

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

LUCAS BAIROS ETCHEVERRIA DE LIMA

**Avaliação de um Método de Correlação para
Aplicação em um Detector e Classificador de
Acordes**

Porto Alegre

2018

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

LUCAS BAIROS ETCHEVERRIA DE LIMA

**Avaliação de um Método de Correlação para Aplicação
em um Detector e Classificador de Acordes**

Projeto de Diplomação apresentado ao Departamento de Engenharia Elétrica da Escola de Engenharia da Universidade Federal do Rio Grande do Sul, como requisito parcial para Graduação em Engenharia Elétrica

Orientador: Prof. Dr. Adalberto Schuck Jr.

Porto Alegre

2018

LUCAS BAIROS ETCHEVERRIA DE LIMA

**Avaliação de um Método de Correlação para Aplicação
em um Detector e Classificador de Acordes**

Projeto de Diplomação apresentado ao Departamento de Engenharia Elétrica da Escola de Engenharia da Universidade Federal do Rio Grande do Sul, como requisito parcial para Graduação em Engenharia Elétrica

Prof. Dr. Adalberto Schuck Jr.
Orientador - UFRGS

Prof. Dr. Ály Ferreira Flores Filho
Chefe do Departamento de Engenharia Elétrica (DELET) - UFRGS

Aprovado em 17 de janeiro de 2018.

BANCA EXAMINADORA

Prof. Dr. Adalberto Schuck Jr.
UFRGS

Prof. Dr. Marcelo de Oliveira Johann
UFRGS

Prof. Dr. Rodrigo Schramm
UFRGS

Resumo

RESUMO

Desenvolvimento de um sistema de identificação e classificação de frequências sonoras simultâneas (acordes) com aplicação voltada para o ramo musical baseado no coeficiente de correlação e em uma função linear de *threshold*. Utilizando o *software* MATLAB 2016a, são elaboradas funções que calculam a correlação entre um acorde e uma base de dados de notas individuais previamente gerada, contendo 36 notas gravadas de uma guitarra elétrica, utilizando uma interface de áudio para gravação em 96 kHz de amostragem, que é reduzida posteriormente para 48 kHz para fins de processamento. Todas as operações sobre os sinais são feitas no domínio frequência, utilizando a FFT (Fast Fourier Transform) para o tratamento dos sinais de áudio. São desenvolvidos dois tipos de seletores de notas a fim de fornecer a saída na forma de notas detectadas. No primeiro seletor, todas as notas consideradas detectadas são informadas na saída. No segundo seletor, dentre todos os valores de correlação, o sistema escolhe os 6 maiores valores para apontar como prováveis notas executadas no sinal de entrada. São realizados testes com 10 amostras de 20 acordes diferentes para caracterizar os classificadores quanto à sua taxa de acerto, falsos positivos e negativos. Também é feita análise da performance dos classificadores utilizando curvas ROC. De posse dos resultados finais, concluiu-se que o coeficiente de correlação pode ser considerado uma boa ferramenta para a comparação dos sinais na aplicação proposta. O primeiro seletor demonstrou melhor taxa de acerto do que o segundo seletor, por ser menos restritivo e informar todas as notas detectadas. As taxas de acerto máximas para o primeiro seletor variam de 0 a 100 % dependendo do acorde, enquanto as taxas do segundo variam de 0 a 66,6667% de acerto máximo. Foi mostrado que para acordes de tons diferentes e construídos de diferentes intervalos, a taxa de acerto dos testes variou de uma forma sem padrão específico. Todavia, a função utilizada como limiar de detecção para o classificador apresentou performance limitada para os dois seletores desenvolvidos, onde, na maioria dos casos, para altos valores de taxa de acerto, esta retorna alta taxa de falsos positivos e, para baixos valores de falsos positivos, esta obtém baixa taxa de verdadeiros positivos.

Palavras-chave: frequência, classificação, acordes, correlação, FFT, detecção

Abstract

ABSTRACT

Development of a system of detection and classification of simultaneous sound frequencies (chords) with application oriented to the musical field based on the correlation coefficient and a linear function for *threshold*. Using MATLAB 2016a, functions are developed to calculate the correlation between a chord and a previously generated individual notes database containing 36 recorded notes of an electric guitar using an audio interface, sampling at 96 kHz, which is subsequently reduced to 48 kHz for processing purposes. All operations on the signals are done in the frequency domain, using the FFT (Fast Fourier Transform) for the treatment of audio signals. Two types of note selectors are developed to provide output in the form of detected notes. In the first selector, all notes considered detected are reported on the output. In the second selector, of all correlation values, the system chooses the 6 largest values to point as probable notes executed on the input signal. Tests are performed with 10 samples of 20 different chords to characterize the classifiers as to their hit rate, false positives and negatives. Also, the performance of the classifiers was analysed using ROC curve. With the final results, it was concluded that the correlation coefficient can be considered as a good tool for the comparison of the signals in the proposed application. The first selector showed a better hit rate than the second selector, because it was less restrictive, reporting all of the detected notes. The hit rates for the first selector range from 0 to 100 % depending on the chord, while the second hit rates vary from 0 to 66,6667 %. It was shown that for different tones and for different construction of chords, the hit rates vary without a specific pattern. However, the function used as the detection threshold for the classifier showed limited performance for the two developed selectors, where, in most cases, with high values of hit rate, it returns a high rate of false positives and with low values of false positives, it returns a low rate of true positives

Keywords: frequency, classification, chords, correlation, FFT, detection

Lista de Figuras

Figura 1 – Exemplo gráfico da convenção padrão de oitavas de acordo com as teclas de um piano comum, sendo as notas marcadas, respectivamente, C_4 e A_4	20
Figura 2 – Ilustração representativa do procedimento evidenciado pela Equação (5)	23
Figura 3 – Representação gráfica de uma classificação utilizando uma função T linear de <i>threshold</i> para duas características arbitrárias $x[n]$ e $y[n]$	27
Figura 4 – Representação de 5 classificadores como pontos em um espaço ROC	28
Figura 5 – Diagrama de blocos da visão geral do sistema desenvolvido	30
Figura 6 – Interface de áudio utilizada para a geração da base de dados e das amostras de teste	31
Figura 7 – Guitarra elétrica <i>Ibanez RGA7</i> utilizada para a aquisição das amostras	31
Figura 8 – Conjunto de códigos elaborado para os nomes dos arquivos de acordo com a posição no braço da guitarra; Os códigos sem cores, à esquerda, são referentes às cordas soltas	33
Figura 9 – Área de trabalho do <i>software</i> Studio One, indicando uma porção das amostras gravadas e já recortadas para a geração da base de dados	34
Figura 10 – Fluxograma para a rotina de obtenção dos dados de um arquivo de áudio de nota	35
Figura 11 – Diagrama de blocos simplificado indicando o funcionamento do sistema	38
Figura 12 – Fluxograma para a rotina de verificação pela análise de pertinência ao acorde	39
Figura 13 – Diagrama de blocos simplificado indicando o funcionamento do sistema	41
Figura 14 – Fluxograma para a rotina de verificação da seleção pelo maior valor de correlação	42
Figura 15 – Resultado ideal para a aplicação da rotina no acorde C	45
Figura 16 – Espectro em frequência de 30 amostras da nota E_2	47
Figura 17 – Espectro em frequência da média aritmética das 30 amostras da nota E_2	48
Figura 18 – Valores dos coeficientes de correlação por nota para 10 amostras do acorde C	52
Figura 19 – Valores dos coeficientes de correlação para 10 amostras do acorde G	53
Figura 20 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde C	55
Figura 21 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde G	55
Figura 22 – Curva ROC para o acorde C	57
Figura 23 – Curva ROC para o acorde G	57

Figura 24 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde C	58
Figura 25 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde G	59
Figura 26 – Curva ROC para o acorde C	60
Figura 27 – Curva ROC para o acorde G	61
Figura 28 – Representação da função T de <i>threshold</i> aplicada a todos os acordes para detecção no acorde G	64
Figura 29 – Representação de uma função T arbitrária de <i>threshold</i> cuja viabilização provavelmente implicaria em melhores resultados	65
Figura A1 – Valores dos coeficientes de correlação para o acorde $C6$	94
Figura A2 – Valores dos coeficientes de correlação para o acorde $C7$	94
Figura A3 – Valores dos coeficientes de correlação para o acorde $C7M$	94
Figura A4 – Valores dos coeficientes de correlação para o acorde $C7M(\#5)$	95
Figura A5 – Valores dos coeficientes de correlação para o acorde C°	95
Figura A6 – Valores dos coeficientes de correlação para o acorde Cm	95
Figura A7 – Valores dos coeficientes de correlação para o acorde $Cm6$	96
Figura A8 – Valores dos coeficientes de correlação para o acorde $Cm7M$	96
Figura A9 – Valores dos coeficientes de correlação para o acorde D	96
Figura A10–Valores dos coeficientes de correlação para o acorde F	97
Figura A11–Valores dos coeficientes de correlação para o acorde $F7M$	97
Figura A12–Valores dos coeficientes de correlação para o acorde $F\#$	97
Figura A13–Valores dos coeficientes de correlação para o acorde $F\#m7$	98
Figura A14–Valores dos coeficientes de correlação para o acorde $G7$	98
Figura A15–Valores dos coeficientes de correlação para o acorde $G\#7$	98
Figura A16–Valores dos coeficientes de correlação para o acorde $G\#m7(b5)$	99
Figura A17–Valores dos coeficientes de correlação para o acorde A	99
Figura A18–Valores dos coeficientes de correlação para o acorde $A9$	99
Figura A19–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde $C6$	100
Figura A20–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde $C7$	100
Figura A21–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde $C7M$	101
Figura A22–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde $C7M(\#5)$	101
Figura A23–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde Cm	102

Figura A24–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde $Cm6$	102
Figura A25–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde $Cm7M$	103
Figura A26–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde C^o	103
Figura A27–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde D	104
Figura A28–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde F	104
Figura A29–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde $F7M$	105
Figura A30–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde $F\sharp$	105
Figura A31–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde $F\sharp m7$	106
Figura A32–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde $G7$	106
Figura A33–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde $G\sharp7$	107
Figura A34–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde $G\sharp m7(b5)$	107
Figura A35–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde A	108
Figura A36–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde $A9$	108
Figura A37–Pontos das curvas ROC para 10 amostras do acorde $C6$	109
Figura A38–Pontos das curvas ROC para 10 amostras do acorde $C7$	109
Figura A39–Pontos das curvas ROC para 10 amostras do acorde $C7M$	109
Figura A40–Pontos das curvas ROC para 10 amostras do acorde $C7M(\sharp5)$	110
Figura A41–Pontos das curvas ROC para 10 amostras do acorde Cm	110
Figura A42–Pontos das curvas ROC para 10 amostras do acorde $Cm6$	110
Figura A43–Pontos das curvas ROC para 10 amostras do acorde $Cm7M$	111
Figura A44–Pontos das curvas ROC para 10 amostras do acorde C^o	111
Figura A45–Pontos das curvas ROC para 10 amostras do acorde D	111
Figura A46–Pontos das curvas ROC para 10 amostras do acorde F	112
Figura A47–Pontos das curvas ROC para 10 amostras do acorde $F7M$	112
Figura A48–Pontos das curvas ROC para 10 amostras do acorde $F\sharp$	112
Figura A49–Pontos das curvas ROC para 10 amostras do acorde $F\sharp m7$	113

Figura A50–Pontos das curvas <i>ROC</i> para 10 amostras do acorde <i>G7</i>	113
Figura A51–Pontos das curvas <i>ROC</i> para 10 amostras do acorde <i>G#7</i>	113
Figura A52–Pontos das curvas <i>ROC</i> para 10 amostras do acorde <i>G#m7(b5)</i>	114
Figura A53–Pontos das curvas <i>ROC</i> para 10 amostras do acorde <i>A</i>	114
Figura A54–Pontos das curvas <i>ROC</i> para 10 amostras do acorde <i>A9</i>	114
Figura A55–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde <i>C6</i>	115
Figura A56–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde <i>C7</i>	115
Figura A57–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde <i>C7M</i>	116
Figura A58–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde <i>C7M(#5)</i>	116
Figura A59–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde <i>Cm</i>	117
Figura A60–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde <i>Cm6</i>	117
Figura A61–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde <i>Cm7M</i>	118
Figura A62–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde <i>C°</i>	118
Figura A63–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde <i>D</i>	119
Figura A64–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde <i>F</i>	119
Figura A65–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde <i>F7M</i>	120
Figura A66–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde <i>F#</i>	120
Figura A67–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde <i>F#m7</i>	121
Figura A68–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde <i>G7</i>	121
Figura A69–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde <i>G#7</i>	122
Figura A70–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde <i>G#m7(b5)</i>	122
Figura A71–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde <i>A</i>	123

Figura A72–Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde $A9$	123
Figura A73–Pontos das curvas ROC para 10 amostras do acorde $C6$	124
Figura A74–Pontos das curvas ROC para 10 amostras do acorde $C7$	124
Figura A75–Pontos das curvas ROC para 10 amostras do acorde $C7M$	124
Figura A76–Pontos das curvas ROC para 10 amostras do acorde $C7M(\#5)$	125
Figura A77–Pontos das curvas ROC para 10 amostras do acorde Cm	125
Figura A78–Pontos das curvas ROC para 10 amostras do acorde $Cm6$	125
Figura A79–Pontos das curvas ROC para 10 amostras do acorde $Cm7M$	126
Figura A80–Pontos das curvas ROC para 10 amostras do acorde C°	126
Figura A81–Pontos das curvas ROC para 10 amostras do acorde D	126
Figura A82–Pontos das curvas ROC para 10 amostras do acorde F	127
Figura A83–Pontos das curvas ROC para 10 amostras do acorde $F7M$	127
Figura A84–Pontos das curvas ROC para 10 amostras do acorde $F\sharp$	127
Figura A85–Pontos das curvas ROC para 10 amostras do acorde $F\sharp m7$	128
Figura A86–Pontos das curvas ROC para 10 amostras do acorde $G7$	128
Figura A87–Pontos das curvas ROC para 10 amostras do acorde $G\sharp7$	128
Figura A88–Pontos das curvas ROC para 10 amostras do acorde $G\sharp m7(b5)$	129
Figura A89–Pontos das curvas ROC para 10 amostras do acorde A	129
Figura A90–Pontos das curvas ROC para 10 amostras do acorde $A9$	129

Lista de Tabelas

Tabela 1	– Referência de frequências aproximadas para cada nota musical utilizada	20
Tabela 2	– Acordes gravados como amostras para os ensaios aleatórios, especificados pelas respectivas notas por corda tocada	44
Tabela 3	– Desvios padrão para cada nota da base de dados	49
Tabela 4	– Coeficiente de correlação para 8 combinações distintas de sinais	50
Tabela 5	– Coeficiente de correlação entre o sinal de $E_2(6)_{\text{med}}$ e outros sinais médios arbitrários escolhidos	51
Tabela 6	– Coeficiente de correlação entre o sinal do acorde E e os sinais da base de dados das notas individuais desse acorde	51
Tabela 7	– Taxas de acerto máximas para cada acorde, juntamente com o valor de T máximo por onde essa taxa de acerto máxima se manteve, sob o seletor de Análise de Pertinência ao Acorde	56
Tabela 8	– Taxas de acerto máximas para cada acorde, juntamente com o valor de T máximo por onde essa taxa de acerto máxima se manteve, sob o Método de Maior Valor de Correlação	60

Lista de Símbolos

<i>C</i>	Dó (nota)
<i>C♯</i>	Dó Sustenido (nota)
<i>D</i>	Ré (nota)
<i>D♯</i>	Ré Sustenido (nota)
<i>E</i>	Mi (nota)
<i>F</i>	Fá (nota)
<i>F♯</i>	Fá Sustenido (nota)
<i>G</i>	Sol (nota)
<i>G♯</i>	Sol Sustenido (nota)
<i>A</i>	Lá (nota)
<i>B</i>	Si (nota)
<i>C</i>	Dó Maior (acorde)
<i>C6</i>	Dó com Sexta (acorde)
<i>C7</i>	Dó Maior com Sétima (acorde)
<i>C7M</i>	Dó com Sétima Maior (acorde)
<i>Cm</i>	Dó Menor (acorde)
<i>Cm6</i>	Dó Menor com Sexta (acorde)
<i>Cm7M</i>	Dó Menor com Sétima Maior (acorde)
<i>C°</i>	Dó Diminuto (acorde)
<i>D</i>	Ré Maior (acorde)
<i>Dm</i>	Ré Menor (acorde)
<i>E</i>	Mi Maior (acorde)
<i>Em</i>	Mi Menor (acorde)

<i>F</i>	Fá Maior (acorde)
<i>F7M</i>	Fá com Sétima Maior (acorde)
<i>F♯</i>	Fá Sustenido Maior (acorde)
<i>F♯m7</i>	Fá Sustenido Menor com Sétima (acorde)
<i>G</i>	Sol Maior (acorde)
<i>G7</i>	Sol Maior com Sétima (acorde)
<i>G♯7</i>	Sol Sustenido Maior com Sétima (acorde)
<i>G♯m7(b5)</i>	Sol Sustenido Menor com Sétima e Quinta Bemol (acorde)
<i>A</i>	Lá Maior (acorde)
<i>A9</i>	Lá Maior com Nona (acorde)
<i>A♯</i>	Lá Sustenido Maior (acorde)
<i>A♯m</i>	Lá Sustenido Menor (acorde)
<i>B°</i>	Si Diminuto (acorde)

Sumário

1	INTRODUÇÃO	15
2	REVISÃO BIBLIOGRÁFICA	17
3	FUNDAMENTAÇÃO TEÓRICA	19
3.1	Conceitos Musicais	19
3.1.1	As Notas Musicais	19
3.1.2	Harmônicas de uma Nota Musical	21
3.1.3	O Acorde	21
3.2	Conceitos de Processamento Digital de Sinais	22
3.2.1	Transformada de Fourier Discreta (DFT) e o Algoritmo de Transformada Rápida (FFT)	22
3.2.2	O Decimador, ou <i>Downsampling</i>	23
3.3	Conceitos Probabilísticos e Estatísticos	24
3.3.1	A Covariância e o Coeficiente de Correlação	24
3.3.2	Média e Desvio-Padrão Amostrais	25
3.4	Classificadores	26
3.4.1	Casamento por Correlação e a Função de Threshold	26
3.4.2	Característica de Operação do Receptor	27
4	METODOLOGIA	30
4.1	Visão Geral do Sistema	30
4.2	Geração da Base de Dados	30
4.2.1	Materiais para Aquisição de Amostras	31
4.2.1.1	Hardware e Instrumento	31
4.2.1.2	Software	32
4.2.2	Organização da Base de Dados	32
4.2.3	Processo de Aquisição de Amostras	33
4.2.4	Tratamento dos Dados	33
4.3	Processamento dos Dados	36
4.3.1	Implementação do Coeficiente de Correlação	36
4.4	Definição dos Seletores	36
4.4.1	Variação do <i>threshold T</i>	37
4.4.2	Análise de Pertinência ao Acorde	37
4.4.3	Análise pelo Maior Valor de Correlação	40
4.5	Verificação e Validação dos Dados Obtidos	40

4.5.1	Teste Preliminar da Correlação com a Base de Dados	40
4.5.2	Ensaios Aleatórios	41
4.5.2.1	Acordes Utilizados	43
4.5.3	Geração dos Gráficos Comparativos	44
4.5.3.1	Correlação por Nota	44
4.5.3.2	Análise das Taxas de Acerto	45
4.5.3.3	Geração das Curvas ROC	46
5	RESULTADOS E DISCUSSÕES	47
5.1	Base de Dados	47
5.2	Testes Preliminares do Processamento	50
5.3	Testes com Amostras de Acordes	51
5.3.1	Resultados da Análise da Correlação por Nota	52
5.3.2	Resultados da Análise das Taxas de Acerto	54
5.3.2.1	Método de Análise de Pertinência ao Acorde	54
5.3.2.2	Método de Maior Valor de Correlação	58
6	CONCLUSÕES	62
7	PROPOSTAS PARA TRABALHOS FUTUROS	64
	REFERÊNCIAS BIBLIOGRÁFICAS	66
	APÊNDICES	68
	APÊNDICE A – CÓDIGOS DESENVOLVIDOS NO MATLAB . . .	69
	APÊNDICE B – RESULTADOS DE CORRELAÇÃO POR NOTA .	94
	APÊNDICE C – RESULTADOS PELA ANÁLISE DE PERTINÊN-	
	CIA AO ACORDE	100
	APÊNDICE D – RESULTADOS PELO MAIOR VALOR DE COR-	
	RELAÇÃO	115

1 Introdução

Com o avanço dos estudos realizados no ramo de processamento digital de sinais e de sistemas de áudio, a demanda por aplicações mais específicas cresce, junto com a necessidade de aprimoramento dos sistemas já desenvolvidos e estudados. Ao mencionar trabalhos relacionados à área de áudio, é inerente a existência de estudos que envolvam análises no domínio da frequência. Por se tratar de uma área extensa e com diversos níveis e diversas técnicas de aplicação, a necessidade de propor diferentes métodos para desenvolver soluções pode oferecer uma nova visão do assunto, sendo possível abranger mais o leque de possibilidades de ferramentas para aplicações específicas. A motivação para a realização desse estudo é a de poder possibilitar que músicos, engenheiros de som, produtores musicais, entre outros profissionais da área, sejam capazes de identificar acordes no que diz respeito à sua composição por notas individuais. A utilidade dessa ferramenta pode ser ampla, podendo ser utilizada por amadores e profissionais do ramo para diversas aplicações.

Nesse contexto, é realizado o estudo de um sistema detector e classificador de conjuntos de frequências sonoras - chamados comumente de acordes, conjuntos de notas musicais - provenientes de um instrumento musical, utilizando um método baseado na função coeficiente de correlação e de uma função de limiar de detecção. Essa aplicação, em especial, costuma ser mais complexa de ser implementada quando comparamos ao problema de identificação individual de notas musicais, pois, em acordes, as harmônicas das notas executadas interferem na análise de amplitudes resultantes no domínio da frequência. Para encontrar apenas frequências fundamentais das notas na análise de um acorde, então, a tarefa torna-se mais desafiadora.

O estudo procede a partir do desenvolvimento de um algoritmo de comparação entre um sinal de entrada composto (um acorde), proveniente de um único instrumento (guitarra elétrica), e uma base de dados de notas individuais gerada previamente com o mesmo instrumento. A geração da base de dados é parte relevante do estudo no que diz respeito à obtenção das características principais do sinal específico proposto para a aplicação. Tais dados são obtidos a partir da realização das FFTs (*Fast Fourier Transform*, ou Transformada Rápida de Fourier) de 30 amostras de cada nota, numa faixa de 36 notas, onde agrega-se um padrão para cada nota possível de ser tocada dentro dessa faixa por meio da média amostral e do desvio padrão amostral.

A implementação do sistema é dada pelo processamento de um acorde de entrada, cuja transformação para o domínio frequência é realizada. A partir dele, a comparação é realizada com toda a base de dados, nota a nota, utilizando um algoritmo de cálculo

do coeficiente de correlação entre o sinal de entrada e os sinais da base de dados. Dessa forma, o sistema retorna valores numéricos de correlação normalizada associados a cada nota da base de dados.

O desenvolvimento dos seletores, por consequência, define os classificadores utilizados a partir da função de *threshold* escolhida. O primeiro seletor se baseia na análise de pertinência das notas ao acorde especificado, cujo resultado informa ao usuário as notas que foram detectadas, sem levar em consideração a ordem ou a quantidade destas. O segundo seletor se baseia na seleção dos seis maiores valores de correlação encontrados para cada corda, a fim de retornar quais foram as seis prováveis notas executadas (a escolha de serem seis valores justifica-se ao ser considerada uma guitarra de seis cordas, que é o caso mais comum).

Os testes são realizados em ambiente restrito a um instrumento e um sistema de aquisição de dados único. Porém, conhecendo os procedimentos desse estudo, será possível utilizar o mesmo método para futuros estudos mais abrangentes envolvendo a mesma aplicação.

Espera-se que os classificadores tenham êxito em detectar as notas mais relevantes, embora seja provável a causa de interferência devido às harmônicas das notas executadas. Para analisar a qualidade dos classificadores, são analisados os dados referentes a suas taxas de acerto, falsos positivos e falsos negativos. Por fim, também é realizada uma segunda análise para verificar a performance dos classificadores utilizando as curvas de Característica de Operação do Receptor (da sigla em inglês, ROC - *Receiver Operating Characteristic*).

2 Revisão Bibliográfica

Atualmente, é possível encontrar algumas soluções no mercado de engenharia de áudio para realizar a tarefa proposta, onde observam-se diferentes métodos para desenvolver essa solução. Além dos produtos vendidos principalmente na área de produção fonográfica, há diversos estudos na área que contemplam o objetivo desse trabalho em outras dimensões. Alguns desses envolvem *template matching*, que é a abordagem tratada no presente trabalho, mas grande parte costuma abordar o problema de maneiras adicionais.

Em trabalhos relativamente recentes, é possível encontrar algumas aplicações similares ao objetivo desse trabalho. Por exemplo, o de transcrição de tablaturas por análise de estimação multitonal (YAZAWA et al, 2013), onde observa-se um método de pré-processamento via *latent harmonic allocation* (alocação harmônica latente), ou LHA, que é um método de aprendizagem de máquina aplicado para estimar múltiplos tons em um espectograma, para depois aplicarem-se estimadores de multitons em arquivos de áudio de diversos estilos musicais a partir de janelas de tempo. Nesse contexto, é considerado um nível de *threshold* para avaliação dos resultados da estimação, com o objetivo de transcrever tablaturas, onde os resultados são apresentados como estimadores-F com resultados de até 0,721 na estimação de frequências fundamentais desses arquivos de áudio. Na mesma linha, pode-se citar o trabalho de transcrição automática de tablaturas num sistema web, chamado *Robotaba* (BURLET et al, 2013), utilizando uma base de dados gerada a partir de um programa sintetizador de sons artificiais de guitarra e analisando os arquivos de áudio resultantes da sintetização a partir de algoritmos de análise polifônica.

Em uma outra abordagem, também foi verificado o trabalho de transcrição automática de acordes acompanhada da digitação correta a partir de arquivos de áudio (BARBANCHO et al, 2012), onde é realizada a extração de características de gravações de guitarra e a posterior aplicação de diferentes modelos de cadeias ocultas de Markov pré treinados para estimar múltiplas frequências fundamentais.

Também há outras alternativas como o uso de cromagramas (LEE, 2006), cuja análise de áudio é feita na frequência e no tempo por meio de *pitch class profiles*, ou perfis de classe de tons, para determinar as notas, e de modelos ocultos de Markov (LEE; SLANEY, 2006), (SHEH; ELLIS, 2003), como visto anteriormente em outro trabalho citado, na análise de arquivos de áudio em tempo real.

Entretanto, apesar de já existirem outras formas de estudar o objetivo do projeto, a identificação de acordes com alta precisão (acima de 70%) ainda se mostra um problema complexo de se resolver e implementar. Propõe-se, portanto, nesse estudo, a validação de um método não muito convencional de resolver essa questão, que é o de realizar o

casamento dos sinais de interesse por meio da correlação normalizada e utilizar uma função linear para detectar os valores mais relevantes. A validação desse método pode servir para complementar algum dos métodos mais convencionais utilizados nessa aplicação, ou para ser utilizado de forma individual, dependendo dos resultados.

3 Fundamentação Teórica

3.1 Conceitos Musicais

3.1.1 As Notas Musicais

Mencionando os fundamentos da área da Música, conforme Nattiez (1990), uma nota musical é definida como o tom e a duração de uma onda mecânica que se propaga no ar. Para exprimir melhor esse conceito em contexto físico, *tom* pode ser interpretado como a frequência na qual essa onda se propaga, gerando, a partir das vibrações mecânicas, seu som característico. Todavia, comumente, a denominação de *tom* é associada a frequências específicas, sendo estas as convencionais notas musicais utilizadas dependendo de alguns fatores, como a fase histórica da música, tradições musicais regionais e tipos de instrumento utilizados.

É importante notar que as notas apresentadas nesse documento são pertinentes a um dos padrões utilizados pela música ocidental. Como citado por Cho (2003), uma das relações matemáticas tradicionalmente definidas entre as frequências subsequentes para essa determinada região do mundo - no caso da música ocidental, é denominada de Escala Igualmente Temperada. A relação entre as notas na escala igualmente temperada pode ser expressa utilizando como exemplo as notas de um piano. Escolhendo-se uma nota de referência P_a , sendo a o número referente à posição da tecla no piano, é possível determinar a frequência P_n relativa à nota posicionada na n -ésima tecla por meio da Equação (1).

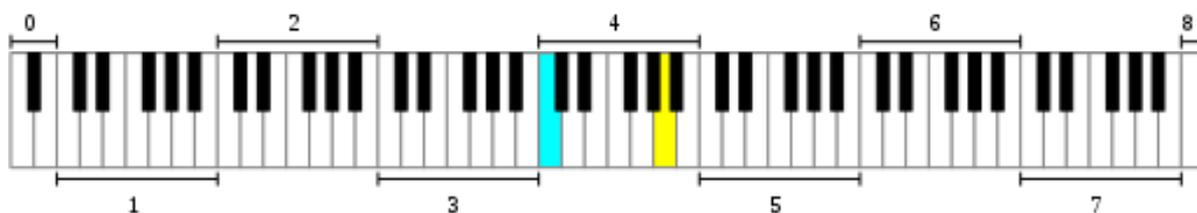
$$P_n = P_a (\sqrt[12]{2})^{(n-a)} \quad (1)$$

O tipo de notação mais utilizado para identificar as notas é o *Scientific Pitch Notation* (em português, Notação Científica de Tons), que é baseado nas oitavas de um piano. Cada oitava é composta por 12 notas, que avançam sequencialmente em intervalos denominados de *semitons*. O intervalo de semitom, por definição, estabelece a razão entre duas frequências P_n e P_{n+1} da forma apresentada pela Equação (2).

$$\frac{P_{n+1}}{P_n} = \sqrt[12]{2} \quad (2)$$

Sendo assim, a definição da notação é baseada na numeração de oitavas conforme a Figura 1, sendo construída a partir da nota tocada acompanhada do número correspondente à oitava. Por exemplo, a nota C (dó) tocada na 4^a oitava é denominada de C_4 a partir desse padrão.

Figura 1 – Exemplo gráfico da convenção padrão de oitavas de acordo com as teclas de um piano comum, sendo as notas marcadas, respectivamente, C_4 e A_4



Ao ser requisitada a frequência da nota de posição $n = 40$, por exemplo, utilizando a tecla $a = 49$ como nota de referência (que, em pianos comuns, corresponde à nota A_4 cuja frequência é de 440 Hz), obtém-se a frequência de $P_n \approx 261,626 \text{ Hz}$, correspondente à nota C_4 , que é a tecla central do piano.

De posse das informações supracitadas, é de interesse, para essa aplicação, obter conhecimento sobre todas as frequências referentes às notas a serem utilizadas. Para esse estudo, a faixa de notas utilizadas compreende da nota E_2 à nota A_4 . As informações de cada uma das notas da faixa especificada com suas frequências correspondentes podem ser verificadas na Tabela 1, de acordo com as informações disponíveis nos estudos de Iazzetta (2008).

A importância da definição desses valores se dá pela necessidade de comparação dos valores gerados na base de dados com os valores padrão para as frequências tonais. Tal verificação auxiliará a determinar se a base de dados foi gerada de forma adequada, bem como a determinar se o sinal de entrada está em condições adequadas para o processo de comparação (em outras palavras, se o instrumento foi afinado corretamente ao gravar as amostras).

Tabela 1 – Referência de frequências aproximadas para cada nota musical utilizada

Nota	E_2	F_2	$F\sharp_2$	G_2	$G\sharp_2$	A_2	$A\sharp_2$	B_2
Frequência (Hz)	82,41	87,31	92,50	98,00	103,82	110,00	116,541	123,471

Nota	C_3	$C\sharp_3$	D_3	$D\sharp_3$	E_3	F_3	$F\sharp_3$	G_3
Frequência (Hz)	130,81	138,59	146,83	155,56	164,81	174,614	185,00	196,00

Nota	$G\sharp_3$	A_3	$A\sharp_3$	B_3	C_4	$C\sharp_4$	D_4	$D\sharp_4$
Frequência (Hz)	207,65	220,00	233,08	246,94	261,62	277,18	293,66	311,12

Nota	E_4	F_4	$F\sharp_4$	G_4	$G\sharp_4$	A_4
Frequência (Hz)	329,63	349,23	369,99	391,99	415,30	440,00

Sendo a nota musical uma onda mecânica, é possível mensurar suas características físicas usando dispositivos transdutores, que realizam a conversão da energia mecânica em energia elétrica. Por exemplo: microfones, captadores elétricos, entre outros.

3.1.2 Harmônicas de uma Nota Musical

Segundo Bingham (2009), a frequência natural mais baixa de um sistema vibrante é denominada de *fundamental*, e as maiores frequências naturais são denominadas de *sobretons*. A partir dessa definição, também pode-se definir que, num sistema que contém uma corda rigidamente presa em suas duas extremidades, os sobretons são chamados de *harmônicas*, termo também utilizado na análise de componentes no espectro da frequência para diversos outros sistemas que envolvem esse tipo de abordagem.

Em geral, pode-se obter os valores das harmônicas de uma frequência fundamental f_0 a partir do produto da fundamental por um valor inteiro n , sendo assim possível obter a n -ésima harmônica relativa àquele sinal, como evidenciado na Equação (3).

$$f_n = n \cdot f_0 \quad (3)$$

Quando uma nota musical é tocada em qualquer instrumento, seja de sopro, de cordas ou piano, ao analisar-se seu espectro em frequência, nota-se a presença das componentes harmônicas do sinal na composição da nota tocada. A relação entre a amplitude da frequência fundamental e das harmônicas presentes no sinal compõe o que é chamado de *timbre*, característica singular influenciada pelos diversos aspectos construtivos e mecânicos de cada instrumento. Espera-se, então, que ao analisar notas individuais de um dado instrumento em seu espectro de frequência, estas sejam compostas de diversos picos de amplitude, sendo estes relacionados à frequência fundamental e as harmônicas.

3.1.3 O Acorde

Conforme Benward & Saker (2003), o conceito musical de *acorde* é definido pela execução simultânea de três ou mais notas em qualquer instrumento musical. A execução simultânea de duas notas é denominada de *intervalo*, pois, como o nome já sugere, expressa em som duas frequências com um intervalo padronizado de unidades entre elas. Assim, pode-se dizer que um acorde é formado de dois ou mais intervalos. Por consequência dessa denominação, a notação utilizada para denominar cada tipo de acorde é baseada nos intervalos que compõem o acorde, utilizando uma das notas tocadas como a referência para a construção deste - geralmente chamada de tônica.

As definições e notações referidas são relevantes no que diz respeito à identificação do tipo de acorde que será analisado pelo processo, pois existe a hipótese de que o programa funcione de forma mais adequada para determinados tipos de acorde em relação

a outros. Essa questão é levantada pois, dependendo do conjunto de intervalos executados simultaneamente, é possível obter diferentes influências de uma nota em outra na análise do espectro das frequências.

3.2 Conceitos de Processamento Digital de Sinais

3.2.1 Transformada de Fourier Discreta (DFT) e o Algoritmo de Transformada Rápida (FFT)

Em aplicações que envolvem áudio, é comum certas análises serem realizadas no domínio frequência, devido à facilidade de observação do comportamento e do formato dos sinais no espectro de frequências. No caso, como o assunto abordado requer a identificação da execução de frequências sonoras, tal tipo de análise torna-se mandatório.

Para essa análise poder ser realizada, é necessário realizar uma transformada de Fourier sobre o sinal de interesse, que é fornecido no formato Amplitude x Tempo, no caso. Um dos tipos de transformada de Fourier amplamente utilizados em aplicações computacionais é a Transformada de Fourier Discreta (ou DFT, na sigla em inglês). Por definição, a transformada de Fourier decompõe uma função temporal e expressa esta a partir das componentes em frequência que a compõem. Segundo Oppenheim (1999), a DFT, por sua vez, é a aplicação dessa função no tempo discreto. Sendo assim, seja um sinal discreto $x[n]$, para $n = 0, 1, \dots, N - 1$, o sinal resultante da transformada $X(k)$ deste sinal $x[n]$ é baseado na definição apresentada pela Equação (4), sendo $W_N = e^{-j(2\pi/N)}$.

$$X(k) = \sum_{n=0}^{N-1} x[n]W_N^{kn} \quad (4)$$

Em adição à DFT, também foi utilizado um algoritmo computacional capaz de realizar esse cálculo de forma eficiente em relação à aplicação da expressão fundamental para a transformada de Fourier, denominado de Fast Fourier Transform (FFT). Neste trabalho, foi utilizado o algoritmo Radix-2 para a implementação da FFT. Tal algoritmo é a forma mais comum do algoritmo de *Cooley-Tukey* para o cálculo da FFT, e este consiste na divisão da DFT de tamanho N em duas DFTs de tamanho $N/2$ para cada estágio recursivo. Essa forma simplificada assume que N é uma potência de 2, ou seja, $N = 2^v$, sendo ($v = 0, 1, 2, \dots$) o termo de potenciação. Uma vez que o número de amostras N pode normalmente ser escolhido livremente para a aplicação, o fato de N ser uma potência de 2 não se mostra uma restrição relevante.

A partir das informações fornecidas pelo resultado do algoritmo, é possível obter uma resposta no domínio frequência para a posterior análise. A utilização desse algoritmo será dada no processamento dos sinais tanto para a geração da base de dados quanto para

a aquisição das amostras para testes, de forma a tratar todos os dados diretamente no domínio frequência.

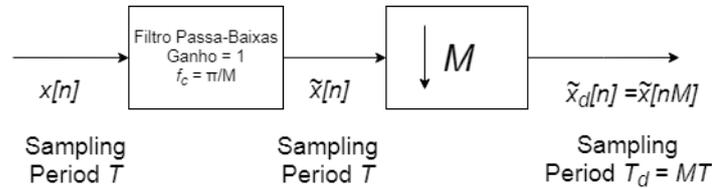
3.2.2 O Decimador, ou *Downsampling*

Frequentemente, é necessário alterar a taxa de amostragem de um sinal discreto para obter uma nova representação de um sinal contínuo, seja para melhorar a qualidade da análise do sinal, para compatibilizar mais de um sinal para uma mesma análise, ou para outros fins. A taxa de amostragem de um sinal pode ser reduzida utilizando funções provenientes dos conceitos de análise de sinais no tempo discreto. Definindo uma nova variável $x_d[n]$ a partir do sinal discreto $x[n]$ previamente filtrado por um filtro passa-baixas, segundo Oppenheim (1999), podemos expressar o resultado desejado da forma apresentada na Equação (5), sendo M o fator de redução da taxa de amostragem e x_c o sinal contínuo.

$$x_d[n] = x[nM] = x_c(nMT) \quad (5)$$

Ao observar-se essa expressão, é possível notar que o sinal $x_d[n]$ é idêntico ao sinal que seria obtido ao amostrar o sinal $x_c(t)$ com período $T_d = MT$. Denotando o sinal filtrado por $\tilde{x}[n]$, pode-se visualizar o modelo de sistema generalizado para essa aplicação na Figura 2.

Figura 2 – Ilustração representativa do procedimento evidenciado pela Equação (5)



A taxa de amostragem pode ser reduzida por um fator M sem causar *aliasing* se a taxa de amostragem original for no mínimo M vezes a taxa de Nyquist ou se a largura de banda da sequência é primeiramente reduzida por um fator M por filtragem em tempo discreto. No geral, qualquer operação de redução de taxa de amostragem é denominada de *downsampling*. Uma das maneiras de implementar o decimador em uma aplicação computacional é utilizando um filtro FIR, capaz de computar cada M -ésima saída do sinal de forma direta. Por exemplo, a aplicação do filtro FIR no caso especificado pode ser expressa pela Equação (6), onde K é o tamanho do sinal e $h[n]$ é a resposta impulsiva.

$$x_d[n] = \sum_{k=0}^{K-1} x[nM - k] \cdot h[k] \quad (6)$$

É relevante o conhecimento do método de *downsampling* para essa aplicação, pois, ao longo do procedimento, é necessário reduzir a taxa de amostragem da base de dados gerada para realizar cálculos envolvendo as amostras de teste.

3.3 Conceitos Probabilísticos e Estatísticos

3.3.1 A Covariância e o Coeficiente de Correlação

Segundo Bendat & Piersol (2010), ao considerar-se uma variável aleatória arbitrária contínua $x(t)$ que assume valores entre $(-\infty, +\infty)$, define-se o conceito de valor esperado $E[x(t)]$ como sendo equivalente ao valor médio desta variável, notado como μ_x , como indicado pela Equação (8), sendo $p(x)$ a função de densidade de probabilidade associada à variável $x(t)$.

$$E[x(t)] = \int_{-\infty}^{+\infty} xp(x)dx = \mu_x \quad (7)$$

Essa definição é estabelecida pelo mesmo princípio pelo qual é deduzida a expressão para média aritmética de $x(k)$, podendo ser aplicada tanto em um domínio de valores discretos como contínuos. Por consequência, o valor esperado de qualquer função contínua de variável única $g(x)$ da variável $x(k)$ será dado conforme a Equação (8).

$$E[g(x(t))] = \int_{-\infty}^{+\infty} g(x)p(x)dx = \mu_x \quad (8)$$

De posse dessa informação, é possível determinar a variância de $x(t)$, que é definida como o quadrado médio de $x(t)$ sobre seu valor médio. Tal definição é evidenciada pela Equação (9).

$$E[(x(t) - \mu_x)^2] = \int_{-\infty}^{+\infty} (x(t) - \mu_x)^2 p(x)dx = \sigma_x^2 \quad (9)$$

Por definição, portanto, o desvio-padrão σ_x associado a $x(t)$ é a raiz quadrada da variância σ_x^2 . A abordagem desses conceitos permite a dedução do valor esperado de uma função de duas variáveis. Assim, sendo $g(x, y)$ uma função de duas variáveis aleatórias contínuas arbitrárias $x(t)$ e $y(t)$, pode-se expressar seu valor esperado $E[g(x, y)]$ a partir da Equação (10), onde $p(x, y)$ é a função densidade de probabilidade associada às duas variáveis.

$$E[g(x, y)] = \iint_{-\infty}^{+\infty} g(x, y)p(x, y)dxdy \quad (10)$$

Por exemplo, se $g(x, y) = (x(t) - \mu_x)(y(t) - \mu_y)$, onde μ_x e μ_y são os valores médios de $x(t)$ e $y(t)$, respectivamente, isso define a **covariância** C_{xy} entre $x(t)$ e $y(t)$. Esta é evidenciada conforme as Equações (11) e (12).

$$C_{xy} = E[(x(t) - \mu_x)(y(t) - \mu_y)] = E[x(t)y(t)] - E[x(t)]E[y(t)] \quad (11)$$

$$C_{xy} = \iint_{-\infty}^{+\infty} (x(t) - \mu_x)(y(t) - \mu_y)p(x, y)dxdy \quad (12)$$

É importante notar que $C_{xx}(0)$ equivale a σ_x^2 , que é a variância de $x(t)$. A partir disso, é possível inferir uma relação entre a covariância C_{xy} de $x(t)$ e $y(t)$ e os desvios-padrão σ_x e σ_y conforme apresentado na Equação (13).

$$|C_{xy}| \leq \sigma_x \sigma_y \quad (13)$$

Pode-se, então, obter a quantidade normalizada denominada de **coeficiente de correlação** ρ_{xy} , definido pela Equação (14).

$$\rho_{xy} = \frac{C_{xy}}{\sigma_x \sigma_y} \quad (14)$$

Tal quantidade assume valores entre -1 e 1 , que exprimem o fato de os sinais serem *correlacionados* entre si ou não. Uma correlação de valor $\rho_{xy} = 0$ indica que os dois sinais são *não-correlacionados*. Essa será a principal função a ser utilizada como base nas comparações realizadas pelo sistema proposto.

3.3.2 Média e Desvio- Padrão Amostrais

Analisando novamente uma variável x como a determinada na Seção 3.3.1, sendo o índice t omitido por simplicidade, é possível determinar os dois parâmetros principais que especificam sua tendência central e sua dispersão, sendo estes a média μ_x e a variância σ_x^2 , conforme calculado anteriormente pelas Equações (7) e (9). Esses dois parâmetros, entretanto, não podem ser calculados com exatidão na prática, pois requerem o conhecimento prévio exato sobre a função de densidade de probabilidade $p(x, y)$, o que geralmente não é fornecido ou determinado. Sendo assim, deve-se estabelecer uma estimativa da média e da variância a partir de um número finito de valores observados.

Um dos métodos possíveis de estimar a média e a variância de x baseando-se em N observações independentes é determinado conforme as Equações (15) e (16).

$$\bar{x} = \hat{\mu}_x = \frac{1}{N} \sum_{i=1}^N x_i \quad (15)$$

$$s_b^2 = \hat{\sigma}_x^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (16)$$

sendo \bar{x} e s_b^2 definidas aqui como a média e a variância amostrais. Aqui, s_b^2 representa um estimador tendencioso da variância, o que é visto como um fator que influencia na eficiência deste. Para estimadores não tendenciosos, a relação entre o valor esperado para o parâmetro estabelecido deve ser de equivalência, o que é representado pela Equação (17).

$$E[\hat{\phi}] = \phi \quad (17)$$

Para tal condição ser atendida em um estimador para a variância, pode ser comprovado que o estimador não tendencioso s^2 atende a essa especificação, cuja expressão é descrita pela Equação (18).

$$s^2 = \hat{\sigma}_x^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (18)$$

Conseqüentemente, o desvio-padrão não tendencioso s é determinado por $s = \sqrt{s^2}$.

A determinação desses parâmetros torna possível, então, a utilização da função coeficiente de correlação ρ_{xy} baseada em um conjunto finito de amostras, realizando o cálculo de forma análoga ao caso contínuo, porém substituindo-se μ_x por \bar{x} , σ_x^2 por s^2 e σ_x por s .

3.4 Classificadores

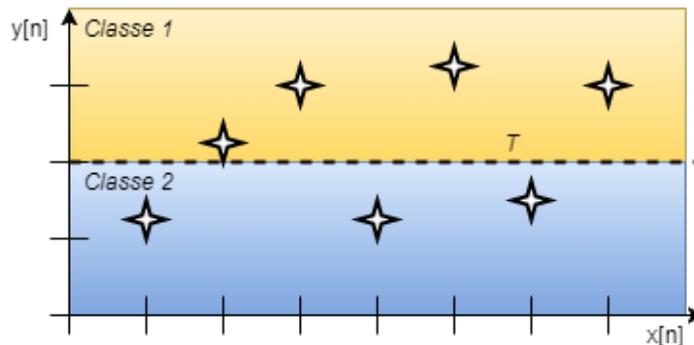
3.4.1 Casamento por Correlação e a Função de Threshold

Como fundamentado anteriormente, a correlação será ferramenta primordial para esse estudo, e será utilizada para definir a similaridade entre dois sinais. Conforme é descrito na obra de Gonzales (2010), é possível utilizar o coeficiente normalizado de correlação para o casamento de sinais distintos de interesse. A correlação máxima, baseando-se na Equação (14), é o indicador de sinais idênticos, onde $\rho_n = 1$, enquanto para o caso de $\rho_n = -1$, temos sinais inversamente correlacionados.

De posse de todos os pontos resultantes da aplicação da correlação, pode-se definir uma função T denominada de limiar, ou *threshold*, que será a função classificadora. Em

outras palavras, a partir da aplicação dessa função, é possível separar um dado sinal em classes diferentes. Para esse estudo, será utilizada uma função T linear de classificação binária. Uma representação da classificação utilizando uma função T linear é apresentada na Figura 3.

Figura 3 – Representação gráfica de uma classificação utilizando uma função T linear de *threshold* para duas características arbitrárias $x[n]$ e $y[n]$



3.4.2 Característica de Operação do Receptor

Para analisar um sistema de classificação binária, é importante obter informações sobre a qualidade de seus resultados. Para isso, a análise por Característica de Operação do Receptor (em inglês, *Receiver Operation Characteristic*, abreviado comumente para ROC) fornece indicativos de relevância baseados na taxa de acerto e de erro do sistema.

Conforme descrito por Hopley (2011), esse tipo de análise se baseia em quatro tipos de resultados que o sistema pode fornecer: verdadeiro positivo (V_P), verdadeiro negativo (V_N), falso positivo (F_P) e falso negativo (F_N). V_P e V_N são os sinais esperados por um sistema ideal, onde um classificador consegue distinguir perfeitamente quando uma amostra faz parte de uma dada classe ou não. Por exemplo, como introduz Hopley (2011), em um sistema que testa um paciente para detectar alguma doença, V_P indica que o paciente possui a doença quando ele realmente possui, enquanto V_N indicaria que o paciente não possui a doença quando ele realmente não possui. Já F_P indica que o paciente teria a doença na situação em que ele não tem, enquanto F_N indicaria que ele não possui a doença numa situação onde ele possui.

Esses dados obtidos da resposta do sistema são utilizados para gerar a curva de Característica de Operação do Receptor, que utiliza os dados da taxa de positivos verdadeiros (TVP) e da taxa de negativos verdadeiros (TVN), denominando-os respectivamente de "Sensibilidade" e "Especificidade". Essas quantidades são definidas como mostram as

Equações (19) e (20). A Taxa de Falsos Positivos é utilizada para a geração da curva e sua expressão é mostrada pela Equação (21).

$$\text{Sensibilidade} = TPV = \frac{V_P}{V_P + F_N} \quad (19)$$

$$\text{Especificidade} = TNV = \frac{V_N}{V_N + F_P} \quad (20)$$

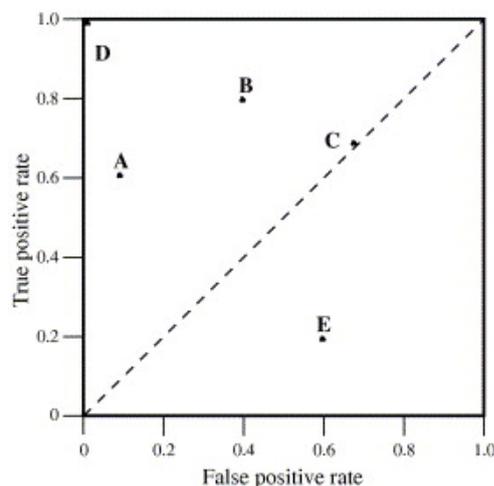
$$TFP = 1 - TNV = \frac{F_P}{V_N + F_P} \quad (21)$$

Gerando o gráfico TPV (Sensibilidade) x TFP , é possível obter indicadores da performance do sistema de detecção, bem como estimar erros associados à detecção. Como os dados utilizados para a geração do gráfico dependem da função de *threshold* escolhida, o *plot* será feito em função da variação do valor de *threshold*, indicando a performance do detector para cada nível.

Segundo Fawcett (2006), um ponto em uma curva ROC é melhor avaliado do que outro se estiver mais para o noroeste do gráfico (TPV maior, TFP é menor, ou ambos). Por exemplo, classificadores que são indicados à esquerda no gráfico ROC e na parte inferior, perto do eixo X, fazem classificações positivas apenas com muita evidência, assim cometem poucos erros de falso positivo, porém, isso implica em baixa taxa de verdadeiros positivos também. Classificadores ao lado superior direito de um gráfico ROC fazem classificações positivas com pouca evidência, classificando corretamente quase todos os casos, mas também com altas taxas de falso positivo.

Na Figura 4 são representados 5 classificadores pontuais em um espaço ROC, denominados de A, B, C, D e E.

Figura 4 – Representação de 5 classificadores como pontos em um espaço ROC



Fonte: Fawcett, 2006

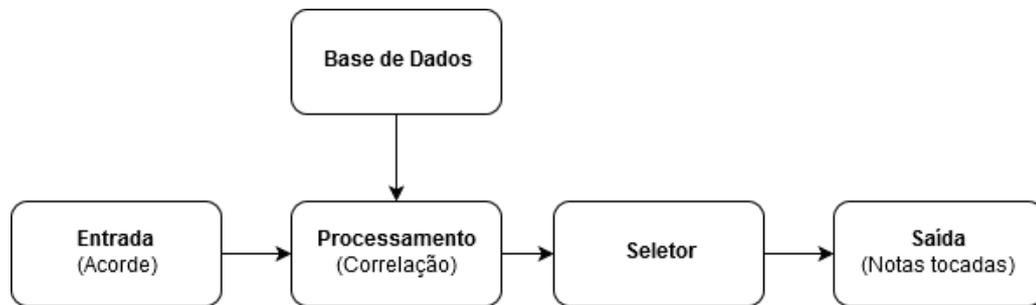
Nesse caso, o classificador D representaria o ponto ótimo de classificação, onde a TPV é unitária e a TFP é igual a zero. O classificador E, por exemplo, seria um classificador pobre, pois teria alta TFP em comparação com sua TPV .

4 Metodologia

4.1 Visão Geral do Sistema

O sistema elaborado deve ser capaz de processar dados de comparação entre um sinal de entrada, composto por um acorde com diferentes notas somadas a suas harmônicas, e todos os sinais do banco de dados, escolhendo os melhores resultados para informar, na saída, qual foi o acorde reproduzido no sinal de entrada. Para a elaboração do sistema, compreendem-se os passos ilustrados pelo diagrama de blocos da Figura 5.

Figura 5 – Diagrama de blocos da visão geral do sistema desenvolvido



De acordo com o diagrama de blocos, o funcionamento do sistema consiste em receber um acorde em um formato compatível com o tipo de processamento estabelecido - baseado no coeficiente de correlação - onde será comparado nota a nota com a base de dados gerada previamente. Assim, os resultados da comparação serão enviados para um seletor, que fará a distinção entre as notas detectadas e as notas não detectadas. Por fim, retorna-se a saída, onde espera-se que os valores informados sejam compatíveis com a real execução do acorde.

Os passos para a elaboração e implementação do sistema são detalhados pelas Seções de 4.2 a 4.5.

4.2 Geração da Base de Dados

Primeiramente, é necessário gerar a base de dados. A base de dados é o conjunto de todos os dados associados a cada nota - sendo estes a média amostral dos 30 espectros de FFT das amostras obtidas e o valor do desvio padrão correspondente da FFT média - dentro da faixa de notas incluídas no sistema proposto.

4.2.1 Materiais para Aquisição de Amostras

A escolha dos materiais utilizados foi feita de acordo com a disponibilidade destes, não sendo necessariamente os dispositivos ou *softwares* de melhor performance para o estudo. Porém, são suficientes para a obtenção adequada dos resultados desejados.

4.2.1.1 Hardware e Instrumento

A obtenção das amostras foi realizada utilizando uma guitarra da marca *Ibanez*, modelo *RGA7*, conectada à interface de áudio *Fast Track Pro*, da marca M-Audio®, representada na Figura 6. Apesar de ser uma guitarra de 7 cordas, a sétima corda não foi utilizada para nenhuma das aquisições, mantendo a proposta de utilizar apenas 6 cordas, como uma guitarra comum.

Figura 6 – Interface de áudio utilizada para a geração da base de dados e das amostras de teste



Figura 7 – Guitarra elétrica *Ibanez RGA7* utilizada para a aquisição das amostras



A guitarra possui três modos de combinação de captadores elétricos, intercambiáveis por meio de uma chave *switch*, e duas configurações de equalização, também intercambiáveis por meio de um *switch*. Essas possibilidades de som permitem que a base de dados possa ser gravada de forma mais abrangente, pois uma variação desses modos e configurações implica em uma variação do timbre do som enviado para a placa de áudio. Assim, o sinal médio resultante terá a influência de diversos timbres distintos, tornando a base de dados mais diversificada no que diz respeito ao tipo de sinal, o que agregará à qualidade

da solução. Da mesma forma, uma maneira de influenciar no timbre é tocar as notas utilizando o dedo ou uma palheta, o que também é realizado no momento da gravação de amostras para a base de dados.

A interface de áudio possui dois canais de entrada, sendo estes configuráveis em diferentes taxas de amostragem. Para os sinais gravados para a base de dados, foi utilizada a configuração de taxa de amostragem em 96 kHz (a serem posteriormente processados para 48 kHz), enquanto os sinais dos acordes de teste foram gravados utilizando a taxa de amostragem de 48 kHz .

4.2.1.2 Software

No que diz respeito à geração da base de dados, foi utilizado o *software* Studio One 3[®] para a gravação de áudio. O *software* possibilita a gravação de múltiplos canais em um mesmo arquivo de projeto, além de atender às configurações de taxa de amostragem da interface de áudio. Entretanto, para a geração de acordes de teste, foi utilizado o *software* REAPER[®], por ser menos complexo de se realizar as gravações, facilitando a geração de arquivos de teste.

O *software* MATLAB[®], na versão 2016a, foi o escolhido para o desenvolvimento de funções, da base de dados e de rotinas de execução, por ser uma ferramenta de matemática computacional que contém todos os recursos necessários para tais fins.

4.2.2 Organização da Base de Dados

O conjunto de notas explorado contempla os dados relativos às amostras das 6 primeiras notas de cada uma das 6 cordas da guitarra (incluindo a nota de cada corda solta), totalizando $N_n = 36$ notas com $N_a = 30$ amostras cada.

Para a devida classificação de cada correspondente forma de onda resultante com sua respectiva nota, é elaborado o sistema de códigos representado pela Figura 8 para nomear as variáveis da base de dados, de acordo com a nota e sua posição no braço da guitarra.

No sistema elaborado, os primeiros dois caracteres indicam o nome da nota, de acordo com a notação previamente definida, e o terceiro, após o *underline*, indica a corda correspondente, sendo "6" referente à 6^a corda (a mais grave e mais espessa) e "1" referente à 1^a corda (a mais aguda e menos espessa). Essa distinção é relevante pois, na guitarra, duas notas iguais nas mesmas oitavas contém diferentes componentes harmônicas relacionadas à corda em que estão sendo tocadas, sabendo-se que as cordas são fabricadas em diâmetros diferentes e em dois tipos de material distintos. Logo, por exemplo, as formas de onda esperadas para as FFTs de $G_3(4)$ e $G_3(3)$ serão distintas devido à presença das componentes harmônicas, que variam conforme o material de fabricação, a espessura da

Figura 8 – Conjunto de códigos elaborado para os nomes dos arquivos de acordo com a posição no braço da guitarra; Os códigos sem cores, à esquerda, são referentes às cordas soltas

E4_1	F4_1	F#4_1	G4_1	G#4_1	A4_1
B3_2	C4_2	C#4_2	D4_2	D#4_2	E4_2
G3_3	G#3_3	A3_3	A#3_3	B3_3	C4_3
D3_4	D#3_4	E3_4	F3_4	F#3_4	G3_4
A2_5	A#2_5	B2_5	C3_5	C#3_5	D3_5
E2_6	F2_6	F#2_6	G2_6	G#2_6	A2_6

corda e a posição na qual a nota está sendo tocada no braço, mesmo que as duas sejam referentes à nota G_3 . Portanto, para a classificação de múltiplas notas simultâneas, é relevante determinar a corda em que a nota está sendo tocada.

São criadas, então, no *software* Studio One 3[®], 36 trilhas de gravação referentes às 36 notas, todas devidamente nomeadas de acordo com a convenção proposta.

4.2.3 Processo de Aquisição de Amostras

Para a aquisição de amostras, primeiramente, antes de qualquer seção de gravação, a guitarra é afinada com o auxílio de um *plugin* nativo do Studio One, denominado *Tuner*. Tal ação é necessária para garantir que as notas da base de dados estarão gravadas corretamente, com seus picos de frequência o mais próximo possíveis das frequências corretas das notas, de acordo com o modelo adotado. As gravações realizadas são feitas de forma aleatória no que diz respeito à execução da nota, utilizando diferentes configurações de captadores e equalização, alternando entre dedo e palheta e variando a duração das notas gravadas, a fim de obter a maior variedade de amostras possível. As gravações são realizadas de forma contínua para cada nota, onde são executadas 30 vezes por sessão de gravação, separadas por intervalos de tempo arbitrários. O procedimento é repetido, então, para as 36 notas, totalizando 1080 amostras. Pela Figura 9, é possível observar o ambiente de trabalho do programa após a gravação das amostras.

4.2.4 Tratamento dos Dados

Essa etapa se dá pela edição dos arquivos de áudio no Studio One e pelo desenvolvimento de funções no MATLAB para dividir o arquivo de áudio completo em amostras de 10 segundos, realizar a FFT de cada amostra e, por fim, realizar os cálculos estatísticos.

O procedimento de edição realizado no *software* de gravação foi o de separar todas as repetições gravadas para as notas em janelas de 10 segundos, com o fim de facilitar

Figura 9 – Área de trabalho do *software* Studio One, indicando uma porção das amostras gravadas e já recortadas para a geração da base de dados



a aquisição dos dados pelo MATLAB. Tendo realizado a divisão em todas as trilhas, os arquivos são exportados em formato **.wav** - na configuração de 24 bit PCM, com taxa de amostragem de 96 *kHz* e em canal *mono* - para serem carregados por uma rotina criada no MATLAB.

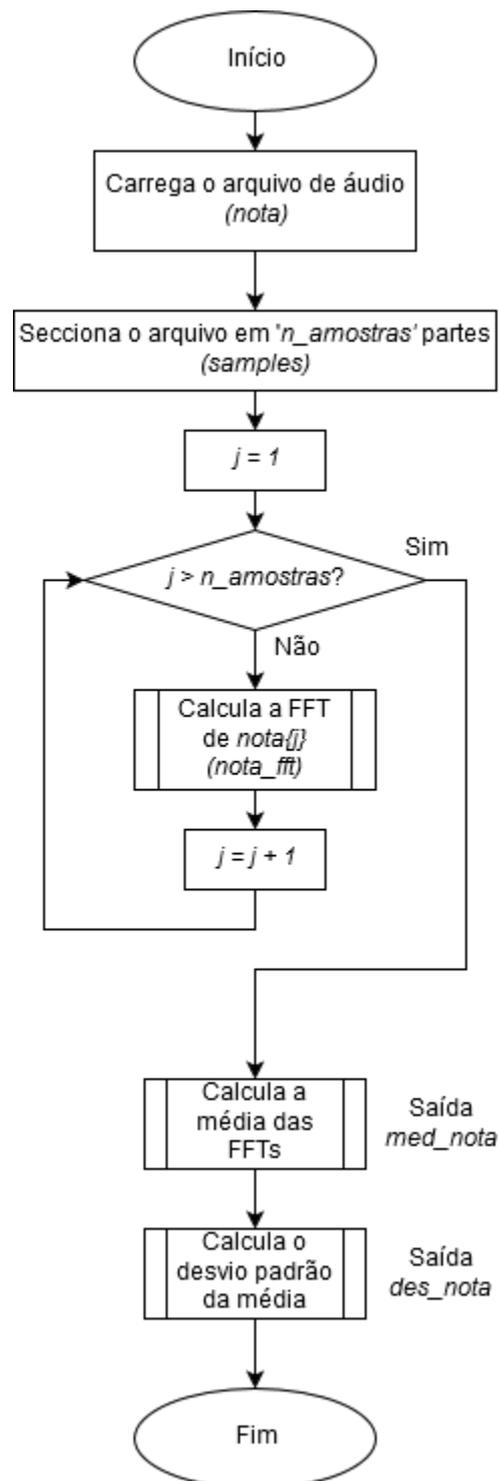
Ao longo do desenvolvimento da base de dados, foi decidido realizar o *downsampling* em todas as amostras para 48 *kHz* no momento do tratamento dos dados, visto que a base de dados seria muito pesada e talvez tornasse o programa mais lento de ser executado. A rotina utilizada para o processo de redução da taxa de amostragem é baseada na aplicação de um filtro FIR no sinal, seguido de uma filtragem das amostras remanescentes. Essa função é implementada pelo Código 1, presente no Apêndice A.

Por sua vez, a função para a obtenção dos parâmetros de média e desvio padrão amostrais de cada nota pode ser verificada nos Códigos 2 e 3, também presentes no Apêndice A. Esta realiza a leitura do arquivo de áudio, seccionando-o de 10 em 10 segundos e armazenando os vetores de cada seção em uma célula (ou seja, um vetor composto de vetores) de tamanho 960000 x 1. Posteriormente, realiza o *downsampling* para tornar os vetores de tamanho 480000 x 1, e então calcula a FFT de todos os sinais. Por fim, a média das FFTs é calculada utilizando a implementação da Equação (15), vista anteriormente na Seção 3.3, resultando em um espectro médio - que é a soma ponto-a-ponto das 30 FFTs dividida por 30 - e o desvio padrão por meio da função nativa do MATLAB *std*, que é uma aplicação direta da raiz de s^2 , descrito anteriormente na Equação (16). O

fluxograma descritivo da implementação da obtenção de média e desvio padrão das FFTs é apresentado na Figura 10.

Assim, os dados resultantes são armazenados em um arquivo de formato **.asv**, que armazena os dados do *workspace* do MATLAB para poderem ser carregados mesmo após o programa ser fechado e reaberto.

Figura 10 – Fluxograma para a rotina de obtenção dos dados de um arquivo de áudio de nota



4.3 Processamento dos Dados

4.3.1 Implementação do Coeficiente de Correlação

O sistema proposto se baseia fundamentalmente no resultado do cômputo do coeficiente de correlação entre o sinal do acorde de entrada e as notas da base de dados. Dessa forma, é necessário implementar a função que calcule os coeficientes de correlação ρ_{xy} entre o sinal de entrada e cada uma das notas da base de dados já definida. Baseando-se na definição da Equação (14), foi desenvolvida a função observada no Código 4, presente no Apêndice A.

Espera-se que a aplicação dessa função nesse contexto retorne valores entre 0 e 1, uma vez que os vetores resultantes das FFTs médias não assumem valores negativos de amplitude, não podendo ser inversamente correlacionados. Analisando os resultados, será possível determinar a similaridade entre os dois sinais inseridos na função proporcionalmente ao valor desse parâmetro, onde 1 representa correlação total e 0 representa total não-correlação. Vale notar que o resultado final da aplicação da correlação será normalizado pelo valor de pico para a etapa de classificação, pois só assim será possível utilizar o tipo de classificador projetado nessa aplicação, uma vez que os cálculos não apresentaram uniformidade nos valores de correlação entre as comparações, embora os valores em si, dentro de uma própria comparação, tenham sido aproximadamente coerentes com o resultado esperado.

Os testes para a validação da função desenvolvida são apresentados em duas etapas. Na primeira, realizam-se testes de formas distintas com notas individuais. Primeiramente, são comparados dois sinais idênticos, onde espera-se que o valor retornado pela função seja igual a 1. Outro tipo de teste é realizado com dois sinais distintos da mesma nota, que devem resultar em valores elevados em relação aos testes com notas diferentes, visto que é esperado que sejam muito semelhantes. Por fim, serão testadas duas notas distintas (não oitavas, sabendo que a oitava é uma harmônica par de uma dada nota, tendo espectro similar em frequência), onde espera-se um resultado numérico relativamente baixo, mais próximo de 0.

Na segunda etapa, serão realizados testes semelhantes, porém fazendo uso dos seletores desenvolvidos.

4.4 Definição dos Seletores

Conforme exemplificado no diagrama de funcionamento do sistema proposto, representado pela Figura 5, o Seletor é definido como a parte do algoritmo que realiza a seleção das notas mais relevantes dentre todas as 36 notas para compor o resultado final. Dentro dessa definição, são propostas duas versões do seletor para a análise de resultados:

a versão de Análise de Pertinência ao Acorde e a versão de análise pelo Maior Valor de Correlação.

Anteriormente à definição dos seletores, é realizada a normalização por picos do vetor resultante de correlações, com o objetivo de ser possível utilizar uma função linear que possa ser aplicada a todos os vetores resultantes para separar o que pode ser visto como uma detecção de nota. Essa função é definida aqui como o *threshold* T , sendo este um valor numérico entre $T = 0$ e $T = 1$ que será comparado com os valores de correlação resultantes para definir o que é considerado como uma nota detectada ou não detectada.

A normalização referida é realizada conforme a Equação (22), sendo $\rho_n(n)$ o vetor normalizado de valores de correlação entre o acorde e a n -ésima nota da base de dados.

$$\rho_n(n) = \frac{\rho(n)}{\max\{\rho(n)\}} \quad (22)$$

onde $\max\{\rho(n)\}$ é o valor de pico do vetor de correlações $\rho(n)$. Dessa forma, será possível aplicar um mesmo valor de *threshold* para a obtenção de resultados a fim de detectar as notas por similaridade relativa.

4.4.1 Variação do *threshold* T

Como descrito anteriormente, na Seção 3.4.1, *threshold* é, no contexto desse estudo, um limiar, representado por uma função linear, utilizado para ser comparado com os valores de correlação resultantes, de forma a definir o que será considerado pelo programa como uma nota detectada ou não detectada.

A variação do *threshold* é um procedimento de fundamental importância na análise realizada, uma vez que é buscada a melhor combinação de parâmetros para a obtenção da maior taxa de acerto de notas possível. Sendo assim, uma das etapas da rotina principal de cálculo de taxa de acerto consiste em variar o nível desse limiar em intervalos constantes e gerar resultados referentes a cada nível. Em termos numéricos, o *threshold* assumirá valores entre 0, 1 e 1, espaçados uniformemente de 0, 1 em 0, 1, para cada subsequente execução da rotina.

Os resultados da execução repetitiva das rotinas de cálculos e detecção em função da variação do *threshold* serão analisados pelos gráficos comparativos da taxa de acerto, falsos positivos e falsos negativos.

4.4.2 Análise de Pertinência ao Acorde

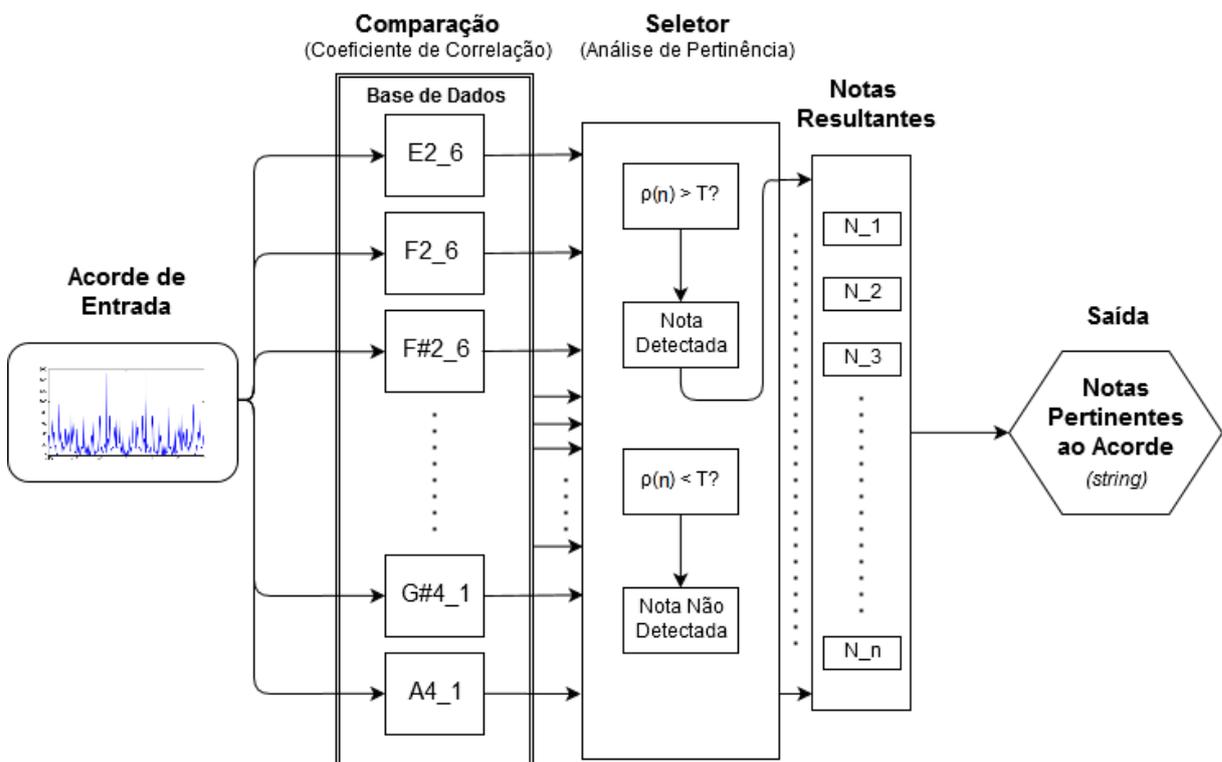
A primeira proposta de seletor é definida no intuito de fornecer como saída todas as prováveis notas pertencentes ao acorde de entrada filtrando as notas consideradas não pertencentes por meio da função T . Por meio dessa análise, pode-se comprovar a eficiência

do método utilizado no que diz respeito à identificação geral das notas, sem preocupação com a ordem de como estas foram tocadas.

É importante observar que é possível estimar o acorde tocado a partir da informação das notas individuais que o compõem, mesmo que fora de ordem. Porém, nesse caso, é necessária uma segunda determinação que depende de conhecimento prévio por parte do usuário para determinar a nomenclatura exata do acorde. Portanto, nesse caso, o usuário terá a informação sobre o conjunto de notas que compõem o acorde que foi tocado, embora seu nome não seja diretamente fornecido.

A descrição do algoritmo em formato funcional pode ser verificada no fluxograma da Figura 11.

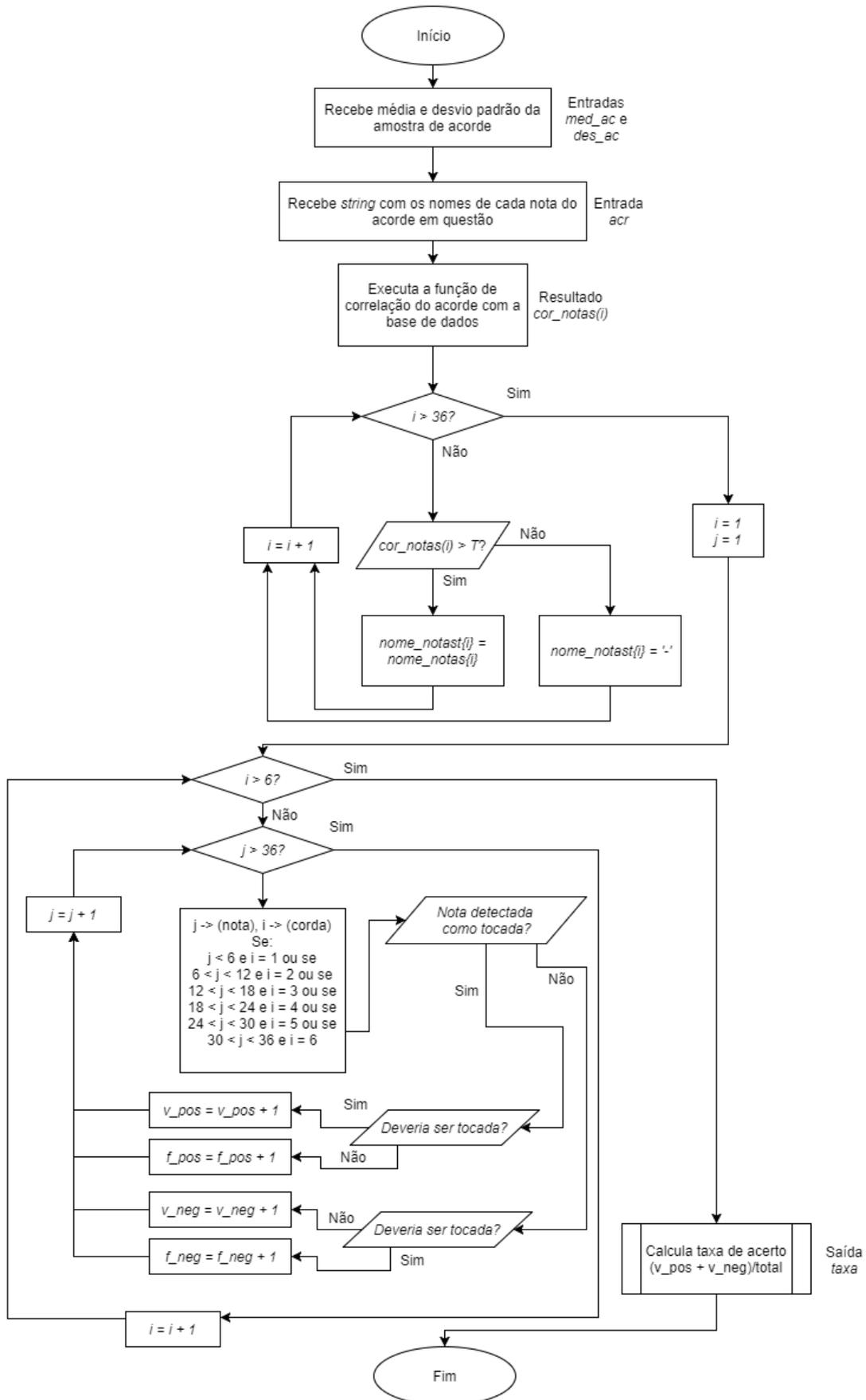
Figura 11 – Diagrama de blocos simplificado indicando o funcionamento do sistema



Com base nesse modelo, portanto, é desenvolvida uma rotina para obter os resultados das notas resultantes. O fluxograma referente ao desenvolvimento da rotina pode ser observado na Figura 12.

Para a implementação do classificador, foi utilizada a lógica apresentada pelo fluxograma em função dos valores de T , obtendo diferentes resultados na detecção de notas e gerando vetores com os resultados de valores de V_P , V_N , F_P e F_N para esse seletor.

Figura 12 – Fluxograma para a rotina de verificação pela análise de pertinência ao acorde



4.4.3 Análise pelo Maior Valor de Correlação

A segunda versão proposta contém algumas etapas adicionais em relação à primeira versão que tornam o algoritmo um classificador por ordem de maior valor de correlação. No caso, esse seletor detecta as seis notas mais relevantes dentre todas as notas consideradas detectadas da base de dados. É importante observar que, para selecionar os seis maiores valores de correlação de forma coerente, o algoritmo deve realizar as comparações corda a corda. Isso se justifica pelo fato de, por exemplo, duas notas não poderem ser tocadas na mesma corda. Logo, o algoritmo não pode fornecer uma saída contendo duas notas da mesma corda. Da mesma forma, como há notas que são equivalentes mesmo estando em cordas distintas, há uma grande probabilidade de, por exemplo, apenas uma delas ter sido tocada, mesmo que os valores de correlação das duas se apresentem elevados. Isso é justificado pela disposição dos tipos de acorde utilizados comumente. Numa tentativa de solucionar esse caso, decidiu-se comparar os segundos maiores valores de correlação das duas cordas em que foram detectadas notas equivalentes, selecionando a nota de maior valor de correlação dessa comparação. Vale notar que se houvesse mais notas do braço na base de dados, seria possível o caso de três até seis notas iguais em cordas diferentes, o que teria de ser levado em consideração para esse tipo de solução.

A descrição do algoritmo em formato funcional pode ser verificada no diagrama da Figura 13.

Dessa maneira, o algoritmo deve retornar 6 resultados, referentes a cada corda, utilizando essa comparação. Logo, sabendo-se que há 6 notas por corda na base de dados gerada, deve-se dividir a base de dados em 6 blocos e extrair o maior valor de correlação de cada bloco, a fim de ser possível compor o acorde completo com os 6 dados resultantes.

Com base nesse modelo, portanto, é desenvolvida uma rotina para obter os resultados das notas resultantes. O fluxograma referente ao desenvolvimento da rotina pode ser observado na Figura 14.

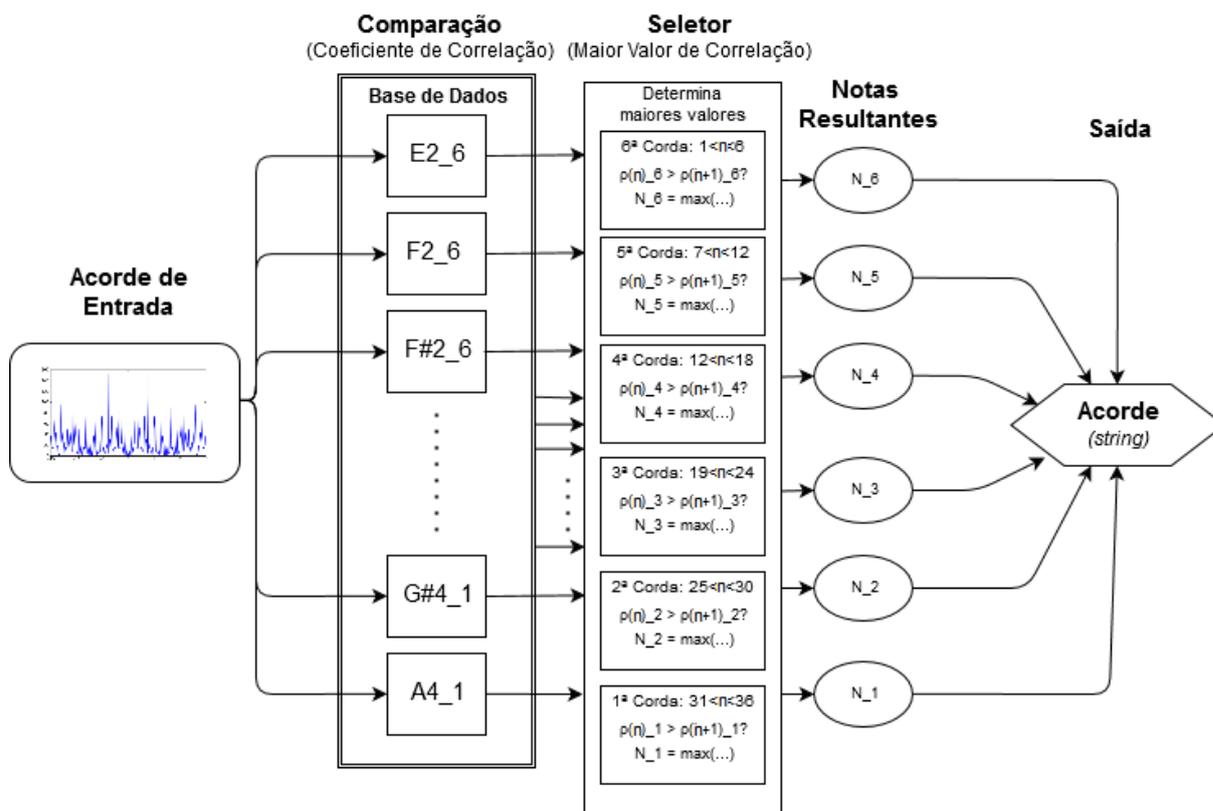
No fluxograma, a verificação dos maiores valores foi feita separadamente pela função *max_cor_notas*, disponível no Apêndice A, no Código 9, e a lógica apresentada já compreende a variação do limiar T , expresso pela variável m que varia de 1 a 10, representando a variação de 0,1 a 1,0 de T .

4.5 Verificação e Validação dos Dados Obtidos

4.5.1 Teste Preliminar da Correlação com a Base de Dados

Após toda a geração da base de dados, é realizada uma etapa de verificação dos dados por meio de testes preliminares da função ρ_{xy} elaborada. Essa etapa consiste em inserir como entrada do sistema a média obtida para uma nota qualquer e observar

Figura 13 – Diagrama de blocos simplificado indicando o funcionamento do sistema



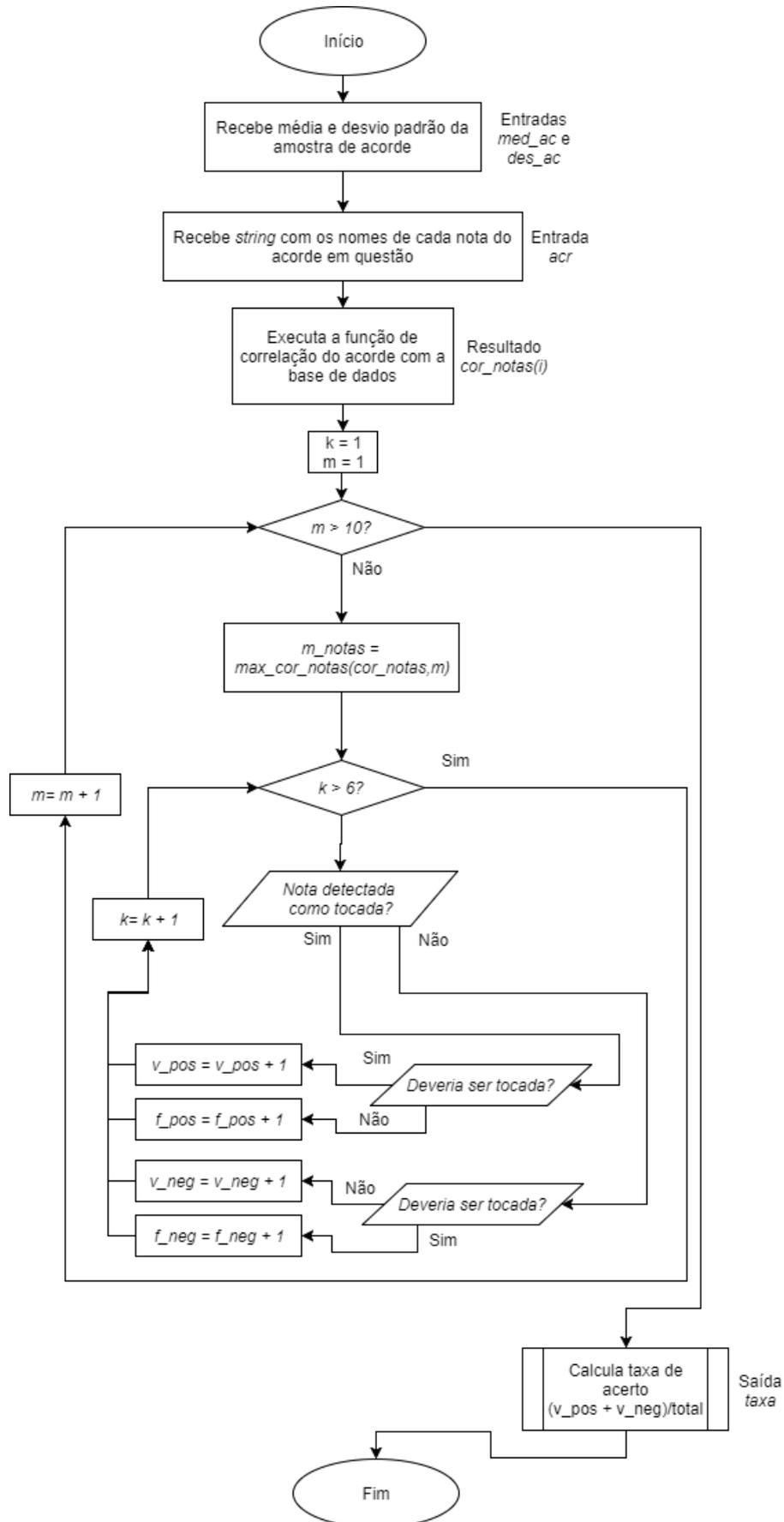
seu valor de correlação considerando amostras específicas da mesma nota na base de dados. De acordo com a definição do coeficiente de correlação, para as mesmas notas, os coeficientes devem retornar valores elevados, ou seja, próximos de 1. Para notas que são equivalentes a harmônicas do sinal da base de dados, espera-se que os coeficientes de correlação apresentem valores menores, porém ainda significativos de alguma forma. Por fim, realiza-se a correlação com notas distintas da base de dados, onde espera-se valores mais próximos de 0 para o coeficiente de correlação em relação aos outros casos.

4.5.2 Ensaios Aleatórios

Após terem sido definidos todos os passos referentes ao processo de comparação, é realizada a geração de amostras de acordes para testes. Para isso, é necessário escolher um determinado número de acordes a serem gravados para serem utilizados como entrada no processo.

Sabendo-se dessas especificações, são gravadas 10 execuções de cada acorde escolhido, que serão analisadas em conjunto após a aplicação das rotinas. A gravação foi realizada no *software* REAPER, utilizando a configuração de 48 kHz de taxa de amostragem de entrada para a interface de áudio *Fast Track Pro*, uma vez que a base de dados foi consolidada, também, nessa taxa de amostragem de gravação (após ser realizado

Figura 14 – Fluxograma para a rotina de verificação da seleção pelo maior valor de correlação



o procedimento de *downsampling*). O procedimento de gravação para cada acorde foi baseado na execução de formas variadas - utilizando dedos, palheta, tocando as notas sequencialmente ou tocando-as todas ao mesmo tempo - de forma aleatória, em intervalos de 10 segundos, até totalizar 10 amostras. A rotina de tratamento dos dados dos acordes pode ser conferida no Código 5, presente no Apêndice A.

4.5.2.1 Acordes Utilizados

A escolha dos acordes foi realizada no intuito de apresentar a maior variedade possível de amostras. Ou seja, tornando a validação do método consistente e obtendo resultados em que é possível analisar se há relação entre a taxa de acerto de determinados acordes e os intervalos pelos quais eles são construídos.

Dessa forma, os acordes utilizados para os ensaios são os indicados na Tabela 2, onde é indicada cada nota tocada em cada corda pelos códigos de S_1 a S_6 - sendo S o símbolo referenciando *string*, tradução em inglês para 'corda', acompanhado do número da corda correspondente - onde S_1 é a corda mais aguda, sendo as próximas cordas as cordas subsequentes, e sendo S_6 a corda mais grave, por consequência. Observa-se que os acordes escolhidos estão entre os mais variados tipos, contendo de 4 a 6 notas, tendo sido escolhidos de forma arbitrária.

A realização efetiva do processo em acordes com menos de 6 notas sendo tocadas também é fator relevante para se adicionar à utilidade da ferramenta e para a detecção de verdadeiros negativos. Essa característica é considerada para o desenvolvimento das rotinas, como pode ser visto no fluxograma da Figura 14, que assume a presença de verdadeiros negativos.

Para fins de comparações e classificação nas rotinas, os acordes definidos nota a nota na Tabela 2 foram implementados como conjuntos de *strings* no Código 11, disponível no Apêndice A.

Como é possível notar, optou-se por acordes de todos os tons possíveis, variando também os tipos de intervalos analisados. Por exemplo, para o tom de C , foram escolhidos diversos tipos de acorde para descobrir se o uso de intervalos diferentes influencia na detecção. Ao mesmo tempo, foram escolhidos outros tons para os mesmos tipos de acordes, de forma arbitrária, a fim de determinar se a mudança de tom também influencia na detecção. Deve-se notar, entretanto, que tanto as amostras da base de dados quanto as amostras dos acordes com intervalos equivalentes e tons diferentes podem não possuir a mesma relação de amplitudes para as frequências (por exemplo, provavelmente a nota E_2 da base de dados não tem os mesmos valores de picos de amplitude do que os da nota A_2), o que já aponta para uma provável diferença nos resultados para tons diferentes.

Tabela 2 – Acordes gravados como amostras para os ensaios aleatórios, especificados pelas respectivas notas por corda tocada

Acorde	S_6	S_5	S_4	S_3	S_2	S_1
C	-	C_3	E_3	G_3	C_4	E_4
$C6$	-	C_3	E_3	A_3	C_4	E_4
$C7$	-	C_3	G_3	$A\sharp_3$	E_4	G_4
$C7M$	-	C_3	G_3	B_3	E_4	G_4
$C7M(\sharp 5)$	-	C_3	E_3	$G\sharp_3$	B_3	E_4
Cm	-	C_3	G_3	C_4	$D\sharp_4$	G_4
$Cm6$	-	C_3	$D\sharp_3$	$G\sharp_3$	C_4	G_4
$Cm7M$	-	C_3	G_3	B_3	$D\sharp_4$	G_4
C°	-	C_3	$F\sharp_3$	A_3	$D\sharp_4$	-
D	-	-	D_3	A_3	D_4	$F\sharp_4$
Dm	-	-	D_3	A_3	D_4	F_4
E	E_2	B_2	E_3	$G\sharp_3$	B_3	E_4
Em	E_2	B_2	E_3	G_3	B_3	E_4
F	F_2	C_3	F_3	A_3	C_4	F_4
$F7M$	F_2	-	E_3	A_3	C_4	-
$F\sharp$	$F\sharp_2$	$C\sharp_3$	$F\sharp_3$	$A\sharp_3$	$C\sharp_4$	$F\sharp_4$
$F\sharp m7$	$F\sharp_2$	-	E_3	A_3	$C\sharp_4$	-
G	G_2	B_2	D_3	G_3	D_4	G_4
$G7$	G_2	D_3	F_3	A_3	D_4	G_4
$G\sharp 7$	$G\sharp_2$	-	$F\sharp_3$	C_4	$D\sharp_4$	-
$G\sharp m7(b5)$	$G\sharp_2$	-	$F\sharp_3$	A_3	D_4	-
A	-	A_2	E_3	A_3	$C\sharp_4$	E_4
$A9$	-	A_2	E_3	B_3	$C\sharp_4$	E_4
$A\sharp$	-	$A\sharp_2$	F_3	$A\sharp_3$	D_4	F_4
$A\sharp m$	-	$A\sharp_2$	F_3	$A\sharp_3$	$C\sharp_4$	F_4
B°	-	B_2	F_3	$G\sharp_3$	D_4	-

4.5.3 Geração dos Gráficos Comparativos

Tendo os resultados para os acordes de teste sido calculados pelas rotinas, é necessária a observação gráfica destes, para ser possível obter conclusões a respeito da performance dos classificadores. A geração dos resultados gráficos contempla três formatos principais: os gráficos de correlação por nota, os gráficos referentes à análise das taxas de acerto e as curvas ROC, detalhados nas seguintes Seções 4.5.3.1, 4.5.3.2 e 4.5.3.3. Além de serem as ferramentas principais para visualizar a eficiência do método adotado, também servem para demonstrar se há relação entre o tipo de acorde escolhido e o resultado na saída do sistema.

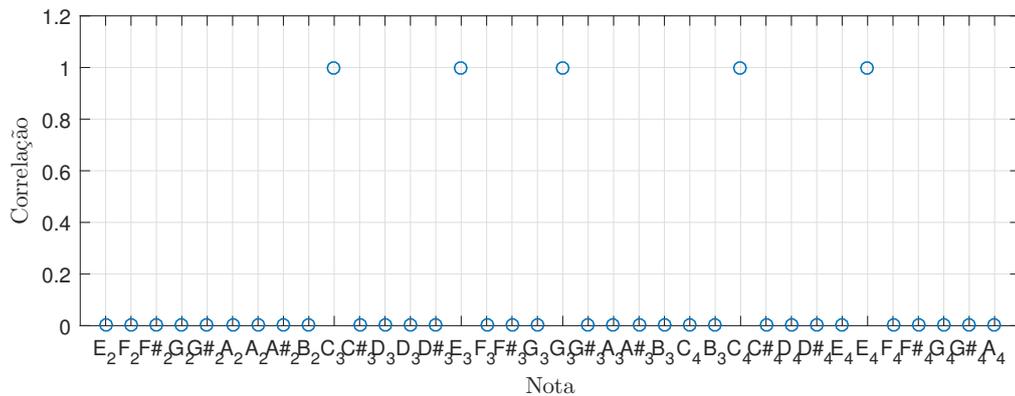
4.5.3.1 Correlação por Nota

Utilizando os resultados obtidos na execução da rotina que calcula a correlação do acorde de entrada com toda a base de dados, pode-se esboçar o gráfico que contém

os valores de correlação individuais para cada nota considerada na faixa especificada. A rotina de geração desses gráficos pode ser vista no Código 6, presente no Apêndice A.

Por exemplo, para o acorde C , num cenário ideal para a aplicação - ou seja, desprezando a influência dos harmônicos, do ruído intrínseco a cada gravação e das diferenças entre as amplitudes e as formas de onda - o resultado esperado deveria ser apresentado conforme indicado na Figura 15. Sabe-se, porém, que todos esses fatores influenciam para que o sinal resultante não seja conforme o apresentado na Figura 15.

Figura 15 – Resultado ideal para a aplicação da rotina no acorde C



A rotina de geração do gráfico da Figura 15 também pode ser vista no Apêndice A, pelo Código 7.

Esse resultado contempla, portanto, o valor de correlação máximo para as notas pertencentes ao acorde C : C_3 , E_3 , G_3 , C_4 e E_4 . É importante salientar que a ordem de notas indicadas no eixo x é disposta conforme a ordem crescente das notas por corda dentro da faixa especificada, havendo repetições de algumas notas na passagem de uma corda para outra. Esse fator resultará em níveis de correlação muito próximos entre duas notas iguais em cordas diferentes, o que está sendo desconsiderado nesse esboço, que seria um resultado ideal.

Os resultados desse formato de análise são especialmente relevantes para definir as notas pertencentes aos acordes, função realizada posteriormente pelos seletores, e para verificar se o casamento por correlação é um bom indicador de similaridade para essa aplicação.

4.5.3.2 Análise das Taxas de Acerto

A partir da implementação dos seletores especificados nas Seções 4.4.2 e 4.4.3, é possível gerar gráficos que exprimem a taxa de acerto do classificador em função da

variação do *threshold*. A taxa de acerto T_A é definida pela Equação (23), onde V_P e V_N são, respectivamente, os verdadeiros positivos e verdadeiros negativos.

$$T_A = \frac{V_P + V_N}{6} \cdot 100\% \quad (23)$$

Junto à taxa de acerto, como parâmetros de análise, também são contabilizados os falsos positivos F_P e falsos negativos F_N . Tais parâmetros são relevantes para determinar o valor de T onde o sistema responde da melhor forma possível, que seria um ponto ótimo entre uma alta taxa de acerto e a menor quantidade de falsos positivos e falsos negativos, simultaneamente.

A implementação da detecção de falsos positivos procede de forma que o valor da k -ésima posição do vetor de detecções será considerado um falso positivo se ele for ao mesmo tempo diferente das notas pertencentes ao acorde teórico e diferente do indicador ' - ', que simboliza a não detecção de uma nota (logo, assumindo que ela não foi tocada). De forma análoga, se, ao mesmo tempo, quando a k -ésima posição do vetor de nomes das notas coincidir com uma nota a ser tocada e o vetor de detecções indicar a falta de detecção, ou seja, o ' - ', este será considerado um falso negativo. Todos os procedimentos de comparação são realizados por comparação de *strings* no código, pois foi escolhido receber como entrada uma string das notas do acorde. Também para poder identificar notas mesmo que sejam iguais em lugares diferentes e para facilitar a obtenção de resultados das notas já com o nome respectivo.

Tendo definido os parâmetros, o formato proposto de resultado para análise é o de um *plot* contendo os três gráficos, referentes aos três parâmetros - Taxa de Acerto (%), F_P e F_N - em função do *threshold* T .

Sabendo os conceitos desse tipo de análise envolvendo falsos positivos e negativos, como consequência, espera-se que o *threshold* assumira seu valor ótimo em um ponto intermediário entre os valores determinados para comparação. Isso se dá pelo fato de que um sistema com valor de *threshold* muito baixo, provavelmente detectará informações excessivas, ou seja, será sensível demais para a aplicação desejada, enquanto um sistema com valor de *threshold* muito elevado detectará informação insuficiente. Portanto, resta verificar se os resultados serão coerentes com tais prerrogativas.

4.5.3.3 Geração das Curvas ROC

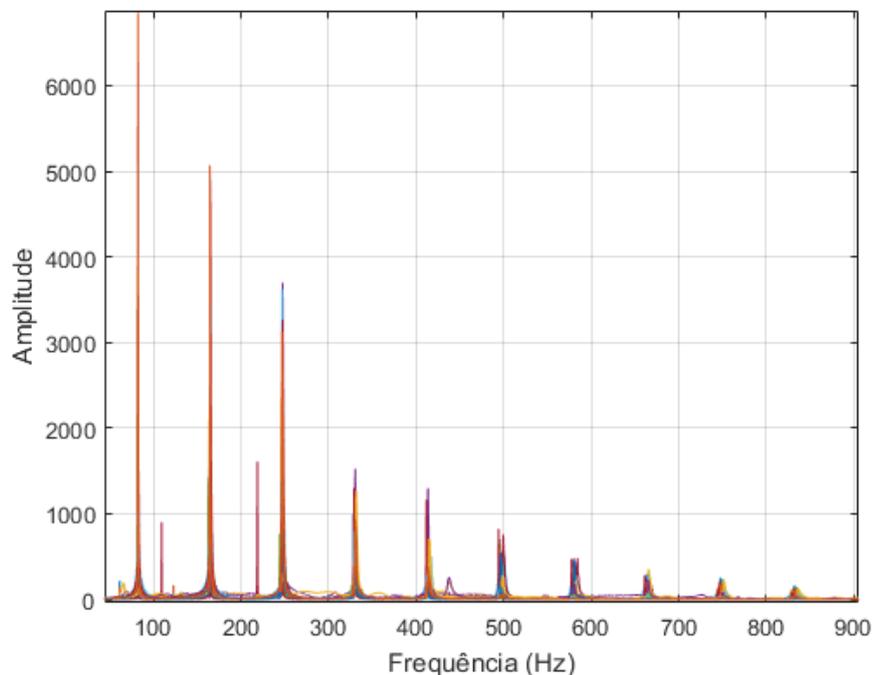
Com a implementação dos seletores já definida, resta obter as curvas ROC a partir dos resultados destes. Como descrito na Seção 3.4.2, a curva ROC pode ser obtida a partir do *plot* da taxa TVP pela taxa TFP , calculadas pelas Equações (19) e (21), definidas na Seção 3.4.2. A implementação da geração das curvas pode ser conferida nas funções dos dois seletores, nos Códigos 8 e 10 do Apêndice A.

5 Resultados e Discussões

5.1 Base de Dados

A aquisição de amostras para a composição da base de dados pôde ser contemplada para todas as notas da faixa escolhida. Realizando os procedimentos especificados na Seção 4.2, é possível obter as formas de onda das FFTs e os valores de desvio-padrão associados. Como exemplo, pode-se observar o esboço de todas as FFTs da nota E_2 superpostas na Figura 16. Percebe-se que há claras distinções entre os sinais da mesma nota, devido aos diferentes tipos de execução ao serem gravados, como descrito na Seção 4.2.1. Também é possível observar o pico de amplitude na frequência fundamental da nota ($f_{E_2} \approx 82,41 \text{ Hz}$), além de suas harmônicas, ou seja, os subseqüentes múltiplos inteiros da frequência fundamental (164,82 Hz ; 247,23 Hz ; 329,64 Hz , etc). Nesse caso, a amplitude das frequências não foi normalizada.

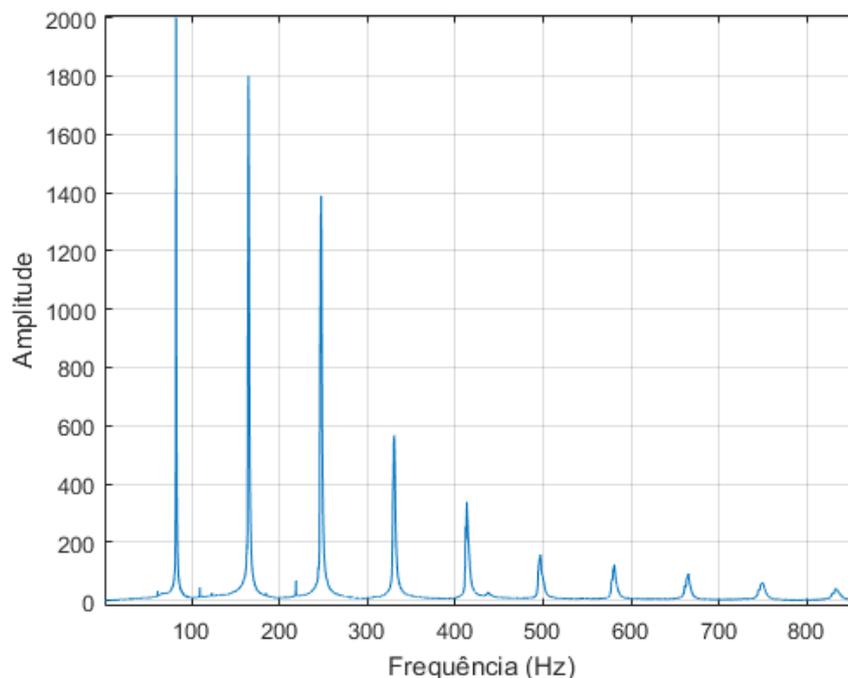
Figura 16 – Espectro em frequência de 30 amostras da nota E_2



De posse das 30 FFTs da nota, é possível calcular a FFT média a ser utilizada como informação fundamental da base de dados. O resultado é dado pelo espectro apresentado na Figura 17. É importante observar que a forma de onda média proporciona uma métrica de comparação muito mais abrangente e precisa do que a forma de onda de uma simples amostra proporcionaria. Quanto mais amostras diferentes de um mesmo

sinal forem contabilizadas no espectro médio, mais abrangente será a base de dados para o processamento de diferentes sinais de entrada. Um dos motivos de o sinal resultante ser adequado para a aplicação é o de que o cálculo da média funciona como um filtro, atenuando uma série de ruídos no sinal, tornando as informações de interesse mais explícitas e homogêneas.

Figura 17 – Espectro em frequência da média aritmética das 30 amostras da nota E_2



Da mesma forma, então, os vetores contendo as FFTs de todas as 36 notas foram gerados, incluindo seus respectivos valores de desvio padrão. Estes são, então, nomeados e armazenados no arquivo de variáveis do *workspace* do MATLAB (de formato **.asv**) para serem carregados como base de dados pelas rotinas.

Os valores resultantes para os desvios padrão de cada FFT média relativa a cada nota da base de dados podem ser observados na Tabela 3, onde os números em parêntesis representam a corda sendo tocada (sendo 6 referente à sexta corda e as outras numerações analogamente referenciadas). É possível observar certo padrão entre os valores de desvio padrão, uma vez que quanto mais aguda (ou seja, com frequência mais alta) é a nota da base de dados, mais baixo tende a ser o valor do desvio padrão. Esse padrão de diminuição é observado acentuadamente principalmente a partir da nota $B_3(2)$. Isso porque, no caso da execução em um instrumento de cordas, a potência do sinal é mais elevada em sinais mais graves, uma vez que a energia mecânica necessária para realizar a vibração das cordas mais graves é maior do que a necessária para realizar a vibração nas cordas mais agudas.

Um outro fato a ser observado é o de que não foram analisadas notas mais agudas além da 5ª casa (5ª posição no braço da guitarra) sendo tocadas nas cordas mais graves. Tal fato, de acordo com a afirmação anterior, deveria confirmar a constatação de que a potência do sinal das cordas mais graves é elevada em comparação às cordas mais agudas.

Tabela 3 – Desvios padrão para cada nota da base de dados

<i>Nota</i>	σ_N
$E_2(6)$	22, 2169
$F_2(6)$	25, 7226
$F\sharp_2(6)$	22, 7907
$G_2(6)$	25, 4801
$G\sharp_2(6)$	25, 4364
$A_2(6)$	24, 8676
$A_2(5)$	22, 1035
$A\sharp_2(5)$	21, 8765
$B_2(5)$	20, 4315
$C_3(5)$	17, 0199
$C\sharp_3(5)$	22, 4146
$D_3(5)$	21, 5181
$D_3(4)$	15, 3087
$D\sharp_3(4)$	23, 4543
$E_3(4)$	15, 0733
$F_3(4)$	32, 3001
$F\sharp_3(4)$	27, 6062
$G_3(4)$	23, 1398
$G_3(3)$	18, 0802
$G\sharp_3(3)$	14, 3185
$A_3(3)$	28, 2185
$A\sharp_3(3)$	29, 7793
$B_3(3)$	24, 4330
$C_4(3)$	32, 1497
$B_3(2)$	8, 0250
$C_4(2)$	14, 6455
$C\sharp_4(2)$	14, 6543
$D_4(2)$	16, 1196
$D\sharp_4(2)$	12, 7671
$E_4(2)$	11, 0892
$E_4(1)$	3, 4113
$F_4(1)$	4, 0534
$F\sharp_4(1)$	5, 1543
$G_4(1)$	5, 5426
$G\sharp_4(1)$	6, 2890
$A_4(1)$	5, 9523

Os resultados para os valores médios das FFTs de cada nota foram obtidos e armazenados em vetores de tamanho $K = 480000$, o que equivale a 10 segundos de um sinal amostrado a 48 kHz .

5.2 Testes Preliminares do Processamento

Tendo obtido o espectro médio em frequência das notas, foi possível realizar uma bateria de testes inicial para verificar a validade da utilização do coeficiente de correlação para o casamento de espectros. Os resultados de oito combinações diferentes de sinais comparados com a nota E_2 a partir do sinal médio e de sinais singulares da base de dados podem ser conferidos na Tabela 4.

Tabela 4 – Coeficiente de correlação para 8 combinações distintas de sinais

$x(k)$	$y(k)$	ρ_{xy}
E_{2med}	E_{2med}	1,0000
E_{2med}	$E_{2(1)}$	0,9298
E_{2med}	$E_{2(2)}$	0,7888
E_{2med}	$E_{2(3)}$	0,6130
E_{2med}	$E_{2(4)}$	0,9317
E_{2med}	$E_{2(5)}$	0,8013
E_{2med}	$E_{2(6)}$	0,9135
E_{2med}	$E_{2(7)}$	0,8999

Como esperado, a comparação entre dois sinais idênticos resultou no valor $\rho_{xy} = 1,000$, o que é um indicador de que o algoritmo do cálculo da correlação foi implementado de forma correta. Entretanto, observa-se uma faixa de valores para as outras combinações, que varia de 0,6130 a 0,9317. Tal resultado mostra a similaridade de diversos sinais com o sinal médio, com uma faixa de aplicação de aproximadamente 4 pontos decimais de variação. Esse resultado, nesse caso, indica que a aplicação de um *threshold* T maior do que 0,6130, por exemplo, implicaria na não detecção dessa nota em alguns casos.

Outro teste realizado contemplou a comparação entre o sinal de E_{2med} e alguns outros sinais médios já obtidos, esperando-se valores baixos de correlação em comparação ao teste anterior. O resultado pode ser conferido na Tabela 5.

Pode-se observar, a partir desses resultados, que foi possível obter certa coerência nas comparações entre duas notas individuais, uma vez que nenhum dos resultados da Tabela 5 apresentou valores muito próximos dos valores de correlação entre os próprios sinais de E_2 . Entretanto, é importante observar que o resultado da primeira comparação entre oitavas diferentes da nota E obteve o maior entre todos os valores do coeficiente de correlação, indicando que as comparações das primeiras oitavas dos sinais podem obter um coeficiente de correlação mais elevado em relação a outras comparações.

Tabela 5 – Coeficiente de correlação entre o sinal de $E_2(6)_{\text{med}}$ e outros sinais médios arbitrários escolhidos

$x(k)$	$y(k)$	ρ_{xy}
$E_2(6)_{\text{med}}$	$B_2(5)_{\text{med}}$	0,3707
$E_2(6)_{\text{med}}$	$E_3(4)_{\text{med}}$	0,4606
$E_2(6)_{\text{med}}$	$G\#_3(3)_{\text{med}}$	0,0987
$E_2(6)_{\text{med}}$	$B_3(2)_{\text{med}}$	0,1952
$E_2(6)_{\text{med}}$	$E_4(1)_{\text{med}}$	0,1902

Por fim, realizou-se outro teste, agora inserindo um acorde na entrada - especificamente, o \mathbf{E} (composto por $E_2(6)$, $B_2(5)$, $E_3(4)$, $G\#_3(3)$, $B_3(2)$ e $E_4(1)$) - a fim de compará-lo com os sinais das notas individuais que o compõem. O resultado pode ser observado na Tabela 6.

Tabela 6 – Coeficiente de correlação entre o sinal do acorde \mathbf{E} e os sinais da base de dados das notas individuais desse acorde

$y(k)$	ρ_{xy}
E_2	0,6936
B_2	0,5508
E_3	0,3311
$G\#_3$	0,1854
B_3	0,1904
E_4	0,2062

Com esse teste, é possível observar que os resultados não são significativos no que diz respeito aos valores numéricos ao utilizar esse método, uma vez que para a comparação de $E_2(6)$ pôde-se obter o valor de 0,6936 enquanto que para os outros sinais, os resultados foram extremamente reduzidos. Isso pode ser explicado pela diferença de amplitude entre os picos das frequências, que faz com que o produto dos sinais ponto a ponto não resulte em valores elevados quando as frequências de picos mais altos não estão contidas nos sinais $y(k)$ de comparação.

Uma forma de tentar resolver esse problema para a aplicação do classificador é a normalização do sinal, de forma que os picos de amplitude não influenciem tanto no resultado. Para evitar esse problema, no desenvolvimento dos classificadores, um valor de correlação resultante, por exemplo, é comparado nota a nota com outros valores resultantes dentro da faixa de uma corda, e não em toda a faixa da base de dados.

5.3 Testes com Amostras de Acorde

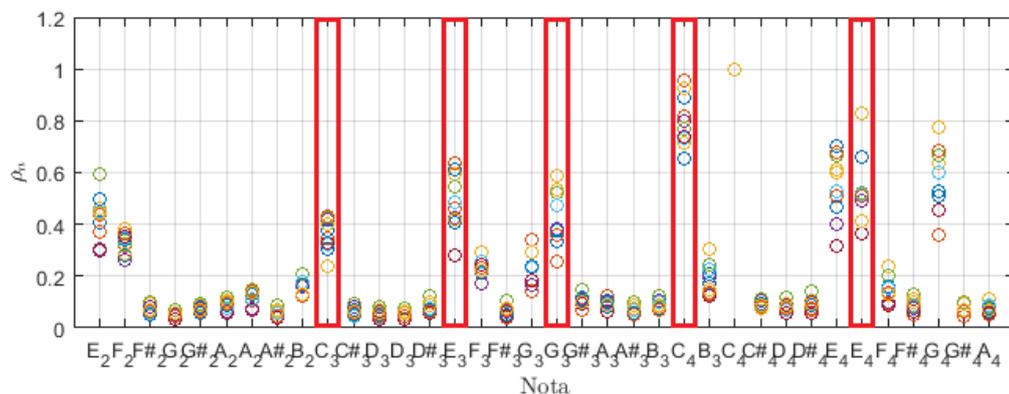
Realizando todos os passos especificados nas Seções 4.3 e 4.4, é possível obter os diferentes resultados de acordo com respectivo tipo de análise. Os resultados foram gerados

com o auxílio do Código 12, disponível no Apêndice A, onde os nomes das variáveis dos acordes foram informados manualmente como parâmetros para a geração de cada *set* de gráficos. Cada tipo de resultado é detalhado nas Seções 5.3.1 e 5.3.2 a seguir. Para a apresentação dos resultados, foram escolhidos os acordes **C** e **G**. Os resultados referentes a todos os outros acordes podem ser conferidos nos Apêndices B, C e D. Devido a um erro persistente no sistema de aquisição de dados que não foi possível de ser resolvido (e prejudicou os arquivos de áudio gravados), não foi possível obter os resultados dos acordes **Dm**, **E**, **Em**, **A♯**, **A♯m** e **B^o** de forma adequada. Portanto, estes foram suprimidos do documento.

5.3.1 Resultados da Análise da Correlação por Nota

Executando a rotina do Código 6 para todos os acordes, obteve-se os gráficos de correlação por nota em função do *threshold* T . Os resultados referentes à correlação por nota para 10 amostras do acorde **C** são demonstrados na Figura 18. As marcações em vermelho indicam as notas reais do acorde em questão para todos os gráficos gerados nessa categoria de resultado. O código pode ser conferido no Apêndice A.

Figura 18 – Valores dos coeficientes de correlação por nota para 10 amostras do acorde **C**

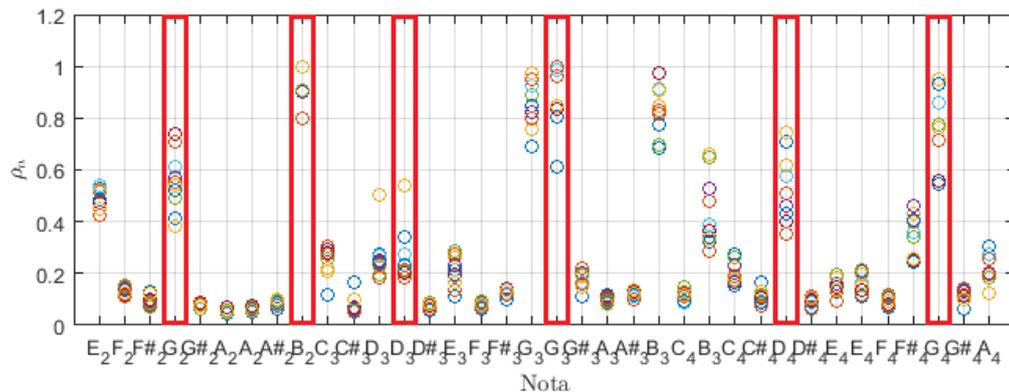


Como pode-se observar pelo resultado do acorde **C**, os valores de correlação por nota obedecem algum nível de coerência, uma vez que a maioria dos maiores valores resultantes coincide com as notas que foram tocadas. Porém, também é possível notar que houve valores de correlação de outras notas que também foram elevados. Analisando o caso do **C**, especificamente, pode-se ver que os altos valores de correlação são apresentados pelas respectivas oitavas de notas já presentes no acorde como, por exemplo, E_2 , E_4 e G_4 , que são oitavas de E_3 e G_3 , respectivamente. Isso pode ser explicado pelo fato de que a segunda harmônica de uma nota individual da base de dados tem um pico elevado na grande parte das vezes, o que influencia fortemente no resultado final.

Como mencionado na Seção 4.4.3, há também valor de correlação relativamente elevado na nota G_3 , que corresponde a uma nota que está presente em duas cordas, como em diversos possíveis casos. Para a classificação, isso é um obstáculo a ser testado pelo algoritmo proposto para o seletor de maior valor de correlação.

Os resultados referentes à correlação por nota para 10 amostras do acorde G são demonstrados na Figura 19

Figura 19 – Valores dos coeficientes de correlação para 10 amostras do acorde G



Da mesma forma como é visto no resultado para o acorde C , pode-se destacar a maioria das notas pertencentes ao acorde dentre os maiores valores de correlação observados. Ainda assim, a influência das oitavas das notas é visível, bem como a de notas repetidas entre cordas. Entretanto, os níveis de correlação da maioria das notas que não pertencem ao acorde está dentro de padrões baixos comparados aos altos níveis de correlação, o que aponta um bom resultado da função de coeficiente de correlação para indicar a similaridade entre os sinais.

É importante lembrar que o valor de correlação ρ_n apresentado pelo gráfico é o coeficiente normalizado de correlação. Como este foi normalizado por pico, o valor $\rho_n = 1$ não é necessariamente correspondente a $\rho_{xy} = 1$, que indicaria correlação total. O valor $\rho_n = 1$ representa o maior valor dentre todos os calculados para cada amostra.

Essa normalização é necessária para a correta aplicação da função T , que assume intervalos fixos de valores. Porém, como pode-se observar em alguns resultados de outros acordes, essa normalização faz com que todos os valores acabem sofrendo alguma alteração. Nos acordes de F e $F7M$ (Figuras A10 e A11), por exemplo, observa-se que os valores de correlação para as notas não pertencentes aos acordes encontram-se em um patamar mais elevado do que os dos acordes C e G . Isso porque o maior valor absoluto de correlação nesses acordes provavelmente foi baixo, o que resulta em uma potencialização de todos os valores no ato da normalização. Tal fato implica em um problema para a aplicação adequada do seletor.

Outra característica que pode ser notada analisando todos os gráficos é a baixa correlação relativa em notas mais agudas que não são oitavas diretas de uma nota já tocada. No resultado para os acordes **Cm6** e **D**, por exemplo, nota-se que as notas mais agudas pertencentes aos acordes apresentaram baixa correlação, a ponto de não terem um nível de provável detecção. Em contrapartida, em acordes como **F#** e **A**, a correlação para as notas mais agudas pertencentes aos acordes foi elevada, mas essas também são oitavas de notas já presentes no acorde. Isso sugere que a potência do sinal é influência direta no valor absoluto de correlação, o que é coerente pois, para essa base de dados, os picos das harmônicas dos sinais das notas mais graves apresentaram valor mais significativo do que os picos das fundamentais de diversas notas mais agudas no espectro da frequência.

5.3.2 Resultados da Análise das Taxas de Acerto

5.3.2.1 Método de Análise de Pertinência ao Acorde

Executando a rotina implementada no Código 8, foi possível obter todos os gráficos contendo a taxa de acerto, os falsos positivos F_P e falsos negativos F_N em função do *threshold* T . Tais gráficos podem ser conferidos no Apêndice C, nas Figuras de A19 a A36. Além disso, pôde-se gerar os gráficos para as curvas ROC de todos os acordes, contempladas nas Figuras de A37 a A54.

Os resultados para o primeiro conjunto de gráficos para 10 amostras do acorde **C** são apresentados pela Figura 20. Os valores possíveis para a taxa de acerto são de 0 %, 16,6667 %, 33,3333 %, 50 %, 66,6667 %, 83,3333 % e 100 %, visto que o sistema pode acertar de 0 a 6 detecções sobre 6 notas totais a serem detectadas corretamente.

Observando o gráfico de taxa de acerto do acorde **C** (que compreende os verdadeiros positivos e verdadeiros negativos), nota-se uma certa dispersão nos valores em função de T . Para valores entre 0,1 e 0,2 e entre 0,7 e 1,0, a taxa de acerto se manteve constante para a maioria das amostras. Porém, é importante notar a alta taxa de falsos positivos para valores baixos de T . Para valores mais altos, por outro lado, é importante notar a alta taxa de falsos negativos. Uma combinação ótima para esse acorde, para algumas amostras, seria entre os pontos de $T = 0,3$ e $T = 0,4$, onde há amostras que ainda apresentam a maior taxa de acerto desse caso (aproximadamente 83,3333%), com valor reduzido de falsos positivos e de falsos negativos.

Os resultados para 10 amostras do acorde **G** são apresentados na Figura 21.

No caso do acorde **G**, foi possível obter taxa de acerto $T_A = 100\%$ para uma faixa de valores de T em algumas amostras. Esse resultado demonstra que o método adotado pode ter detecção total das notas pertencentes ao acorde. Pode-se observar que para $T = 0,5$ uma das amostras ainda se manteve com taxa de acerto $T_A = 100\%$, com poucos casos de falsos positivos e falsos negativos.

Figura 20 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde \mathbf{C}

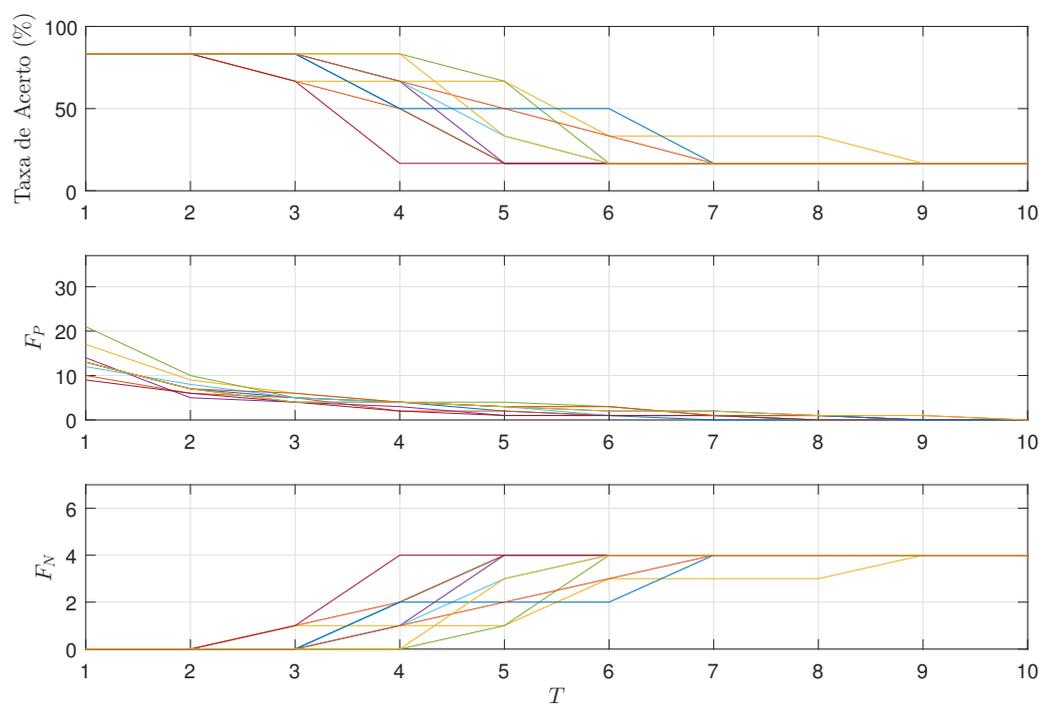
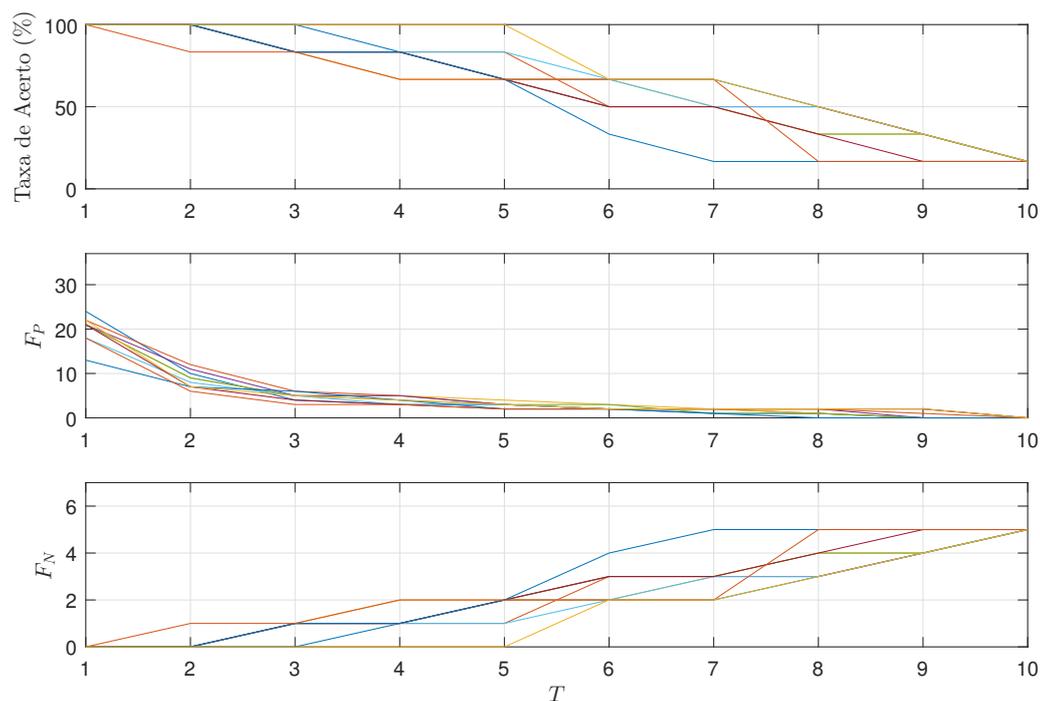


Figura 21 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde \mathbf{G}



No geral, esse acorde apresentou melhores resultados na detecção do que o acorde C , o que confirma a hipótese de que acordes diferentes implicam em diferentes resultados na detecção. Tal afirmação é coerente, pois até nas amostras do mesmo acorde há variações da resposta, por mais que dentro de uma certa faixa para cada quantidade analisada.

Obtendo as informações sobre todos os acordes de teste, pode-se observar ampla variabilidade entre os resultados. Em resultados para alguns acordes se observa taxa de acerto nula por toda a variação de T , enquanto em outros observa-se alta taxa de acerto para os primeiros valores de T .

Os valores máximos de taxa de acerto TAM observados para amostras observadas de todos os acordes, juntamente com o valor de T até onde estes se mantêm são listados na Tabela 7.

Tabela 7 – Taxas de acerto máximas para cada acorde, juntamente com o valor de T máximo por onde essa taxa de acerto máxima se manteve, sob o seletor de Análise de Pertinência ao Acorde

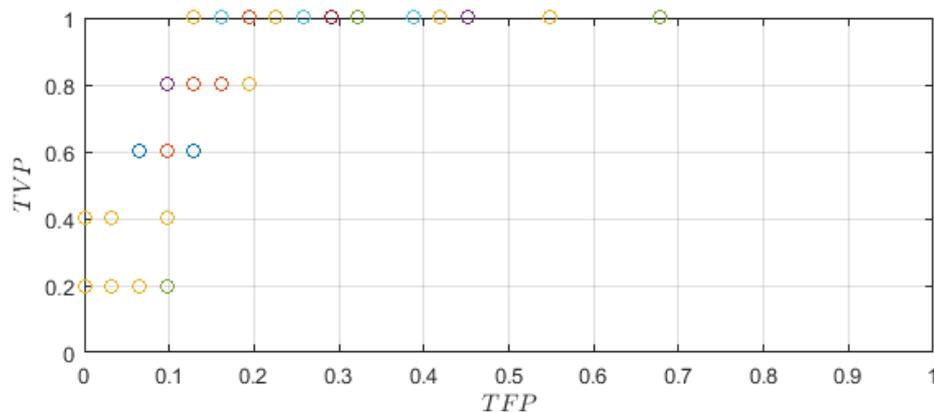
Acorde	TAM (%)	T_{TAM}
C	83,3333	0,4
$C6$	83,3333	0,4
$C7$	50,0000	0,2
$C7M$	83,3333	0,3
$C7M(\#5)$	50,0000	0,1
Cm	16,6667	1,0
$Cm6$	66,6667	0,1
$Cm7M$	33,3333	0,1
C°	16,6667	0,1
D	16,6667	0,7
F	100,0000	0,2
$F7M$	66,6667	0,3
$F\#$	33,3333	0,1
$F\#m7$	16,6667	0,1
G	100,0000	0,5
$G7$	100,0000	0,1
$G\#7$	0,0000	0,1
$G\#m7(b5)$	33,3333	0,2
A	33,3333	0,4
$A9$	33,3333	0,4

As taxas de acerto na análise de pertinência ao acorde não demonstraram relação direta com os intervalos pelos quais os acordes são compostos. Por exemplo, os acordes A , C , D , F e G são construídos a partir das mesmas relações intervalares e mesmo assim não apresentaram resultados similares. Da mesma forma, poucos acordes apresentaram alta taxa de acerto para valores elevados de T , o que indica que essa função de *threshold*

aplicada a esse resultado de correlações por nota pode não ser a melhor forma de classificar as notas que pertencem ao acorde.

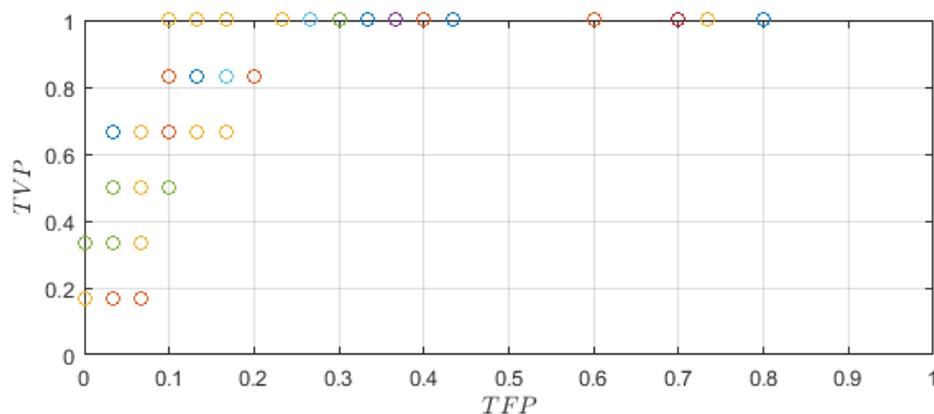
Como previsto pelo método apresentado na Seção 3.4.2, a geração da curva ROC fornecerá outra visão sobre a performance do classificador desenvolvido. A curva ROC para 10 amostras do acorde C pode ser visualizada na Figura 22, onde cada cor dos círculos representa uma das amostras testadas. Nessa análise, toma-se como ótimo o caso onde $TVP = 1$ e $TFP = 0$ para um detector ideal. No caso do acorde C , nota-se que alguns pontos apresentam uma performance razoável, com TVP entre 0,6 e 0,8 ao mesmo tempo que a TFP está entre 0,1 e 0,2. Esses casos apontam uma relação entre verdadeiros positivos e falsos positivos que pode ser de interesse. Porém, há diversos casos para diversas amostras onde o ponto $TVP = 1$ é atingido conjuntamente de uma alta TFP , o que é indesejado, pois por mais que uma alta TVP seja desejada, o preço a ser pago para obtê-la, nesse caso, é de uma alta TFP . Para valores muito baixos de TFP , observa-se diversos casos onde TVP também é baixo, o que caracteriza classificadores com pouca taxa de acerto.

Figura 22 – Curva ROC para o acorde C



A curva ROC para 10 amostras do acorde G pode ser visualizada na Figura 23.

Figura 23 – Curva ROC para o acorde G



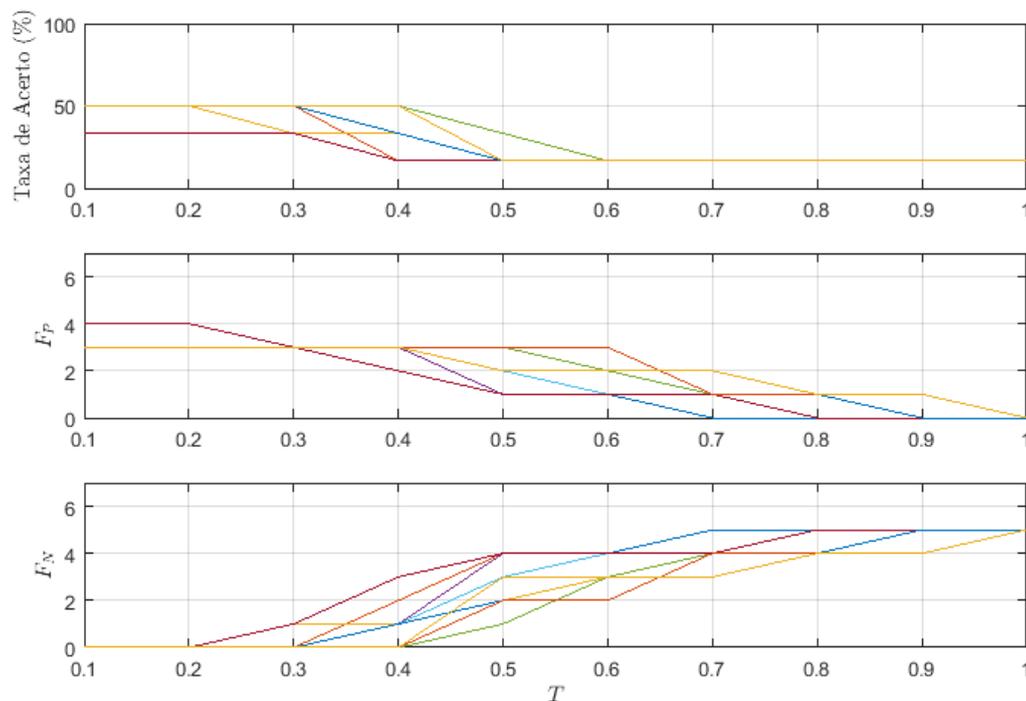
O resultado para o acorde \mathbf{G} pode ser visto tanto como melhor que o resultado para \mathbf{C} e como pior, pois por mais que haja mais pontos próximos de 0 para os valores de TFP , quando $TVP = 1$, tem-se uma elevada TFP para diversas amostras, o que caracteriza os classificadores como frágeis, ou seja, apontam diversos verdadeiros positivos conjuntamente com diversos falsos positivos.

5.3.2.2 Método de Maior Valor de Correlação

Executando a rotina implementada no Código 10, foi possível obter todos os gráficos contendo a taxa de acerto, os falsos positivos F_P e falsos negativos F_N em função do *threshold* T . O restante dos gráficos pode ser conferido no Apêndice D, nas Figuras de A55 a A72. Além disso, pôde-se gerar os gráficos para as curvas ROC de todos os acordes, contempladas nas Figuras de A73 a A90.

Os resultados para o primeiro conjunto de gráficos para 10 amostras do acorde \mathbf{C} são apresentados pela Figura 24.

Figura 24 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde \mathbf{C}

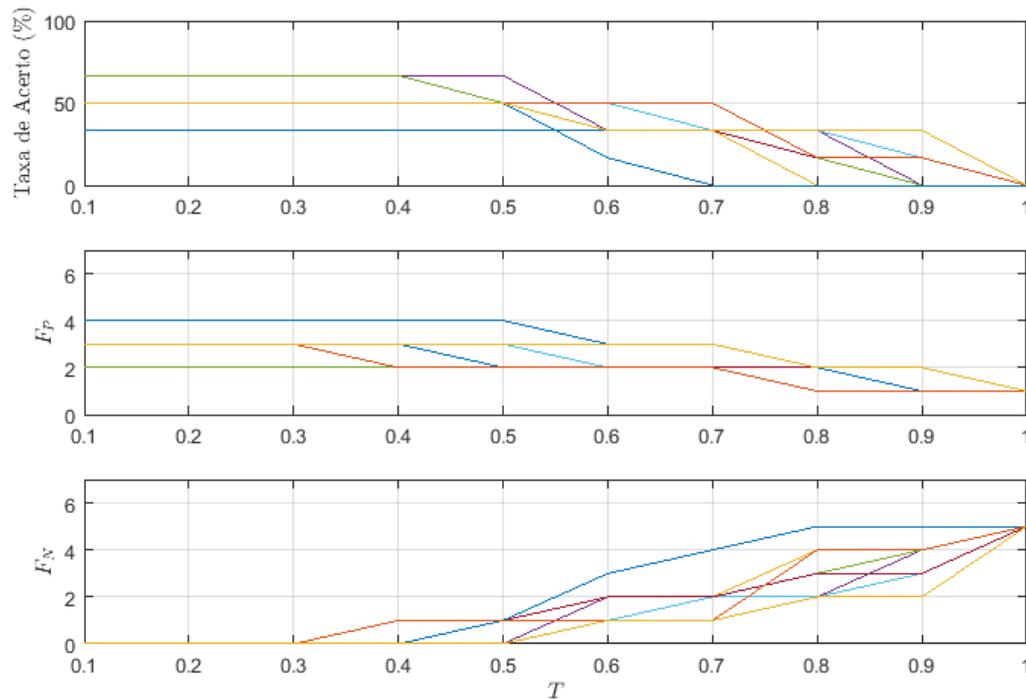


Analisando os gráficos, pode-se notar uma menor eficiência do resultado para o acorde \mathbf{C} em relação ao método anterior, visto que a taxa de acerto máxima utilizando esse método é de 50 %. Da mesma forma, atinge-se o valor de 5 falsos negativos para a

análise de algumas amostras, indicando a detecção de uma nota apenas. A análise de mais acordes pode informar mais sobre as características desse método.

Os resultados para o primeiro conjunto de gráficos para 10 amostras do acorde G são apresentados pela Figura 25.

Figura 25 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde G



Assim como no caso do acorde C , nota-se uma redução geral na taxa de acerto, que, no melhor dos casos para esse acorde, atinge a marca de 66,6667 %. A aplicação desse método de seleção resulta em menos falsos positivos possíveis por ser mais restritiva e informar apenas 6 notas detectadas, porém o impacto da restrição sobre a taxa de acerto é aparente. A partir dos dados de todos os outros gráficos gerados, pôde-se elaborar uma tabela análoga à Tabela 7 para esse caso. Os valores máximos de taxa de acerto TAM observados para amostras observadas de todos os acordes, juntamente com o valor de T até onde estes se mantêm, para o seletor por maior valor de correlação, são listados na Tabela 8.

Assim como para o caso do seletor analisado anteriormente, as taxas de acerto não demonstraram relação direta com os intervalos pelos quais os acordes são compostos. Percebe-se uma diminuição geral nos valores de taxa de acerto para esse método, o que indica que o classificador por maior valor de correlação aponta as notas corretas dos

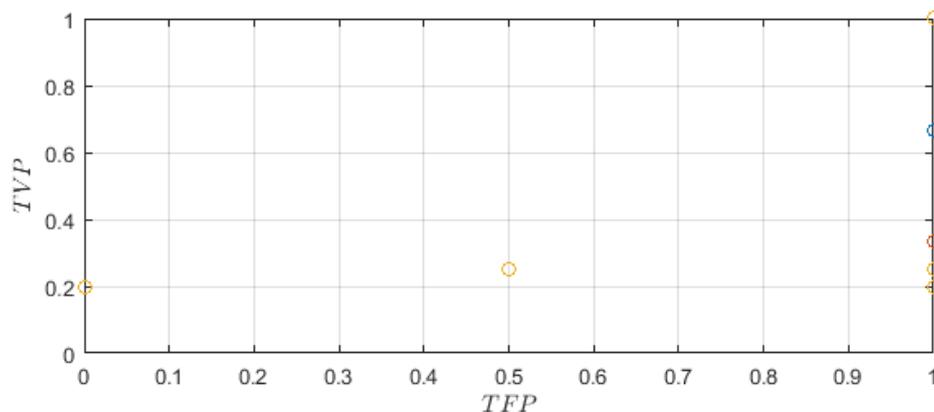
Tabela 8 – Taxas de acerto máximas para cada acorde, juntamente com o valor de T máximo por onde essa taxa de acerto máxima se manteve, sob o Método de Maior Valor de Correlação

Acorde	TAM (%)	T_{TAM}
<i>C</i>	50,0000	0,4
<i>C6</i>	50,0000	0,3
<i>C7</i>	16,6667	0,9
<i>C7M</i>	16,6667	1,0
<i>C7M(#5)</i>	50,0000	0,4
<i>Cm</i>	33,3333	0,6
<i>Cm6</i>	33,3333	0,6
<i>Cm7M</i>	16,6667	1,0
<i>C°</i>	16,6667	1,0
<i>D</i>	16,6667	0,8
<i>F</i>	0,0000	1,0
<i>F7M</i>	0,0000	1,0
<i>F#</i>	66,6667	0,3
<i>F#m7</i>	16,6667	0,9
<i>G</i>	66,6667	0,5
<i>G7</i>	66,6667	0,4
<i>G#7</i>	16,6667	0,9
<i>G#m7(b5)</i>	16,6667	0,5
<i>A</i>	16,6667	0,4
<i>A9</i>	16,6667	0,4

acordes com menos frequência. Uma análise final sobre sua performance pode ser verificada pelas curvas ROC geradas.

A curva ROC para 10 amostras do acorde ***C*** pode ser visualizada na Figura 26.

Figura 26 – Curva ROC para o acorde ***C***

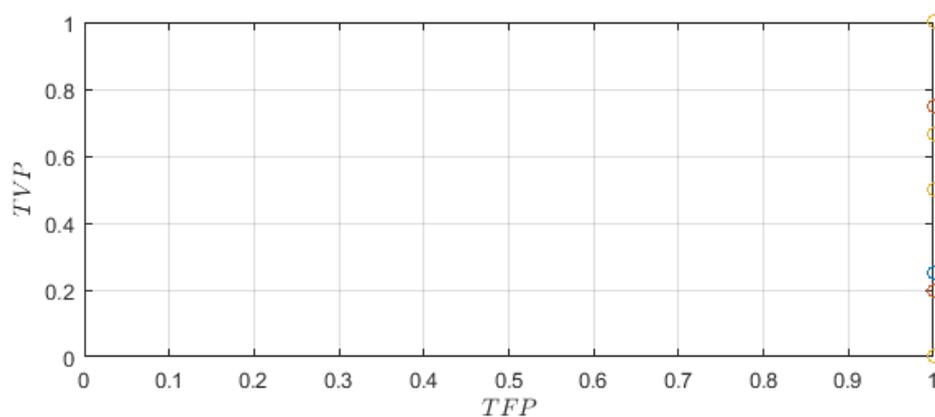


Nota-se que no resultado dessa curva ROC, o classificador se apresentou de performance muito inferior, uma vez que pontos localizados nas regiões à direita do gráfico

indicam alta taxa de falsos positivos, sendo que em poucos casos a taxa de verdadeiros positivos é alta. Adicionalmente, valores elevados da taxa de verdadeiros positivos nunca são acompanhados de valores baixos da taxa de falsos positivos.

A curva ROC para 10 amostras do acorde G pode ser visualizada na Figura 27.

Figura 27 – Curva ROC para o acorde G



Assim como na curva ROC para o acorde C , pode-se observar o mesmo fenômeno de performance inferior nesse gráfico, onde todos os pontos estão localizados em $TFP = 1$. Observando as curvas ROC para todos os acordes restantes, nota-se que o padrão de resultados cuja TFP é igual a 1 é mantido na grande maioria das vezes. Isso indica que esse classificador apresentou resultados que sempre serão acompanhados de uma alta taxa de falsos positivos, com poucos casos onde a taxa de verdadeiros positivos é elevada.

6 Conclusões

Realizando todos os procedimentos propostos e fazendo a análise dos resultados obtidos, foi possível gerar a base de dados de 36 notas compatível com os propósitos definidos. No processamento, verificou-se que a função de coeficiente de correlação é um bom indicador de similaridade para o casamento dos sinais das FFTs, resultando em valores próximos de 1 para sinais similares e em valores próximos de 0 para valores muito distintos. Isso indica que os procedimentos utilizados para a geração da base de dados podem ser reproduzidos em outras escalas e de outras formas para um tipo similar de aplicação.

Foi possível desenvolver o detector e classificador proposto, com faixa de taxa de acerto variável para os dois modos propostos em função de cada acorde analisado. A taxa de acerto para o método de Análise de Pertinência ao Acorde apresentou valores entre 0% e 100%, sendo os casos onde a taxa de acerto foi inferior a 16,6667 % a minoria destes. Já a taxa de acerto para o método de Maior Valor de Correlação apresentou valores entre 0% e 66,6667%, sendo os casos onde a taxa de acerto foi inferior a 16,6667 % a maioria destes, demonstrando uma performance reduzida. Com o auxílio das curvas ROC, foi possível determinar que o primeiro seletor possui pontos de operação de melhor performance do que o segundo seletor, embora seus propósitos finais sejam diferentes.

Como também era esperado, pôde-se observar a influência das harmônicas das notas na amplitude de outras notas, o que gera resultados elevados de correlação também para notas que não estão sendo tocadas. Nesse contexto, o algoritmo implementado para selecionar a nota mais provável entre duas notas equivalentes detectadas pôde incrementar a performance dos classificadores, evitando algumas classificações específicas de serem feitas de forma equivocada. A utilização de várias amostras tanto na base de dados quanto na base de testes foi fator essencial para obter resultados consistentes, uma vez que a análise de amostras individuais traria resultados não confiáveis.

Considerando os resultados do método de Análise de Pertinência ao Acorde, nota-se que é uma ferramenta de performance limitada para a classificação, mas que muitas vezes consegue apontar notas corretas do acorde que foram tocadas baseando-se nos valores de correlação e utilizando a função T para definir o que provavelmente foi tocado ou não. Já o método de Maior Valor de Correlação, por basear-se na mesma lógica e ser ainda mais restritivo, acaba sendo ainda menos eficiente. Com o auxílio do algoritmo de seleção de notas equivalentes detectadas, foi possível obter alguma melhora nos resultados, mas não o suficiente para alcançar valores elevados de taxa de acerto no segundo método.

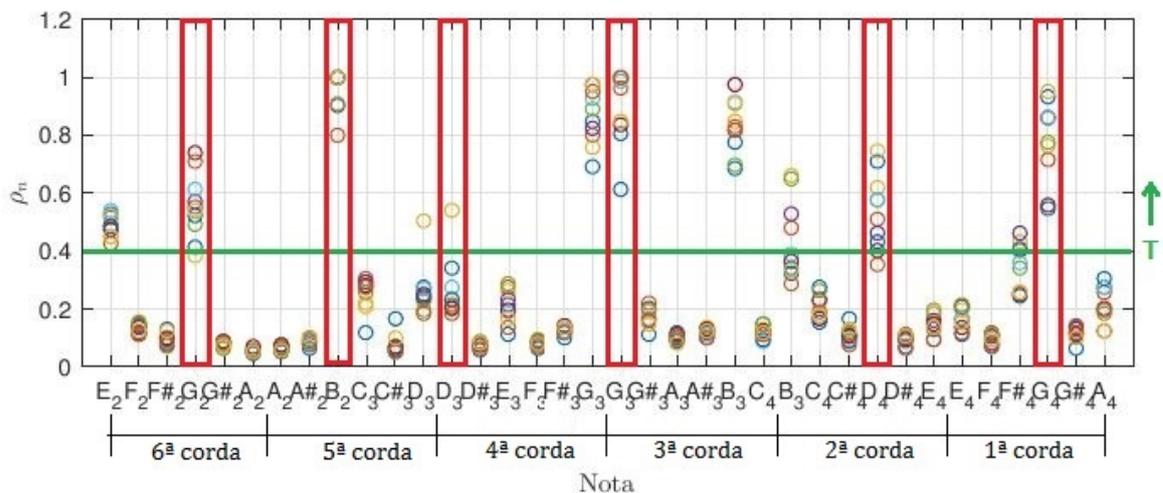
A variabilidade dos resultados indica que a aplicação da função T no vetor de correlação normalizada pode ser aprimorada, tanto com a utilização de outro tipo de função T quanto com um melhor pré-processamento da base de dados, considerando que a variação dos picos e a potência dos sinais influencia na normalização. Tais procedimentos evitariam que baixos valores de correlação ocorressem para sinais de baixa potência e poderiam apresentar melhor classificação, visto que uma função linear para T não é similar ao formato observado pelo resultado das correlações por nota.

7 Propostas para Trabalhos Futuros

Problemas apontados pela análise dos resultados incluem fatores no pré-processamento dos sinais, nos algoritmos de seleção e na função de *threshold*. Para trabalhos futuros, é válido analisar uma forma de normalizar o resultado da função coeficiente de correlação em função da potência dos sinais, para evitar o problema do resultado da correlação ser afetado por baixos picos em altas frequências e por altos picos em baixas frequências. Outra sugestão seria utilizar mais de um instrumento para gerar a base de dados, aumentando ainda mais o alcance de variabilidade dos sinais de entrada para o classificador. Também é importante analisar o classificador utilizando uma base de dados gerada a partir das 132 notas da guitarra para uma análise completa.

Essas sugestões provavelmente sejam proveitosas para melhorar a performance dos classificadores atuais. Entretanto, após analisar o sistema desenvolvido como um todo, é possível notar um dos fatores principais que prejudica os classificadores é a função de *threshold* linear adotada. Um exemplo visual da aplicação da função da forma como foi feita pode ser visto na Figura 28. A linha verde representa a função linear, onde é definido que os valores acima desta são considerados como detectados e os valores abaixo como não detectados.

Figura 28 – Representação da função T de *threshold* aplicada a todos os acordes para detecção no acorde G

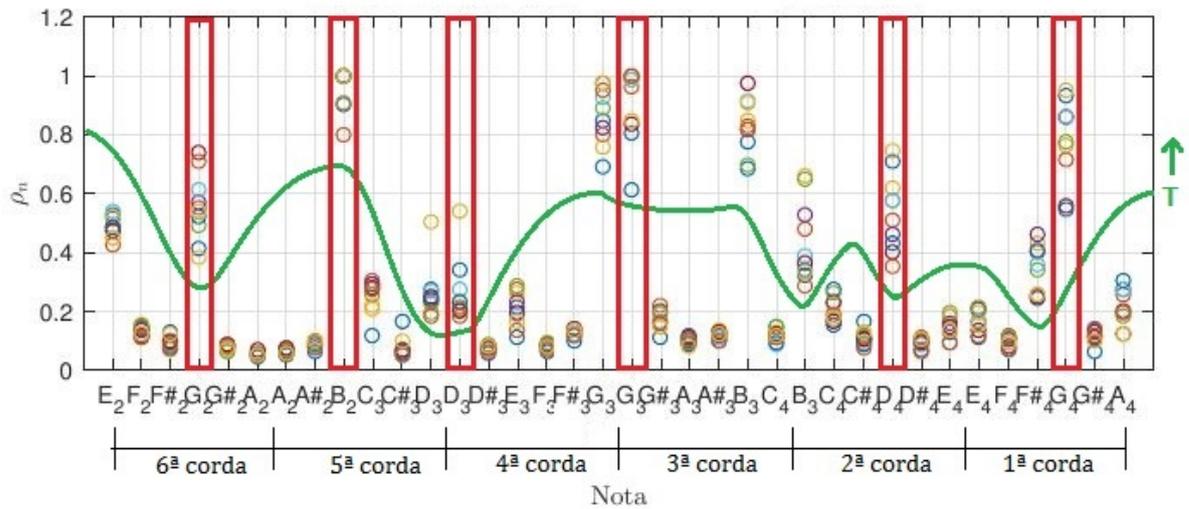


É possível ver que, se fossem utilizadas técnicas para fazer com que a função T seja adaptável à forma média dos resultados, seria possível obter um classificador mais eficiente. Uma sugestão para implementar essa ideia seria utilizar técnicas de Redes Neurais

Artificiais (ou RNAs) para cada acorde. Sendo assim, cada acorde teria sua função T associada com o formato do resultado das correlações por nota de acordo com as notas desejadas para classificação.

Um exemplo da aplicação de uma função T arbitrária não-linear com caráter adaptativo pode ser conferido na Figura 29.

Figura 29 – Representação de uma função T arbitrária de *threshold* cuja viabilização provavelmente implicaria em melhores resultados



Dessa forma, fica evidente que a classificação seria realizada com menos falsos positivos, menos falsos negativos e uma maior taxa de acerto.

Referências Bibliográficas

- LEE, K. *Automatic Chord Recognition from Audio Using Enhanced Pitch Class Profile*. ISMIR, 2006.
- LEE, K.; SLANEY, M. *Automatic Chord Recognition from Audio Using an HMM with Supervised Learning*. ISMIR, 2006.
- SHEH, A.; ELLIS, D. P. *Chord segmentation and recognition using EM-trained hidden Markov models*. ISMIR, 2003.
- YAZAWA, K.; SAKAUE, D.; NAGIRA, K.; ITOYAMA, K.; OKUNO, H. G. *Audio-based guitar tablature transcription using multipitch analysis and playability constraints*. IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, pp. 196-200, 2013.
- BARBANCHO, A. M.; KLAPURI, A.; TARDON, L. J.; BARBANCHO, I. *Automatic Transcription of Guitar Chords and Fingering From Audio* in IEEE Transactions on Audio, Speech, and Language Processing, vol. 20, no. 3, pp. 915-921, 2012.
- BURLET, G., FUJINAGA, I. *Robotaba Guitar Tablature Transcription Framework*. ISMIR. 2013.
- NATTIEZ, J. J. *Music and discourse: Toward a semiology of music..* Princeton, N.J: Princeton University Press, 1990.
- CHO, G. J. *The discovery of musical equal temperament in China and Europe in the sixteenth century*. Lewiston, NY [u.a.: Mellen], 2003.
- IAZZETTA, F. *Tabela de Frequências, Períodos e Comprimentos de Onda*. Em: Tutoriais de Áudio e Acústica, 2008. Disponível em <<http://www2.eca.usp.br/prof/iazzetta/tutor/acustica/introducao/tabela1.html>>.
- YOUNG, R. W. *Terminology for Logarithmic Frequency Units*, Journal of the Acoustical Society of America, Volume 11, Number 1, p. 134–139, Jul de 1939.
- BENWARD, B.; SAKER, M. *Music: In Theory and Practice*, Vol. I, pp. 67, 359. Seventh Edition, 2003.
- BINGHAM, R. P. *HARMONICS: Understanding the Facts*. Dranetz Technologies, 2009. Disponível em <<https://web.archive.org/web/20091229052719/http://www.dranetz-bmi.com/pdf/harmonicsall.pdf>>.
- LEONARD, H. *Picture Chord Encyclopedia: Photos & Diagrams for Over 2,600 Guitar Chords*. Hal Leonard Corporation, 2000.
- FRIGO, M.; JOHNSON S. G. *FFTW: An Adaptive Software Architecture for the FFT*. Proceedings of the International Conference on Acoustics, Speech, and Signal Processing. Vol. 3, p. 1381-1384, 1998.

- OPPENHEIM, A. V. *Discrete-time signal processing*. Upper Saddle River, N.J: Prentice Hall, 1999.
- BENDAT, J. S.; PIERSOL, A. G. *Random Data - Analysis and Measurement Procedures*. 4th ed, Wiley, 2010.
- GONZALES, R. C.; Woods, R. C. *Processamento Digital de Imagens*. 3ª edição, p. 571-575, 2010.
- HOPLEY, L. *The magnificent ROC (Receiver Operating Characteristic) curve*. 2011. Disponível em <<http://www.anaesthetist.com/mnm/stats/roc/Findex.htm>>
- FAWCETT, T. *An introduction to ROC analysis*. In: Pattern Recognition Letters, Volume 27, Issue 8, 2006, Pages 861-874, ISSN 0167-8655 Disponível em: <<https://doi.org/10.1016/j.patrec.2005.10.010>>.

Apêndices

APÊNDICE A – Códigos desenvolvidos no MATLAB

Código 1: Função de utilizada para reduzir o número de amostras dos vetores da base de dados via *downsampling*

```

%% Codigo 1 - Up/Downsampler

% Autor: Prof. Adalberto Schuck Jr.
% Alteracoes por Lucas Bairros Etcheverria de Lima

% Funcao para realizar up-sampling, down-sampling ou ambos simultaneamente
em um determinado sinal

% Expande o sinal x adicionando L-1 zeros entre as amostras,
% onde eh feita a interpolacao com um filtro FIR
% Depois, elimina N amostras entre cada amostra

function y = up_down_sampler(x,L,N)

%% Expansao do sinal (upsampling)

comp = length(x);
xe = zeros(1,L*comp);
xe(1) = x(1);

for i = 0:(comp-1),
    % Atribui valores ao novo vetor
    xe((L*i)+1) = x (i+1);
end

% Projeta o filtro FIR interpolador: Filtro simetrico do tipo Bartlet
compfilt = (2*L)-1;
b = zeros(1,compfilt);

for i = 1:L,
    b(i) = i;
end
for i = 1:L-1
    b(L+i) = b(L-i);
end
b = b/L;

```

```
% Filtra o sinal
y1 = filter(b,1,xe);

%% Etapa decimadora (downsampling)

% Projeto do filtro FIR decimador:
b = fir1(32,1/N);

% Filtragem do sinal:
aux = filter(b,1,y1);

fim = length(y1);

% Elimina as amostras
y=aux(1:N:fim);
```

Código 2: Função desenvolvida pra carregar as notas gravadas em 96 *kHz* na base de dados e retornar os parâmetros de FFT média e desvio-padrão associados

```
%% Codigo 2: Funcao para o processamento das notas da base de dados

% Autor: Lucas Bairros Etcheverria de Lima

% Carrega arquivos de audio amostrados a 96 kHz, realiza downsampling
% e determina fft media e desvio padrao da fft media

function [med_nota,des_nota] = dados_base(filename)

%% Inicializacao de variaveis

n_samples = 30;          % Numero de amostras
fs_original = 48000;    % Frequencia de amostragem a ser usada ao fim da
    rotina

sec10 = 10*fs_original; % Numero de pontos correspondentes a 10 segundos de
    sinal

nota = cell(1,n_samples);
%% Leitura dos arquivos de audio

fs_arquivos = 96000;

seg10 = fs_arquivos*10;

samples = cell(1,n_samples);
```

```
for i = 1:1:n_samples
    samples{i} = [(i-1)*seg10+1,i*seg10];
end

for k = 1:1:n_samples
    [nota{k},~] = audioread(filename,samples{k});
end

%% Downsampling do audio.

nota_down = cell(1,n_samples);

for k = 1:1:n_samples
    nota_down{k} = up_down_sampler(nota{k},1,2);
end

%% Obtendo media e desvio padrao das FFTs

% Realizando a FFT sobre os sinais.
nota_fft = cell(1,n_samples);

for j = 1:1:n_samples
    nota_fft{j} = abs(fft(nota_down{j}));
end

% Calculando a media das FFTs resultantes.
sum = 0;
med_nota = 0;

for n = 1:1:30
    sum = sum + nota_fft{n};
    if n == n_samples
        med_nota = sum./n;
    end
end

% Calculando o desvio padrao.
des_nota = std(med_nota);
```

Código 3: Função desenvolvida pra carregar as notas gravadas em 48 *kHz* na base de dados e retornar os parâmetros de FFT média e desvio padrão associados

```
%% Codigo 3: Funcao para o processamento das notas gravadas em 48kHz da
base de dados
```

```
% Autor: Lucas Bairros Etcheverria de Lima

% Carrega arquivos de audio amostrados a 48 kHz e determina fft media e
% desvio padrao da fft media

function [med_nota,des_nota] = dados_base48(filename)

%% Inicializacao de variaveis

n_samples = 30;          % Numero de amostras
fs_original = 48000;    % Frequencia de amostragem a ser usada

sec10 = 10*fs_original; % Numero de pontos correspondentes a 10 segundos de
    sinal

nota = cell(1,n_samples);
%% Leitura dos arquivos de audio

fs_arquivos = 48000;

seg10 = fs_arquivos*10;

samples = cell(1,n_samples);

for i = 1:1:n_samples
    samples{i} = [(i-1)*seg10+1,i*seg10];
end

for k = 1:1:n_samples
    [nota{k},~] = audioread(filename,samples{k});
end

%% Obtendo media e desvio padrao das FFTs

% Realizando a FFT sobre os sinais.
nota_fft = cell(1,n_samples);

for j = 1:1:n_samples
    nota_fft{j} = abs(fft(nota{j}));
end

% Calculando a media das FFTs resultantes.
sum = 0;
med_nota = 0;

for n = 1:1:30
    sum = sum + nota_fft{n};
```

```

    if n == n_samples
        med_nota = (sum./n).';
    end
end

%% Tornando a celula de vetores um vetor unico para o calculo do desvio
%% padrao.
% vet = cell2mat(nota_fft);

% Calculando o desvio padrao.
des_nota = std(med_nota);

```

Código 4: Implementação da função Coeficiente de Correlação

```

%% Codigo 4 - Funcao Coeficiente de Correlacao

% Autor: Lucas Bairros Etcheverria de Lima

% Funcao que implementa a expressao do coeficiente de correlacao a ser
% aplicada para duas entradas especificadas por suas FFTs medias e desvios
% padrao. A funcao tambem necessita da frequencia de amostragem dos sinais

function y = correlacao(m_sinal1,d_sinal1,m_sinal2,d_sinal2,fs)

% Calcula as medias dos sinais medios fornecidos
med_1 = mean(m_sinal1);
med_2 = mean(m_sinal2);

% Subtrai a media de todos os pontos dos vetores
m_sinal1 = m_sinal1 - med_1;
m_sinal2 = m_sinal2 - med_2;

% Definicao da variavel de saida
corr = 0;

% Especificacao do numero de amostras em 10 segundos de audio
sec10 = fs*10;

% Para todo 'i' entre 1 e o numero total de amostras, realizar produto
% ponto a ponto dos sinais, ja com suas medias subtraidas
for i = 1:sec10
    corr = corr + (m_sinal1(i)*m_sinal2(i));
end

% Calcula, por fim, a saida normalizada
y = corr/(d_sinal1*d_sinal2*(sec10-1));

```

Código 5: Função desenvolvida pra carregar as amostras de acordes gravadas em 48 *kHz* e retornar os parâmetros de FFT média e desvio-padrão associados

```
% Codigo 5: Funcao para o processamento das amostras de acordes

% Autor: Lucas Bairros Etcheverria de Lima

% Funcao que carrega os dados de media da fft e desvio padrao da fft das
% 10 amostras gravadas de um acorde

function [med_ac,des_ac] = dados_amostras48(filename)

%% Inicializacao de variaveis

n_samples = 10;          % Numero de amostras
fs_original = 48000;    % Frequencia de amostragem a ser usada

seg10 = 10*fs_original; % Numero de pontos correspondentes a 10 segundos de
    sinal

nota = cell(1,n_samples);
%% Leitura dos arquivos de audio

fs_arquivos = 48000;

seg10 = fs_arquivos*10;

samples = cell(1,n_samples);

for i = 1:1:n_samples
    samples{i} = [(i-1)*seg10+1,i*seg10];
end

for k = 1:1:n_samples
    [nota{k},~] = audioread(filename,samples{k});
end

%% Obtendo media e desvio padrao das FFTs

% Realizando a FFT sobre os sinais.
nota_fft = cell(1,n_samples);

for j = 1:1:n_samples
    nota_fft{j} = abs(fft(nota{j}));
end
```

```
med_ac = nota_fft;  
  
% Calculando o desvio padrao.  
  
des_ac = zeros(1,n_samples);  
  
for i = 1:n_samples  
    des_ac(i) = std(med_ac{i});  
end
```

Código 6: Função desenvolvida para realizar a correlação de um acorde de entrada com toda a base de dados, retornando os valores associados a cada nota da base e o gráfico correspondente

```
%% Codigo 6: Correlacao com a base de dados  
% Autor: Lucas Bairros Etcheverria de Lima  
  
% Funcao que carrega a base de dados por inteiro e realiza o calculo  
% do coeficiente de correlacao com um dado acorde de entrada de FFT media  
% 'sinal_m' e desvio padrao 'sinal_d'.  
  
function y = corr_total(sinal_m,sinal_d)  
%% Carregando a base de dados  
  
% Atribui nome do arquivo da base de dados a uma variavel 'filename'  
filename = 'base_notas_2.mat';  
  
% Cria objeto 'matfile' para referenciar as variaveis de dentro do arquivo  
% de extensao '.mat'  
m = matfile(filename);  
  
% Carregando FFTs medias da base de dados  
  
E26med = m.E26med;  
F26med = m.F26med;  
Fs26med = m.Fs26med;  
G26med = m.G26med;  
Gs26med = m.Gs26med;  
A26med = m.A26med;  
  
A25med = m.A25med;  
As25med = m.As25med;  
B25med = m.B25med;  
C35med = m.C35med;  
Cs35med = m.Cs35med;  
D35med = m.D35med;
```

```
D34med = m.D34med;  
Ds34med = m.Ds34med;  
E34med = m.E34med;  
F34med = m.F34med;  
Fs34med = m.Fs34med;  
G34med = m.G34med;
```

```
G33med = m.G33med;  
Gs33med = m.Gs33med;  
A33med = m.A33med;  
As33med = m.As33med;  
B33med = m.B33med;  
C43med = m.C43med;
```

```
B32med = m.B32med;  
C42med = m.C42med;  
Cs42med = m.Cs42med;  
D42med = m.D42med;  
Ds42med = m.Ds42med;  
E42med = m.E42med;
```

```
E41med = m.E41med;  
F41med = m.F41med;  
Fs41med = m.Fs41med;  
G41med = m.G41med;  
Gs41med = m.Gs41med;  
A41med = m.A41med;
```

```
% Carregando desvios padrao
```

```
E26des = m.E26des;  
F26des = m.F26des;  
Fs26des = m.Fs26des;  
G26des = m.G26des;  
Gs26des = m.Gs26des;  
A26des = m.A26des;
```

```
A25des = m.A25des;  
As25des = m.As25des;  
B25des = m.B25des;  
C35des = m.C35des;  
Cs35des = m.Cs35des;  
D35des = m.D35des;
```

```
D34des = m.D34des;  
Ds34des = m.Ds34des;
```

```
E34des = m.E34des;
F34des = m.F34des;
Fs34des = m.Fs34des;
G34des = m.G34des;

G33des = m.G33des;
Gs33des = m.Gs33des;
A33des = m.A33des;
As33des = m.As33des;
B33des = m.B33des;
C43des = m.C43des;

B32des = m.B32des;
C42des = m.C42des;
Cs42des = m.Cs42des;
D42des = m.D42des;
Ds42des = m.Ds42des;
E42des = m.E42des;

E41des = m.E41des;
F41des = m.F41des;
Fs41des = m.Fs41des;
G41des = m.G41des;
Gs41des = m.Gs41des;
A41des = m.A41des;

% Armazenando resultados em vetor de coeficientes de correlacao

% Define frequencia de amostragem para o calculo do coef de correlacao
fs = 48000;

% Cria vetor para armazenar resultados
mat_corr = zeros(1,36);

mat_corr(1) = correlacao(sinal_m,sinal_d,E26med,E26des,fs);
mat_corr(2) = correlacao(sinal_m,sinal_d,F26med,F26des,fs);
mat_corr(3) = correlacao(sinal_m,sinal_d,Fs26med,Fs26des,fs);
mat_corr(4) = correlacao(sinal_m,sinal_d,G26med,G26des,fs);
mat_corr(5) = correlacao(sinal_m,sinal_d,Gs26med,Gs26des,fs);
mat_corr(6) = correlacao(sinal_m,sinal_d,A26med,A26des,fs);

mat_corr(7) = correlacao(sinal_m,sinal_d,A25med,A25des,fs);
mat_corr(8) = correlacao(sinal_m,sinal_d,As25med,As25des,fs);
mat_corr(9) = correlacao(sinal_m,sinal_d,B25med,B25des,fs);
mat_corr(10) = correlacao(sinal_m,sinal_d,C35med,C35des,fs);
mat_corr(11) = correlacao(sinal_m,sinal_d,Cs35med,Cs35des,fs);
mat_corr(12) = correlacao(sinal_m,sinal_d,D35med,D35des,fs);
```

```
mat_corr(13) = correlacao(sinal_m, sinal_d, D34med, D34des, fs);
mat_corr(14) = correlacao(sinal_m, sinal_d, Ds34med, Ds34des, fs);
mat_corr(15) = correlacao(sinal_m, sinal_d, E34med, E34des, fs);
mat_corr(16) = correlacao(sinal_m, sinal_d, F34med, F34des, fs);
mat_corr(17) = correlacao(sinal_m, sinal_d, Fs34med, Fs34des, fs);
mat_corr(18) = correlacao(sinal_m, sinal_d, G34med, G34des, fs);

mat_corr(19) = correlacao(sinal_m, sinal_d, G33med, G33des, fs);
mat_corr(20) = correlacao(sinal_m, sinal_d, Gs33med, Gs33des, fs);
mat_corr(21) = correlacao(sinal_m, sinal_d, A33med, A33des, fs);
mat_corr(22) = correlacao(sinal_m, sinal_d, As33med, As33des, fs);
mat_corr(23) = correlacao(sinal_m, sinal_d, B33med, B33des, fs);
mat_corr(24) = correlacao(sinal_m, sinal_d, C43med, C43des, fs);

mat_corr(25) = correlacao(sinal_m, sinal_d, B32med, B32des, fs);
mat_corr(26) = correlacao(sinal_m, sinal_d, C42med, C42des, fs);
mat_corr(27) = correlacao(sinal_m, sinal_d, Cs42med, Cs42des, fs);
mat_corr(28) = correlacao(sinal_m, sinal_d, D42med, D42des, fs);
mat_corr(29) = correlacao(sinal_m, sinal_d, Ds42med, Ds42des, fs);
mat_corr(30) = correlacao(sinal_m, sinal_d, E42med, E42des, fs);

mat_corr(31) = correlacao(sinal_m, sinal_d, E41med, E41des, fs);
mat_corr(32) = correlacao(sinal_m, sinal_d, F41med, F41des, fs);
mat_corr(33) = correlacao(sinal_m, sinal_d, Fs41med, Fs41des, fs);
mat_corr(34) = correlacao(sinal_m, sinal_d, G41med, G41des, fs);
mat_corr(35) = correlacao(sinal_m, sinal_d, Gs41med, Gs41des, fs);
mat_corr(36) = correlacao(sinal_m, sinal_d, A41med, A41des, fs);

y = mat_corr;

% Normaliza resultados pelo pico
y = y./max(y);

figure(1);

% Variaveis do tamanho da janela para o plot
x0=10;
y0=10;
width=580;
height=200;

%Plota os valores dos coeficientes de correlacao normalizados
plot(y, 'o');
grid on;
hold on;
ylabel('$\rho_n$', 'Interpreter', 'latex');
```

```

xlabel('Nota','Interpreter','latex');
xlim([0 37]);
ylim([0 1.2]);
set(gca, 'XTick', [1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
    23 24 25 26 27 28 29 30 31 32 33 34 35 36]);
set(gca, 'XTickLabel', {'E_2', 'F_2', 'F#_2', 'G_2', 'G#_2', 'A_2', 'A_2', 'A#_2', '
    B_2', 'C_3', 'C#_3', 'D_3', 'D_3', 'D#_3', 'E_3', 'F_3', 'F#_3', 'G_3', 'G_3', 'G#
    _3', 'A_3', 'A#_3', 'B_3', 'C_4', 'B_3', 'C_4', 'C#_4', 'D_4', 'D#_4', 'E_4', 'E_4'
    , 'F_4', 'F#_4', 'G_4', 'G#_4', 'A_4'});
set(gcf, 'units', 'points', 'position', [x0,y0,width,height])
end

```

Código 7: Rotina para a geração da Figura 15, exemplificando um caso ideal para os resultados do coeficiente de correlação para o acorde C .

```

% Codigo 7: Plot do exemplo ideal de resultado para o acorde C

% Autor: Lucas Bairros Etcheverria de Lima

% Plota o resultado ideal para um caso de correlacao com um acorde

% Nova figura
figure(1);

% Define o vetor de coeficientes de correlacao unitarios
y = [0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0
    0];

% Variaveis do tamanho da janela para o plot
x0=10;
y0=10;
width=580;
height=200;

% Plota os valores dos coeficientes de correlacao em funcao da nota
plot(y, 'o');
grid on;
hold on;
title('','Interpreter','latex');
ylabel('Correla\c{c}\~ao','Interpreter','latex');
xlabel('Nota','Interpreter','latex');
xlim([0 37]);
ylim([0 1.2]);
set(gca, 'XTick', [1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
    23 24 25 26 27 28 29 30 31 32 33 34 35 36]);

```

```

set(gca,'XTickLabel',{'E_2','F_2','F#_2','G_2','G#_2','A_2','A_2','A#_2','
    B_2','C_3','C#_3','D_3','D_3','D#_3','E_3','F_3','F#_3','G_3','G_3','G#
    _3','A_3','A#_3','B_3','C_4','B_3','C_4','C#_4','D_4','D#_4','E_4','E_4'
    ,'F_4','F#_4','G_4','G#_4','A_4'});
set(gcf,'units','points','position',[x0,y0,width,height])

```

Código 8: Função desenvolvida com a lógica do seletor de Análise de Pertinência ao Acorde, cuja saída fornece um *plot* da taxa de acerto, dos falsos positivos e dos falsos negativos em função do nível de *threshold*, e da curva ROC.

```

%% Codigo 8: Funcao de Analise de Pertinencia ao Acorde
% Autor: Lucas Bairros Etcheverria de Lima

% Funcao que compara um dado acorde teorico com sua respectiva amostra para
% informar se as notas detectadas sao as notas do acorde em funcao do valor
% do threshold. Fornece como saida a taxa de acerto, falsos positivos,
% falsos negativos e a curva ROC.

function taxas_acerto_ap(acr,med_ac,des_ac)

% Invertendo a ordem das notas informadas no vetor para um pensamento mais
% intuitivo
ac = fliplr(acr);

% Vetor para definir os limiaries de aceitacao
limiar = 0.1:0.1:1;

% String contendo os nomes de cada nota analisada
not_nomes = {'E','F','F#','G','G#','A','A','A#','B','C','C#','D','D','D#','
    E','F','F#','G','G','G#','A','A#','B','C','B','C','C#','D','D#','E','E',
    'F','F#','G','G#','A'};

toca = zeros(1,36);

% Chama a funcao corr_total para obter os valores de correlacao de cada
% nota da base de dados com a amostra fornecida
cor_notas = corr_total(med_ac,des_ac);

% Matriz de comparacao
comp = zeros(1,36);

nome_notast = cell(10,36);

% Deteccao de notas tocadas a partir do nivel de threshold
for i = 1:36
    for j = 1:10

```

```
    if cor_notas(i) < limiar(j)
        comp(j,i) = 0;
        nome_notast{j,i} = '-';
    else if cor_notas(i) >= limiar(j)
        comp(j,i) = 1;
        nome_notast{j,i} = not_nomes{i};
    end
end
end
end

cord = zeros(1,36);

% Definicao das variaveis de resultado
v_pos = zeros(1,10);
v_neg = zeros(1,10);
f_pos = zeros(1,10);
f_neg = zeros(1,10);

% Laco iterativo de comparacoes corda a corda das notas do acorde teste com
% as notas detectadas
for m = 1:10
    for i = 1:6
        for j = 1:36
            if j <= 6 && i == 6
                toca(j) = strcmp(ac(i),not_nomes{j});
                cord(j) = strcmp(ac(i),nome_notast{m,j});
                if strcmp(ac(i),nome_notast{m,j}) == 0 && strcmp(
                    nome_notast{m,j}, '-') == 0
                    f_pos(m) = f_pos(m) + 1;
                end
                if toca(j) == 1 && strcmp(nome_notast{m,j}, '-') == 1
                    f_neg(m) = f_neg(m) + 1;
                end
                if cord(j) == 1 && strcmp(nome_notast{m,j}, '-') == 0
                    v_pos(m) = v_pos(m) + 1;
                end
                if toca(j) == 0 && strcmp(nome_notast{m,j}, '-') == 1
                    v_neg(m) = v_neg(m) + 1;
                end
            end
        end
        if j > 6 && j <=12 && i == 5
            toca(j) = strcmp(ac(i),not_nomes{j});
            cord(j) = strcmp(ac(i),nome_notast{m,j});
            if cord(j) == 1 && strcmp(nome_notast{m,j}, '-') == 0
                v_pos(m) = v_pos(m) + 1;
            end
        end
    end
end
```

```
    if toca(j) == 0 && strcmp(nome_notast{m,j}, '-') == 1
        v_neg(m) = v_neg(m) + 1;
    end
    if toca(j) == 1 && strcmp(nome_notast{m,j}, '-') == 1
        f_neg(m) = f_neg(m) + 1;
    end
    if strcmp(nome_notast{m,j}, ac(i)) == 0 && strcmp(
        nome_notast{m,j}, '-') == 0
        f_pos(m) = f_pos(m) + 1;
    end
end
end
if j > 12 && j <= 18 && i == 4
    toca(j) = strcmp(ac(i), not_nomes{j});
    cord(j) = strcmp(ac(i), nome_notast{m,j});
    if cord(j) == 1 && strcmp(nome_notast{m,j}, '-') == 0
        v_pos(m) = v_pos(m) + 1;
    end
    if toca(j) == 0 && strcmp(nome_notast{m,j}, '-') == 1
        v_neg(m) = v_neg(m) + 1;
    end
    if toca(j) == 1 && strcmp(nome_notast{m,j}, '-') == 1
        f_neg(m) = f_neg(m) + 1;
    end
    if strcmp(nome_notast{m,j}, ac(i)) == 0 && strcmp(
        nome_notast{m,j}, '-') == 0
        f_pos(m) = f_pos(m) + 1;
    end
end
end
if j > 18 && j <= 24 && i == 3
    toca(j) = strcmp(ac(i), not_nomes{j});
    cord(j) = strcmp(ac(i), nome_notast{m,j});
    if cord(j) == 1 && strcmp(nome_notast{m,j}, '-') == 0
        v_pos(m) = v_pos(m) + 1;
    end
    if toca(j) == 0 && strcmp(nome_notast{m,j}, '-') == 1
        v_neg(m) = v_neg(m) + 1;
    end
    if toca(j) == 1 && strcmp(nome_notast{m,j}, '-') == 1
        f_neg(m) = f_neg(m) + 1;
    end
    if strcmp(nome_notast{m,j}, ac(i)) == 0 && strcmp(
        nome_notast{m,j}, '-') == 0
        f_pos(m) = f_pos(m) + 1;
    end
end
end
if j > 24 && j <= 30 && i == 2
    toca(j) = strcmp(ac(i), not_nomes{j});
```

```

        cord(j) = strcmp(ac(i), nome_notast{m,j});
        if cord(j) == 1 && strcmp(nome_notast{m,j}, '-') == 0
            v_pos(m) = v_pos(m) + 1;
        end
        if toca(j) == 0 && strcmp(nome_notast{m,j}, '-') == 1
            v_neg(m) = v_neg(m) + 1;
        end
        if toca(j) == 1 && strcmp(nome_notast{m,j}, '-') == 1
            f_neg(m) = f_neg(m) + 1;
        end
        if strcmp(nome_notast{m,j}, ac(i)) == 0 && strcmp(
            nome_notast{m,j}, '-') == 0
            f_pos(m) = f_pos(m) + 1;
        end
    end
    if j > 30 && j <= 36 && i == 1
        toca(j) = strcmp(ac(i), not_nomes{j});
        cord(j) = strcmp(ac(i), nome_notast{m,j});
        if cord(j) == 1 && strcmp(nome_notast{m,j}, '-') == 0
            v_pos(m) = v_pos(m) + 1;
        end
        if toca(j) == 0 && strcmp(nome_notast{m,j}, '-') == 1
            v_neg(m) = v_neg(m) + 1;
        end
        if toca(j) == 1 && strcmp(nome_notast{m,j}, '-') == 1
            f_neg(m) = f_neg(m) + 1;
        end
        if strcmp(nome_notast{m,j}, ac(i)) == 0 && strcmp(
            nome_notast{m,j}, '-') == 0
            f_pos(m) = f_pos(m) + 1;
        end
    end
end
end
end

% Definindo variavel para taxa de acerto
taxa = zeros(1,10);
tvp = zeros(1,10);
tfp = zeros(1,10);

% Computo da taxa de acerto
for n = 1:10
    taxa(n) = (v_pos(n)+v_neg(n))/6*100;
    tvp(n) = v_pos(n)/(v_pos(n)+f_neg(n));
    tfp(n) = f_pos(n)/(v_neg(n)+f_pos(n));
end
end

```

```
x0 = 10;
y0 = 10;
width = 600;
height = 380;
width2 = 500;
height2 = 200;

% Plotando os graficos da taxa de acerto, falsos positivos e falsos
% negativos

figure(2);
subplot(3,1,1);
plot(limiar,taxa);
ylabel('Taxa de Acerto (\%)', 'Interpreter', 'latex');
ylim([0 100]);
hold on;
grid on;
subplot(3,1,2);
plot(limiar,f_pos);
ylabel('$F_P$', 'Interpreter', 'latex');
ylim([0 37]);
hold on;
grid on;
subplot(3,1,3);
plot(limiar,f_neg);
ylabel('$F_N$', 'Interpreter', 'latex');
ylim([0 7]);
hold on;
grid on;
xlabel('$T$', 'Interpreter', 'latex');
set(gcf, 'units', 'points', 'position', [x0,y0,width,height])

% Plotando o grafico da caracteristica de operacao do receptor

figure(3);
plot(tfp,tvp, 'o');
xlim([0 1]);
ylim([0 1]);
xlabel('$TFP$', 'Interpreter', 'latex');
ylabel('$TVP$', 'Interpreter', 'latex');
grid on;
hold on;
set(gcf, 'units', 'points', 'position', [x0,y0,width2,height2])
end
```

Código 9: Função desenvolvida para retornar as 6 notas com os maiores valores de correlação determinados pela função *corr_total.m* na análise de um acorde. Essa função também tenta contornar o problema das notas equivalentes em cordas diferentes.

```
% Codigo 9: Notas de Maior Correlacao
% Autor: Lucas Bairros Etcheverria de Lima

% Funcao que aponta as notas de maior valor de correlacao, tentando
% contornar o problema de duas notas iguais detectadas em cordas diferentes

function Nmaior = max_cor_notas(corr,lim)

% Definindo vetores para armazenar os nomes (strings)
% e os valores numericos das notas de maior correlacao

Nmaior = cell(1,6);
Cmaior = zeros(1,6);

% Aplicando o threshold linear para deteccao

for i = 1:36
    if corr(i) < lim
        corr(i) = 0;
    end
end

% Separando o vetor de valores de correlacao em blocos correspondentes a
% cada corda da guitarra

C6 = corr(1:6);
C5 = corr(7:12);
C4 = corr(13:18);
C3 = corr(19:24);
C2 = corr(25:30);
C1 = corr(31:36);

% Define o nome de cada nota da corda
corda6 = {'E','F','F#','G','G#','A'};
% Se foi detectada...
if max(C6) > 0
    % Nmaior recebe a nota de maior valor de correlacao para essa corda
    [Cmaior(6),maior6] = max(C6);
    Nmaior{6} = corda6{maior6};
end
% Se nao foi detectada
if max(C6) == 0
```

```
% Recebe o código '-'
Nmaior{6} = '-';
end

% Mesmo procedimento aplicado nas outras cordas

corda5 = {'A', 'A#', 'B', 'C', 'C#', 'D'};
if max(C5) > 0
    [Cmaior(5), maior5] = max(C5);
    Nmaior{5} = corda5{maior5};
end
if max(C5) == 0
    Nmaior{5} = '-';
end

corda4 = {'D', 'D#', 'E', 'F', 'F#', 'G'};
if max(C4) > 0
    [Cmaior(4), maior4] = max(C4);
    Nmaior{4} = corda4{maior4};
end
if max(C4) == 0
    Nmaior{4} = '-';
end

corda3 = {'G', 'G#', 'A', 'A#', 'B', 'C'};
if max(C3) > 0
    [Cmaior(3), maior3] = max(C3);
    Nmaior{3} = corda3{maior3};
end
if max(C3) == 0
    Nmaior{3} = '-';
end

corda2 = {'B', 'C', 'C#', 'D', 'D#', 'E'};
if max(C2) > 0
    [Cmaior(2), maior2] = max(C2);
    Nmaior{2} = corda2{maior2};
end
if max(C2) == 0
    Nmaior{2} = '-';
end

corda1 = {'E', 'F', 'F#', 'G', 'G#', 'A'};
if max(C1) > 0
    [Cmaior(1), maior1] = max(C1);
    Nmaior{1} = corda1{maior1};
end
```

```
if max(C1) == 0
    Nmaior{1} = '-';
end

% Fazendo a escolha para casos em que as mesmas notas sao tocadas

% Definindo vetores temporarios
C6t = C6;
C5t = C5;
C4t = C4;
C3t = C3;
C2t = C2;
C1t = C1;

% Se duas notas em cordas subseqüentes foram iguais
if Nmaior{6} == 'A' & Nmaior{5} == 'A'
    C6t(maior6) = NaN;
    C5t(maior5) = NaN;
    % Compara os segundos maiores valores de correlacao das cordas
    % e o que for maior sera o novo maior valor daquela corda
    if max(C6t) > max(C5t)
        [Cmaior(6),maior6] = max(C6t);
        Nmaior{6} = corda6{maior6};
    end
    if max(C6t) < max(C5t)
        [Cmaior(5),maior5] = max(C5t);
        Nmaior{5} = corda5{maior5};
    end
end
end

% Redefine o vetor temporario para a proxima comparacao
C5t = C5;

% Mesmo procedimento aplicado nas outras cordas

if strcmp(Nmaior{5}, 'D') == 1 && strcmp(Nmaior{4}, 'D') == 1
    C5t(maior5) = NaN;
    C4t(maior4) = NaN;
    if max(C5t) > max(C4t)
        [Cmaior(5),maior5] = max(C5t);
        Nmaior{5} = corda5{maior5};
    end
    if max(C5t) < max(C4t)
        [Cmaior(4),maior4] = max(C4t);
        Nmaior{4} = corda4{maior4};
    end
end
end
```

```
C4t = C4;

if strcmp(Nmaior{4}, 'G') == 1 && strcmp(Nmaior{3}, 'G') == 1
    C4t(maior4) = NaN;
    C3t(maior3) = NaN;
    if max(C4t) > max(C3t)
        [Cmaior(4), maior4] = max(C4t);
        Nmaior{4} = corda4{maior4};
    end
    if max(C4t) < max(C3t)
        [Cmaior(3), maior3] = max(C3t);
        Nmaior{3} = corda3{maior3};
    end
end
end

C3t = C3;

if strcmp(Nmaior{3}, 'B') == 1 && strcmp(Nmaior{2}, 'B') == 1
    C3t(maior3) = NaN;
    C2t(maior2) = NaN;
    if max(C3t) > max(C2t)
        [Cmaior(3), maior3] = max(C3t);
        Nmaior{3} = corda3{maior3};
    end
    if max(C3t) < max(C2t)
        [Cmaior(2), maior2] = max(C2t);
        Nmaior{2} = corda2{maior2};
    end
end

end

C2t = C2;

if strcmp(Nmaior{2}, 'E') == 1 && strcmp(Nmaior{1}, 'E') == 1
    C2t(maior2) = NaN;
    C1t(maior1) = NaN;
    if max(C2t) > max(C1t)
        [Cmaior(2), maior2] = max(C2t);
        Nmaior{2} = corda2{maior2};
    end
    if max(C2t) < max(C1t)
        [Cmaior(1), maior1] = max(C1t);
        Nmaior{1} = corda1{maior1};
    end
end
end
```

end

Código 10: Função desenvolvida com a lógica do seletor de maior valor de correlação, cuja saída fornece um *plot* da taxa de acerto, dos falsos positivos e dos falsos negativos em função do nível de *threshold*, e da curva ROC.

```

%% Codigo 10: Funcao de Analise pelo Maior Valor de Correlacao
% Autor: Lucas Bairros Etcheverria de Lima

% Funcao que compara um dado acorde teorico com sua respectiva amostra para
% informar se as 6 maiores notas detectadas sao as notas reais do acorde em
% funcao do threshold. Fornece como saida a taxa de acerto, falsos
% positivos, falsos negativos e a curva ROC.

function taxas_acerto_mv(acr,med_ac,des_ac)

% Invertendo a ordem das notas informadas no vetor para um pensamento mais
% intuitivo
ac = fliplr(acr);

% Vetor para definir os limiares de aceitacao
limiar = 0.1:0.1:1;

% Chama a funcao corr_total para obter os valores de correlacao de cada
% nota da base de dados com a amostra fornecida
cor_notas = corr_total(med_ac,des_ac);

% Definicao das variaveis de resultado
v_pos = zeros(1,10);
v_neg = zeros(1,10);
f_pos = zeros(1,10);
f_neg = zeros(1,10);

% Laco iterativo de comparacoes corda a corda das notas do acorde teste com
% as notas detectadas
for m = 1:10
    m_notas = max_cor_notas(cor_notas,limiar(m));
    m_notas_c = fliplr(m_notas);
    for k = 1:6
        if strcmp(ac(k),m_notas_c(k)) == 1 && strcmp(m_notas_c{k},'-') == 0
            v_pos(m) = v_pos(m) + 1;
        end
        if strcmp(ac(k),m_notas_c(k)) == 1 && strcmp(m_notas_c{k},'-') == 1
            v_neg(m) = v_neg(m) + 1;
        end
        if strcmp(ac(k),m_notas_c(k)) == 0 && strcmp(m_notas_c{k},'-') == 0

```

```

        f_pos(m) = f_pos(m) + 1;
    end
    if strcmp(ac(k),m_notas_c(k)) == 0 && strcmp(m_notas_c{k}, '-') == 1
        f_neg(m) = f_neg(m) + 1;
    end
end
end

% Definindo variavel para taxa de acerto
taxa = zeros(1,10);
tvp = zeros(1,10);
tfp = zeros(1,10);

% Computo da taxa de acerto
for n = 1:10
    taxa(n) = (v_pos(n)+v_neg(n))/6*100;
    tvp(n) = v_pos(n)/(v_pos(n)+f_neg(n));
    tfp(n) = f_pos(n)/(v_neg(n)+f_pos(n));
end

x0 = 10;
y0 = 10;
width = 600;
height = 380;
width2 = 500;
height2 = 200;

% Plotando os graficos da taxa de acerto, falsos positivos e falsos
% negativos

figure(1);
subplot(3,1,1);
plot(limiar,taxa);
ylabel('Taxa de Acerto (\%)', 'Interpreter', 'latex');
ylim([0 100]);
hold on;
grid on;
subplot(3,1,2);
plot(limiar,f_pos);
ylabel('$F_P$', 'Interpreter', 'latex');
ylim([0 7]);
hold on;
grid on;
subplot(3,1,3);
plot(limiar,f_neg);
ylabel('$F_N$', 'Interpreter', 'latex');
ylim([0 7]);

```

```

hold on;
grid on;
xlabel('$T$', 'Interpreter', 'latex');
set(gcf, 'units', 'points', 'position', [x0,y0,width,height])

% Plotando o grafico da caracteristica de operacao do receptor

figure(1);
plot(tfp, tvp, 'o');
xlabel('$TFP$', 'Interpreter', 'latex');
ylabel('$TVP$', 'Interpreter', 'latex');
xlim([0 1]);
ylim([0 1]);
grid on;
hold on;
set(gcf, 'units', 'points', 'position', [x0,y0,width2,height2])

end

```

Código 11: Rotina para determinação de variáveis a serem utilizadas como gabarito de verificação dos acertos pela função **taxas_acerto**

```

%% Codigo 11: Definicao das strings com notas dos acordes
% Autor: Lucas Bairros Etcheverria de Lima

% Definicao de strings contendo notas respectivas de cada acorde
% das amostras geradas para teste

Ac_A      = {'x', 'A', 'E', 'A', 'C#', 'E'};
Ac_A9     = {'x', 'A', 'E', 'B', 'C#', 'E'};
Ac_As     = {'x', 'A#', 'F', 'A#', 'D', 'F'};
Ac_Asm    = {'x', 'A#', 'F', 'A#', 'C#', 'F'};
Ac_Bd     = {'x', 'B', 'F', 'G#', 'D', 'x'};
Ac_C      = {'x', 'C', 'E', 'G', 'C', 'E'};
Ac_C6     = {'x', 'C', 'E', 'A', 'C', 'E'};
Ac_C7     = {'x', 'C', 'G', 'A#', 'E', 'G'};
Ac_C7M    = {'x', 'C', 'G', 'B', 'E', 'G'};
Ac_C7Ms5  = {'x', 'C', 'E', 'G#', 'B', 'E'};
Ac_Cm     = {'x', 'C', 'G', 'C', 'D#', 'G'};
Ac_Cm6    = {'x', 'C', 'D#', 'G#', 'C', 'G'};
Ac_Cm7M   = {'x', 'C', 'G', 'B', 'D#', 'G'};
Ac_Cd     = {'x', 'C', 'F#', 'A', 'D#', 'x'};
Ac_D      = {'x', 'x', 'D', 'A', 'D', 'F#'};
Ac_Dm     = {'x', 'x', 'D', 'A', 'D', 'F'};
Ac_E      = {'E', 'B', 'E', 'G#', 'B', 'E'};
Ac_Em     = {'E', 'B', 'E', 'G', 'B', 'E'};

```

```

Ac_F      = {'F', 'C', 'F', 'A', 'C', 'F'};
Ac_F7M    = {'F', 'x', 'E', 'A', 'C', 'x'};
Ac_Fs     = {'F#', 'C#', 'F#', 'A#', 'C#', 'F#'};
Ac_Fsm7   = {'F#', 'x', 'E', 'A', 'C#', 'x'};
Ac_G      = {'G', 'B', 'D', 'G', 'D', 'G'};
Ac_G7     = {'G', 'D', 'F', 'A', 'D', 'G'};
Ac_Gs7    = {'G#', 'x', 'F#', 'C', 'D#', 'x'};
Ac_Gsm7b5 = {'G#', 'x', 'F#', 'A', 'D', 'x'};

```

Código 12: Rotina para carregamento dos parâmetros de todos os acordes de teste e para aplicação das funções **corr_total** e **taxas_acerto** a todas as amostras de cada acorde (feita de forma manual, mudando os nomes das variáveis e executando o respectivo bloco de cada função separadamente para gerar os gráficos

```

%% Codigo 12: Rotina de carregamento das amostras e geracao de resultados
% Autor: Lucas Bairros Etcheverria de Lima

```

```

% Rotina que carrega todas as 10 amostras de cada acorde especificado e
% permite a geracao de resultados pela execucao das funcoes de analise
% definidas.

```

```

[amed, ades] = dados_amostras48('A.wav');
[a9med, a9des] = dados_amostras48('A9.wav');
%[asmed, asdes] = dados_amostras48('A#.wav');
%[asmmmed, asmdes] = dados_amostras48('A#m.wav');
%[bdmed, bddes] = dados_amostras48('Bd.wav');
[cmed, cdes] = dados_amostras48('C.wav');
[c6med, c6des] = dados_amostras48('C6.wav');
[c7med, c7des] = dados_amostras48('C7.wav');
[c7mmed, c7mdes] = dados_amostras48('C7M.wav');
[c7ms5med, c7ms5des] = dados_amostras48('C7M(#5).wav');
[cmmed, cmdes] = dados_amostras48('Cm.wav');
[cm6med, cm6des] = dados_amostras48('Cm6.wav');
[cm7mmed, cm7mdes] = dados_amostras48('Cm7M.wav');
[cdmed, cddes] = dados_amostras48('Cd.wav');
[dmed, ddes] = dados_amostras48('D.wav');
%[dmmed, dmdes] = dados_amostras48('Dm.wav');
%[emed, edes] = dados_amostras48('E.wav');
%[emmed, emdes] = dados_amostras48('Em.wav');
[fmed, fdes] = dados_amostras48('F.wav');
[f7mmed, f7mdes] = dados_amostras48('F7M.wav');
[fsmed, fsdes] = dados_amostras48('F#-001.wav');
[fsm7med, fsm7des] = dados_amostras48('F#m7.wav');
[gmed, gdes] = dados_amostras48('G.wav');
[g7med, g7des] = dados_amostras48('G7.wav');
[gs7med, gs7des] = dados_amostras48('G#7.wav');

```

```
[gsm7b5med, gsm7b5des] = dados_amostras48('G#m7b5.wav');

%% Secao de geracao dos resultados de correlacao por nota

close all;

figure(1)

for i = 1:10
    corr_total(cmed{i}, cdes(i));
end

%% Secao de geracao dos resultados da analise de pertinencia ao acorde

close all;

for i = 1:10
    taxas_acerto_ap(Ac_C, cmed{i}, cdes(i));
end

%% Secao de geracao dos resultados da analise por maior valor de correlacao

close all;

for i = 1:10
    taxas_acerto_mv(Ac_C, cmed{i}, cdes(i));
end
```

APÊNDICE B – Resultados de Correlação por Nota

Figura A1 – Valores dos coeficientes de correlação para o acorde $C6$

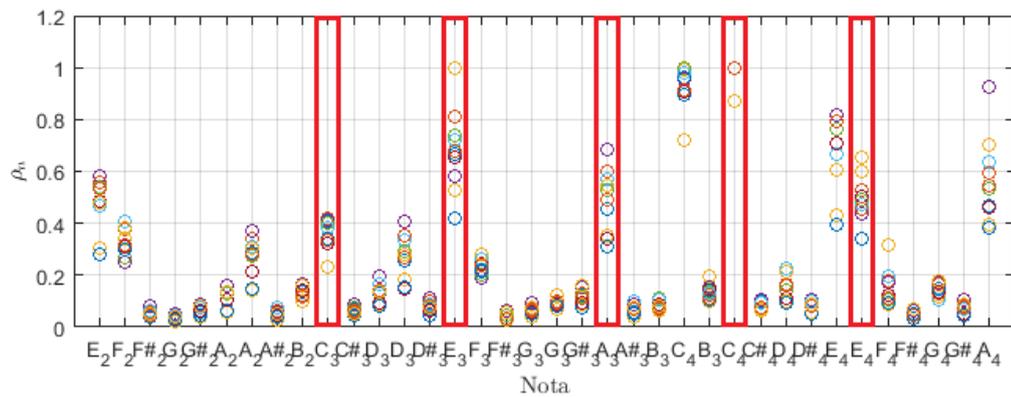


Figura A2 – Valores dos coeficientes de correlação para o acorde $C7$

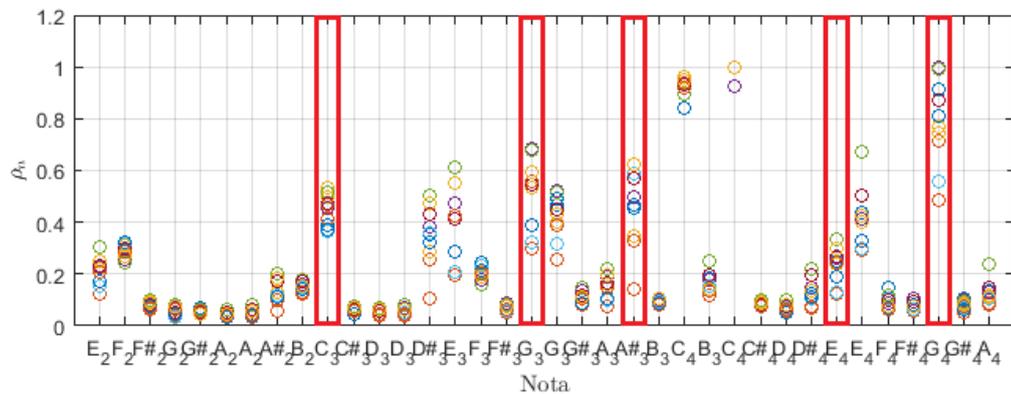


Figura A3 – Valores dos coeficientes de correlação para o acorde $C7M$

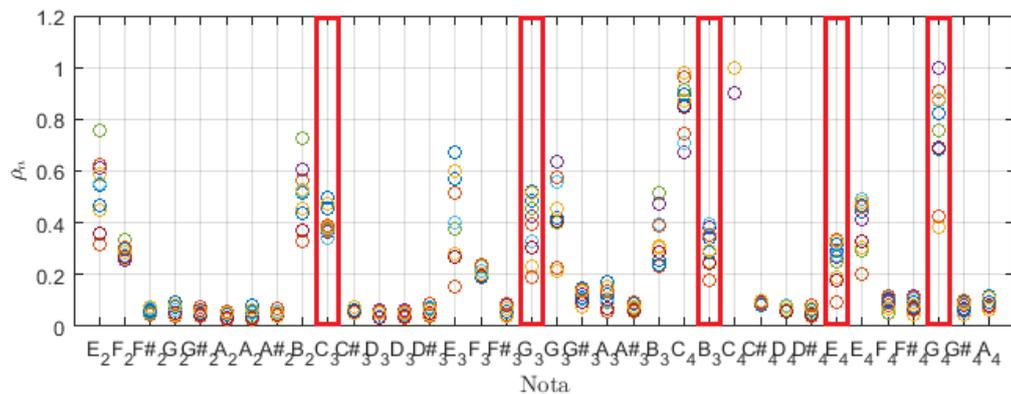


Figura A4 – Valores dos coeficientes de correlação para o acorde $C7M(\#5)$

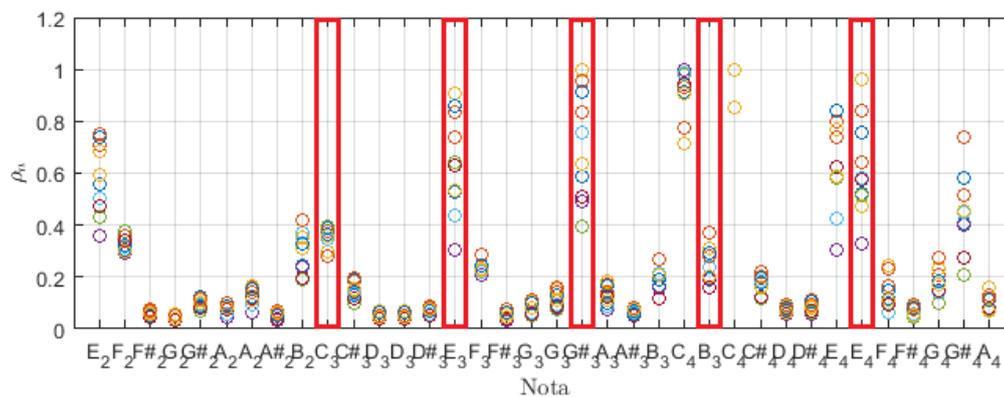


Figura A5 – Valores dos coeficientes de correlação para o acorde C°

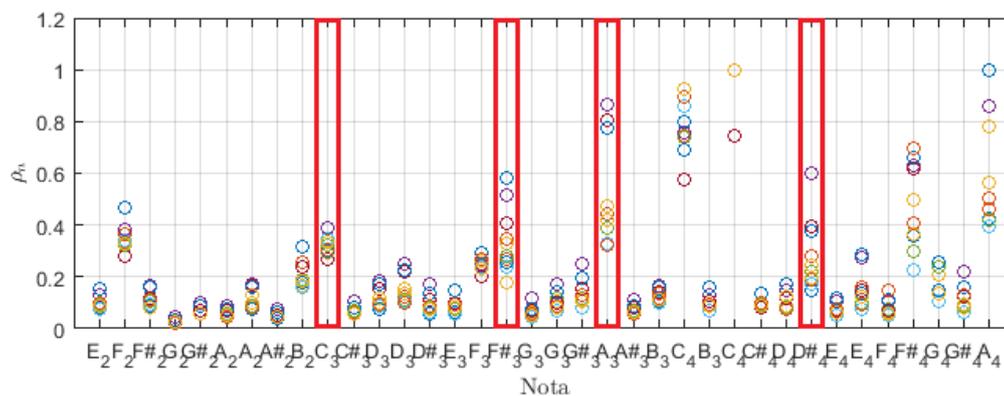


Figura A6 – Valores dos coeficientes de correlação para o acorde Cm

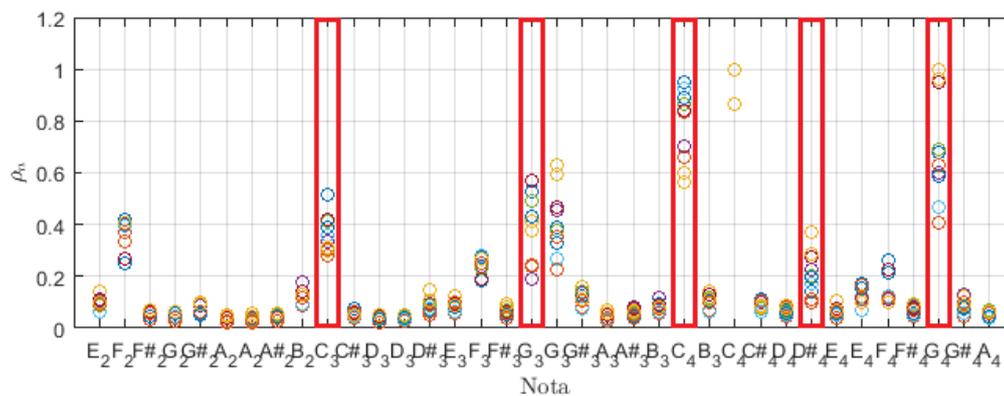


Figura A7 – Valores dos coeficientes de correlação para o acorde $Cm6$

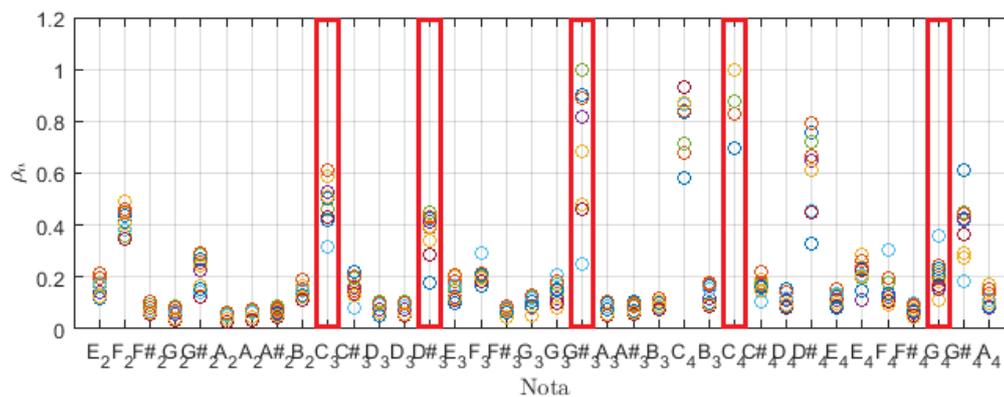


Figura A8 – Valores dos coeficientes de correlação para o acorde $Cm7M$

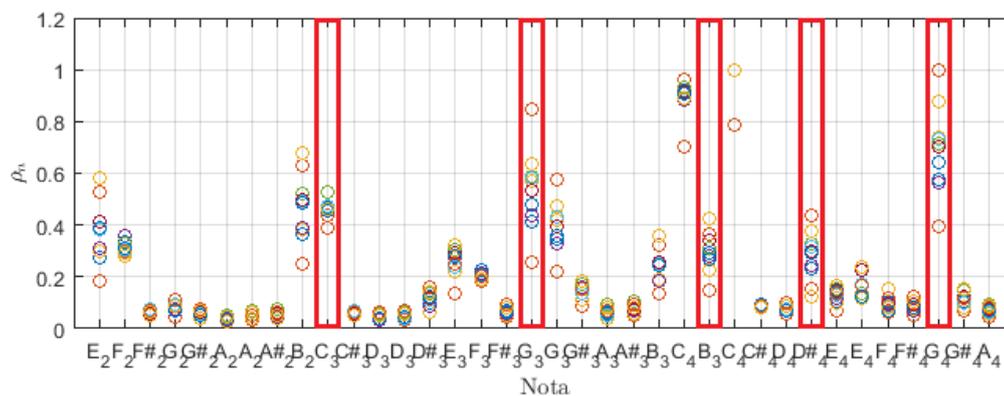


Figura A9 – Valores dos coeficientes de correlação para o acorde D

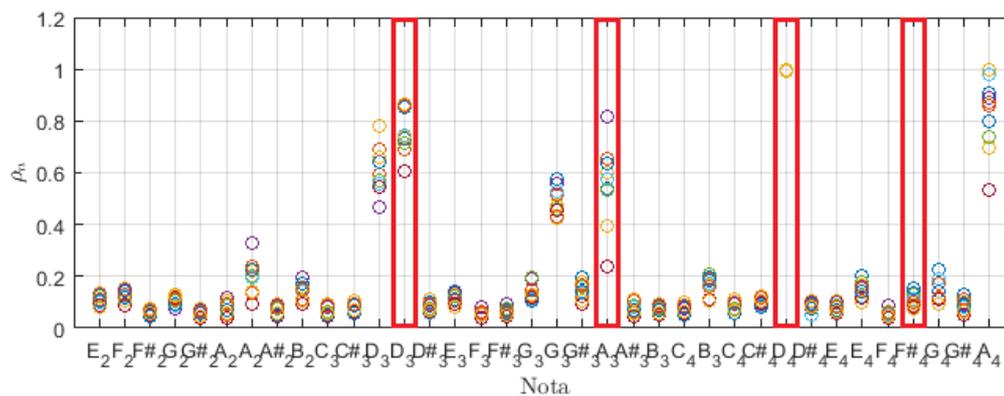


Figura A10 – Valores dos coeficientes de correlação para o acorde F

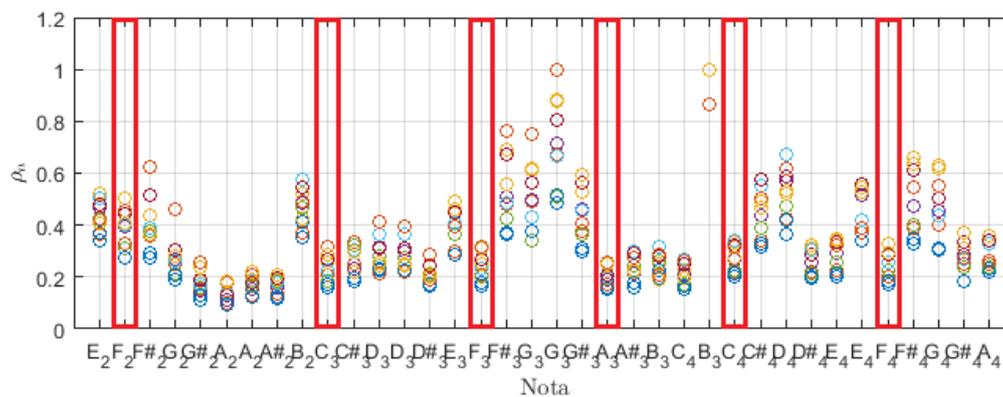


Figura A11 – Valores dos coeficientes de correlação para o acorde $F7M$

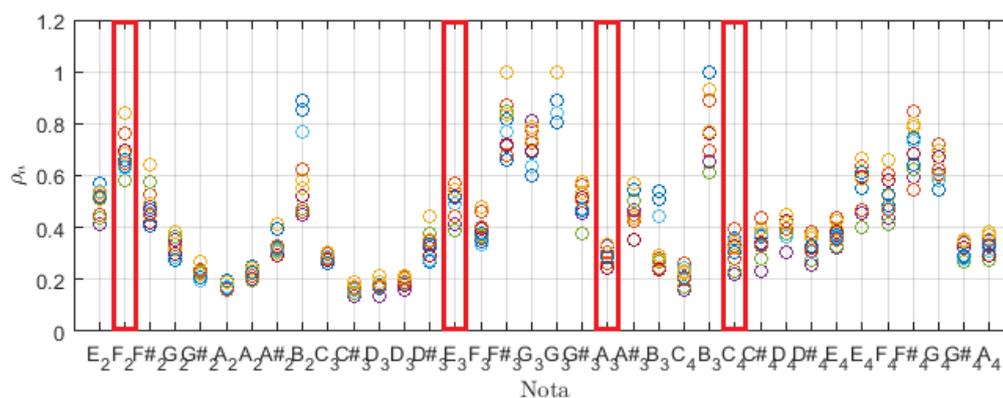


Figura A12 – Valores dos coeficientes de correlação para o acorde $F\#$

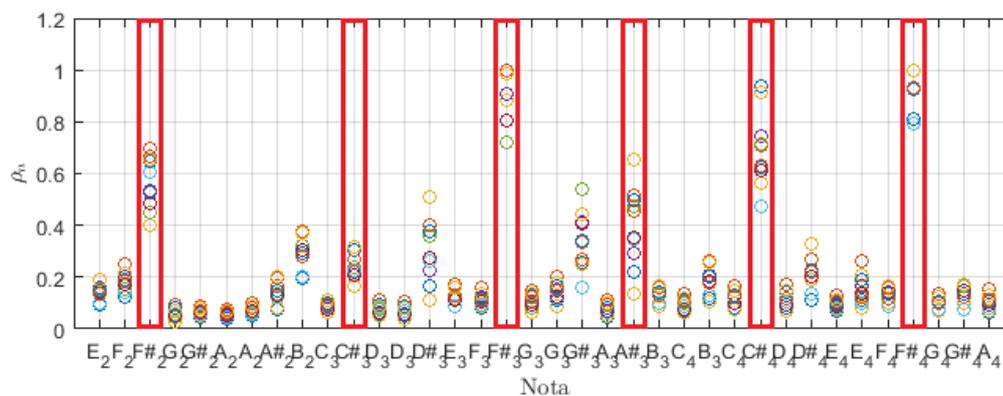


Figura A13 – Valores dos coeficientes de correlação para o acorde $F\#m7$

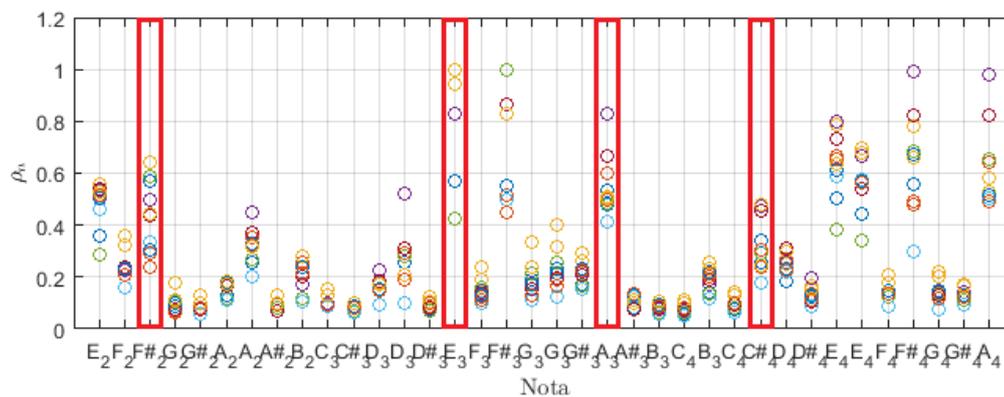


Figura A14 – Valores dos coeficientes de correlação para o acorde $G7$

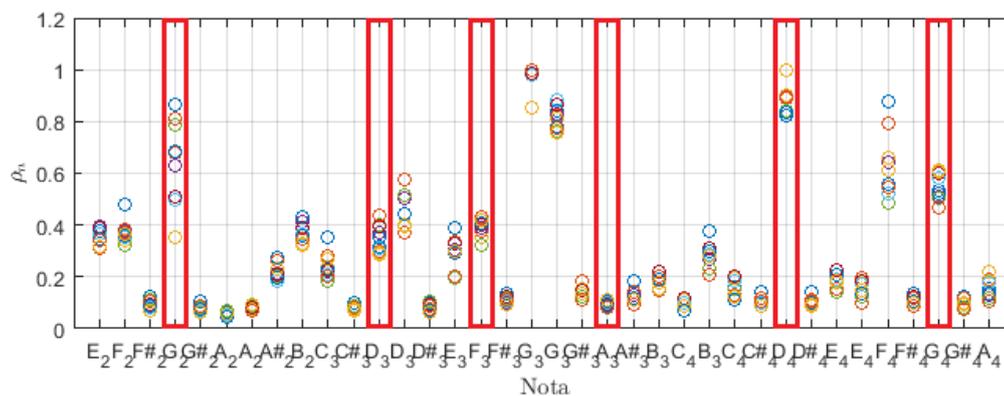


Figura A15 – Valores dos coeficientes de correlação para o acorde $G\#7$

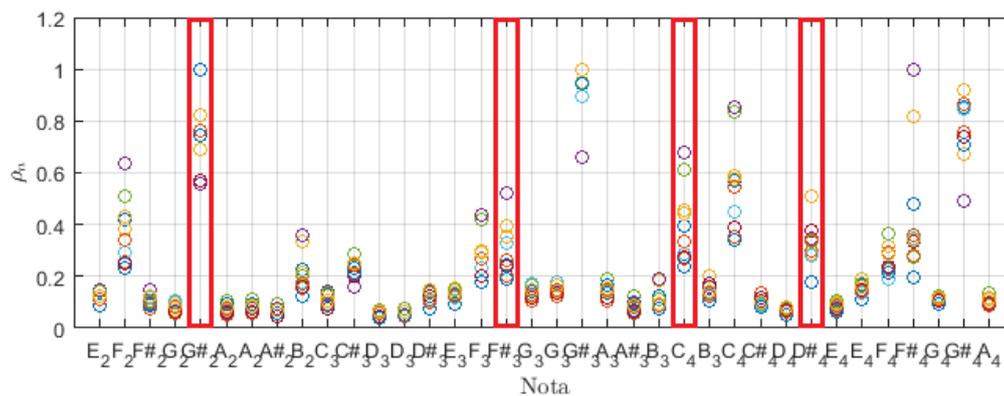


Figura A16 – Valores dos coeficientes de correlação para o acorde $G\#m7(b5)$

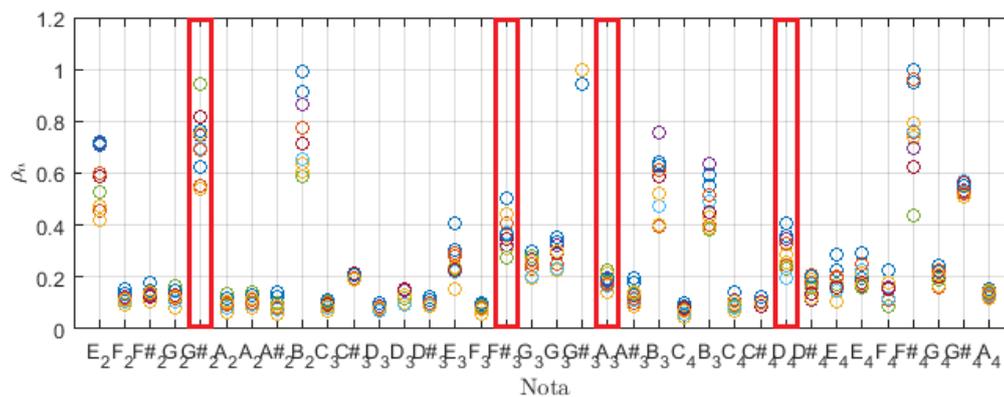


Figura A17 – Valores dos coeficientes de correlação para o acorde A

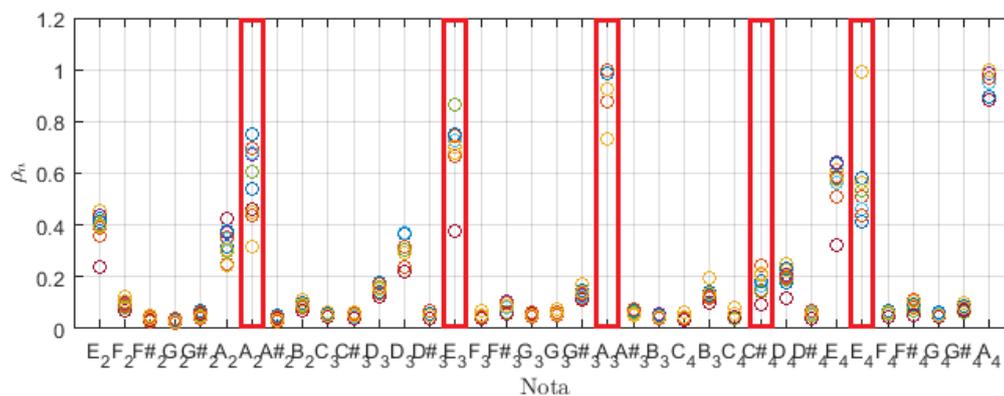
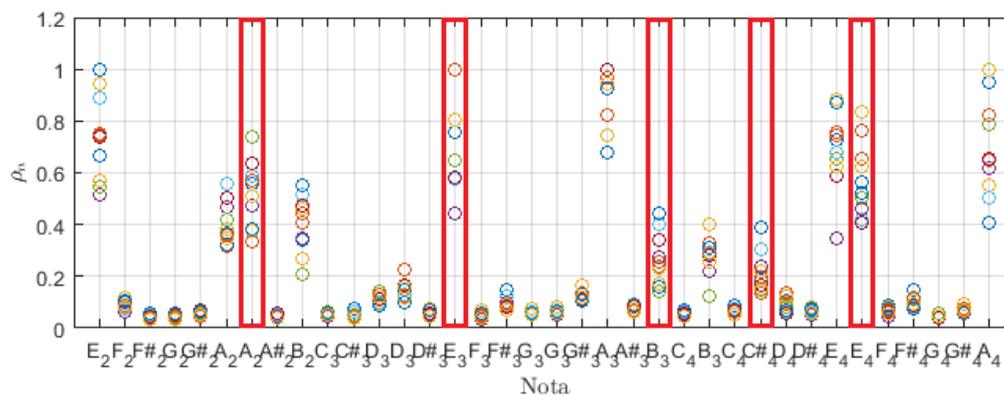


Figura A18 – Valores dos coeficientes de correlação para o acorde $A9$



APÊNDICE C – Resultados pela Análise de Pertinência ao Acorde

Figura A19 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde $C6$

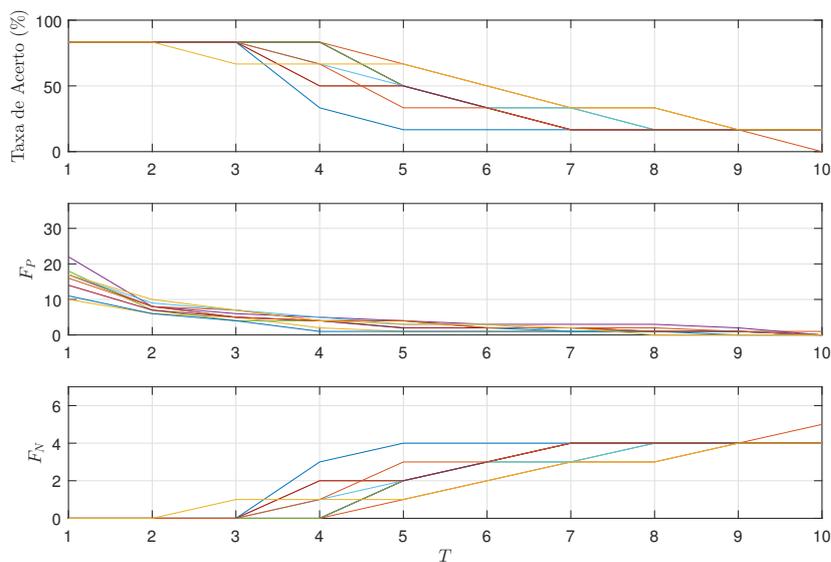


Figura A20 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde $C7$

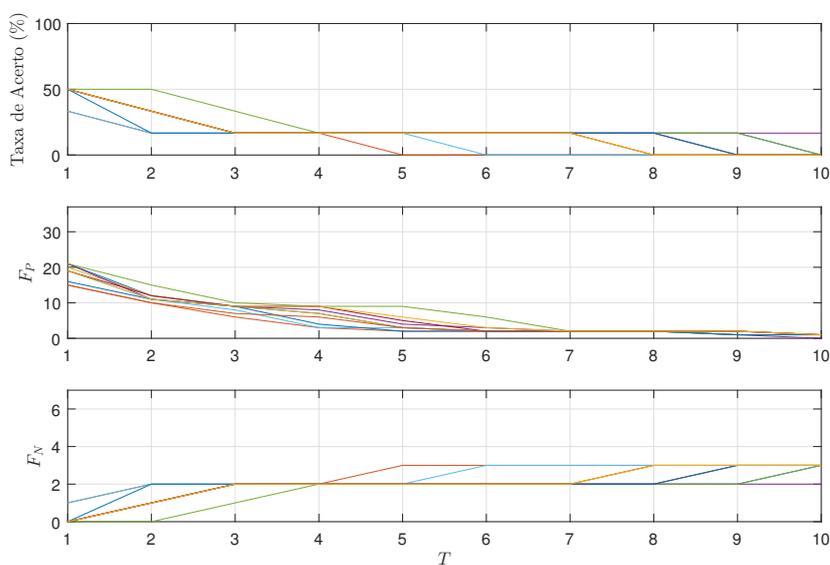


Figura A21 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde C7M

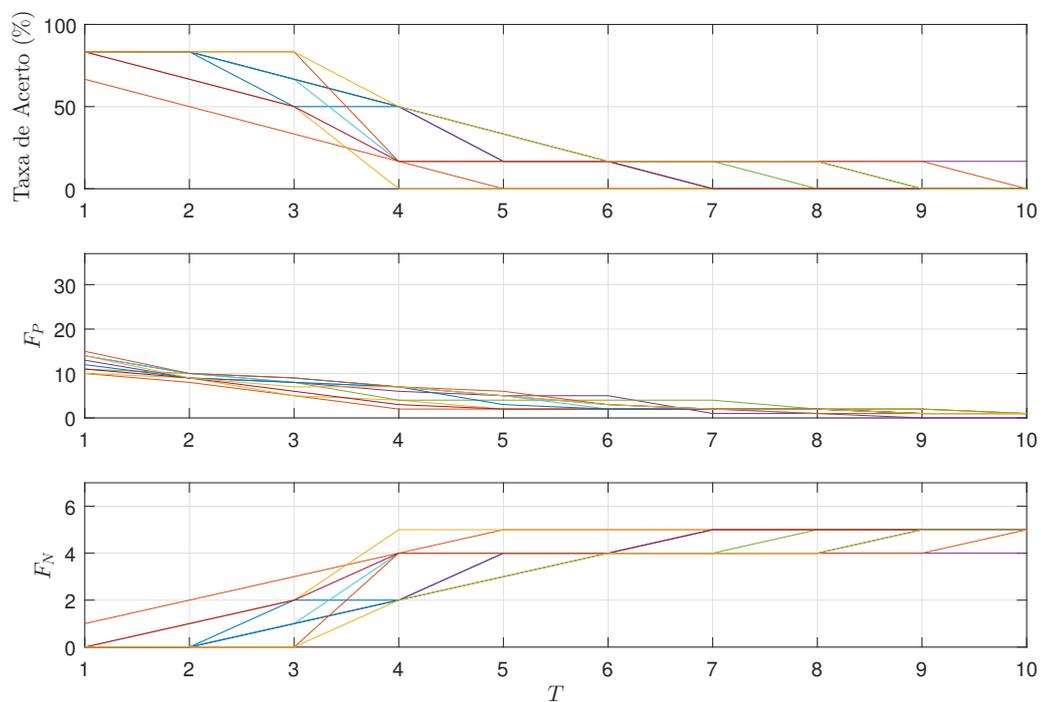


Figura A22 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde C7M(#5)

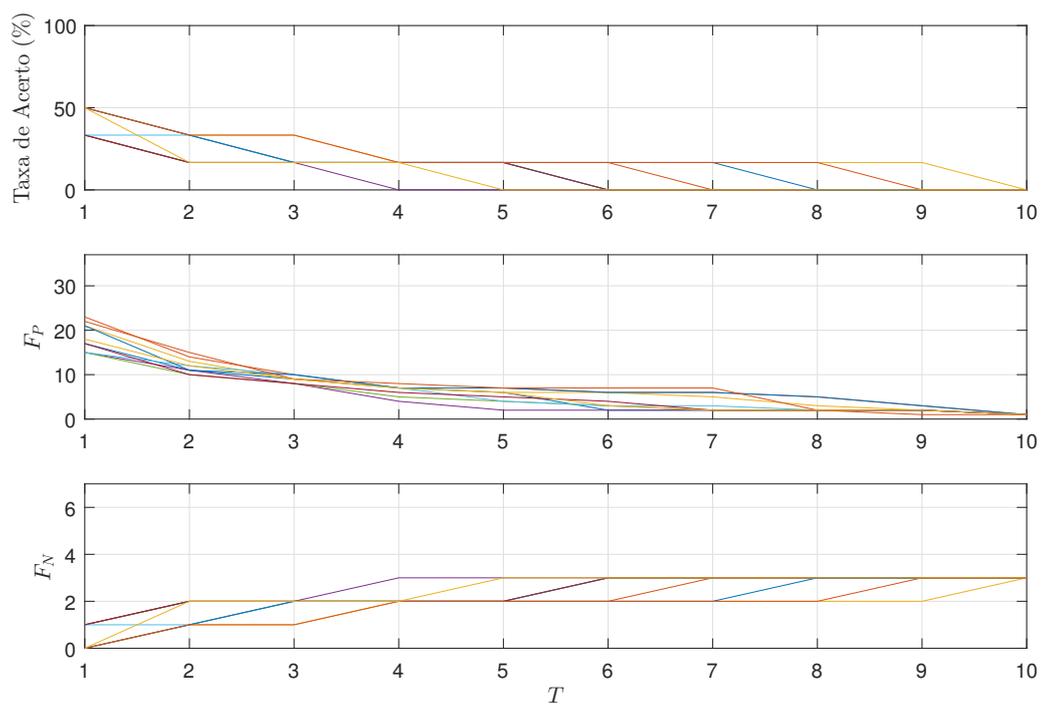


Figura A23 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde Cm

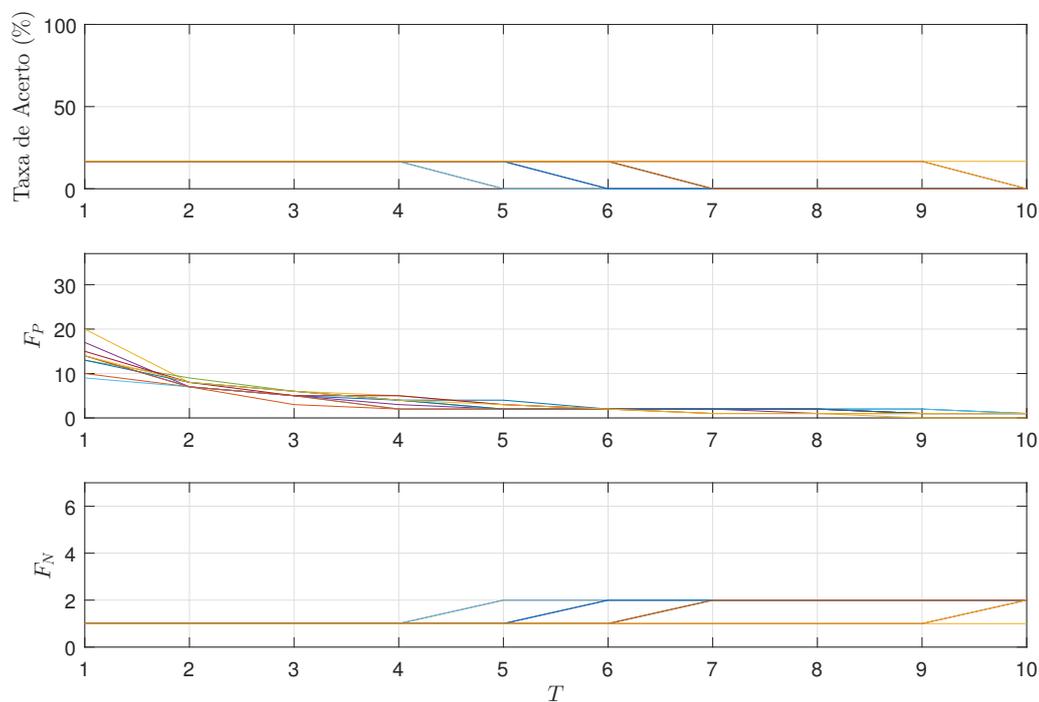


Figura A24 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde Cm6

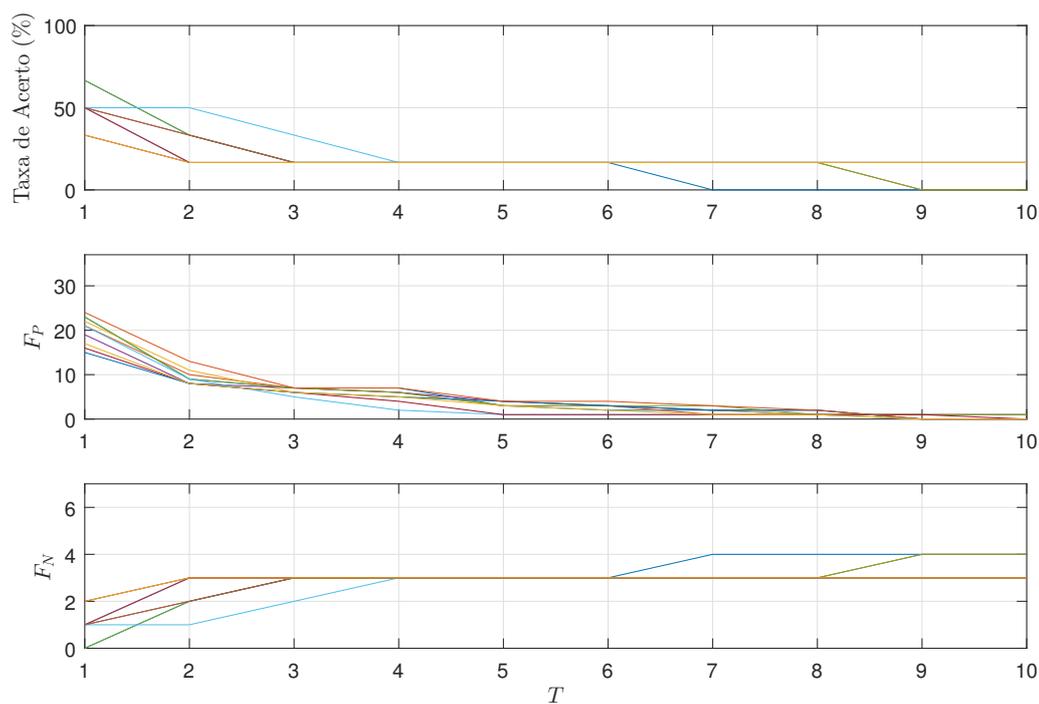


Figura A25 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde $Cm7M$

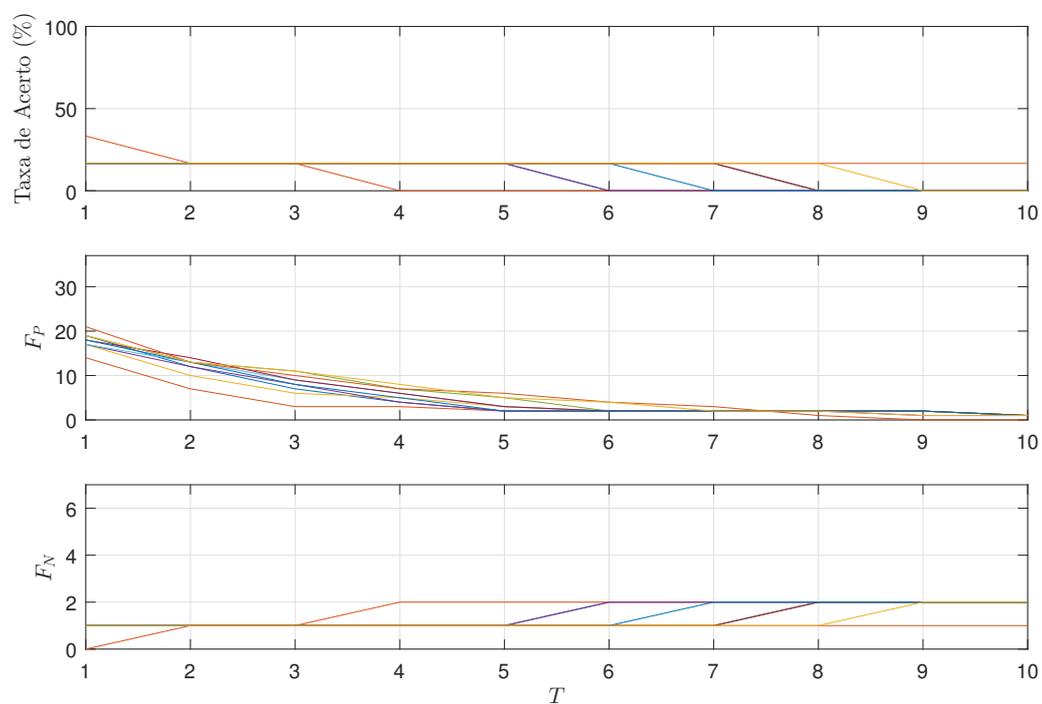


Figura A26 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde C^o

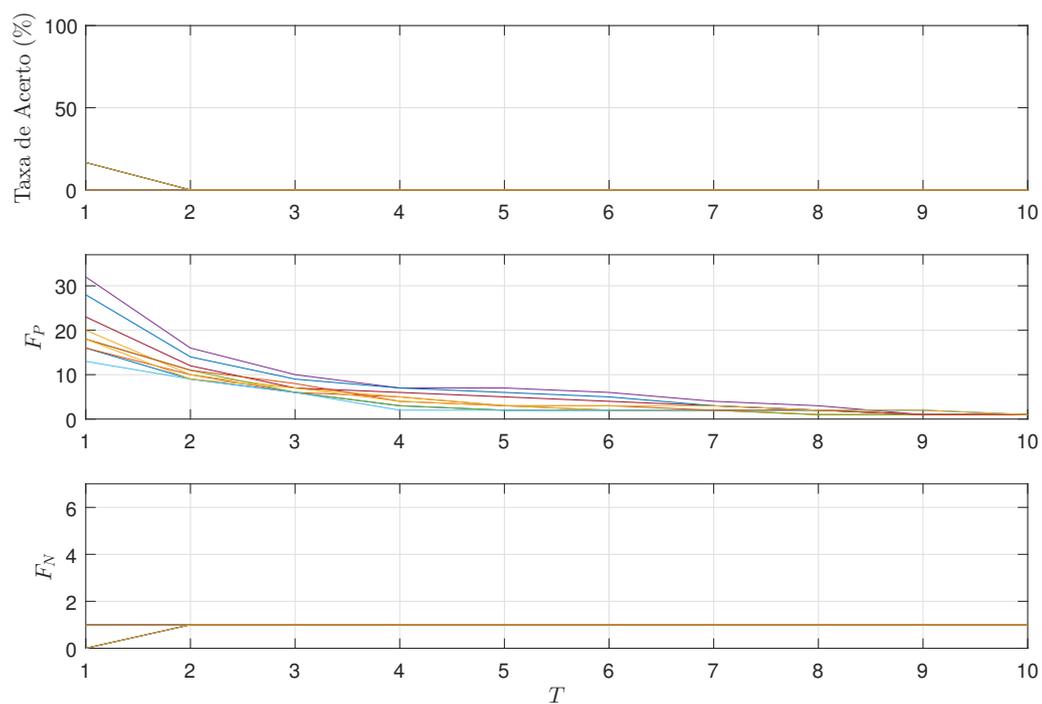


Figura A27 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde D

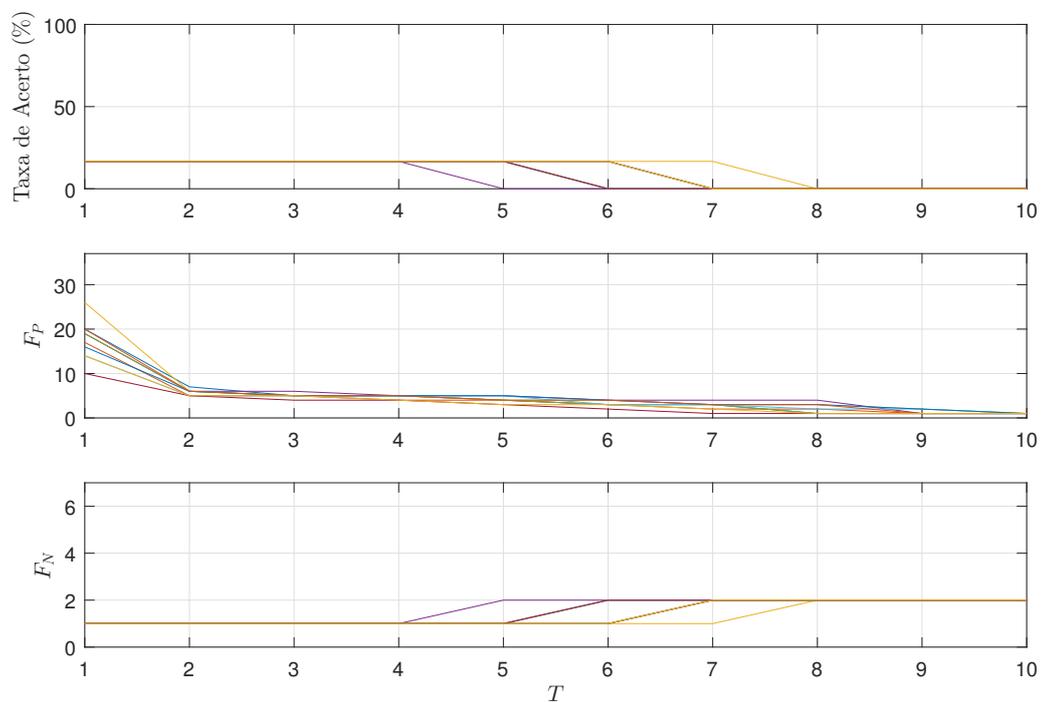


Figura A28 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde F

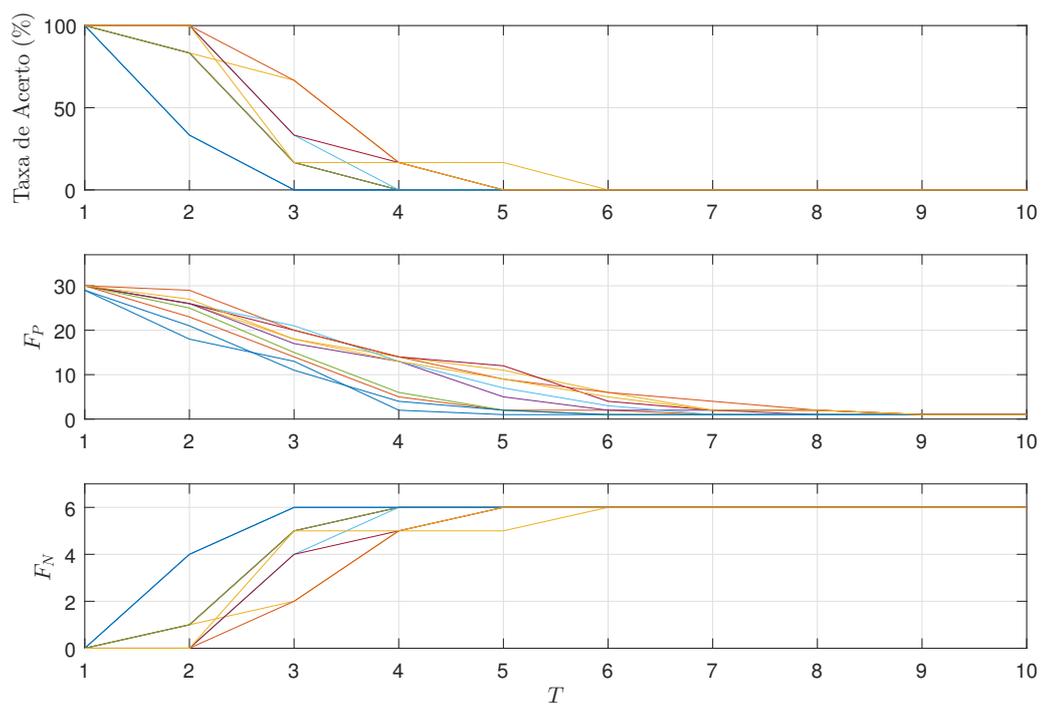


Figura A29 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde $F7M$

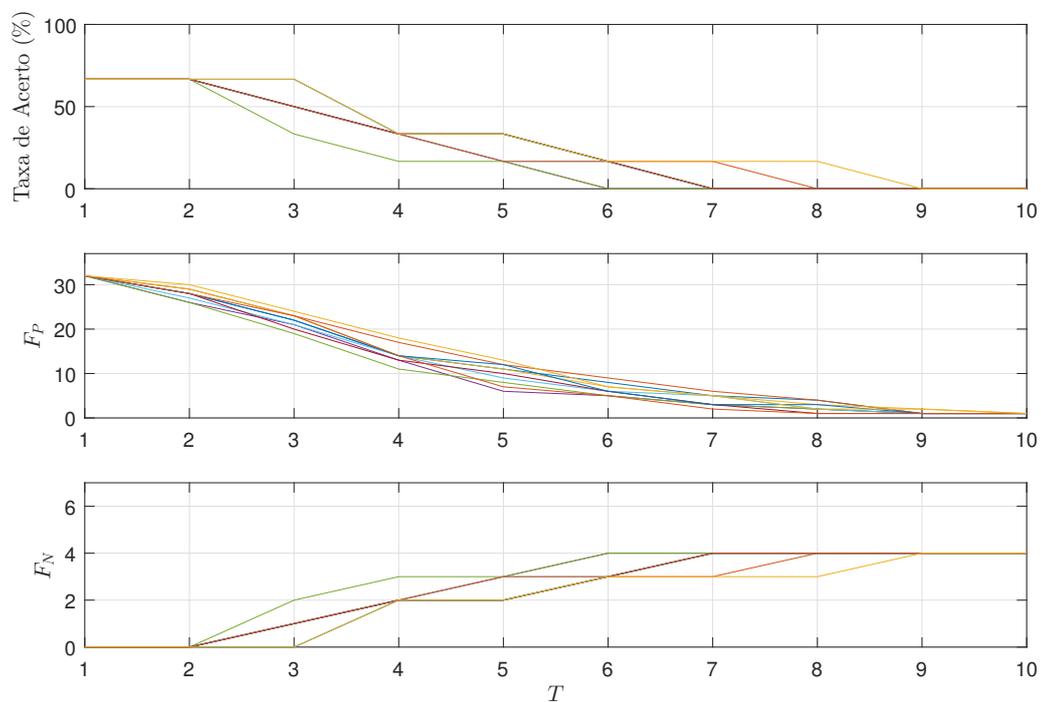


Figura A30 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde $F\sharp$

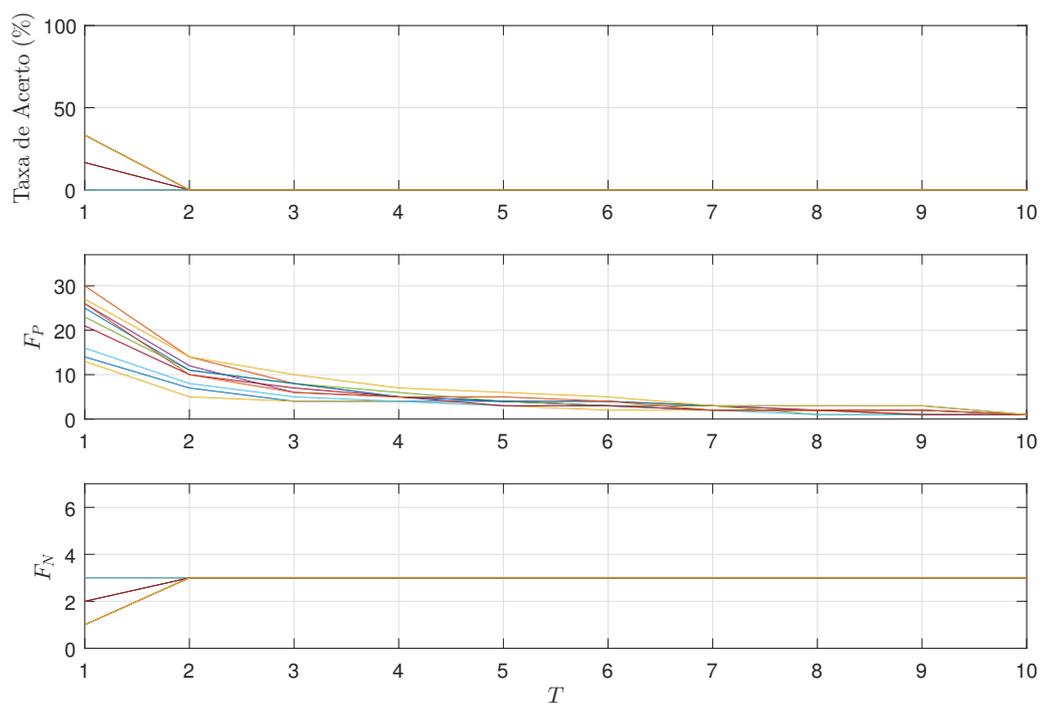


Figura A31 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde $F\sharp m7$

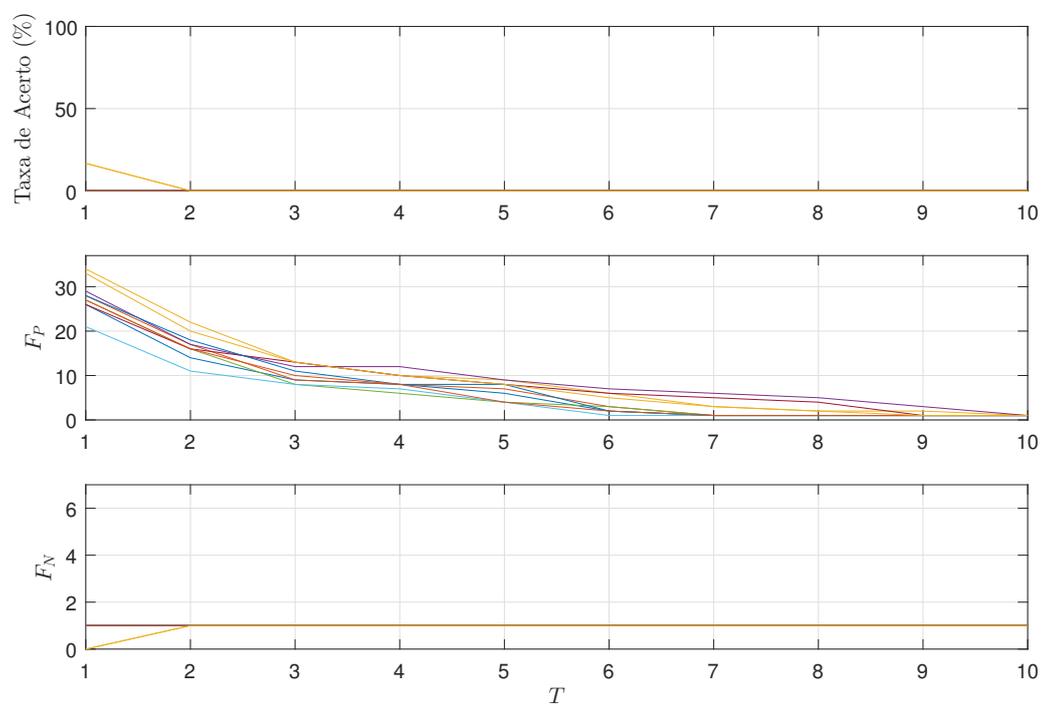


Figura A32 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde $G7$

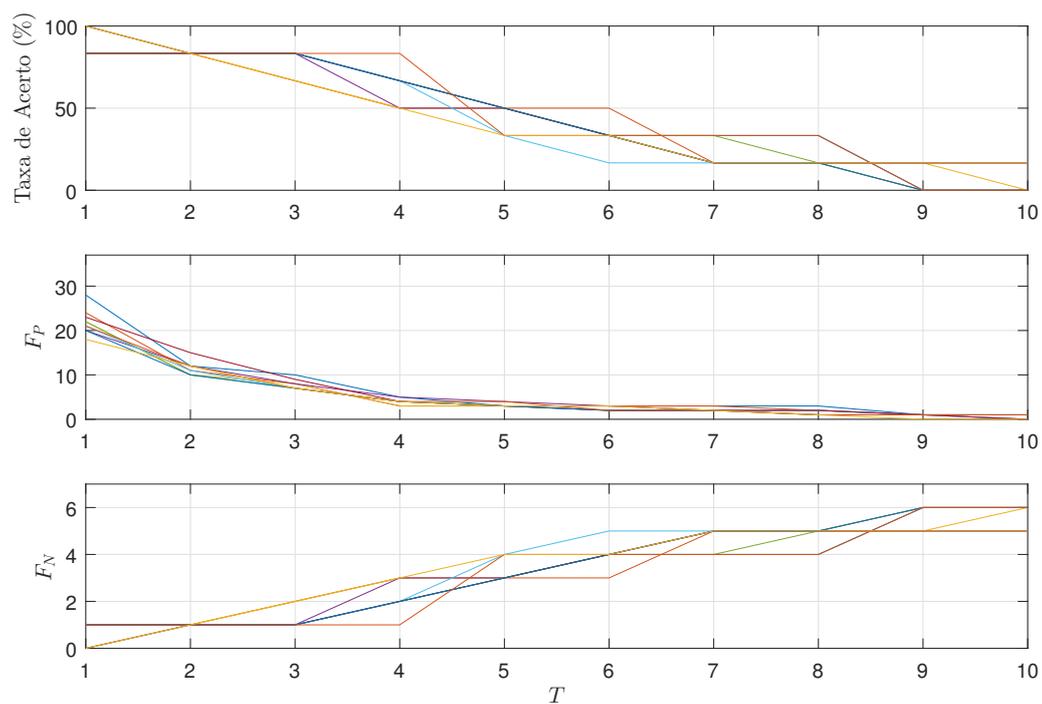


Figura A33 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde G#7

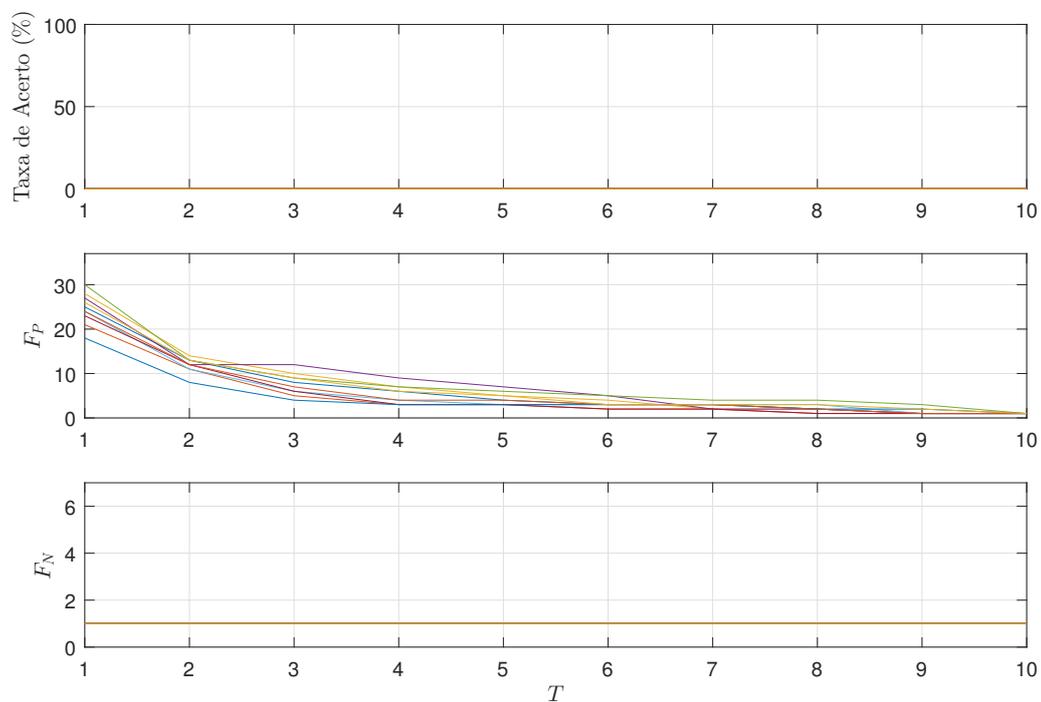


Figura A34 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde G#m7(b5)

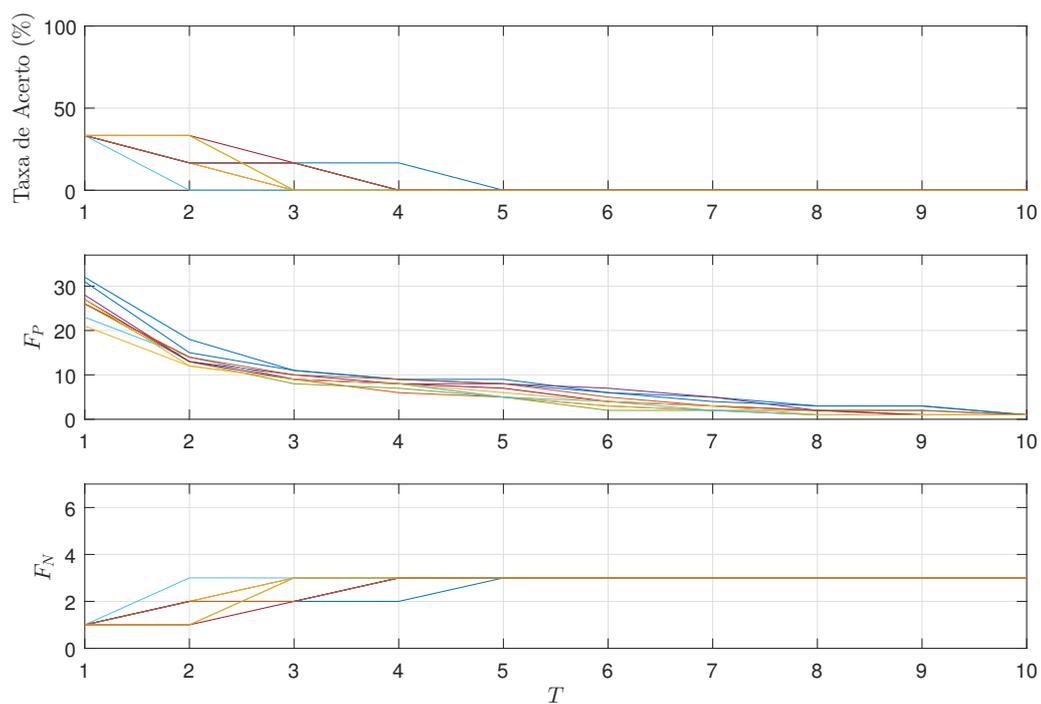


Figura A35 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde A

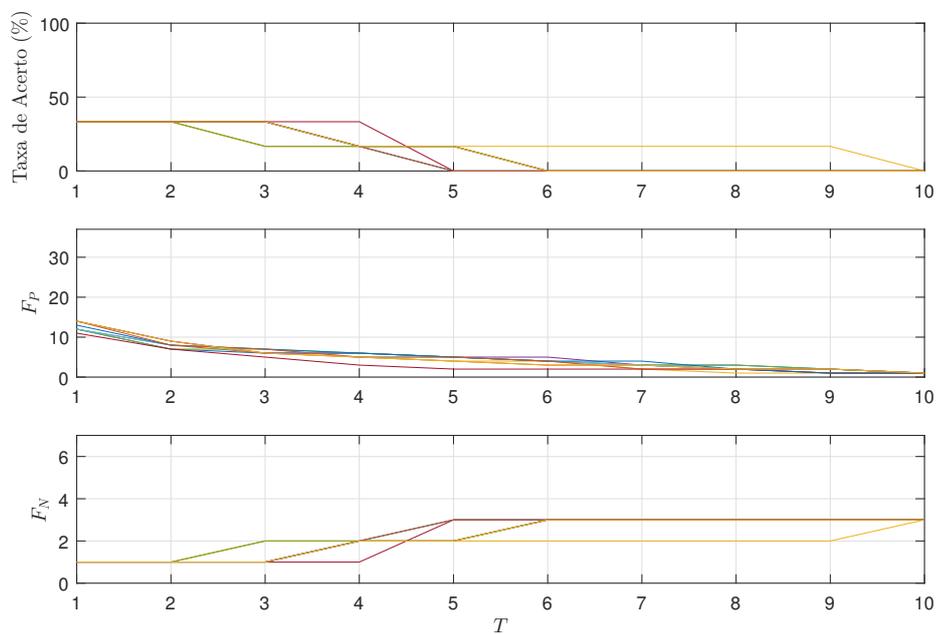


Figura A36 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para o acorde A9

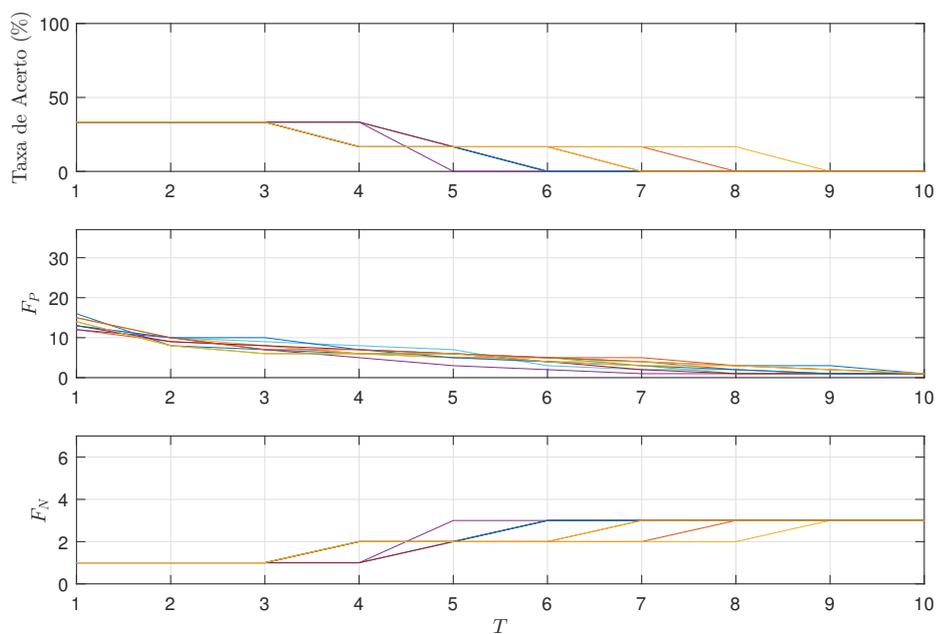


Figura A37 – Pontos das curvas ROC para 10 amostras do acorde C6

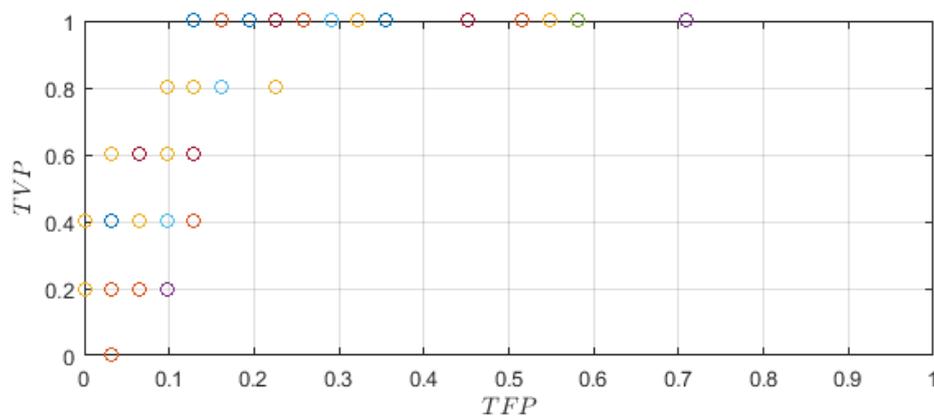


Figura A38 – Pontos das curvas ROC para 10 amostras do acorde C7

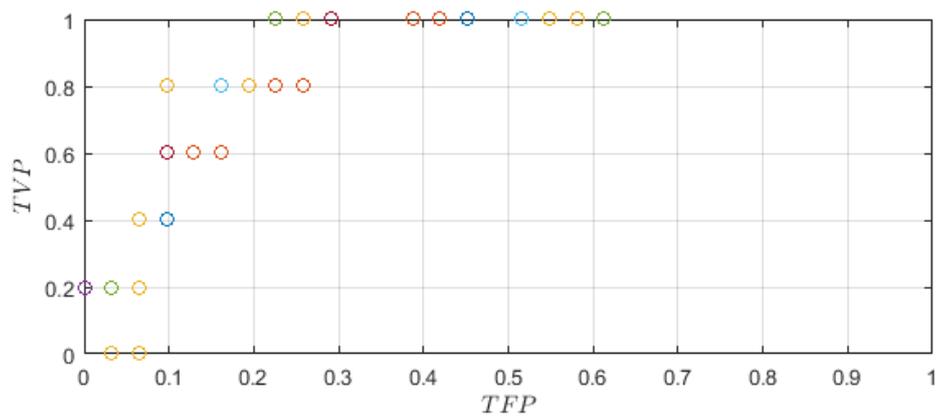


Figura A39 – Pontos das curvas ROC para 10 amostras do acorde C7M

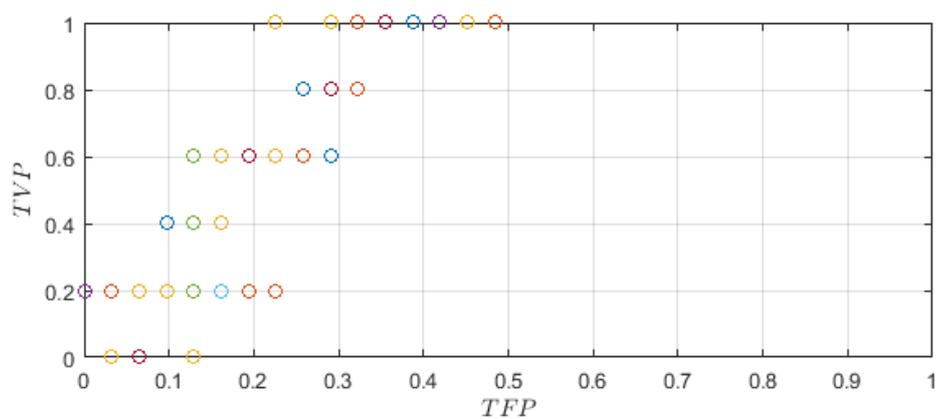


Figura A40 – Pontos das curvas *ROC* para 10 amostras do acorde *C7M*(#5)

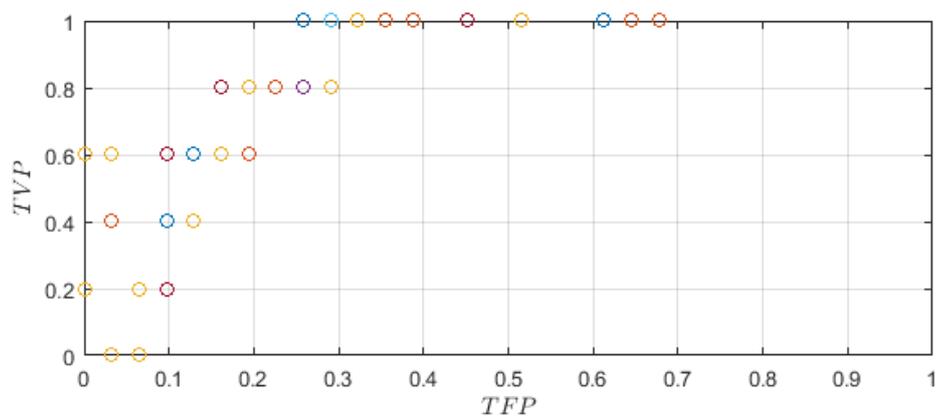


Figura A41 – Pontos das curvas *ROC* para 10 amostras do acorde *Cm*

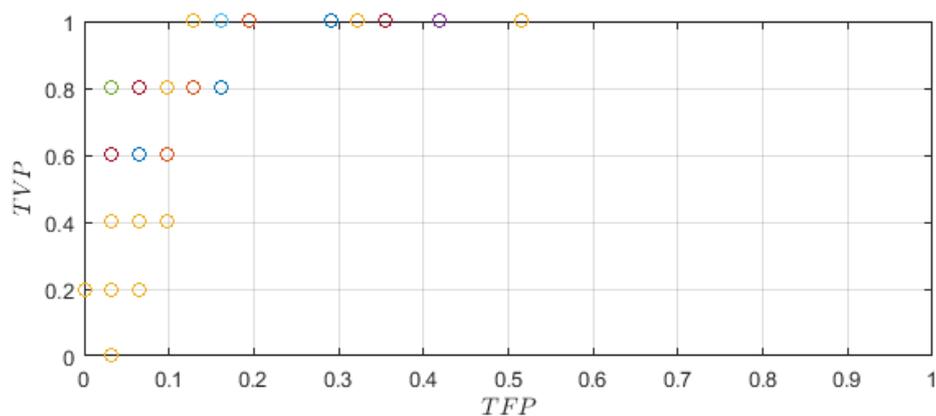


Figura A42 – Pontos das curvas *ROC* para 10 amostras do acorde *Cm6*

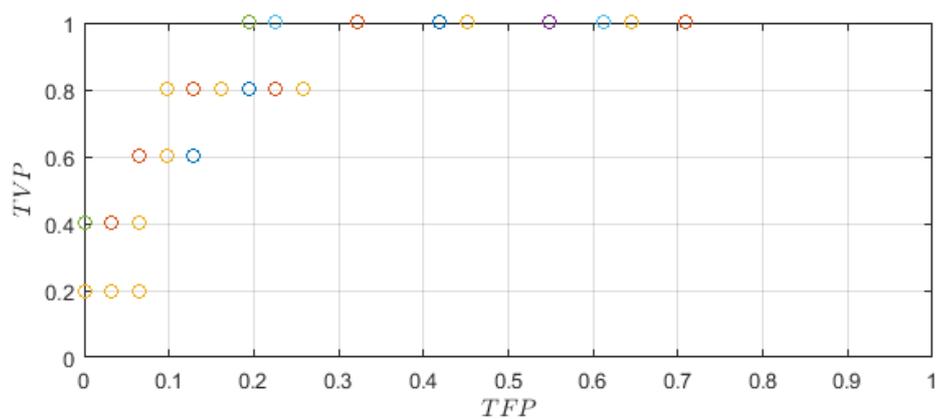


Figura A43 – Pontos das curvas *ROC* para 10 amostras do acorde *Cm7M*

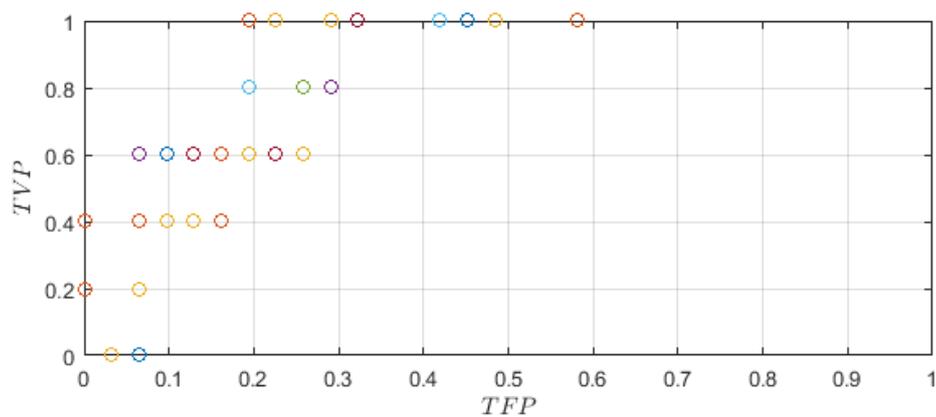


Figura A44 – Pontos das curvas *ROC* para 10 amostras do acorde *C°*

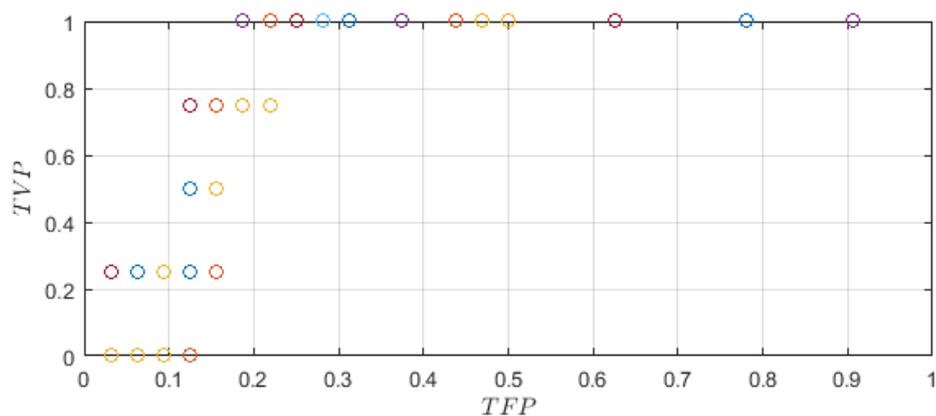


Figura A45 – Pontos das curvas *ROC* para 10 amostras do acorde *D*

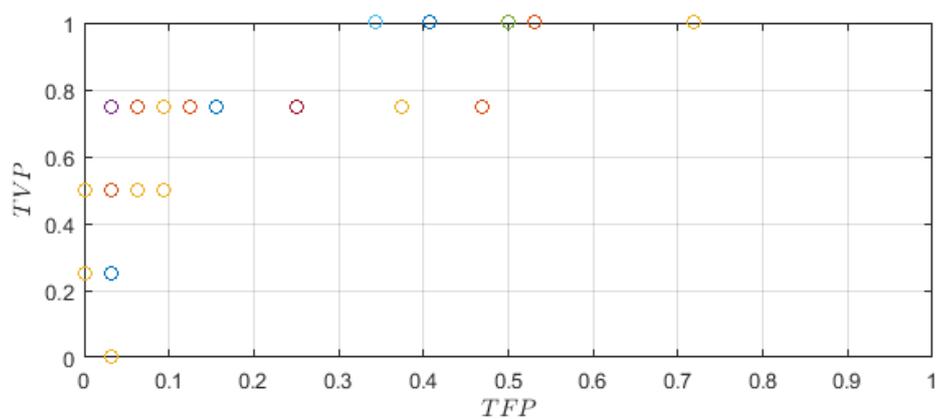


Figura A46 – Pontos das curvas ROC para 10 amostras do acorde F

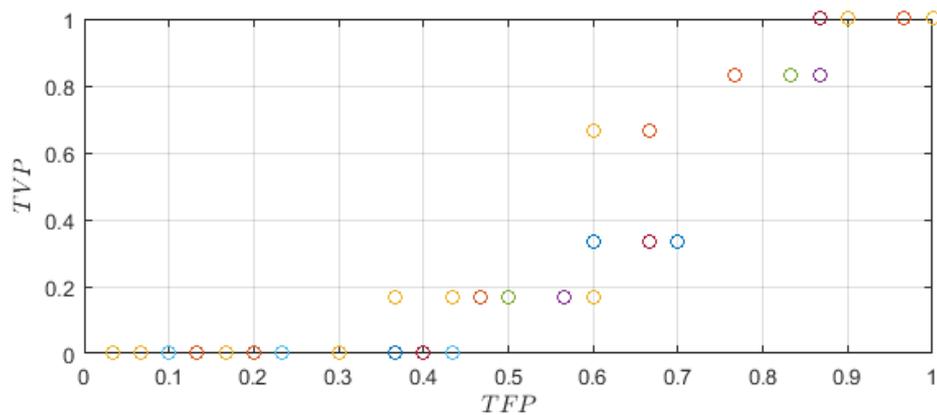


Figura A47 – Pontos das curvas ROC para 10 amostras do acorde $F7M$

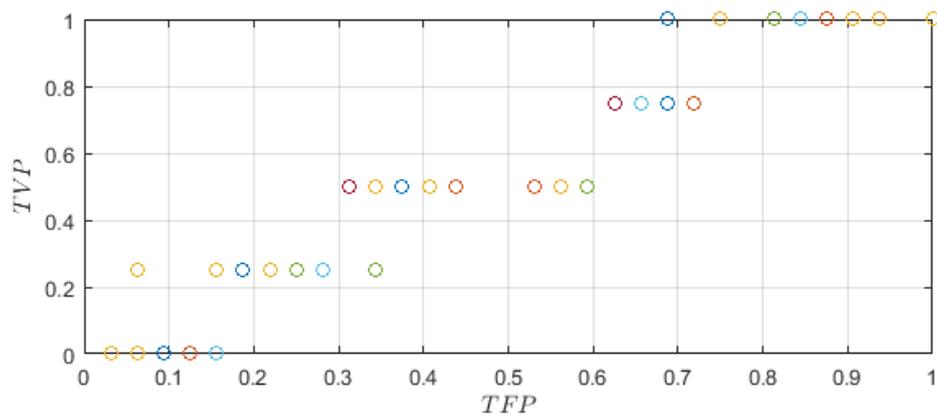


Figura A48 – Pontos das curvas ROC para 10 amostras do acorde $F\sharp$

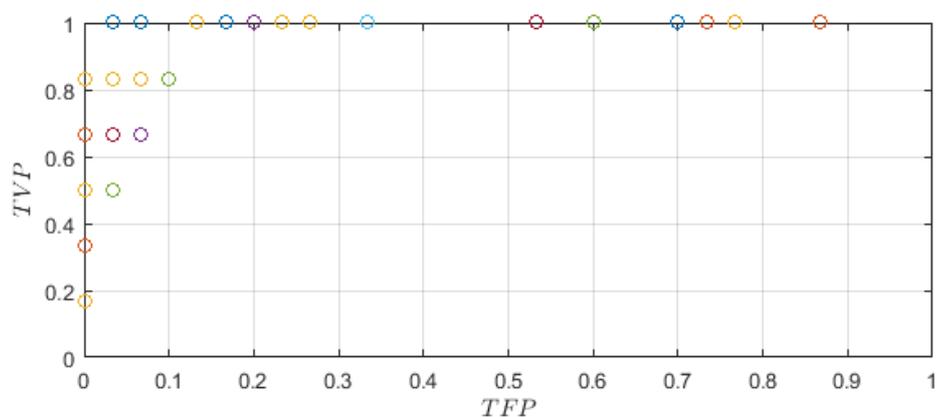


Figura A49 – Pontos das curvas *ROC* para 10 amostras do acorde $F\sharp m7$

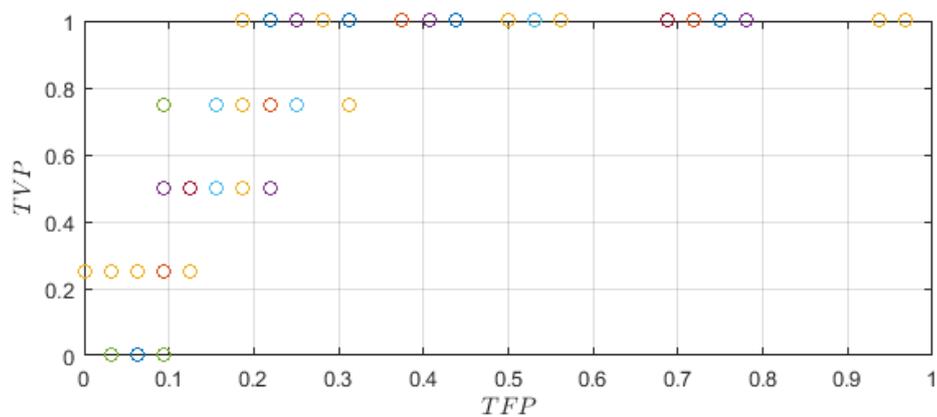


Figura A50 – Pontos das curvas *ROC* para 10 amostras do acorde $G7$

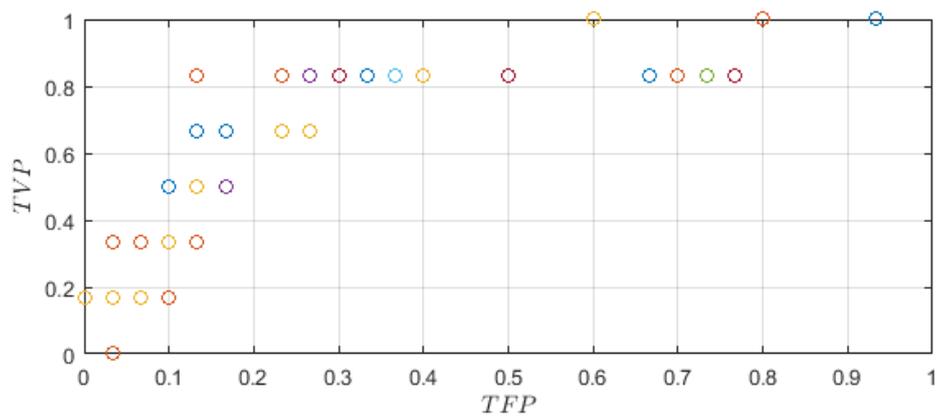


Figura A51 – Pontos das curvas *ROC* para 10 amostras do acorde $G\sharp7$

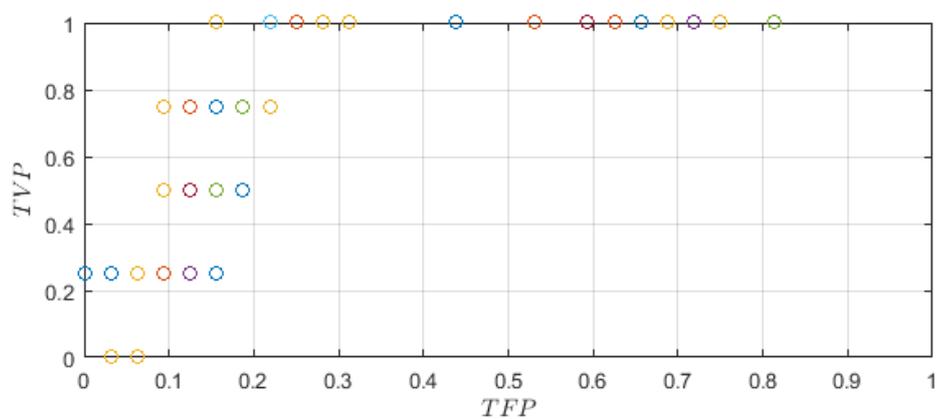


Figura A52 – Pontos das curvas *ROC* para 10 amostras do acorde $G\sharp m7(b5)$

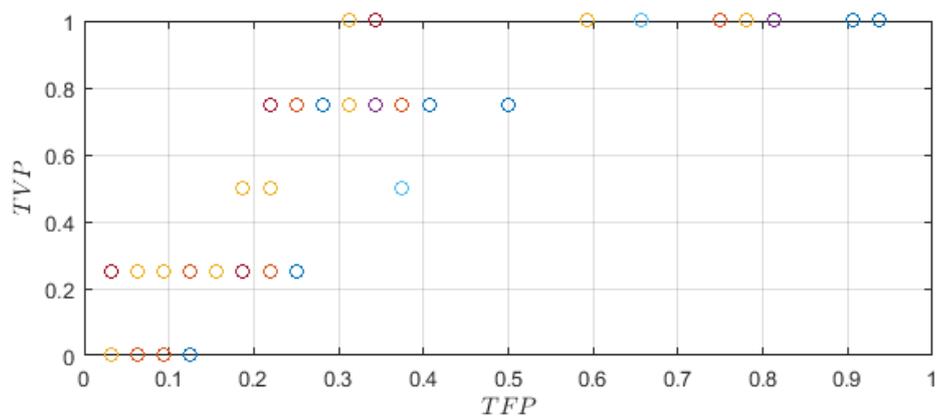


Figura A53 – Pontos das curvas *ROC* para 10 amostras do acorde *A*

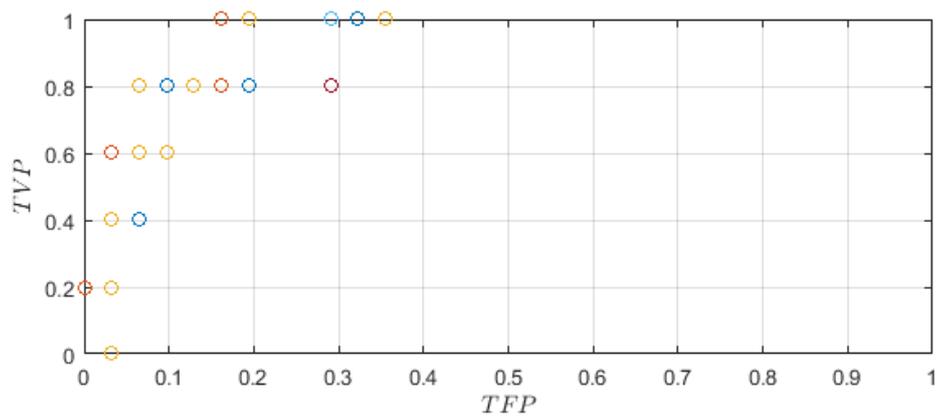
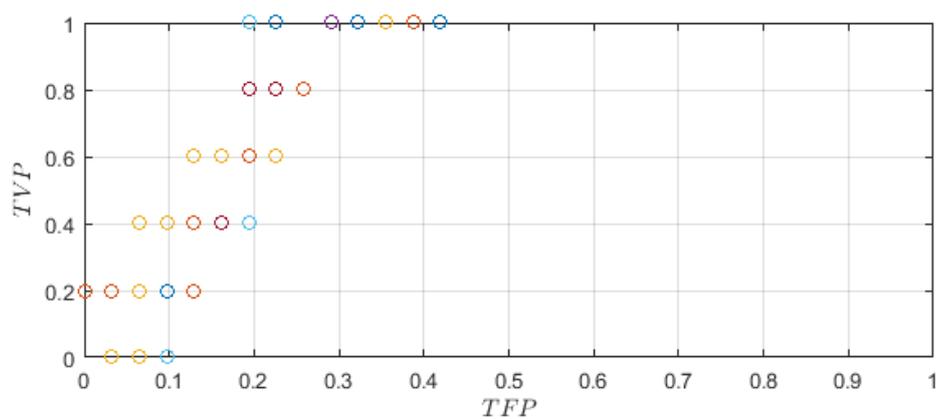


Figura A54 – Pontos das curvas *ROC* para 10 amostras do acorde *A9*



APÊNDICE D – Resultados pelo Maior Valor de Correlação

Figura A55 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde $C6$

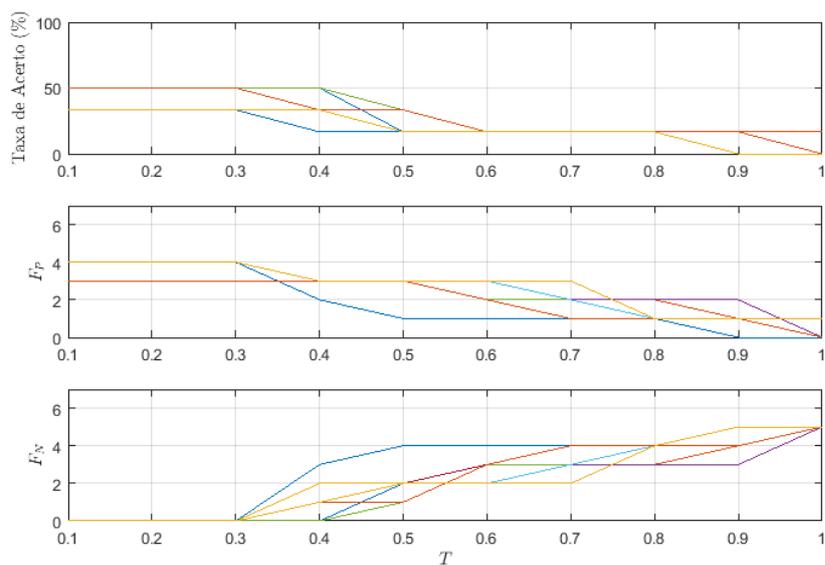


Figura A56 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde $C7$

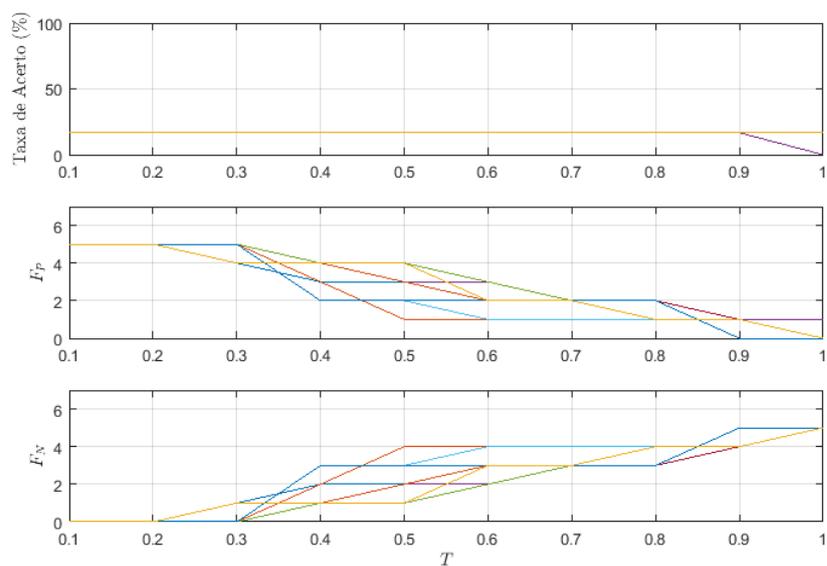


Figura A57 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde C7M

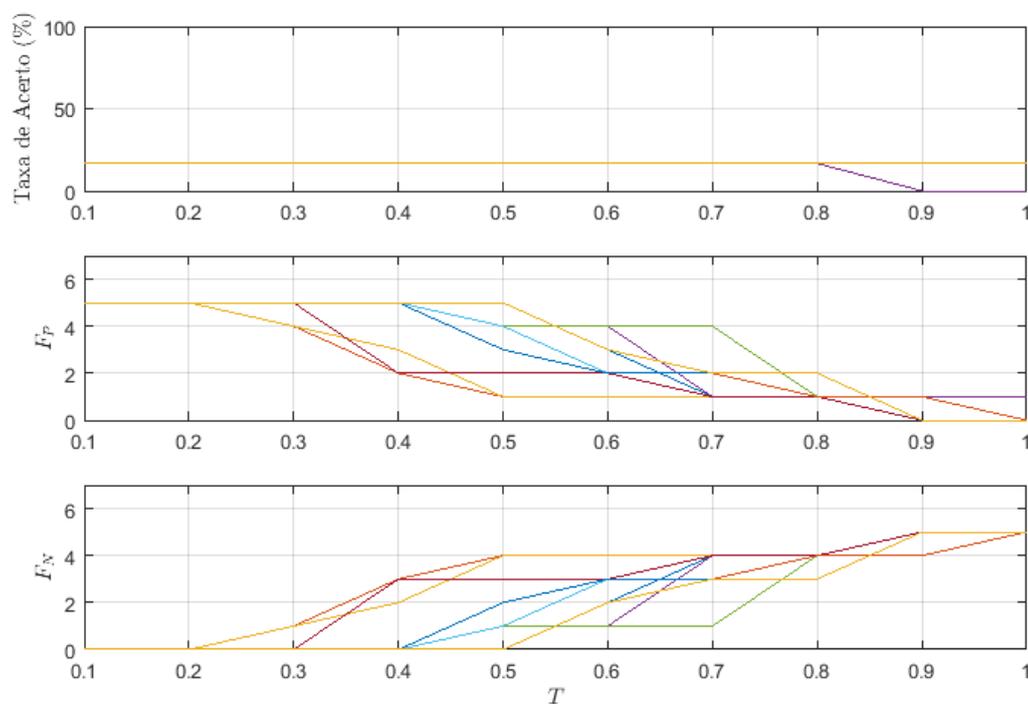


Figura A58 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde C7M(#5)

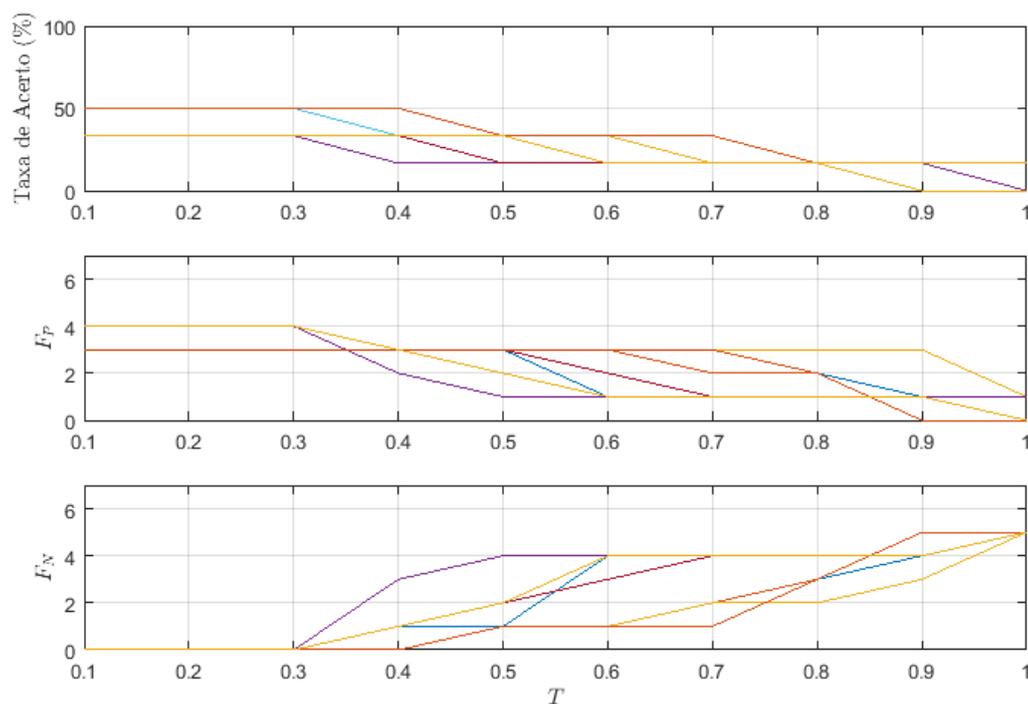


Figura A59 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde Cm

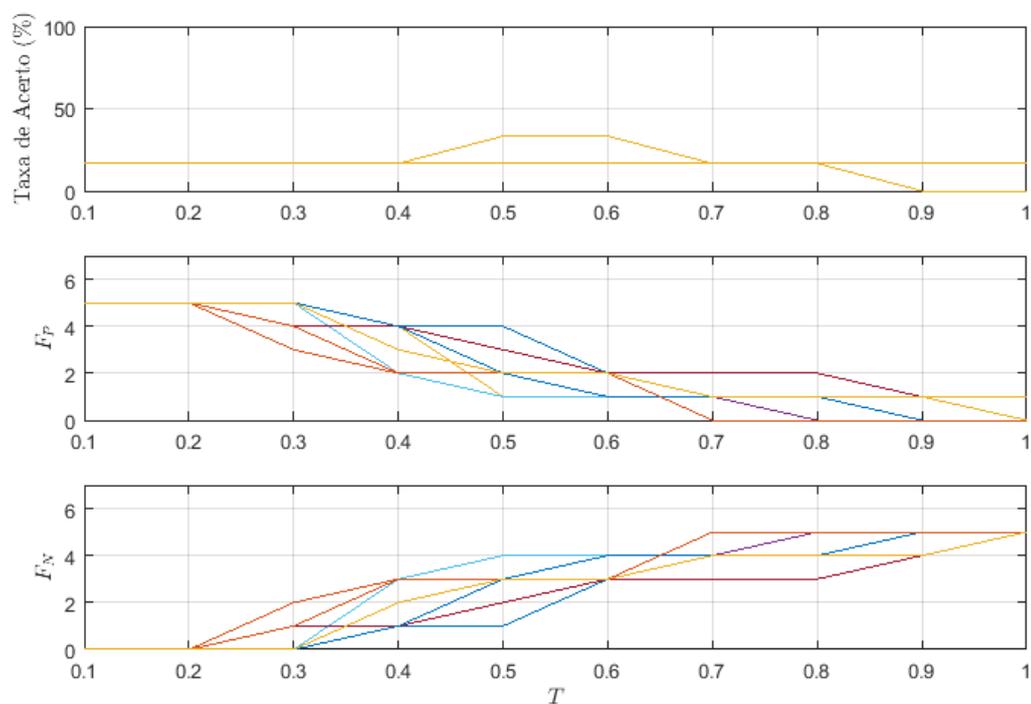


Figura A60 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde $Cm6$

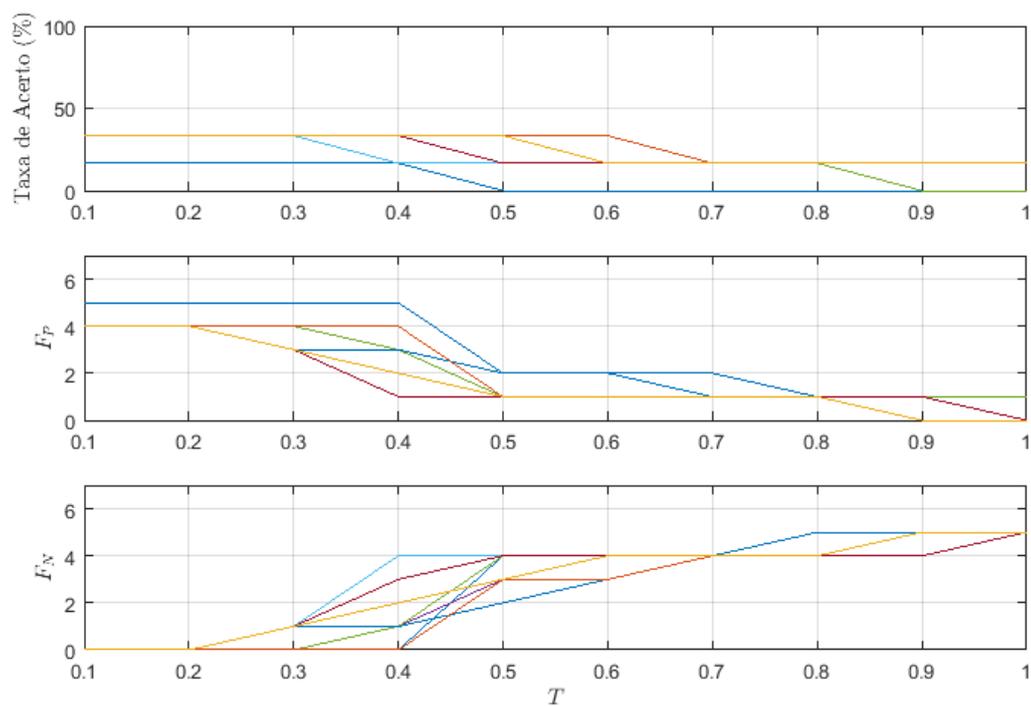


Figura A61 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde $Cm7M$

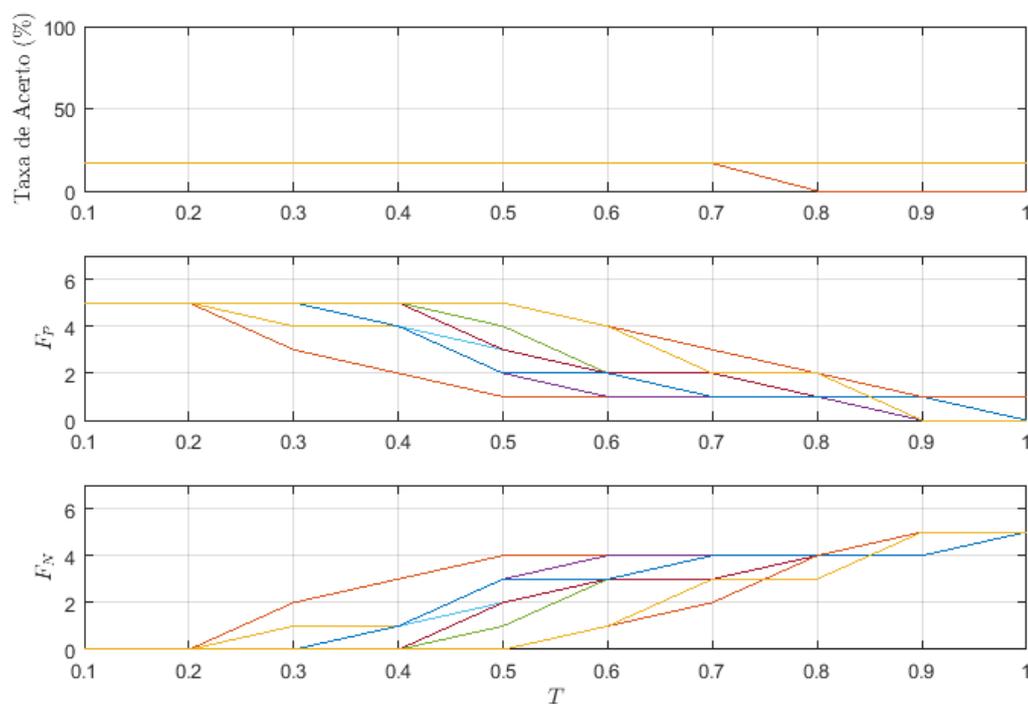


Figura A62 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde C^o

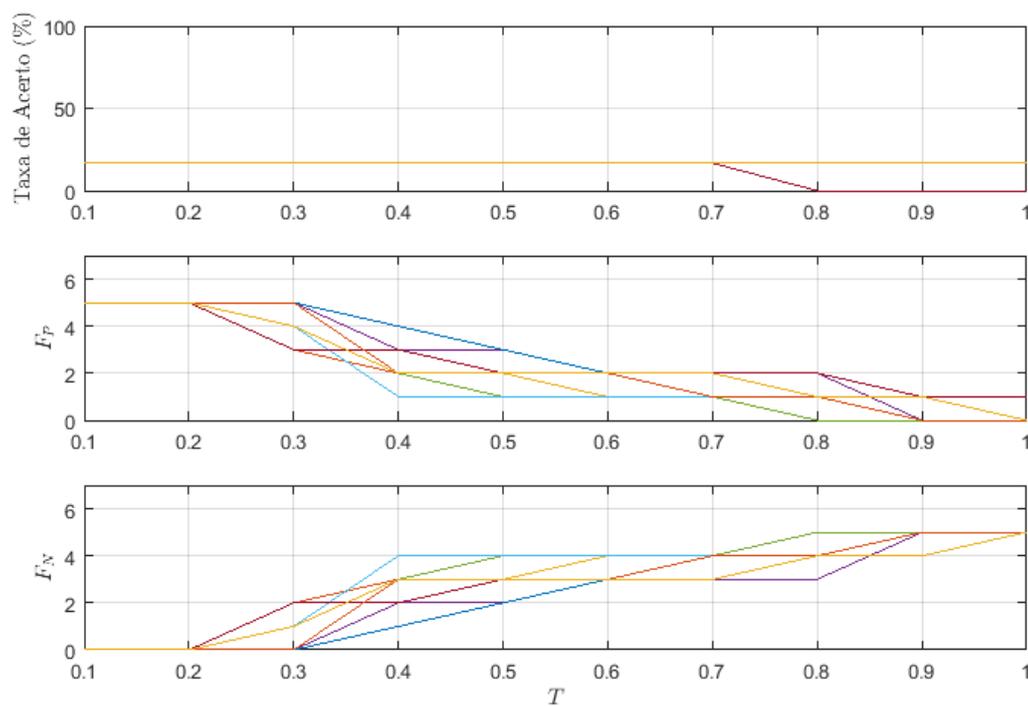


Figura A63 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde D

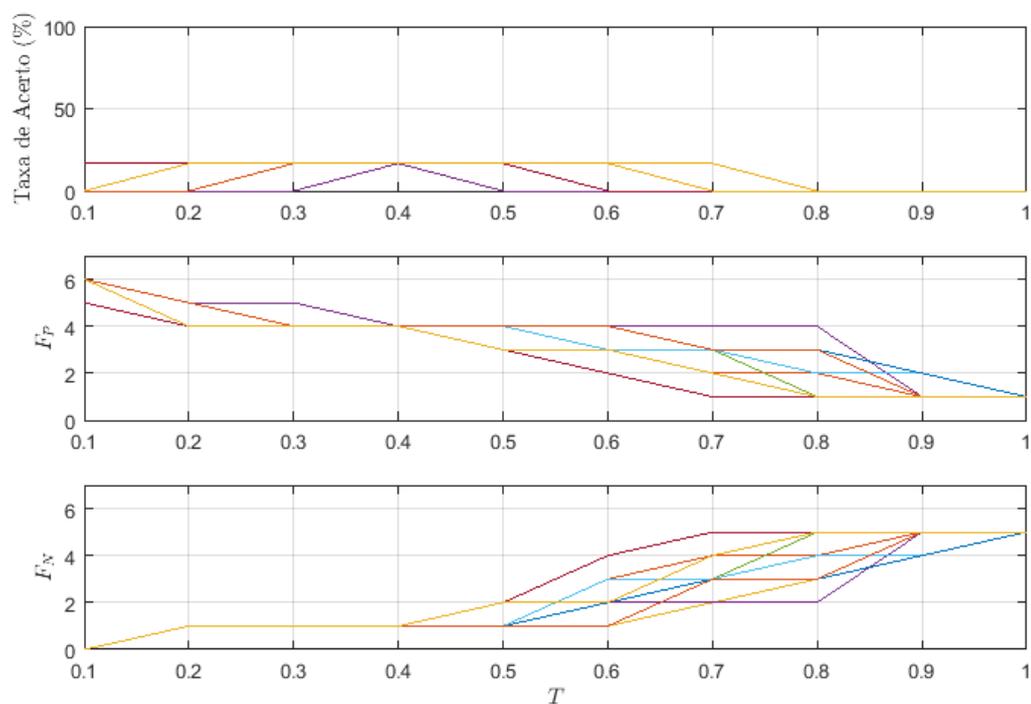


Figura A64 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde F

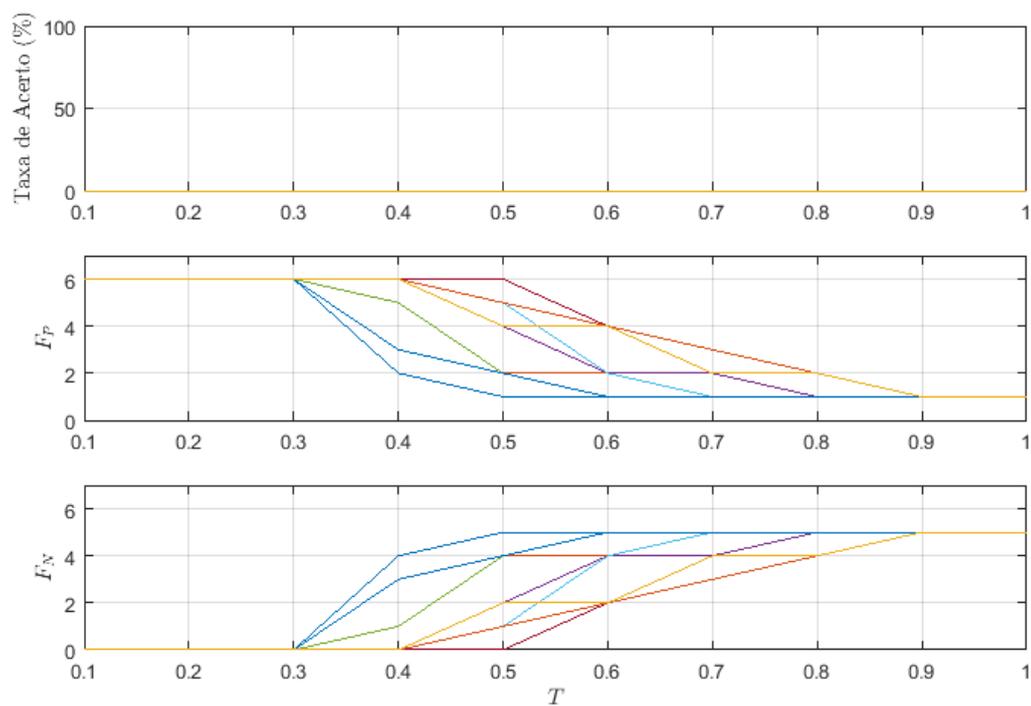


Figura A65 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde $F7M$

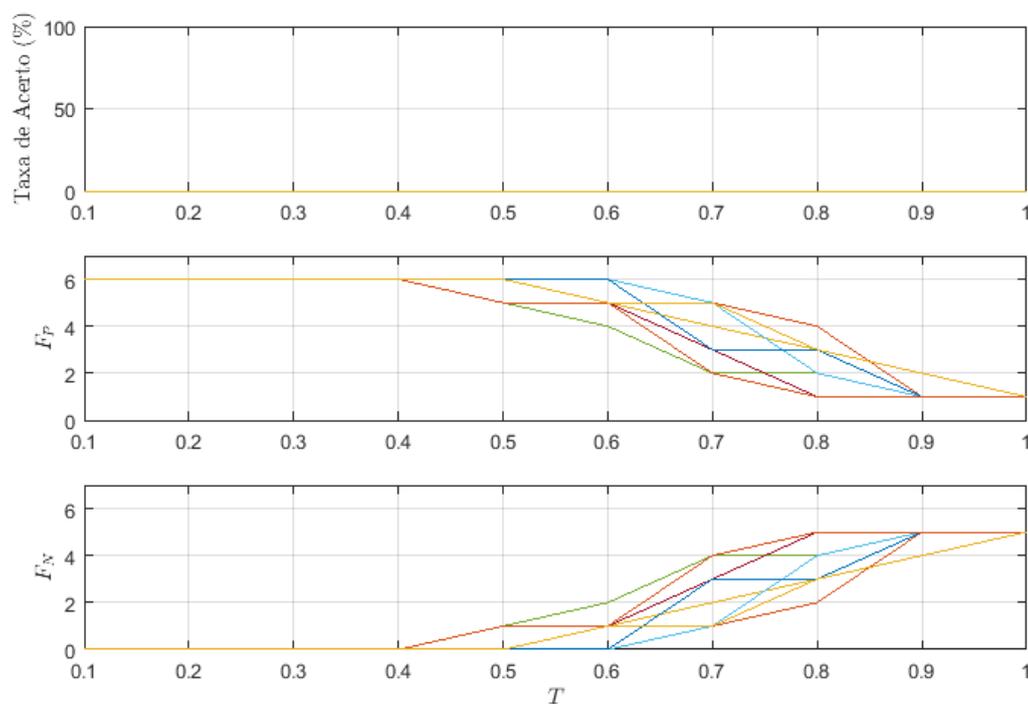


Figura A66 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde $F\sharp$

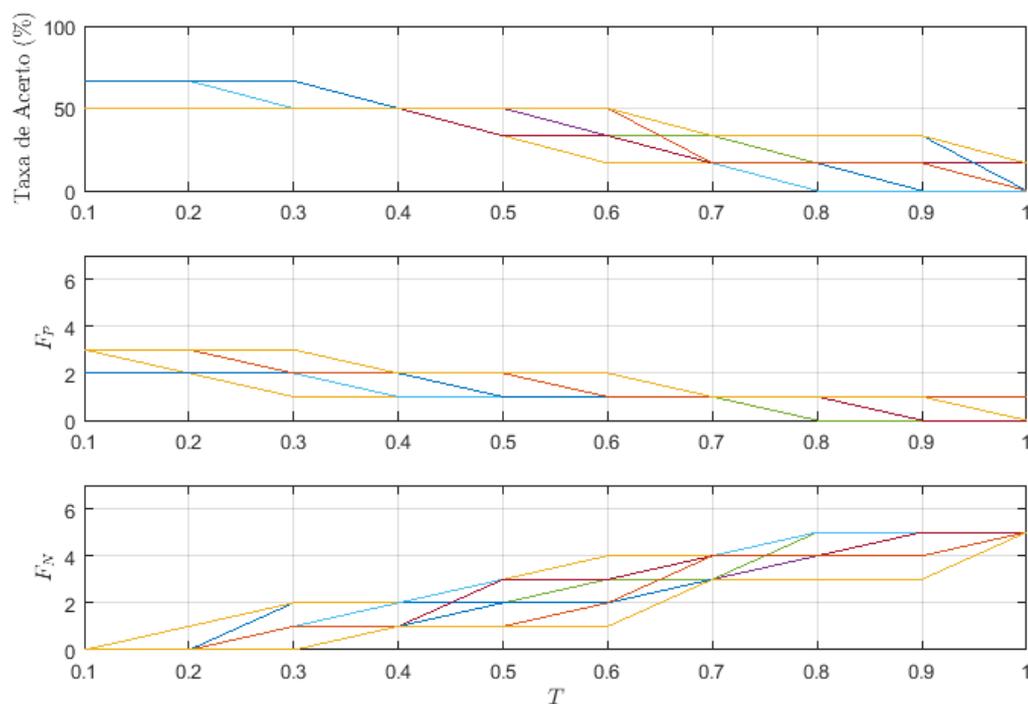


Figura A67 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde $F\#m7$

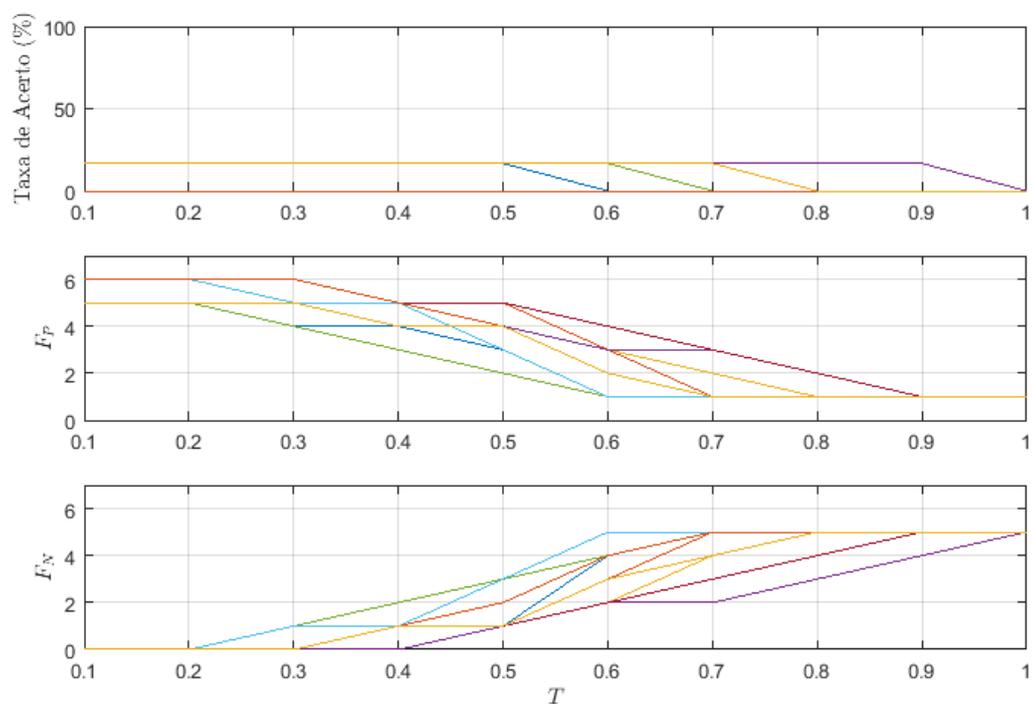


Figura A68 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde $G7$

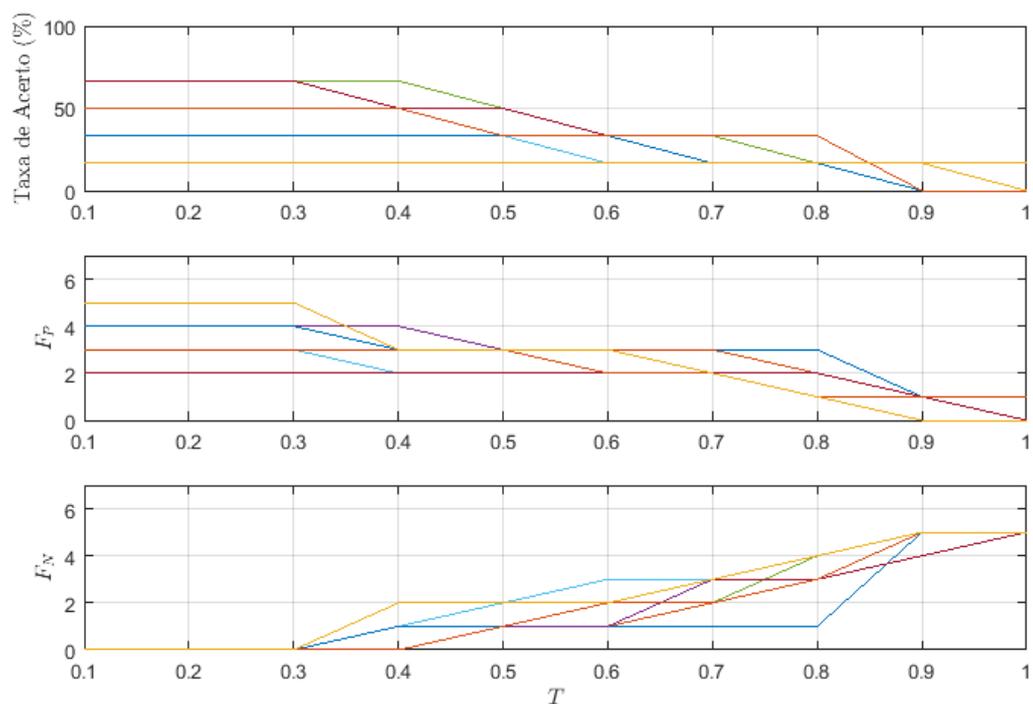


Figura A69 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde $G\#7$

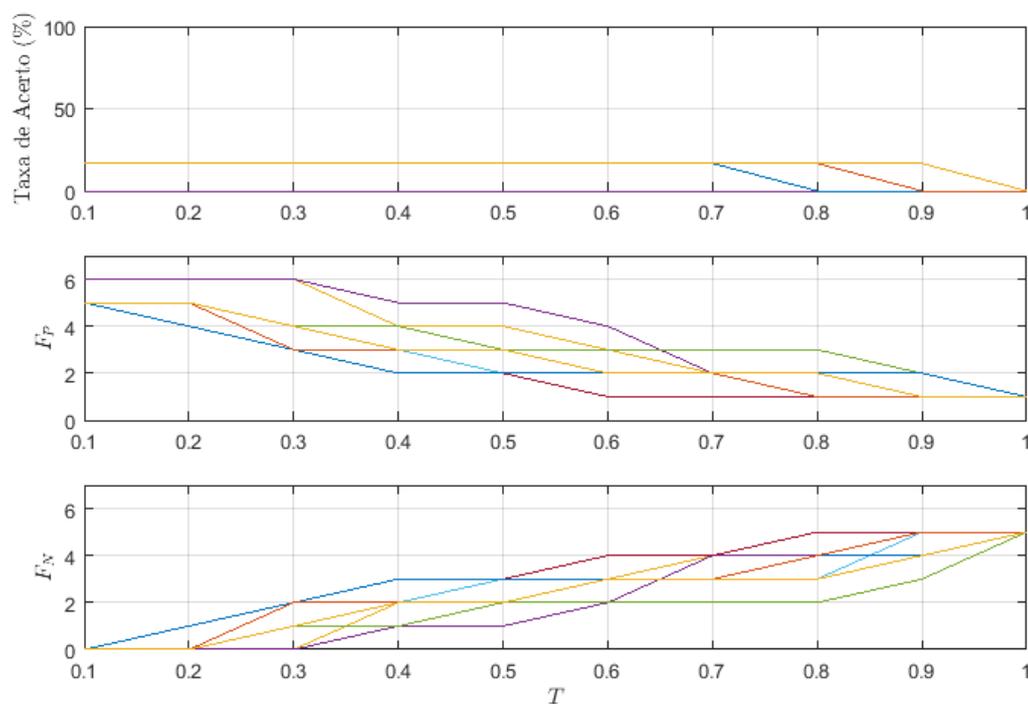


Figura A70 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde $G\#m7(b5)$

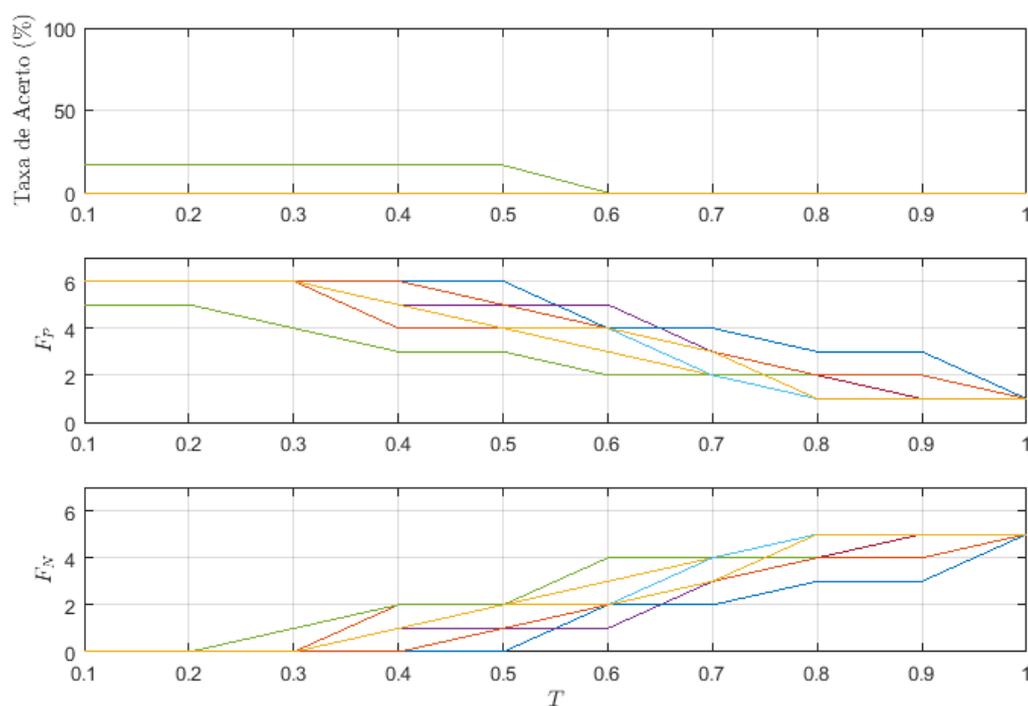


Figura A71 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde A

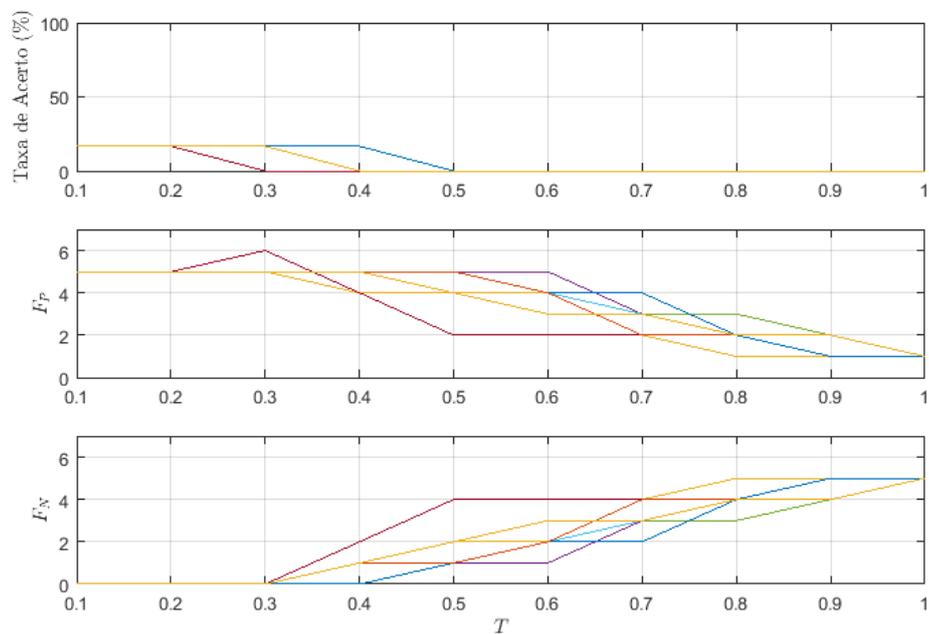


Figura A72 – Taxa de acerto, Falsos Positivos e Falsos Negativos em função de T para 10 amostras do acorde A9

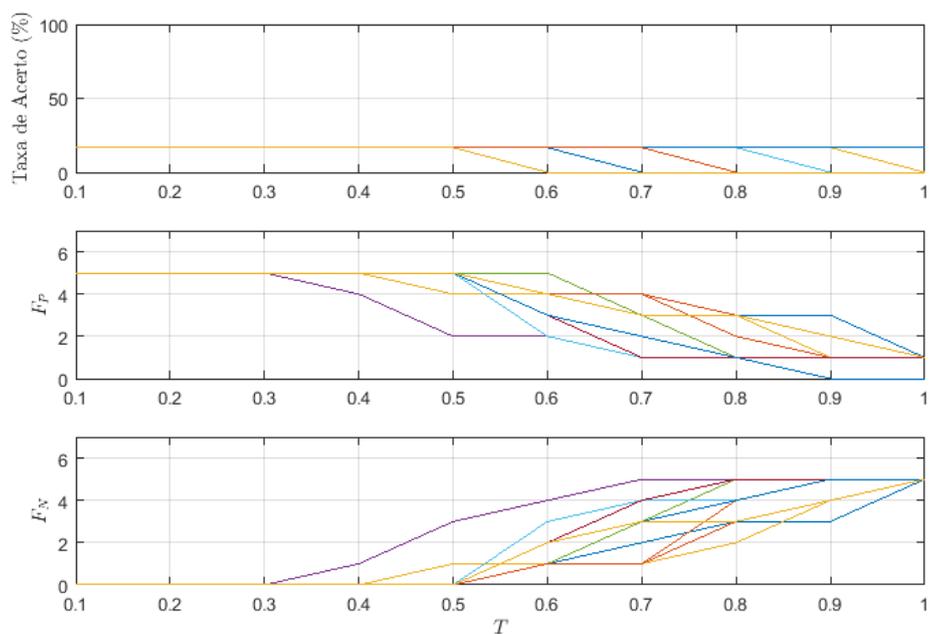


Figura A73 – Pontos das curvas ROC para 10 amostras do acorde C6

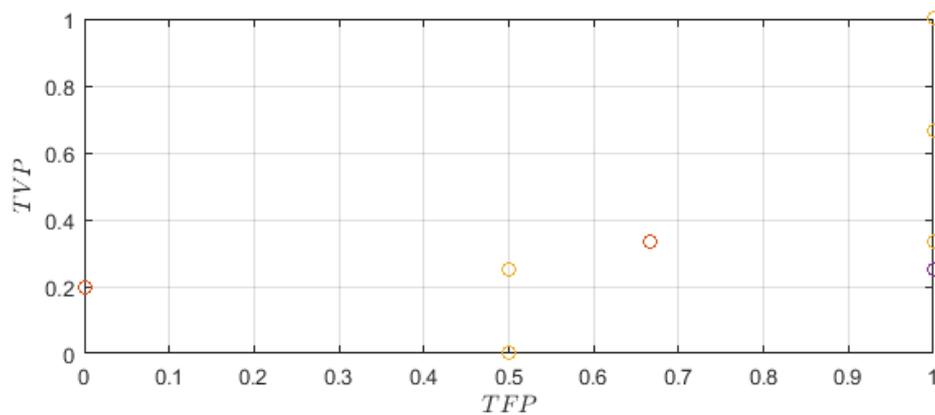


Figura A74 – Pontos das curvas ROC para 10 amostras do acorde C7

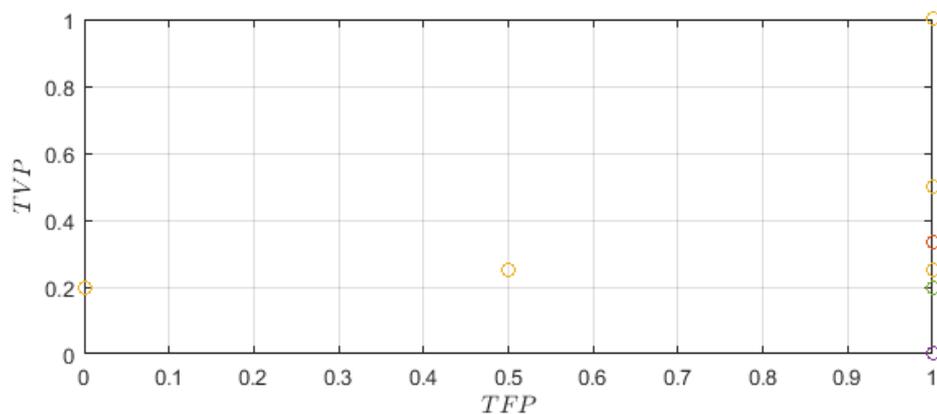


Figura A75 – Pontos das curvas ROC para 10 amostras do acorde C7M

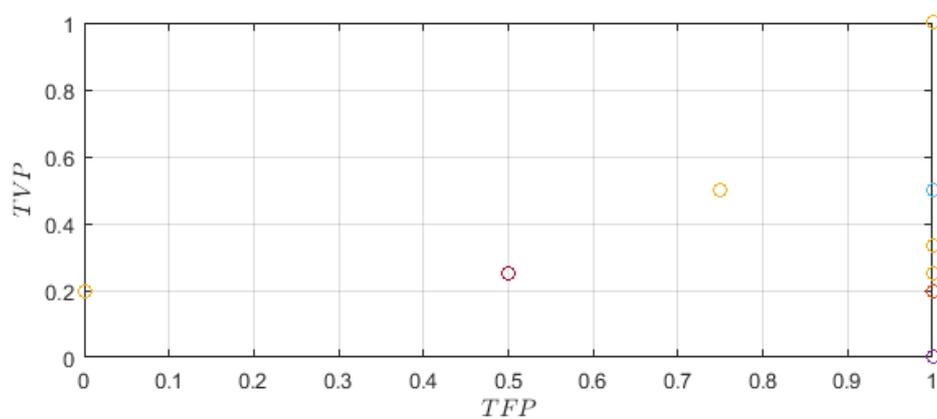


Figura A76 – Pontos das curvas *ROC* para 10 amostras do acorde *C7M*(#5)

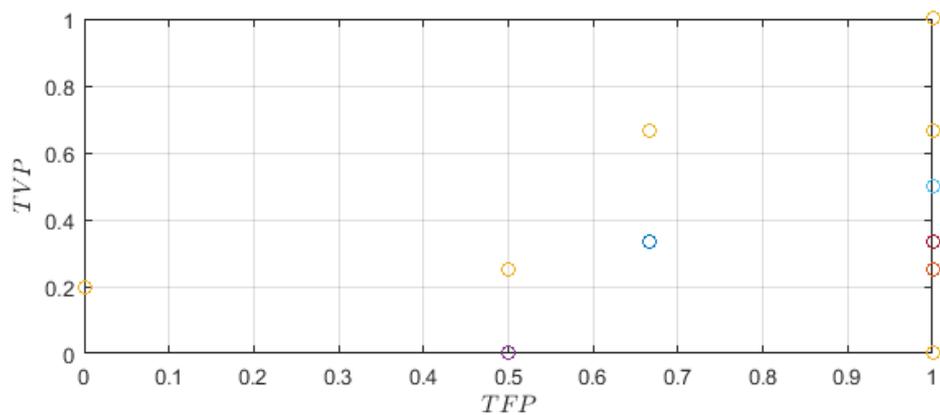


Figura A77 – Pontos das curvas *ROC* para 10 amostras do acorde *Cm*

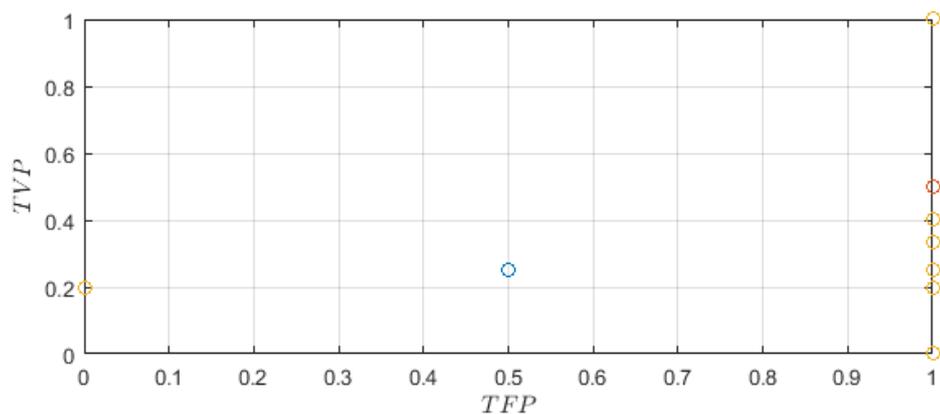


Figura A78 – Pontos das curvas *ROC* para 10 amostras do acorde *Cm6*

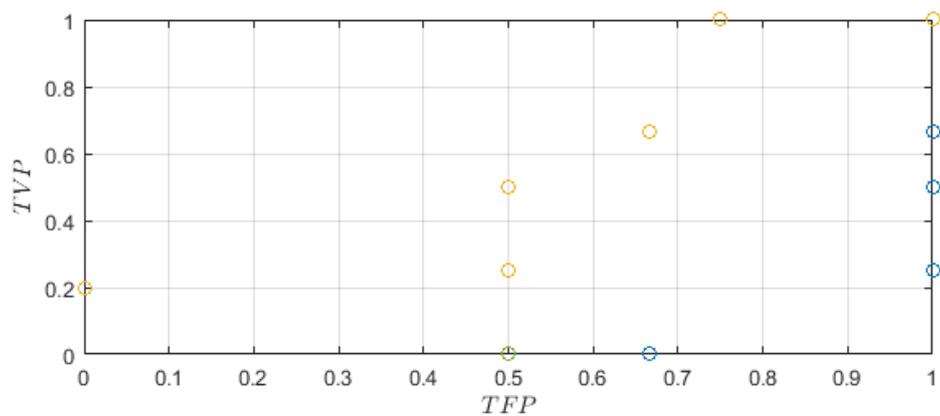


Figura A79 – Pontos das curvas *ROC* para 10 amostras do acorde *Cm7M*

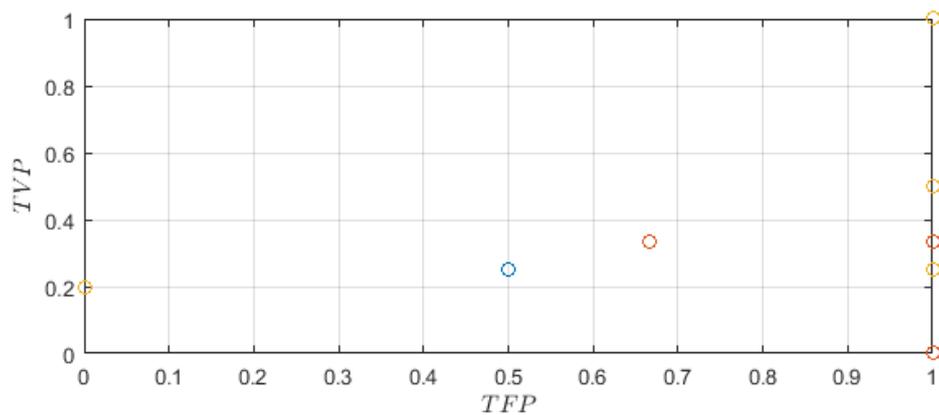


Figura A80 – Pontos das curvas *ROC* para 10 amostras do acorde *C°*

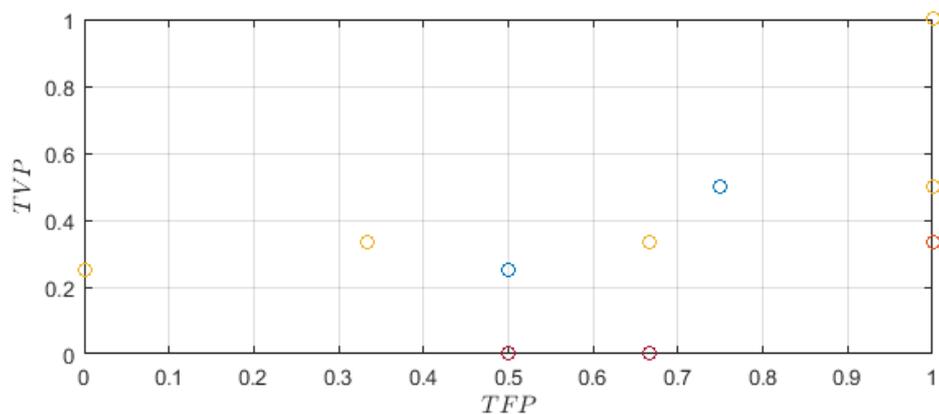


Figura A81 – Pontos das curvas *ROC* para 10 amostras do acorde *D*

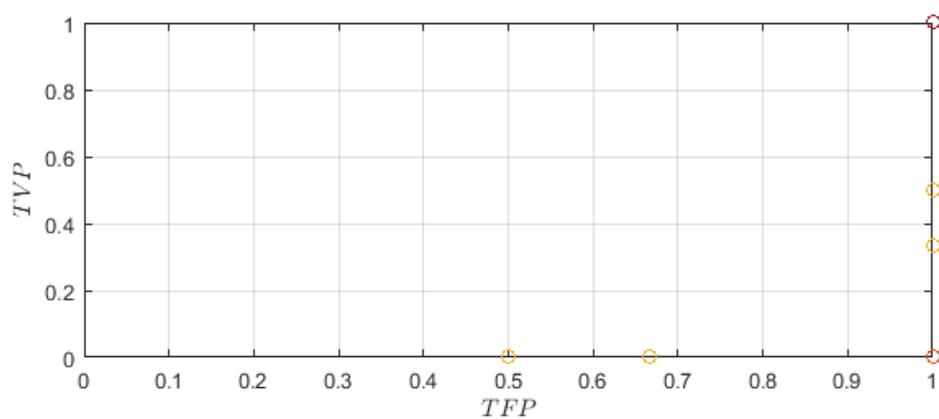


Figura A82 – Pontos das curvas *ROC* para 10 amostras do acorde *F*

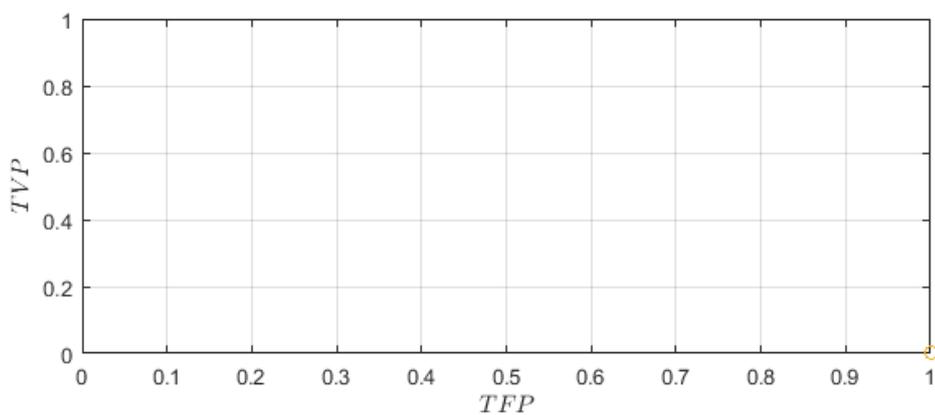


Figura A83 – Pontos das curvas *ROC* para 10 amostras do acorde *F7M*

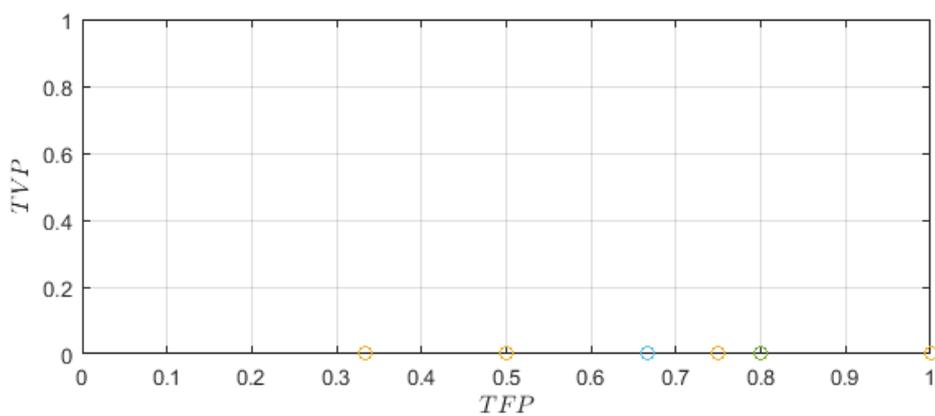


Figura A84 – Pontos das curvas *ROC* para 10 amostras do acorde *F#*

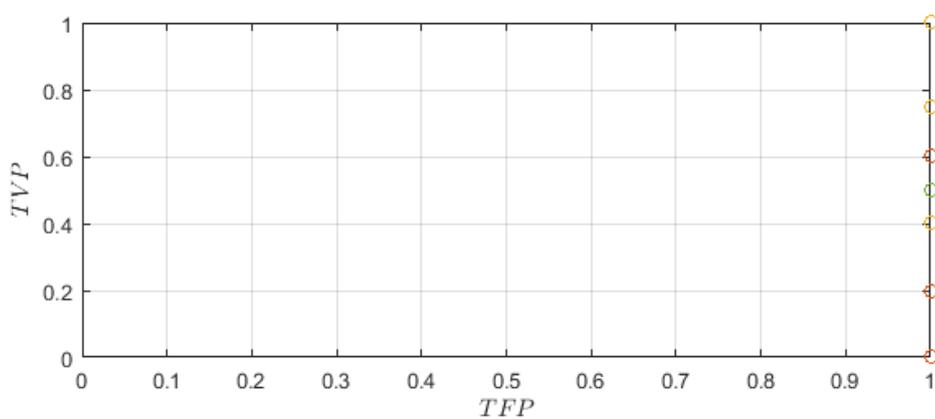


Figura A85 – Pontos das curvas *ROC* para 10 amostras do acorde $F\#m7$

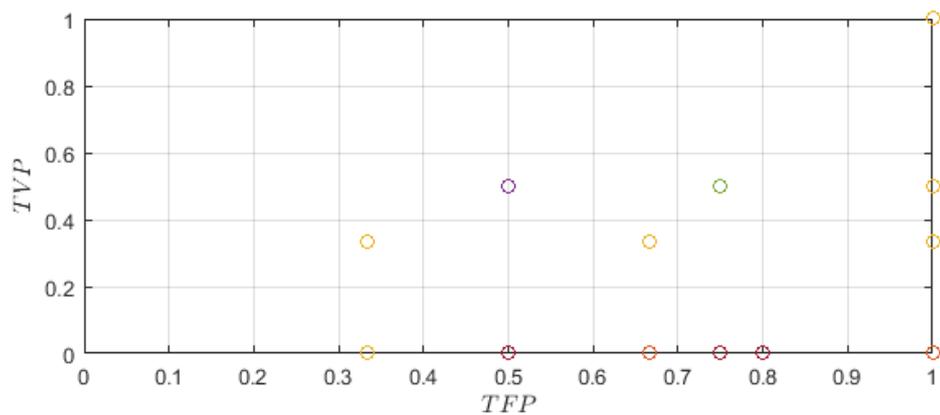


Figura A86 – Pontos das curvas *ROC* para 10 amostras do acorde $G7$

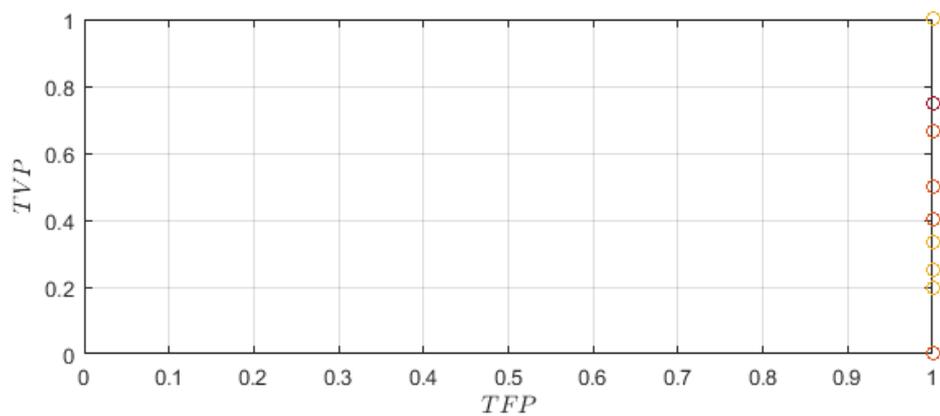


Figura A87 – Pontos das curvas *ROC* para 10 amostras do acorde $G\#7$

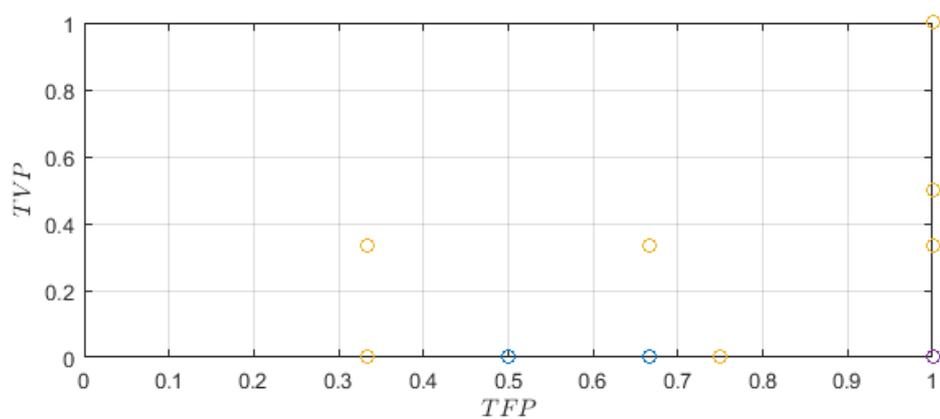


Figura A88 – Pontos das curvas *ROC* para 10 amostras do acorde $G\sharp m7(b5)$

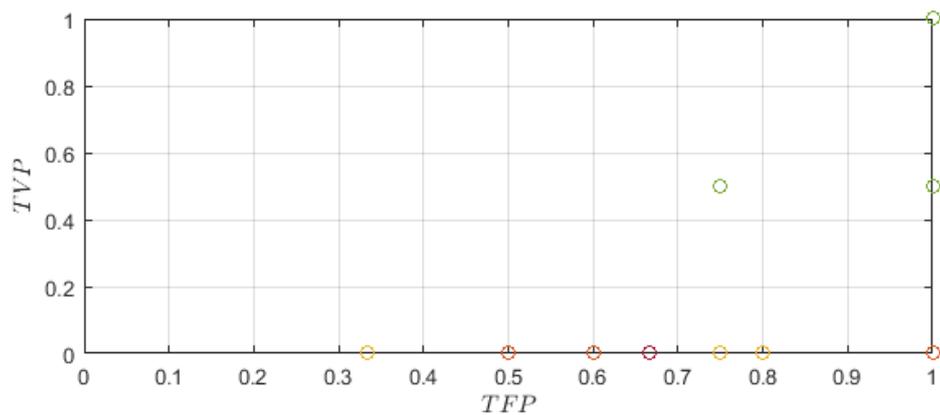


Figura A89 – Pontos das curvas *ROC* para 10 amostras do acorde *A*

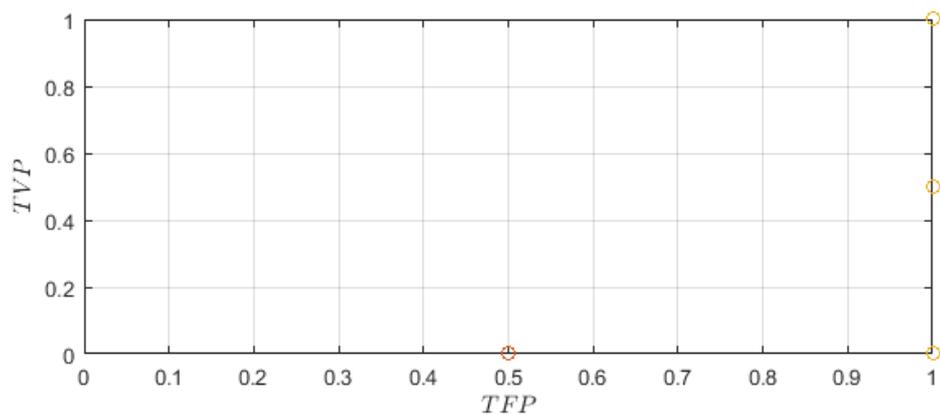


Figura A90 – Pontos das curvas *ROC* para 10 amostras do acorde *A9*

