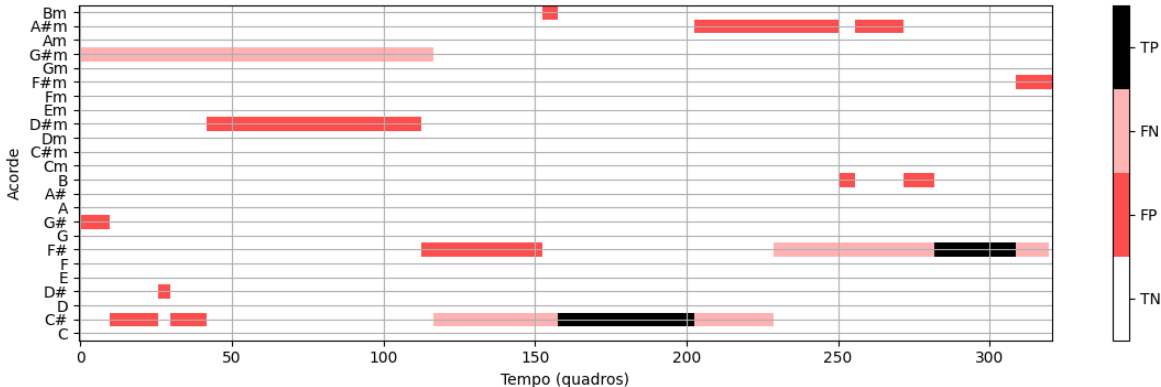
Fonte: Testes MPB do (PITA, 2024).

Música: Começar De Novo [STFT; HMM] (TP=73, FP=248, FN=247, P=0.227, R=0.228, F=0.228) Matriz: Uniforme 24x24 (p=0.15)



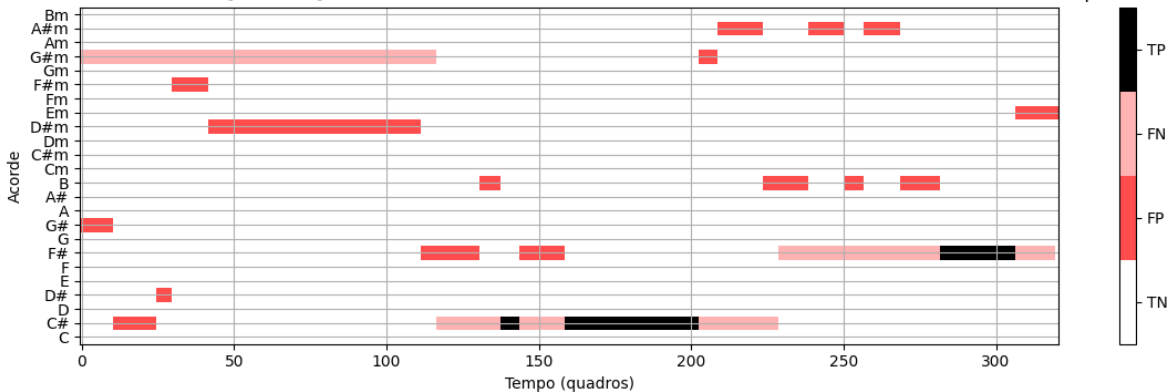
(a) Matriz Uniforme + STFT

Música: Começar De Novo [CQT; HMM] (TP=72, FP=249, FN=248, P=0.224, R=0.225, F=0.225) Matriz: Uniforme 24x24 (p=0.15)



(b) Matriz Uniforme + CQT

Música: Começar De Novo [IIR; HMM] (TP=75, FP=246, FN=245, P=0.234, R=0.234, F=0.234) Matriz: Uniforme 24x24 (p=0.15)

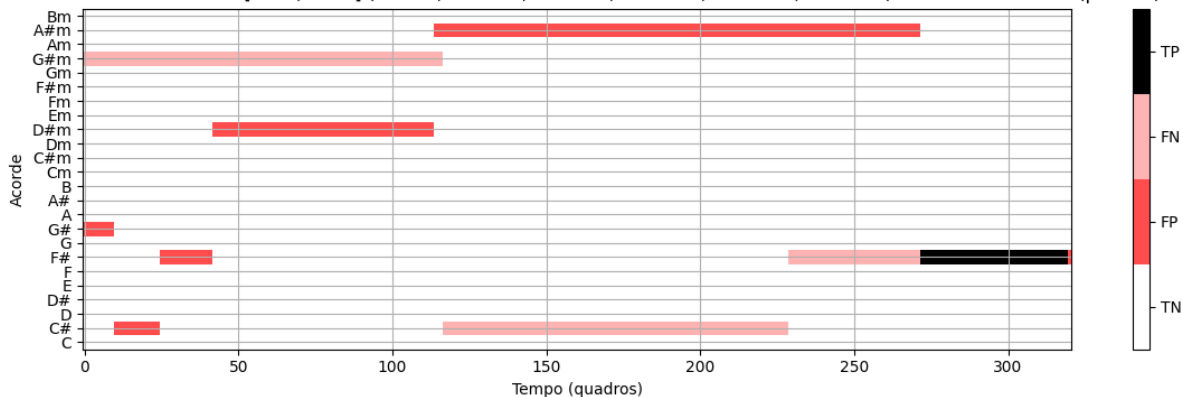


(c) Matriz Uniforme + IIR

Figura 77 – Resultados com a Matriz Uniforme para a música *Começar de novo*.

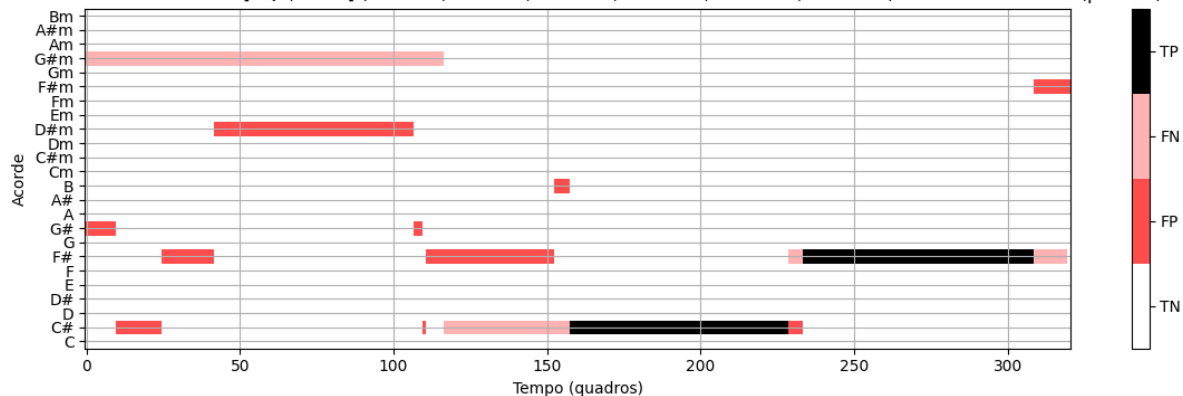
Fonte: Testes MPB do (PITA, 2024).

Música: Começar De Novo [STFT; HMM] (TP=48, FP=273, FN=272, P=0.150, R=0.150, F=0.150) Matriz: MPB 24x24 (p=0.50)



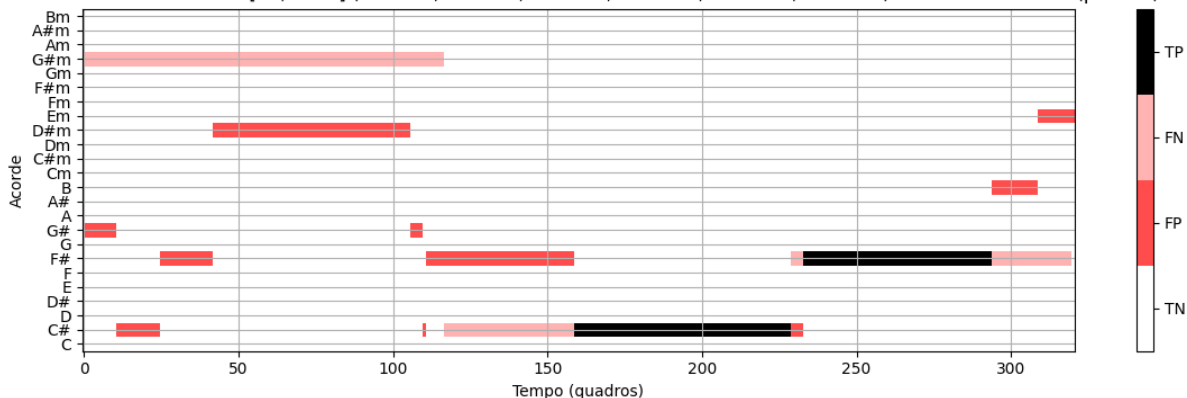
(a) Matriz MPB + STFT

Música: Começar De Novo [CQT; HMM] (TP=146, FP=175, FN=174, P=0.455, R=0.456, F=0.456) Matriz: MPB 24x24 (p=0.50)



(b) Matriz MPB + CQT

Música: Começar De Novo [IIR; HMM] (TP=131, FP=190, FN=189, P=0.408, R=0.409, F=0.409) Matriz: MPB 24x24 (p=0.50)

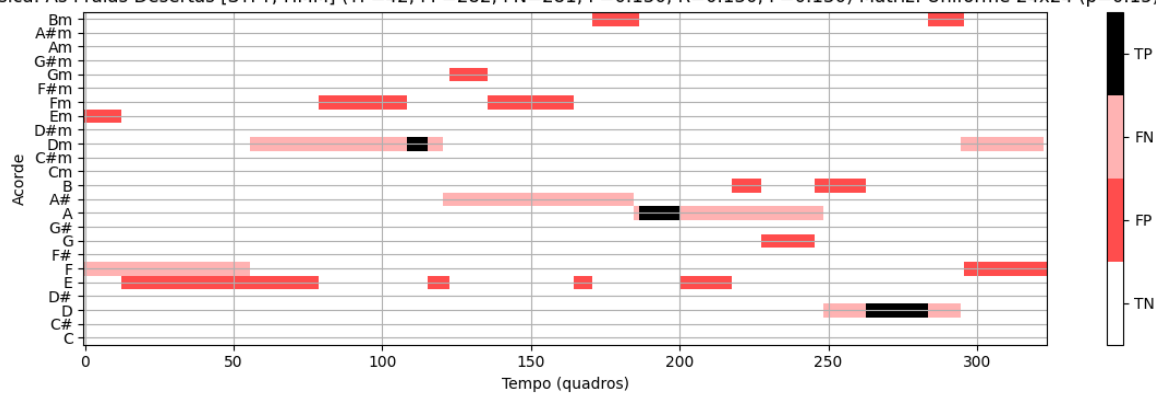


(c) Matriz MPB + IIR

Figura 78 – Resultados com a Matriz Triádica Alterada para a música *Começar de novo*.

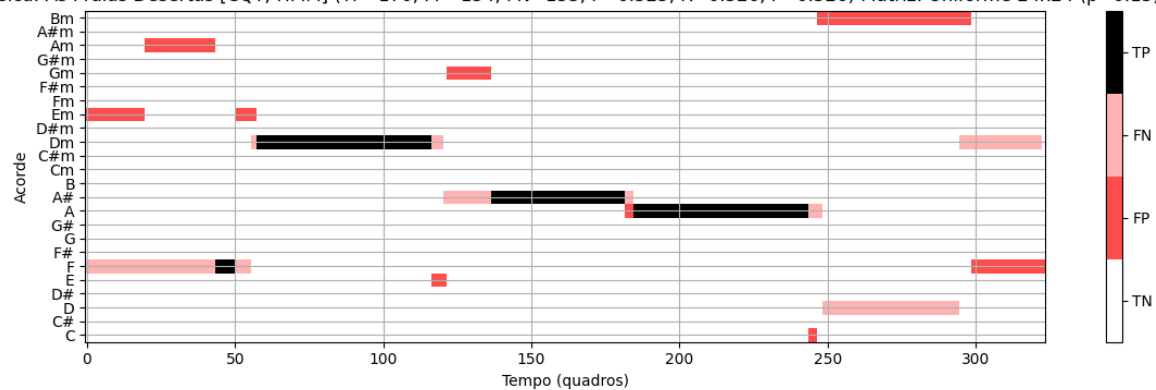
Fonte: Testes MPB do (PITA, 2024).

Música: As Praias Desertas [STFT; HMM] (TP=42, FP=282, FN=281, P=0.130, R=0.130, F=0.130) Matriz: Uniforme 24x24 ($p=0.15$)



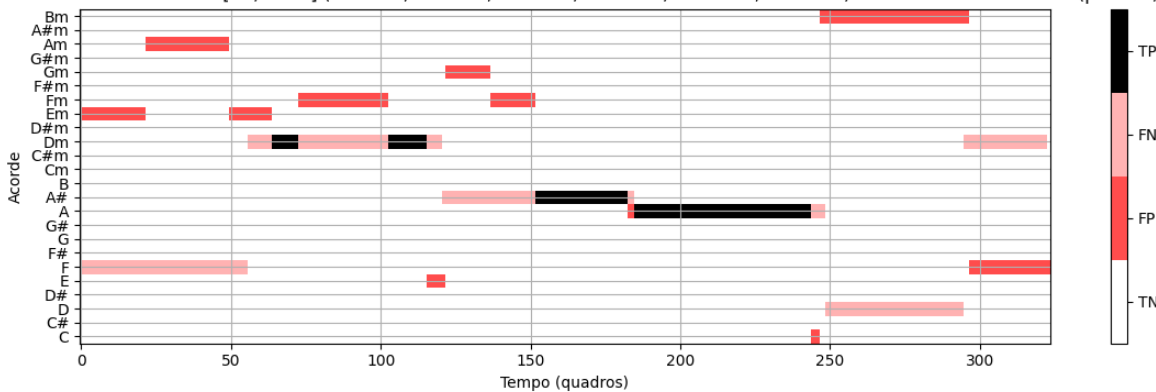
(a) Matriz Uniforme + STFT

Música: As Praias Desertas [CQT; HMM] (TP=170, FP=154, FN=153, P=0.525, R=0.526, F=0.526) Matriz: Uniforme 24x24 ($p=0.15$)



(b) Matriz Uniforme + CQT

Música: As Praias Desertas [IIR; HMM] (TP=112, FP=212, FN=211, P=0.346, R=0.347, F=0.346) Matriz: Uniforme 24x24 ($p=0.15$)

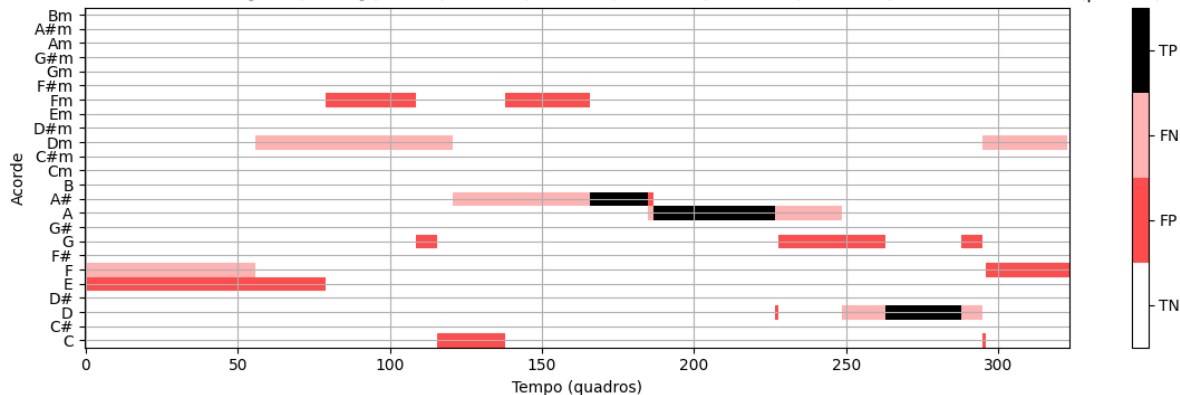


(c) Matriz Uniforme + IIR

Figura 79 – Resultados com a Matriz Uniforme para a música *As praias desertas*.

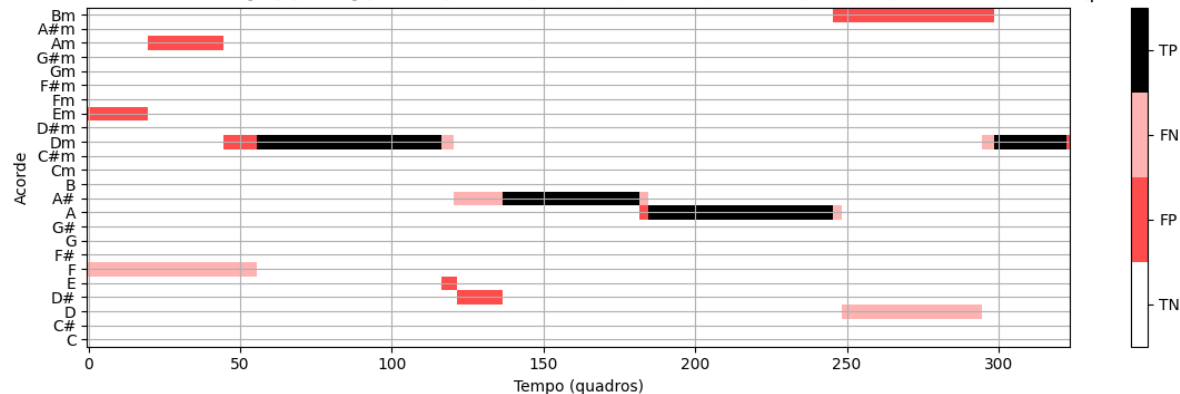
Fonte: Testes MPB do (PITA, 2024).

Música: As Praias Desertas [STFT; HMM] (TP=84, FP=240, FN=239, P=0.259, R=0.260, F=0.260) Matriz: MPB 24x24 (p=0.50)



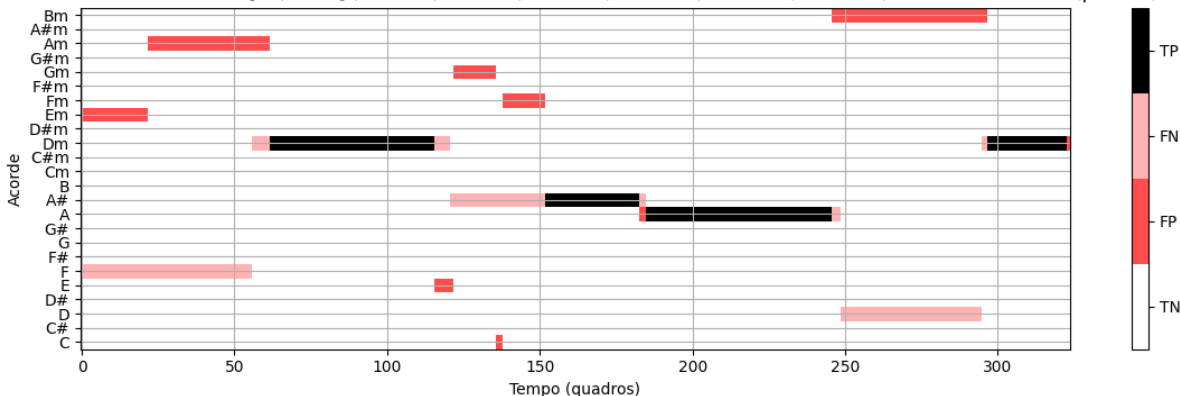
(a) Matriz MPB + STFT

Música: As Praias Desertas [CQT; HMM] (TP=191, FP=133, FN=132, P=0.590, R=0.591, F=0.590) Matriz: MPB 24x24 (p=0.50)



(b) Matriz MPB + CQT

Música: As Praias Desertas [IIR; HMM] (TP=172, FP=152, FN=151, P=0.531, R=0.533, F=0.532) Matriz: MPB 24x24 (p=0.50)

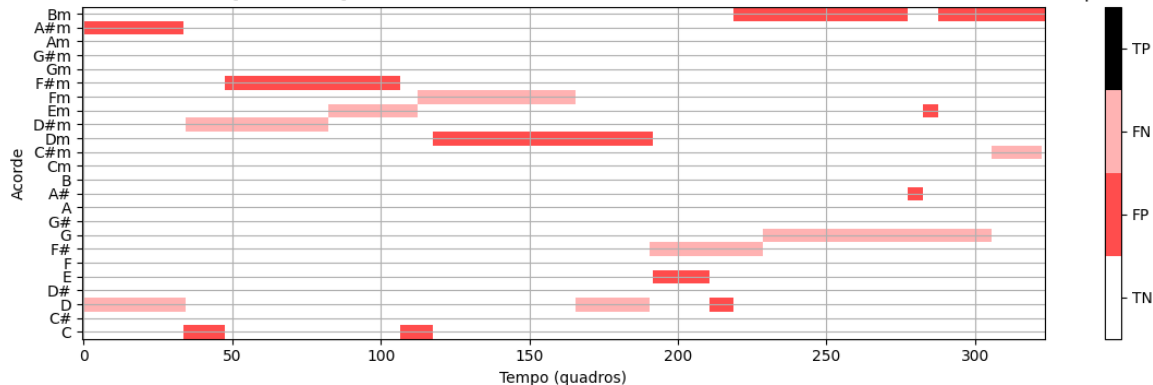


(c) Matriz MPB + IIR

Figura 80 – Resultados com a Matriz Triádica Alterada para a música *As praias desertas*.

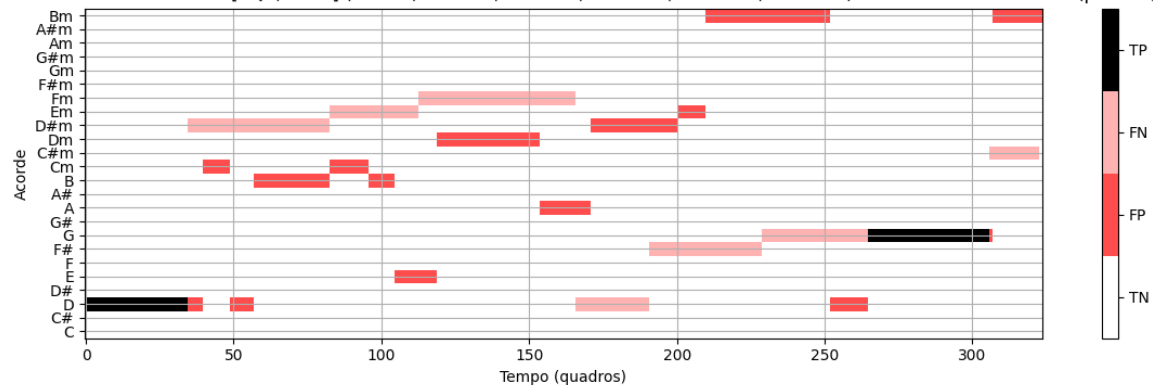
Fonte: Testes MPB do (PITA, 2024).

Música: Aula De Matematica [STFT; HMM] (TP=0, FP=324, FN=323, P=0.000, R=0.000, F=0.000) Matriz: Uniforme 24x24 (p=0.15)



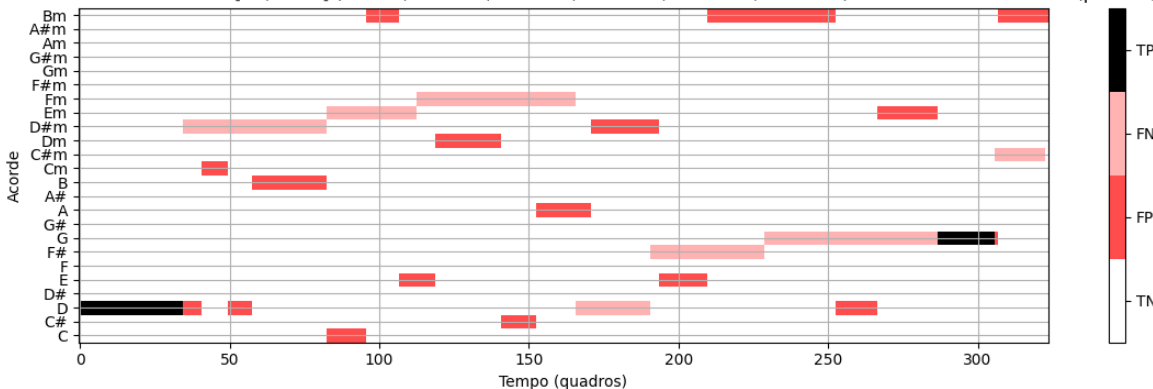
(a) Matriz Uniforme + STFT

Música: Aula De Matematica [CQT; HMM] (TP=76, FP=248, FN=247, P=0.235, R=0.235, F=0.235) Matriz: Uniforme 24x24 (p=0.15)



(b) Matriz Uniforme + CQT

Música: Aula De Matematica [IIR; HMM] (TP=54, FP=270, FN=269, P=0.167, R=0.167, F=0.167) Matriz: Uniforme 24x24 (p=0.15)

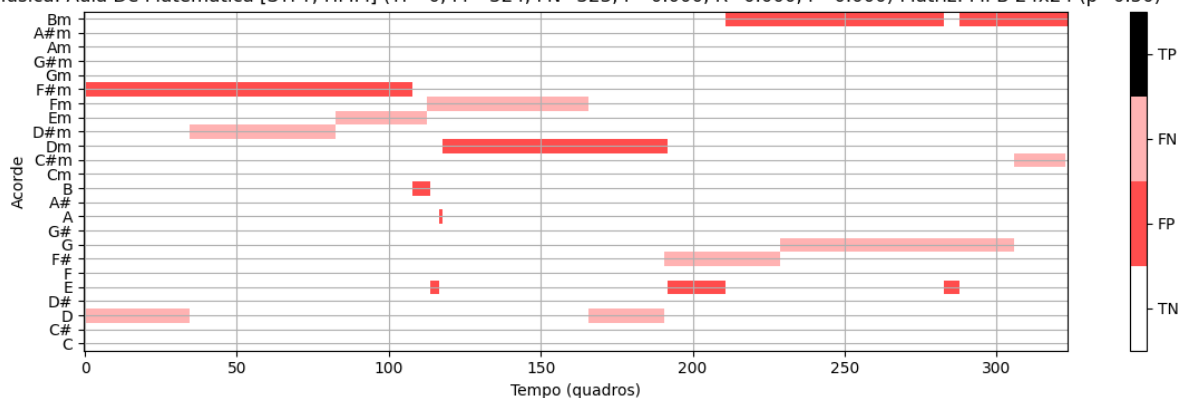


(c) Matriz Uniforme + IIR

Figura 81 – Resultados com a Matriz Uniforme para a música *Aula de matemática*.

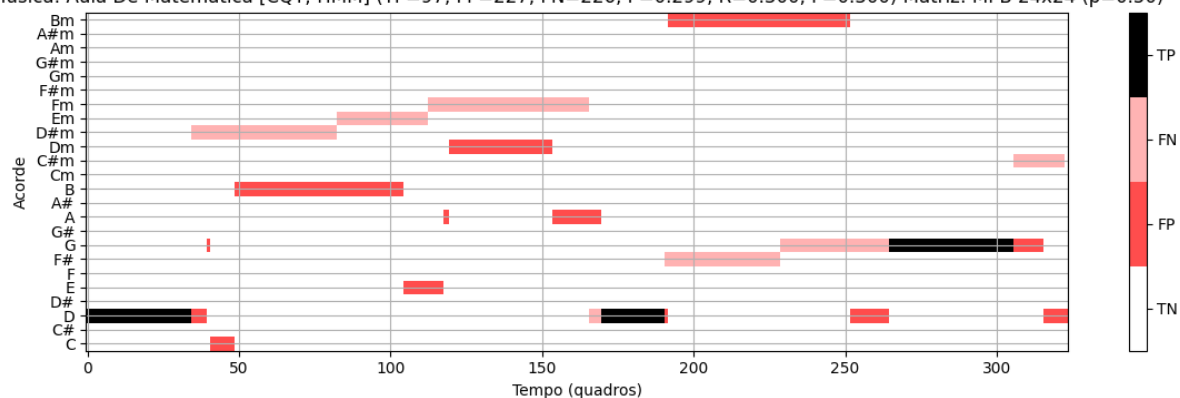
Fonte: Testes MPB do (PITA, 2024).

Música: Aula De Matematica [STFT; HMM] (TP=0, FP=324, FN=323, P=0.000, R=0.000, F=0.000) Matriz: MPB 24x24 ($p=0.50$)



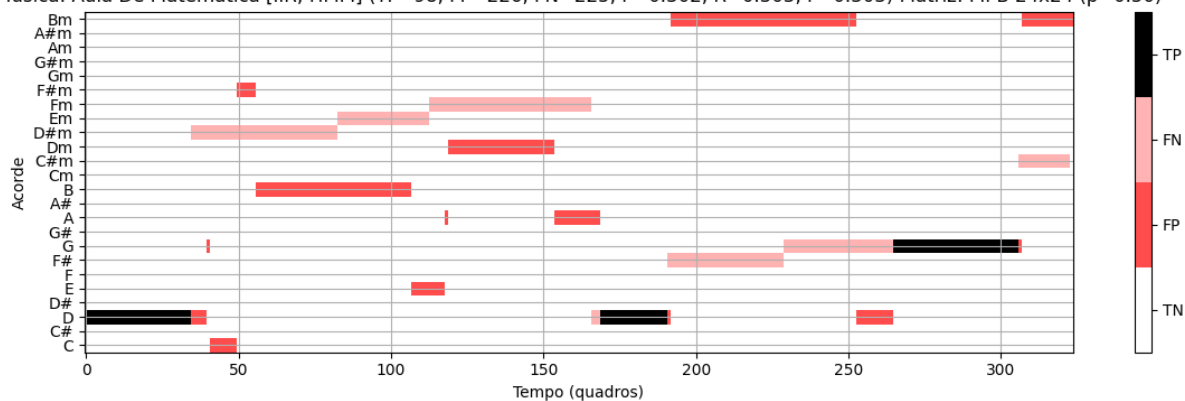
(a) Matriz MPB + STFT

Música: Aula De Matematica [CQT; HMM] (TP=97, FP=227, FN=226, P=0.299, R=0.300, F=0.300) Matriz: MPB 24x24 ($p=0.50$)



(b) Matriz MPB + CQT

Música: Aula De Matematica [IIR; HMM] (TP=98, FP=226, FN=225, P=0.302, R=0.303, F=0.303) Matriz: MPB 24x24 ($p=0.50$)



(c) Matriz MPB + IIR

Figura 82 – Resultados com a Matriz Triádica Alterada para a música *Aula de matemática*.

Fonte: Testes MPB do (PITA, 2024).