

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
DEPARTAMENTO ACADÊMICO DE ELETROTÉCNICA
CURSO DE ENGENHARIA ELÉTRICA

ALVARO ALCEU FERREIRA SANTOS
LEONARDO AUGUSTO PIOVEZAN
RAFAEL COUTO DE OLIVEIRA

**ANÁLISE DE SINAIS DE ÁUDIO E IMPACTOS NA COLETA, CONVERSÃO E
CODIFICAÇÃO DE SINAIS**

TRABALHO DE CONCLUSÃO DE CURSO

CURITIBA
2018

ALVARO ALCEU FERREIRA SANTOS
LEONARDO AUGUSTO PIOVEZAN
RAFAEL COUTO DE OLIVEIRA

**ANÁLISE DE SINAIS DE ÁUDIO E IMPACTOS NA COLETA, CONVERSÃO E
CODIFICAÇÃO DE SINAIS**

Trabalho de Conclusão de Curso de Graduação, do curso de Engenharia Elétrica do Departamento Acadêmico de Eletrotécnica (DAELT) da Universidade Tecnológica Federal do Paraná (UTFPR), como requisito parcial para obtenção do título de Engenheiro Eletricista.

Orientador: Prof. Dr. Antonio Carlos Pinho

CURITIBA

2018

ALVARO ALCEU FERREIRA SANTOS
LEONARDO AUGUSTO PIOVEZAN
RAFAEL COUTO DE OLIVEIRA

**ANÁLISE DE SINAIS DE ÁUDIO E IMPACTOS NA COLETA, CONVERSÃO E
CODIFICAÇÃO DE SINAIS**

Este Trabalho de Conclusão de Curso de Graduação foi julgado e aprovado como requisito parcial para a obtenção do Título de Engenheiro Eletricista, do curso de Engenharia Elétrica do Departamento Acadêmico de Eletrotécnica (DAELT) da Universidade Tecnológica Federal do Paraná (UTFPR).

Curitiba, 23 de novembro de 2018.

Prof. Antonio Carlos Pinhor, Dr.
Coordenador de Curso
Engenharia Elétrica

Profa. Annemarlen Gehrke Castagna, Ma.
Responsável pelos Trabalhos de Conclusão de Curso
de Engenharia Elétrica do DAELT

ORIENTAÇÃO

Antonio Carlos Pinho, Dr.
Universidade Tecnológica Federal do Paraná
Orientador

BANCA EXAMINADORA

Antonio Carlos Pinho, Dr.
Universidade Tecnológica Federal do Paraná

Glauber Gomes de Oliveira Brante, Dr.
Universidade Tecnológica Federal do Paraná

Marcelo de Oliveira Rosa, Dr.
Universidade Tecnológica Federal do Paraná

AGRADECIMENTOS

Gostaríamos de agradecer primeiramente ao excelentíssimo Prof. Dr. Antonio Carlos Pinho por ter aceitado nos orientar nesse trabalho e além disso nos emprestado seus equipamentos e casa para a realização da parte prática deste trabalho, nos acolhendo sempre com muita atenção. Também agradecemos a todos os nossos amigos e familiares que direta ou indiretamente nos deram apoio para a conclusão, não só desse trabalho, mas de todo o percurso de nossa graduação.

*“Só existe uma maneira de
se viver pra sempre, (..),
Que é compartilhando a sabedoria adquirida
E exercitando a gratidão, sempre.
É o homem entender que ele
é parte do todo.”*

Leandro Roque De Oliveira.

RESUMO

OLIVEIRA, Rafael, PIOVEZAN, Leonardo, SANTOS. Álvaro. Análise de sinais de áudio e impactos na coleta, conversão e codificação de sinais. Trabalho de Conclusão de Curso – Graduação em Engenharia Elétrica, Universidade Tecnológica Federal do Paraná. Curitiba, 2018.

Este trabalho consiste na pesquisa sobre áudios de alta fidelidade e se suas características técnicas são percebidas qualitativamente. Para tal, foram gravados áudios com o padrão normal de CDs, 48 kHz de frequência de amostragem e 16 bits de quantização, e os mesmos áudios com uma frequência de amostragem de 192 kHz e 24 bits de quantização. Foram realizadas duas análises, a primeira por meio de *software* comparando as diferenças matemáticas do sinal, onde foram encontradas diferenças na ordem de 0,0017% de potência. A segunda foi realizada por meio de um teste discriminativo, aplicando a distribuição qui-quadrada. Os resultados do teste levaram a conclusão de que a hipótese nula não foi rejeitada, sendo assim, não existe diferença significativa na percepção dos áudios.

Palavras-chave: Alta fidelidade, *Redbook*, teste sensorial, áudio.

ABSTRACT

OLIVEIRA, Rafael, PIOVEZAN, Leonardo, SANTOS. Álvaro. Audio Signals Analysis and impact on recording, conversion and coding of its signals. Trabalho de Conclusão de Curso – Graduação em Engenharia Elétrica, Universidade Tecnológica Federal do Paraná. Curitiba, 2018.

This work consists on a research on high fidelity audios and if its technical properties are perceived qualitatively. Audios in standard quality were recorded, 48 kHz for sampling frequency and 16 bits for quantization, and those same audios with 192 kHz for sampling frequency and 24 bits for quantization. Two analyzes were performed, the first using software comparing the mathematical differences of the signal, where differences were found in the order of 0.0017% power. the second was performed using discriminative test, and applying the chi-square distribution in the test results, was possible to conclude that the null hypothesis was not rejected, concluding that there is no significant difference in auditory perception.

Keywords: High fidelity, *Redbook*, sensory test, audio.

LISTA DE FIGURAS

Figura 1 - Fluxo de gravação analógica	16
Figura 2 - <i>Zoom</i> de uma forma de onda gravada	25
Figura 3 - Amostragem por meio de um trem de impulsos.....	26
Figura 4 - Espectro do sinal original e espectro do trem de impulsos.	28
Figura 5 - Quantização de uma senóide com 2 bits.	29
Figura 6 - Quantização de uma senóide com 3 bits.	30
Figura 7 - Diferença de resolução	30
Figura 8 - Diagrama de blocos representando a quantização de um sinal.....	32
Figura 9 - Processo de PCM.....	32
Figura 10 - Sinal sendo amostrado.	32
Figura 11 - Tabela de quantização e codificação.....	33
Figura 12 - Sinal amostrado, seu equivalente binário e o sinal PCM.....	34
Figura 13 - Esquema do microfone/e de bobina móvel.....	38
Figura 14 - Esquema do microfone de capacitor.....	39
Figura 15 - Esquemático da mesa de som.....	40
Figura 16 - Forma de onda obtida de um diapasão.....	42
Figura 17 - Forma de onda obtida de um clarinete.....	42
Figura 18 - Partes do Saxofone Alto	44
Figura 19 - Representação do aparelho auditivo.....	47
Figura 20 - Representação do órgão de Corti.	49
Figura 21 - Níveis de Probabilidade teste qui-quadrado.	51
Figura 22 - Formulário de julgamento	52
Figura 23 - Formulário de favorabilidade.....	53
Figura 24 – Diagrama de blocos das etapas da gravação	54
Figura 25 – Fotografia das etapas de gravação	54
Figura 26 - Microfone Behringer ECM8000®	55
Figura 27 - Diagrama direcional do microfone Behringer ECM8000®.....	56
Figura 28 - Mesa de Som Yamaha® AG06.....	56
Figura 29 - Escaleta	57
Figura 30 – Gravação do áudio da escaleta.....	57
Figura 31 - Saxofone Alto.....	57

Figura 32 - Gravação do áudio do saxofone	58
Figura 33 – Guitarra acústica clássica de 6 cordas.....	58
Figura 34 - Gravação do áudio do violão	59
Figura 35 – Imagens do ambiente de gravação	60
Figura 36 – Formato de onda do áudio da escaleta em 24 bits e 192 kHz e diferença entre as ondas sobrepostas (<i>zoom</i>).	64
Figura 37 – Formato de onda da diferença entre o áudio em 24 bits e o em 16 bits.	65
Figura 38 – Espectrograma do áudio da escaleta em 24 bits e 192 kHz	66
Figura 39 – Espectrograma do áudio da escaleta em 16 bits e 192 kHz	66
Figura 40 – Formato de onda do áudio do violão em 24 bits e 192 kHz e diferença entre as ondas sobrepostas (<i>zoom</i>).	68
Figura 41 – Formato de onda da diferença entre o áudio em 24 bits e o em 16 bits.	69
Figura 42 – Espectrograma do áudio do violão em 24 bits e 192 kHz	69
Figura 43 – Espectrograma do áudio do violão em 16 bits e 192 kHz	69
Figura 44 – Formato de onda do áudio do saxofone em 24 bits e 192 kHz e diferença entre as ondas sobrepostas (<i>zoom</i>).	71
Figura 45 – Formato de onda da diferença entre o áudio em 24 bits e o em 16 bits.	71
Figura 46 – Espectrograma do áudio do saxofone em 24 bits e 192 kHz.	72
Figura 47 – Espectrograma do áudio do saxofone em 16 bits e 192 kHz.	72
Figura 48 – Formato de onda do áudio da escaleta em 24 bits e 192 kHz e diferença entre as ondas sobrepostas (<i>zoom</i>).	74
Figura 49 – Espectrograma do áudio da escaleta em 24 bits e 192 kHz.	75
Figura 50 – Espectrograma do áudio da escaleta em 24 bits e 48 kHz.	75
Figura 51 – Formato de onda do áudio do violão em 24 bits e 192 kHz e diferença entre as ondas sobrepostas (<i>zoom</i>).	77
Figura 52 – Espectrograma do áudio do violão em 24 bits e 192 kHz.	78
Figura 53 – Espectrograma do áudio do violão em 24 bits e 48 kHz.	78
Figura 54 – Forma de onda do áudio do saxofone em 24 bits e 192 kHz e diferença entre as ondas sobrepostas (<i>zoom</i>).	80
Figura 55 – Espectrograma do áudio do saxofone em 24 bits e 192 kHz.	81
Figura 56 – Espectrograma do áudio do saxofone em 24 bits e 48 kHz.	81
Figura 57 – Forma de onda do áudio da escaleta em 24 bits e 192 kHz e diferença entre as ondas sobrepostas (<i>zoom</i>).	83
Figura 58 – Espectrograma do áudio da escaleta em 24 bits e 192 kHz.	84

Figura 59 – Espectrograma do áudio da escaleta em 16 bits e 48 kHz.	84
Figura 60 – Forma de onda do áudio do violão em 24 bits e 192 kHz e diferença entre as ondas sobrepostas (<i>zoom</i>).	85
Figura 61 – Espectrograma do áudio do violão em 24 bits e 192 kHz.	86
Figura 62 – Espectrograma do áudio do violão em 16 bits e 48 kHz.	86
Figura 63 – Forma de onda do áudio do saxofone em 24 bits e 192 kHz e diferença entre as ondas sobrepostas (<i>zoom</i>).	87
Figura 64 – Espectrograma do áudio do saxofone em 24 bits e 192 kHz.	88
Figura 65 – Espectrograma do áudio do saxofone em 16 bits e 192 kHz.	88
Figura 66 – Sala de Audição.	91
Figura 67 – Tabela de Resultados por Grupo	93
Figura 68 – Gráfico de média por Grupo	94

LISTA DE TABELA

Tabela 1 – Lista de <i>scripts</i> criados e sua respectiva função	61
Tabela 2 – Relação dos áudios exportados dos instrumentos.	62
Tabela 3 – Potência do áudio da escaleta.	64
Tabela 4 – Parâmetros da análise da escaleta.	64
Tabela 5 – Potência dos áudios do violão.....	67
Tabela 6 – Parâmetros da análise do violão	67
Tabela 7 – Potência dos áudios do saxofone.....	70
Tabela 8 – Parâmetros da análise dos áudios do saxofone.....	70
Tabela 9 – Potência dos áudios da escaleta com a mesma quantização.	74
Tabela 10 – Potência dos áudios do violão com mesma quantização.	76
Tabela 11 – Potência dos áudios do saxofone com mesma quantização.	79
Tabela 12 - Potência dos áudios da escaleta com quantização e frequência de amostragem diferentes.....	82
Tabela 13 – Potência dos áudios do violão com mesma quantização.	85
Tabela 14 – Potência dos áudios do saxofone com frequência e quantização diferentes.....	87
Tabela 15 – Tabela de Perfil dos avaliadores	93

LISTA DE SIGLAS

A/D	Analógico Digital
BIT	<i>Digito binário, menor parte da informação digital</i>
CD	<i>Compact Disc</i>
CLOCK	Ciclo do processador
DAELT	Departamento Acadêmico de Engenharia Elétrica
DSP	<i>Digital signal processor</i>
DVD	<i>Digital Versatile Disc</i>
IEC	International Electrotechnical Commission
MSE	<i>Mean Square Error</i>
PCM	<i>Pulse Code Modulation</i>
SNR	<i>Signal-to-Noise-Ratio</i>
UTFPR	Universidade Tecnológica Federal do Paraná

SUMÁRIO

1	INTRODUÇÃO	15
1.1	TEMA	15
1.1.1	Gravação Analógica	15
1.1.2	Gravação Digital	16
1.1.3	Padrões de áudio	17
1.1.4	Conversão Analógico Digital.....	18
1.1.5	Delimitação Do Tema	19
1.2	Problemas e Premissas.....	20
1.3	Objetivos.....	20
1.3.1	Objetivo geral	20
1.3.2	Objetivos específicos.....	21
1.4	Justificativa	21
1.5	Procedimentos Metodológicos.....	22
1.6	Estrutura Do Trabalho	23
2	CONVERSÃO ANALÓGICO-DIGITAL	24
2.1	Amostragem	24
2.2	Quantização	29
2.3	PCM	32
2.4	Signal to Noise Ratio (SNR)	34
2.5	Mean Square Error (MSE).....	35
3	DISPOSITIVOS DE GRAVAÇÃO	36
3.1	Microfone.....	36
3.2	Mesa De Som.....	38
3.3	Som e suas propriedades.....	39
3.3.1	Intensidade	39
3.3.2	Altura	40
3.3.3	Timbre	40
3.3.4	Fontes de sons musicais	41
4	TESTE QUALITATIVO.....	44

4.1	Dispositivos de Alta Fidelidade.....	44
4.2	Ouvido Humano.....	45
4.3	Análises sensoriais.....	48
5	GRAVAÇÃO.....	53
5.1	Microfone.....	54
5.2	Mesa de gravação	55
5.3	Instrumentos.....	56
5.4	Ambiente de gravação.....	60
6	ANÁLISE QUANTITATIVA	62
6.1	Matlab.....	62
6.2	Mesma frequência de amostragem e quantização diferente	64
6.2.1	Escaleta.....	65
6.3	Violão	69
6.3.1	Saxofone	72
6.4	Diferentes frequências e mesma quantização.....	76
6.4.1	Escaleta.....	76
6.4.2	Violão	79
6.4.3	Saxofone	82
6.5	Diferentes faixas de frequência e quantização	84
6.5.1	Escaleta.....	84
6.5.2	Violão	87
6.5.3	Saxofone	90
7	ANÁLISE SENSORIAL.....	93
7.1	Hipótese Nula	93
7.2	Audições.....	93
7.2.1	Objetos de teste	93
7.2.2	Avaliadores.....	94
7.2.3	Ambiente de teste.....	94
7.2.4	Realização do teste	95
7.3	Resultados obtidos.....	96
7.3.1	Teste Sensorial.....	96

7.3.2	Teste de favorabilidade	98
8	RESULTADOS E CONCLUSÕES	99
	REFERÊNCIAS.....	101
	Apêndice A – Ficha de identificação	104
	Apêndice B – ficha de avaliação	105
	Apêndice C – Códigos Matlab	106

1 INTRODUÇÃO

1.1 TEMA

1.1.1 Gravação Analógica

O início da ideia de gravação sonora tem seu princípio inicial na tradução da vibração das ondas sonoras naturais para outra linguagem. De acordo com Piccino (2005), o experimento de Thomas Young, no começo do século XVIII, de um aparelho que traduz graficamente as vibrações de um corpo vibrante, foi o ponto de partida para outros instrumentos que viriam a se tornar o que existe hoje em tecnologia de gravação sonora.

De forma notável, o experimento de Thomas Alva Edison, construído em 1877, foi um marco tanto para gravação quanto para reprodução sonora. O fonógrafo era capaz de gravar e após o final da gravação, reproduzir o som gravado, através de um diafragma acoplado ao bocal em forma de cone. Um cilindro, com folha de estanho, era girado manualmente enquanto o som ia sendo gerado, este que movimentava o diafragma e com o auxílio de uma agulha criava um sulco na superfície de estanho acoplada ao cilindro. Ao final da gravação era preciso apenas trocar a agulha (para uma agulha que não criasse o sulco, apenas o seguisse), gira o cilindro no sentido contrário e o dispositivo reproduzia o som que havia sido gravado na folha de estanho.

Segundo Marchi (2005), esse dispositivo sofreu inúmeros aperfeiçoamentos ao longo dos anos. Entretanto, a ideia principal da gravação foi utilizada por muitos anos como principal forma comercial de áudio, através dos discos de vinil. Isso aconteceu até o surgimento dos formatos digitais de gravação e reprodução, mais especificamente os *compact discs* (CDs), em meados dos anos 80.

Considerando o método mais atual, a gravação dos discos de vinil consiste na captação do sinal sonoro, convertido em sinal elétrico, este que é convertido em movimento mecânico e que através de agulhas forma sulcos na superfície do disco, permitindo a reprodução do sinal gravado inicialmente através do processo inverso.

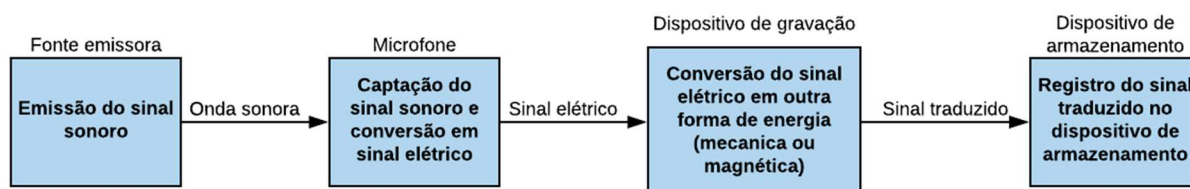
De acordo com Marchi (2005), juntamente com os discos de vinil, as gravações sobre fitas magnéticas eram soluções baratas e mais práticas (sistemas

menos robustos) e davam suporte para inovações tecnológicas, sendo maleáveis e possibilitando cortes e edições. As fitas magnéticas causaram grande impacto na cultura fonográfica, pois possibilitavam gravações não profissionais e domésticas, bem como incrementou a cópia e comercialização de material gravado.

As gravações em fitas magnéticas funcionam através da conversão do sinal sonoro em sinal elétrico, este que é convertido em campo magnético com certa intensidade (relacionada com a intensidade do sinal sonoro que gerou o sinal elétrico em outra intensidade que o gerou) e essa intensidade de magnetização orienta os elementos magnéticos da fita. Uma vez magnetizada, essa fita com certa magnetização é utilizada para reproduzir o sinal elétrico que a gerou (em um efeito inverso), este que por sua vez pode ser usado para gerar a onda sonora original.

Os métodos de gravação citados anteriormente possuem o mesmo princípio. Captam o sinal sonoro, traduzem para informações elétricas e essas são traduzidas aos dispositivos de armazenamento antes de serem reproduzidas, conforme a figura 1.

Figura 1 - Fluxo de gravação analógica



Fonte: Os autores

Apesar do sinal sonoro original sofrer conversões, estas são analógicas, ou seja, a informação armazenada e futuramente reproduzida é de um sinal contínuo, similar ao sinal gravado. Pode-se haver alterações de intensidades e frequências, porém ainda assim são sinais contínuos.

1.1.2 Gravação Digital

Como dito anteriormente, na gravação analógica a representação da onda sonora é armazenada diretamente no dispositivo de armazenamento. Com a evolução das tecnologias até a era digital, novas formas de armazenamento surgiram. Em 1983 o *CD* foi lançado sendo um suporte digital de reprodução, menor e mais leve quando comparado com os discos de vinil. A principal característica da

gravação digital é a representação da onda sonora em bits. Segundo Marchi (2005), essa tecnologia foi aplicada também à gravação magnética, facilitando e barateando ainda mais o processo.

Segundo o Teorema de Nyquist, o valor mínimo da frequência de amostragem deve ser o dobro da máxima frequência que se deseja armazenar. Como a frequência mais alta que o ouvido humano é capaz de perceber está em torno de 20.000 Hz (20 kHz), o valor de amostragem deve ser maior do que 40 kHz.

Segundo Iazzetta¹, relacionar a quantização com a taxa de amostragem permite analisar o tamanho do arquivo armazenado. Uma maior quantização (maior quantidade de bit por amostra) e uma maior taxa de amostragem implicam em arquivos maiores, necessitando de mais espaço alocado para armazenar o arquivo, mais tempo para ser transmitido por vias digitais e processado. O tamanho do arquivo pode ser calculado pela equação 1

$$S = \frac{f_a \cdot Q \cdot C \cdot t}{8}, \quad (1)$$

na qual S é o tamanho do arquivo em bytes, f_a é a frequência de amostragem, Q a quantidade de bits por amostra, C o número de canais utilizados (estéreo são dois canais e mono apenas um) e t é o tempo em segundos.

A representação mais fiel do sinal original seria um sinal digital com a maior taxa de amostragem e maior quantidade de bits por amostra possível. Contudo, pela equação 1, quanto maiores forem esses valores, maior será o espaço de armazenamento necessário e nos dispositivos esse espaço é limitado.

1.1.3 Padrões de áudio

De acordo com Silva (2018), existem os chamados Low-Res e *High-res* áudios. Ambos utilizam em sua definição o padrão *RedBook*. Áudios Low-res são aqueles que estão abaixo do padrão *RedBook*, ou seja, são quantizados com menos de 16 bits e suas frequências de amostragem são menores do que 44.1 kHz. Em

¹ Disponível em: <http://www2.eca.usp.br/prof/iazzetta/tutor/audio/a_digital/a_digital.html>. Acesso em: 22 set. 2018.

contrapartida, *High-res* áudios são definidos como qualquer áudio que tenha seus parâmetros a cima dos padrões *RedBook*.

O padrão *RedBook* é o formato de CDs de áudio. Foi lançado em sua primeira versão pelas empresas Philips e Sony, em 1980. Esse padrão foi aceito pelo *Digital Audio Disc Committee* e ratificado como uma IEC.

Nele todas as características, desde as físicas do CD, até as das informações que serão gravadas, são especificadas. É nesse documento que a taxa de amostragem de 44.1 kHz e a quantização de 16 bits são definidos como padrão. (IEC, 1999)

Assim, *High Resolution Audio* é um termo usado para se referir a áudios digitais que foram amostrados com frequências de amostragem maiores do que 44.1 kHz e/ou quantizados com mais de 16 bits. Ou seja, aumenta as especificações do formato *Redbook*, com o objetivo de guardar mais informações para que a reprodução do áudio seja a mais próxima do original possível.

Como Stuart (2015) menciona, áudio de alta fidelidade é aquele que se assemelha ao áudio produzido na vida real, ou seja, o arquivo de áudio não possui diferenças perceptíveis com relação a sua fonte original.

Ainda, no esforço de definir o que é um áudio de alta fidelidade, a Digital Entertainment, o Group, Consumer Technology Association, e a The Recording Academy se juntaram e o definiram como um áudio sem perdas que é capaz de reproduzir toda a gama de sons de gravações que foram masterizadas a partir de fontes de música melhores do que a qualidade de CD.

1.1.4 Conversão Analógico Digital

Existem várias formas de converter um sinal de analógico para digital, segundo SICA (2006) existem quatro tipos de conversores AD. O primeiro, chamado de “comparador paralelo”, é o mais rápido e o mais caro pois necessita de $2n-1$ comparadores para um conversor de n bits. O segundo conversor é o conversor AD por “aproximação sucessiva”, este que é o mais comum entre os conversores AD e que permite uma gama de 100 ou mais conversões por segundo. Para um sistema de n bits, o tempo de conversão é de n períodos de *clock*. O terceiro conversor é o

“rampa tipo contador” no qual o resultado é incrementado bit por bit até que o mesmo ultrapasse o valor analógico original. Nesse conversor podem ser perdidos inúmeros ciclos de *clock* até que se chegue ao valor aproximado. O último conversor é chamado de “rampa dupla tipo integrador” e é um conversor lento, porém de alta resolução e baixo custo, pois trabalha com comparadores, resistores e capacitores.

Sica (2006) afirma que nos dispositivos atuais apenas as ideias dos conversores citados anteriormente são seguidas, uma vez que é possível haver circuitos mais complexos e robustos e que acabam funcionando de maneira diferente, como por operação bipolar, onde é possível se indicar a polaridade da entrada. Porém, em todos os casos a taxa de conversão acaba dependendo da velocidade de oscilação do *clock* e da resolução desejada. A adição de outros elementos também possibilita a maior imunidade a ruídos e minimização de erros.

1.1.5 Delimitação Do Tema

Tendo em vista que as gravações digitais tomaram o mercado e que, em comparação com as analógicas, elas incorrem em perdas de informações do sinal gravado nas transformações analógico digitais é relevante saber quais configurações de áudio permitem que um sinal digital seja reproduzido fielmente ao sinal original, como uma gravação analógica.

Então, o foco deste trabalho é estudar se existem diferenças significativas entre áudios com frequências de amostragem de 192 kHz, com uma quantização de 24 bits e os áudios gravados em padrão *RedBook*.

Com este estudo entender se as diferenças geram perdas matemáticas mensuráveis entre os formatos e se elas impactam na experiência auditiva do usuário final.

Esse projeto não contempla análises sobre impactos na performance dos dispositivos de reprodução e gravação, tais como consumo energético, vida útil dos dispositivos, entre outros.

1.2 PROBLEMAS E PREMISSAS

A reprodução de um áudio envolve várias etapas, desde a sua captação e processamento até a sua eventual reprodução em algum dispositivo. Os arquivos de áudio gravados com frequências de amostragem maiores que 44,1 kHz e quantizados em taxas mais elevadas que 16 bits são comercialmente vendidos como se fossem melhores. A questão relevante é se realmente esses áudios, chamados de alta fidelidade, possuem diferenças perceptíveis aos humanos.

Para se avaliar tal questão, é necessário além de analisar os sinais quantitativamente, seu impacto qualitativo para quem os escuta. Assim, para realizar testes sobre esses áudios, as cobaias foram consideradas sem treinamento para detecção de diferenças em áudios e não sabiam se o áudio que foi reproduzido era um áudio com especificações normais ou de alta fidelidade.

Parte-se da premissa que áudios gravados com frequências maiores e com mais bits tem mais informações e, portanto, podem causar diferenças entre as sensações de um mesmo áudio gravado em frequências menores.

Não foram consideradas perdas que eventualmente possam existir em cabos e no ambiente e ainda ruídos existentes na hora das gravações, visto que todas as gravações foram feitas nas mesmas condições.

Cabe também ressaltar que para facilidade de comparação, o padrão *Redbook*, que é de 44,1 kHz foi considerado 48 kHz. Essa consideração foi feita para que essa frequência seja um múltiplo inteiro de frequências adotadas pelo mercado como de alta fidelidade, como 96 kHz e 192 kHz.

1.3 OBJETIVOS

1.3.1 Objetivo geral

Estudar as diferenças encontradas em áudios gravados no padrão *RedBook* que é o padrão atual de CDs, e áudios *Hi-Res* (alta qualidade). Para tal, realizar gravações com a qualidade *RedBook* e de alta fidelidade, comparando as duas em aspectos quantitativos e qualitativos.

1.3.2 Objetivos específicos

De forma a atingir o objetivo principal, foram estabelecidos os objetivos específicos a serem cumpridos de forma a se completarem e darem embasamento para o objetivo geral, sendo eles:

- Estudo dos formatos padrões de áudio e de alta fidelidade;
- Estudo das formas de gravação em alta fidelidade utilizando uma frequência de amostragem de 96 kHz e 24 bits ou maior;
- Estudos de equipamentos de gravação;
- Estudo dos conversores analógico digitais;
- Gravações de diferentes instrumentos em alta qualidade;
- Análise matemática dos sinais gravados;
- Comparação qualitativa entre os formatos dos sinais gravados por meio de um teste sensorial.

1.4 JUSTIFICATIVA

No contexto da alta fidelidade, cada etapa da reprodução de um áudio importa, desde a maneira como o áudio foi gravado, suas conversões analógico-digitais, seu processamento e por fim sua reprodução novamente em algum dispositivo.

Dito isso, para se obter uma qualidade alta, o estudo de como se realizar e melhorar cada etapa desse processo é primordial. Nesse contexto, a qualidade dos arquivos de áudio é uma das áreas de estudo de relevância.

Estudar o impacto qualitativo de parâmetros de áudio maiores que os habituais mostra se a diferença técnica se reflete em uma percepção sensorial. Também é relevante para se justificar ou não o uso desses parâmetros aumentados (frequência de amostragem e quantização) em determinadas situações, visto que quanto maiores os parâmetros, maior o tamanho do arquivo de áudio gerado. É importante saber se esse espaço extra agrega valor ou não para a reprodução de um sinal de áudio, mais especificamente nesse caso para a música.

1.5 PROCEDIMENTOS METODOLÓGICOS

Com base nos objetivos gerais e específicos estabelecidos foram organizadas as etapas a seguir:

Etapa 1 - Pesquisa Bibliográfica

Uma revisão bibliográfica sobre áudio e os fatores envolvidos em sua produção, captação e posteriormente reprodução. Também engloba a pesquisa de testes qualitativos para a avaliação sensorial da percepção de áudios.

Etapa 2 - Gravação

Após essa revisão, foram realizados os procedimentos para a gravação de áudios de diferentes instrumentos musicais, variando as suas frequências de amostragem e sua quantização.

Etapa 3 – Análise matemática dos sinais

Por meio do *software* Matlab © foi comparado os áudios gerados em diferentes qualidades a fim de encontrar diferenças nas propriedades de onda.

Etapa 4 – Produção de um teste sensorial

Por meio do *software* Audacity © foram editados os áudios gravados e produzir, utilizando a teoria sobre os testes qualitativos, um teste para comprovar se os áudios com frequências e taxas de quantização maiores conseguiram ser percebidos pelos testados.

Etapa 5 – Aplicação do teste sensorial

Aplicar o teste com o número mínimo estipulado pelo tipo de teste escolhido para que os dados tenham significância estatística desejada.

Etapa 6 – Análise estatística dos resultados

Com os dados do teste realizado, realizar a análise estatística e aceitar ou não a hipótese de que os áudios possuem diferenças significativas entre si ou não.

Etapa 7 – Documentação

Após a análise estatística, as informações foram documentadas, angariando resultados de simulações, modelos e diagramas, além dos resultados das comparações, encerrando o trabalho.

1.6 ESTRUTURA DO TRABALHO

Tendo definido as etapas e comparando-as com os objetivos definidos na seção 1.3, o trabalho fica organizado nos capítulos a seguir:

- Capítulo 1: Introdução do trabalho, seus objetivos, sua estrutura e contexto;
- Capítulo 2: Revisão bibliográfica e fundamentação teórica sobre o conversores AD e o impacto que a variação na taxa de amostragem e na taxa de bits podem gerar nas ondas obtidas pelos mesmos.
- Capítulo 3: Revisão bibliográfica e fundamentação teórica acerca dos principais conceitos do som. Engloba também a teoria sobre gravação de um sinal sonoro.
- Capítulo 4: Revisão bibliográfica sobre testes qualitativos
- Capítulo 5: Explicação do processo de gravação dos áudios
- Capítulo 6: Análise Quantitativa dos áudios obtidos
- Capítulo 7: Explicação do processo de criação de aplicação do teste qualitativo.
- Capítulo 8: Conclusões e considerações finais.

2 CONVERSÃO ANALÓGICO-DIGITAL

Verificando que o sinal original (onda sonora) é um sinal contínuo (analógico) e que se busca representá-lo por meio de uma linguagem digital, é necessária uma conversão do sinal analógico para o sinal digital, em bits. Conforme Zube (2004), o elemento responsável por essa função é o conversor analógico digital (A/D). Este trabalha coletando pequenas amostras periodicamente (samples) em curtos intervalos de tempo. Essas amostras são representadas em bits, numa sequência que se equivale à intensidade do sinal original naquela determinada posição de tempo a qual ele se refere. Esse processo é chamado de quantização e está diretamente relacionada com a resolução do sinal. Outros processos também estão relacionados a conversão A/D, estão são amostragem e codificação. O sinal digitalizado pode ser avaliado por alguns parâmetros. Neste projeto foram utilizados o *Signal to Noise Ratio (SNR)* e *Mean Square Error (MSE)*

2.1 AMOSTRAGEM

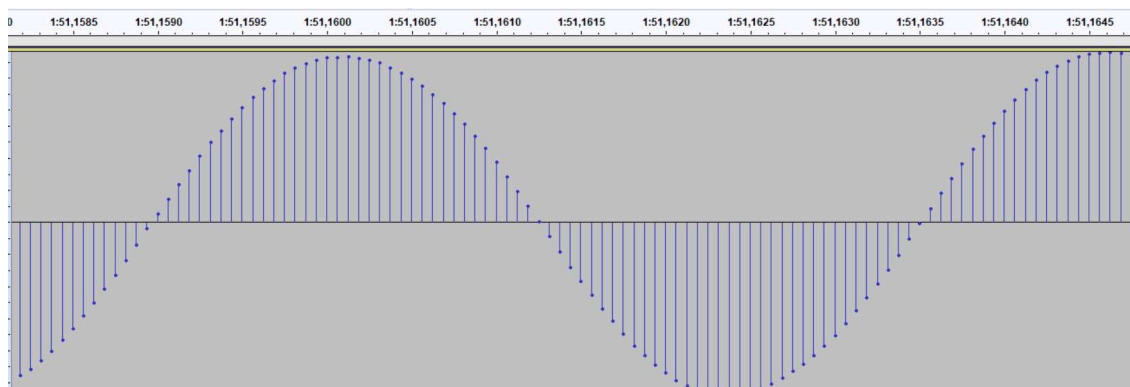
Na teoria de sinais, estes podem ser classificados como contínuos ou discretos. Matematicamente, os contínuos são aqueles que podem assumir qualquer valor real tanto no tempo como em sua magnitude. Já os discretos também podem assumir qualquer valor real em magnitude, porém tem gama de valores possíveis no tempo limitada aos números inteiros.

A maioria dos sinais presentes no mundo tem sua natureza contínua. Entretanto, a análise e processamento de sinais é mais vantajosa utilizando sinais discretos, devido à grande evolução da eletrônica digital e das capacidades, tanto de armazenamento como de processamento dos computadores.

Assim, a conversão de um sinal contínuo, para discreto, mais especificamente de um sinal analógico para um digital, é uma parte primordial na análise de qualquer sistema.

A amostragem pode ser exemplificada através de um *zoom* de uma forma de onda gravada (figura 2), onde é possível verificar os pontos nos quais o sinal original foi amostrado.

Figura 2 - Zoom de uma forma de onda gravada



Fonte: Os autores

O nome dado para o método de transformação de um sinal contínuo para discreto é amostragem, ou em inglês, *sampling*. Ele consiste em retirar uma sequência de amostras, igualmente espaçadas no tempo, de um sinal contínuo e o resultado é utilizado para representar o sinal.

Essa representação pode ou não conter todas as informações do sinal analógico, dependendo das características tanto do sinal original, como do método de amostragem. Isso interfere diretamente na reconstrução do sinal, visto que a falta de informações no sinal amostrado se reflete em distorções quando algum método de conversão digital-analógico for aplicado.

Segundo Oppenheim (1997), o princípio básico da amostragem pode ser modelado utilizando um sinal amostrador $p(t)$, que vai retirar, de tempos em tempos, uma medida do sinal original e gerar o sinal amostrado $x_p(t)$. Representando um sinal contínuo por uma função $x(t)$, tem-se que:

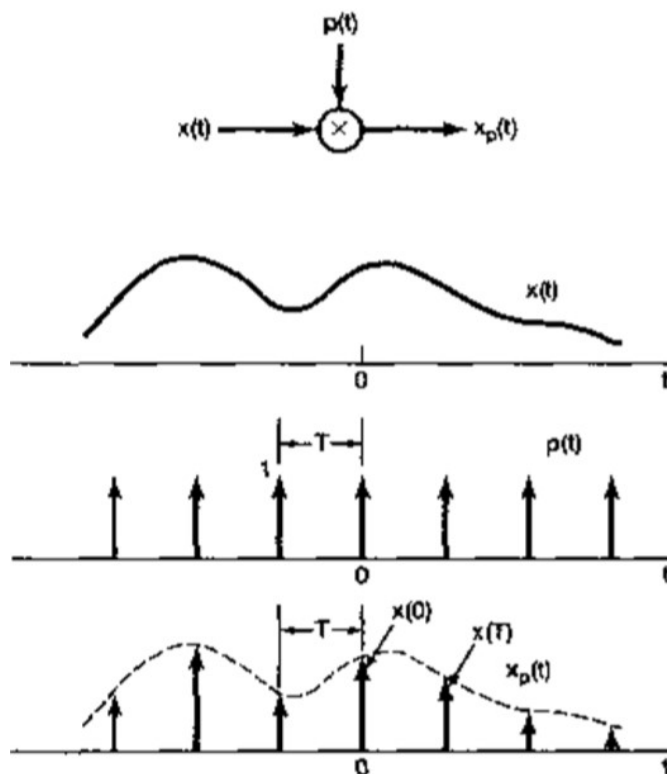
$$x_p(t) = x(t) \cdot p(t) . \quad (2)$$

O sinal $p(t)$ pode ter diversas representações conforme o tipo de amostragem. A representação ideal é a de um trem de impulsos de período T , que ao multiplicar o sinal original gera a sequência de amostras com o mesmo período do sinal amostrador,

$$p(t) = \sum_{n=-\infty}^{+\infty} \delta(t - n \cdot T) . \quad (3)$$

Graficamente, o sistema de amostragem por impulsos é mostrado na figura 3.. Essa modelagem aproveita a propriedade da função de impulso deslocada no tempo, que ao ser multiplicada por outra função, retorna o valor dessa função no tempo do impulso.

Figura 3 - Amostragem por meio de um trem de impulsos.



Fonte: OPPENHEIM (1997)

Substituindo a equação 3 na equação 2:

$$x_p(t) = x(t) \cdot \sum_{n=-\infty}^{+\infty} \delta(t - n \cdot T)$$

$$x_p(t) = \sum_{n=-\infty}^{+\infty} x(n \cdot T) \cdot \delta(t - n \cdot T). \quad (4)$$

Para analisar as características de frequência e energia dos sinais, o domínio da frequência se torna mais conveniente do que o do tempo. Logo, aplicando a transformada de Fourier em ambos os lados da equação 4 e usando as fórmulas e as propriedades da multiplicação:

$$X_p(j\omega) = \frac{1}{2 \cdot \pi} \cdot \sum_{n=-\infty}^{+\infty} X(j\omega) * \frac{2 \cdot \pi}{T} \delta(\omega - k \cdot \omega_s)$$

$$X_p(j\omega) = \frac{1}{T} \sum_{n=-\infty}^{+\infty} X(j\omega) * \delta(\omega - k \cdot \omega_s).$$

Como a convolução de uma função com a função impulso gera um deslocamento dela mesma, tem-se:

$$X_p(j\omega) = \frac{1}{T} \sum_{n=-\infty}^{+\infty} X(j(\omega - k \cdot \omega_s)). \quad (5)$$

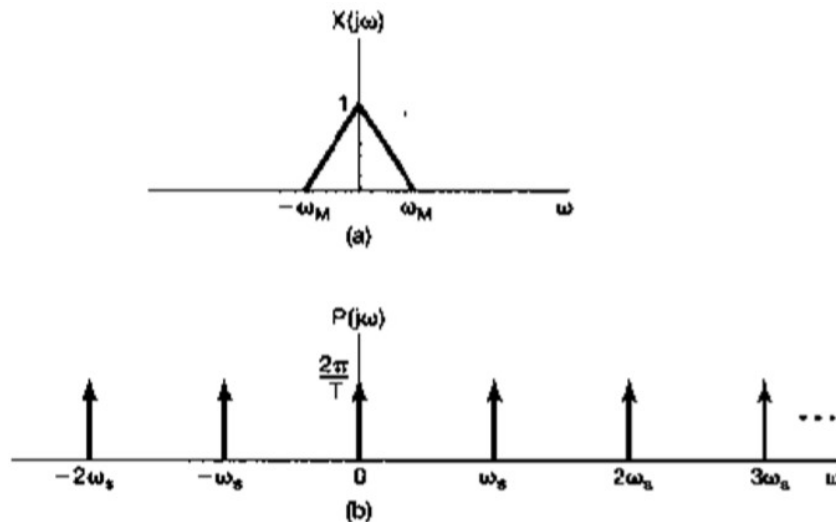
Onde ω_s é a frequência fundamental, ou seja, a frequência de amostragem, que, em termos do período é dada por:

$$\omega_s = \frac{2 \cdot \pi}{T}.$$

Com relação a equação 5, pode-se inferir que a transformada de Fourier do sinal amostrado é uma função periódica que consiste na replicação da transformada do sinal original transladada e escalada num fator de $1/T$.

Essas replicações podem ou não ficarem sobrepostas, o que depende da frequência de amostragem, visto que cada réplica fica deslocada $k \cdot \omega_s$ das outras. Considerando o espectro e o trem de impulsos da figura 4 a frequência final do espectro da ω_M é a constante que determina se haverá ou não sobreposição e por consequência perda de informações ao se amostrar um sinal. Isso só é válido para espectros que tenham sua banda limitada, ou seja, a partir de determinado valor de frequência seu espectro é 0.

Figura 4 - Espectro do sinal original e espectro do trem de impulsos.



Fonte: OPPENHEIN (1997)

Para que os repetidos espectros não se sobreponham:

$$\begin{aligned}\omega_M &< \omega_s - \omega_M \\ 2 \cdot \omega_M &< \omega_s.\end{aligned}\tag{6}$$

Portanto, se a inequação 6 for satisfeita, é possível extrair todo o espectro do sinal original do sinal amostrado. Caso contrário, ocorre a superposição de réplicas do espectro original, o que impede a sua completa extração sem perdas de informação.

Essa inequação é a base do teorema da amostragem ou teorema de Nyquist. Ele define que um sinal contínuo $x(t)$ com banda limitada é determinado unicamente por suas amostras, ou seja, essas amostras caracterizam o sinal caso a inequação 6 e a equação 7 sejam satisfeitas.

$$X(j\omega) = 0.\tag{7}$$

Para,

$$|\omega| > \omega_M,$$

consequentemente, caso o sinal atenda esses requisitos, é possível extrair o sinal original de seu sinal amostrado por meio de um filtro passa-baixas ideal com

ganho T , o que permite ter apenas uma réplica do espectrograma idêntica ao espectrograma do sinal original.

2.2 QUANTIZAÇÃO

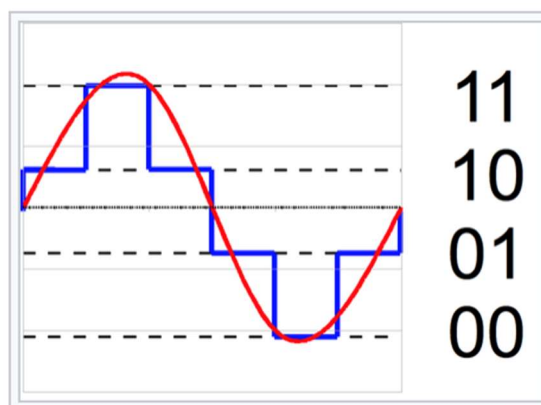
Sinais contínuos ou discretos no tempo podem ter qualquer amplitude. Como DSPs (*Digital signal processor*) e microprocessadores não conseguem trabalhar com amplitudes contínuas, um processo que as torne discretas é necessário, para o devido processamento. O nome dado para essa transformação é quantização.

Nessa transformação, a amplitude do sinal em determinado ponto é representada pelo valor mais próximo que o quantizador puder representar. Esse valor está intrinsecamente ligado a quantidade de bits utilizada para quantizar determinada amostra. O número máximo de diferentes valores que podem ser representados por uma quantização de n bits é representado pela equação 8:

$$m = 2^n. \quad (8)$$

No caso das figura 5 e figura 6, é possível comparar a diferença entre uma quantização com 2 bits e uma com 3 bits, respectivamente. A com 2 bits possui 4 valores possíveis e a com 3 possui 8. Isso faz com que a forma de onda com 3 bits se aproxime mais da onda original do que a com 2 bits.

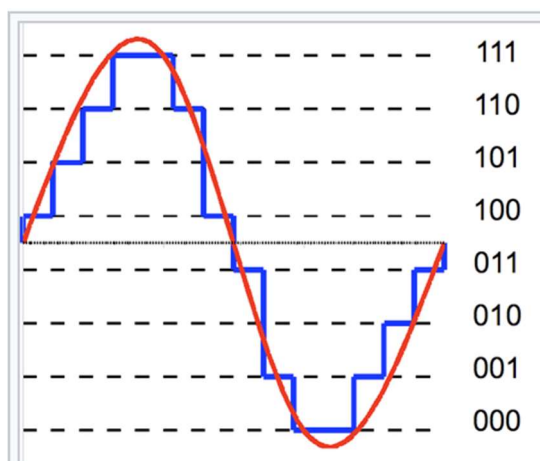
Figura 5 - Quantização de uma senóide com 2 bits.



Fonte: Página da Wikipédia sobre Quantização²

² Disponível em: <[https://en.wikipedia.org/wiki/Quantization_\(signal_processing\)](https://en.wikipedia.org/wiki/Quantization_(signal_processing))>. Acesso em 27 de outubro de 2018

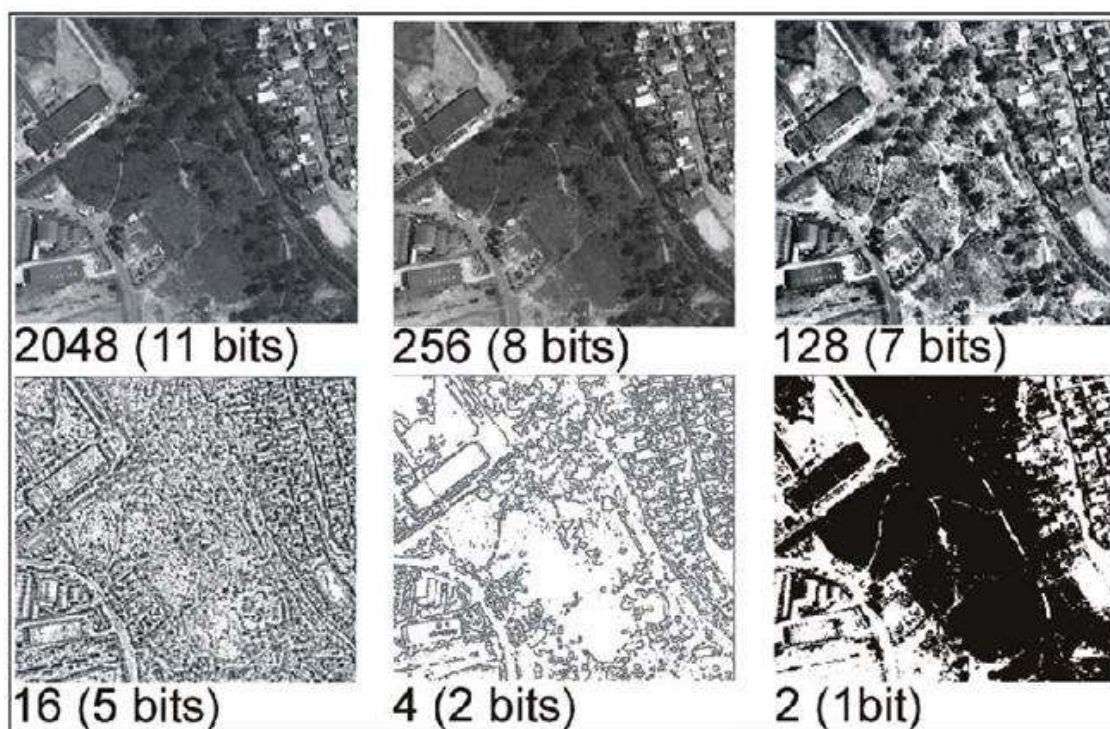
Figura 6 - Quantização de uma senóide com 3 bits.



Fonte: Página da Wikipédia sobre Quantização³

É possível ainda fazer uma analogia da taxa de bits com a captura de imagens e suas resoluções. Na figura 7 tem-se seis capturas diferentes da mesma área geográfica em diferentes resoluções. O número de bits interfere diretamente na quantidade de informações salvas pelo dispositivo de gravação, logo está relacionado também com a fidelidade do dado obtido em relação ao original. É possível verificar que a imagem em 11 bits na figura 7 tem muito mais detalhes e se aproxima mais da possível imagem real que a imagem em 1 bit.

Figura 7 - Diferença de resolução



Fonte: MELO (2002)

³ Mesma referência que a nota de rodapé 2.

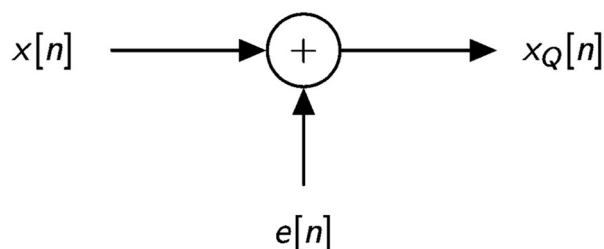
O processo de quantização é irreversível, visto que há uma quantidade limitada de valores que a quantização pode representar, em contrapartida ao número ilimitado de valores que o sinal a ser quantizado pode assumir. Assim, sempre haverá um erro intrínseco a esse processo, visto que um valor quantizado pode representar mais de um valor do sinal original. Definindo o erro de quantização como:

$$e[n] = x[n] - x_Q[n].$$

Logo, é possível ver que o erro de um sinal a ser quantizado depende do próprio sinal, sendo impossível reconstruir o sinal com perfeição após esse processo.

O erro de quantização pode ser considerado como a adição de um sinal ao sinal original, conforme o diagrama de blocos figura 8 mostra.

Figura 8 - Diagrama de blocos representando a quantização de um sinal.



Fonte: Os autores

2.3 PCM

Pulse Code Modulation (PCM) é um método de representação digital de sinais amostrados. Segundo OWEN (1982), PCM é uma técnica que consiste em amostrar o sinal em intervalos regulares de tempo e codificar o resultado em uma sequência de pulsos. O resultado é uma sequência de pulsos que representam valores binários que servem para reconstruir o sinal analógico original.

O processo de PCM é dividido em 3 etapas: amostragem, quantização e codificação. Cada uma delas pode ser feita de diversas maneiras e com variadas tecnologias e métodos. A figura 9 mostra as etapas do processo de PCM.

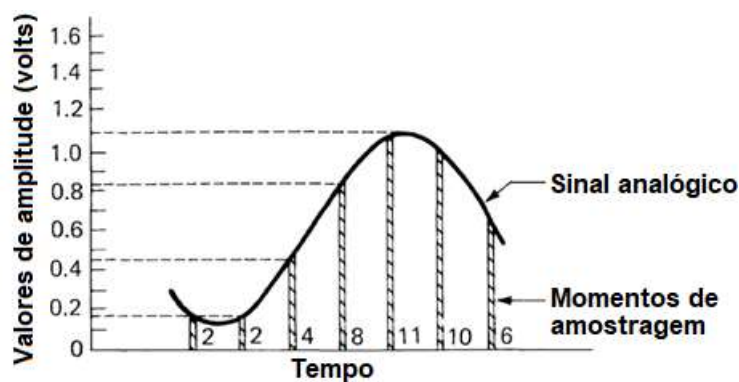
Figura 9 - Processo de PCM



Fonte: Os autores

Na figura 10 um sinal analógico é amostrado.

Figura 10 - Sinal sendo amostrado.



Fonte: Adaptado de OWEN (1982)

Na figura 11, uma quantização de 4 bits, que pode representar 16 valores é feita e cada um dos valores de amplitude foi codificado em um dos valores binários. Então, esses valores foram representados por pulsos, daí o nome de modulação por pulsos.

Figura 11 - Tabela de quantização e codificação.

Valor de amplitude	Código binário equivalente	Sinal "Pulse code modulate"
1	0 0 0 0	
2	0 0 0 1	
3	0 0 1 0	
4	0 0 1 1	
5	0 1 0 0	
6	0 1 0 1	
7	0 1 1 0	
8	0 1 1 1	
9	1 0 0 0	
10	1 0 0 1	
11	1 0 1 0	
12	1 0 1 1	
13	1 1 0 0	
14	1 1 0 1	
15	1 1 1 0	
16	1 1 1 1	

Fonte: Adaptado de OWEN (1982)

Então o sinal amostrado é representado pelo código de pulsos, sendo esse o sinal PCM, como mostrado na figura 12.

Figura 12 - Sinal amostrado, seu equivalente binário e o sinal PCM.

Valores decimais: 2 , 2 , 4 , 8 , 11 , 10 , 6
 Valores binários: 0 0 0 1, 0 0 0 1, 0 0 1 1, 0 1 1 1, 1 0 1 0, 1 0 0 1, 0 1 0 1,
 Sinal PCM:

Fonte: Adaptado de OWEN (1982)

2.4 SIGNAL TO NOISE RATIO (SNR)

Segundo Kieser (2005), a SNR (*Signal to Noise Ratio*) é uma medida que compara o nível de potência de um sinal com a potência do ruído dele. Definida como a razão entre essas duas potências e geralmente dada em decibéis. Essa medida se refere ao ruído gerado da quantização e não considera outros ruídos imputados por outros processos.

$$SNR = \frac{P_{sinal}}{P_{ruído}}$$

$$SNR_{dB} = 10 \log_{10}(SNR)$$

$$SNR_{dB} = 10 \log_{10} \left(\frac{P_{\text{signal}}}{P_{\text{ruído}}} \right).$$

É importante ressaltar que quanto maior for seu valor, menor é o impacto do ruído no sinal, visto que nesse caso a potência do sinal é preponderante sobre a do ruído. Pode-se considerar a diferença entre dois áudios exportados de maneira diferente como o ruído, considerando um deles como o sinal original. Logo, pode-se concluir que se um áudio exportado em uma configuração for exatamente igual a outro com outra configuração, o valor da SNR tenderá ao infinito, pois o ruído será nulo.

Cabe ainda ressaltar o conceito de variação dinâmica de um sinal. Pode ser definida como a razão entre o mais alto valor possível do sinal e a amplitude do menor valor possível, que nesse caso será o valor da amplitude do ruído. Considerando apenas o ruído de quantização

$$SQNR_{dB} = 20 \log_{10} \left(\frac{2^{n-1}}{1/2} \right) = 6,02 * N.$$

2.5 MEAN SQUARE ERROR (MSE)

Uma medida comum para analisar a adequação de um conjunto de pontos em determinada função é a chamada *Mean Square Error* (MSE). Para obter essa medida subtrai-se ponto a ponto os valores do conjunto de dados considerado verdadeiro X_i de um conjunto de valores estimado \hat{X}_i e eleva-se essa diferença, que pode ser considerada o erro absoluto, ao quadrado. Por fim, somam-se todos os erros ao quadrado e faz-se uma média com os N pontos calculados. Os valores são elevados ao quadrado para que erros positivos e negativos não se cancelem entre si. A fórmula abaixo demonstra esse conceito,

$$MSE = \frac{\sum_{i=1}^N (X_i - \hat{X}_i)^2}{N}.$$

No caso de uma análise entre dois áudios, um deles se torna o valor verdadeiro e o outro é comparado com esse.

Como é uma medida que soma diferenças, que nesse caso são consideradas erros, quanto maior o resultado da MSE, maior será a diferença dos áudios entre si.

Para que o valor da MSE possa ser mais significativo, pode-se tirar a raiz quadrada, gerando a *Root Mean Square Error* (RMSE). Com isso o valor perde a característica quadrática e pode ser considerado, ao multiplicar-se por 100, uma porcentagem de erro, sendo um valor mais fácil para comparações diretas.

3 DISPOSITIVOS DE GRAVAÇÃO

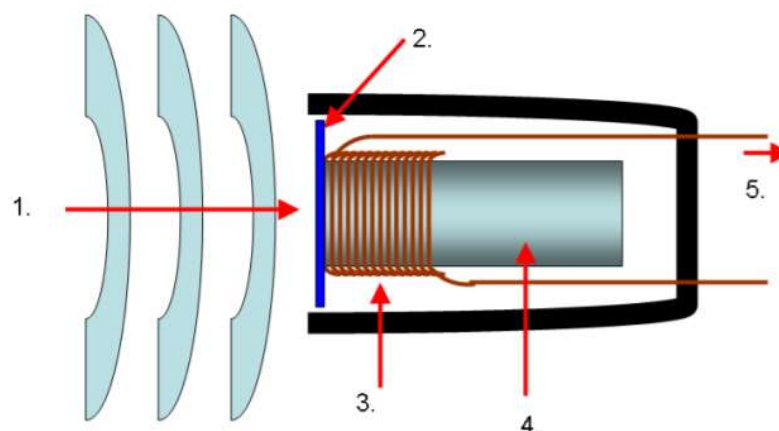
Com o objetivo de montar um teste prático para a comparação entre áudios com diferentes taxas de bits e frequências de amostragem, alguns itens são necessários para tirar conclusões. São eles: o dispositivo de reprodução desses áudios, o funcionamento da audição humana e como este processo ocorre no corpo humano. Com a compreensão dos parâmetros qualitativos envolvidos na dinâmica do teste, um modelo de testes pode então ser escolhido.

3.1 MICROFONE

O ponto em comum entre as gravações citadas anteriormente é a forma de captação do sinal sonoro e sua conversão em sinal elétrico. Segundo Gomes (2012), esse processo é realizado por um dispositivo transdutor eletroacústico, transformando a magnitude física da onda sonora em um sinal elétrico. Esses transdutores podem ser dos mais diversos tipos, sendo os de bobina móvel, piezoelétricos e eletrostáticos os mais usados atualmente. Nas gravações, o dispositivo mais comum para essa função é o microfone. O microfone comumente pode ser constituído de um transdutor eletroacústico de bobina móvel ou eletrostático (capacitor).

O microfone de bobina móvel é constituído de um diafragma acoplado fisicamente a uma bobina que possui em seu interior um ímã permanente, conforme a figura 13. A vibração das ondas sonoras que chega ao diafragma o faz vibrar, juntamente com a bobina. Pelo princípio da indução magnética, o movimento da bobina em relação ao ímã permanente cria uma corrente que é proporcional ao deslocamento da bobina móvel.

Figura 13 - Esquema do microfone/e de bobina móvel.



[1]. Ondas sonoras; [2]. Diafragma; [3]. Bobina móvel; [4]. Ímã permanente; [5]. Corrente elétrica

Fonte: GOMES (2012, pg 32)

Já o microfone de capacitor (ou condensador) é composto de duas placas, uma fixa e outra móvel (diafragma de baixa espessura), conforme a figura 14. A vibração das ondas sonoras que chega ao diafragma o faz vibrar e essa vibração o faz mover em relação à placa do capacitor, alterando a distância entre eles. A carga no capacitor é mantida constante, e quando há alteração da distância, entre as placas, a sua capacidade é alterada, alterando a tensão do capacitor. O elemento de polarização fornece uma tensão constante que é utilizada para comparar a variação ocorrida no capacitor. Esse comportamento pode ser explicado pela definição de capacitor de placas paralelas, conforme a equação 9 e a partir dela definir uma relação da tensão em função da distância entre as placas paralelas.

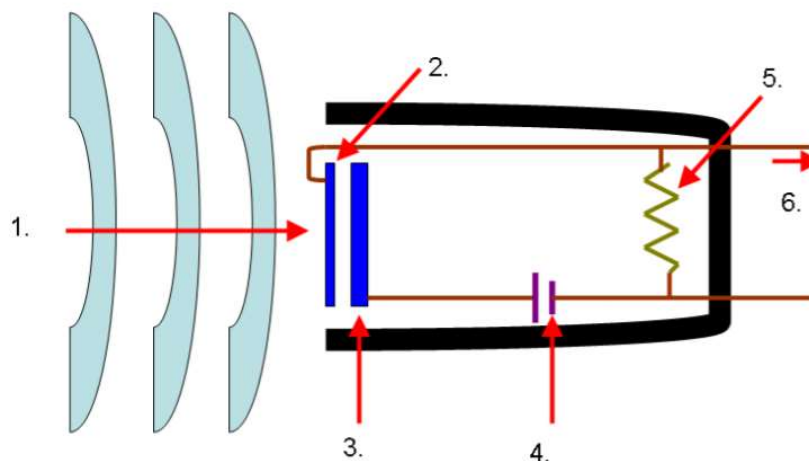
$$C = \frac{\epsilon_0 A}{d} = \frac{Q}{V} \quad (9)$$

$$V = \frac{Q \cdot d}{\epsilon_0 A}, \quad (10)$$

onde C é a capacidade do capacitor, ϵ_0 é a constante de permissividade elétrica no vácuo, A é a área das placas paralelas e d a distância entre elas.

A partir da equação 10 e do funcionamento do transdutor, onde a carga e área das placas são constantes, obtém-se uma função diretamente proporcional da tensão em relação à distância entre as placas.

Figura 14 - Esquema do microfone de capacitor.



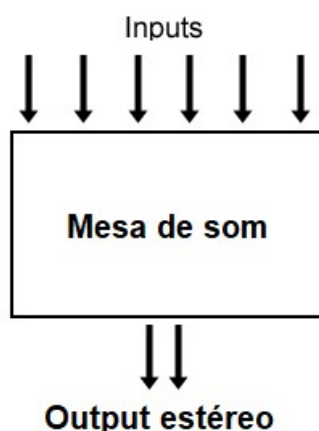
[1]. Ondas sonoras; [2]. Diafragma; [3]. Placa do capacitor; [4]. Polarização;
[5]. Resistência da carga; [6]. Corrente elétrica

Fonte: GOMES (2012, p.32)

3.2 MESA DE SOM

A mesa de som (também chamada de *mixer*) é um equipamento que tem como função primária somar duas ou mais entradas de áudio, provendo um ou mais sinais de saída (figura 15). Também possui inúmeras funções secundárias, como editar o sinal de saída, alterando certos parâmetros como ganho, equalização, adição de efeitos, entre outros efeitos.

Figura 15 - Esquemático da mesa de som



Fonte: Os autores

As mesas de som comumente possuem vários canais, pois elas permitem que duas gravações diferentes sejam feitas ao mesmo momento. Por isso, cada elemento a ser gravado deve ser destinado para uma entrada diferente. Como cada elemento pode possuir características diferentes de som, essa separação permite que para cada entrada se tenha um ajuste diferente, melhor adequando a entrada para o resultado final desejado.

Em casos de microfones condensados (aqueles que necessitam de alimentação para gerar o sinal elétrico) as mesas de som são capazes de gerar o chamado *phantom power*, normalmente na faixa de 48V.

As mesas ainda podem ser analógicas e digitais. As analógicas funcionam apenas com sinais analógicos enquanto as digitais podem funcionar tanto com sinais analógicos quanto digitais.

3.3 SOM E SUAS PROPRIEDADES

Segundo Halliday (2009), as ondas mecânicas precisam de um meio material para se propagar, e essas podem ser transversais ou longitudinais. As ondas transversais têm oscilações perpendiculares à direção de propagação da onda. Já as longitudinais oscilam na direção de propagação da onda. O som é uma onda mecânica longitudinal, produzida através da vibração de corpos. Essa vibração é transferida através dos meios físicos da mesma forma que qualquer onda mecânica. Segundo Martins (2010), o som também pode ser considerado, de forma subjetiva,

como uma sensação, percebida pelo sistema auditivo humano. Este que se propaga segundo uma onda senoidal, determinada por três propriedades objetivas. Essas propriedades são frequência, espectro e nível de pressão.

Através das propriedades objetivas, é possível relacionar propriedades subjetivas que correspondem às objetivas. As propriedades subjetivas são intensidade, altura e timbre.

3.3.1 Intensidade

A intensidade está relacionada com a energia de vibração da fonte emissora (pressão sonora). Quanto maior a pressão exercida, maior a intensidade da onda, logo, maior quantidade de energia é transmitida pela onda. Para diferentes frequências é necessário alterar o nível de pressão para que a mesma intensidade seja percebida.

De acordo com Halliday (2009), a intensidade de uma sonora em uma superfície é a taxa média por unidade de área com a qual a energia contida na onda atravessa ou é absorvida pela superfície, conforme a equação 11,

$$I = \frac{P}{A}, \quad (11)$$

onde P é a potência da onda sonora e A é a área da superfície que intercepta o som.

3.3.2 Altura

A altura é uma característica do som que está relacionada com a frequência da onda sonora. Essa característica que permite distinguir um som agudo (alta frequência) de um som grave (baixa frequência). Geralmente, um som é resultante da composição de diferentes frequências, chamadas de frequências harmônicas, sendo a frequência de menor valor (frequência fundamental) a que corresponde à altura do som.

3.3.3 Timbre

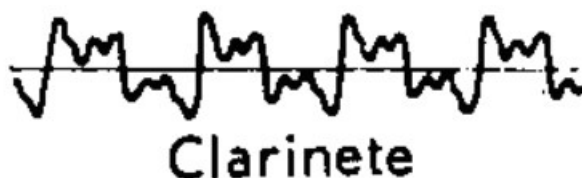
O timbre está relacionado com a característica da fonte emissora. Duas fontes diferentes podem emitir ondas com mesma intensidade e mesma frequência, porém soarem de forma distintas. Essa característica está relacionada a forma de onda ou espectro. Como a forma da onda é resultado da composição de várias frequências diferentes (frequência fundamental mais suas frequências harmônicas), o timbre é diferente para cada onda. Essa diferença entre timbres fica evidenciada na comparação das figura 16 e figura 17, respectivamente de um diapasão e clarinete.

Figura 16 - Forma de onda obtida de um diapasão



Fonte: Os autores

Figura 17 - Forma de onda obtida de um clarinete



Fonte: Os autores

3.3.4 Fontes de sons musicais

De acordo com Harley (2015), sons reais e musicais não são compostos de ondas senoidais simples, mas sim por ondas complexas, que são inúmeras ondas senoidais de diferentes frequências superpostas uma sobre as outras. Porém essa onda sonora terá uma componente de frequência fundamental. As demais frequências que a compõe são chamadas de harmônicas.

Segundo Halliday (2009), os sons musicais podem ser produzidos pelas oscilações de cordas, membranas, colunar de ar, blocos de madeira, barras de aço e muitos outros corpos. O som gerado nos instrumentos causa oscilações que são

propagadas para o ar em determinada frequência produzindo uma onda sonora audível.

Nos sistemas oscilatórios que produzem som musical, a frequência fundamental e um ou mais harmônicos superiores costumam ser gerados simultaneamente, superpondo-se e formando a onda resultante. Então quando diferentes instrumentos tocam a mesma nota, produzindo a mesma frequência fundamental, os harmônicos superiores podem ter intensidades diferentes. Por esse motivo, diferentes instrumentos produzem sons diferentes, mesmo quando tocados na mesma nota. Essa é característica que define o timbre.

Os instrumentos musicais de forma geral possuem três partes principais. A primeira parte é a que reproduz o som, a segunda que amplifica o som gerado e a terceira parte é responsável por modificar a onda gerada de acordo com a vontade do instrumentista. Na figura 18 tem-se um esquemático de um saxofone alto, com suas partes especificadas, onde o número 1 se refere à parte onde o som é reproduzido ou gerado, o número 2 se refere às partes onde o som é amplificado e por fim o número 3 se refere às partes nas quais o instrumentista atua de forma a alterar a onda gerada.

Figura 18 - Partes do Saxofone Alto



Fonte: Adaptado de Imagem disponível na página Blog Mundo Max⁴

Segundo Henrique (2011), os instrumentos ainda podem ser separados pela maneira em que o som é gerado. Os cordofones são instrumentos nos quais o som é gerado através da vibração de cordas tensionadas. A espessura e o comprimento das cordas impactam diretamente no som gerado. Já os aerofones são instrumentos onde som é gerado a partir de uma coluna de ar. O tamanho da coluna de ar é que define qual a característica do som final. Também existem os idiofones, que são instrumentos autorressonadores, ou seja, produzem som a partir da sua própria vibração. Nos membranofones o som é produzido pela vibração de uma membrana tensionada, ao ser percutida, raspada ou friccionada. Por último tem-se os eletrofones, que são instrumentos nos quais o som é gerado por meio de uma corrente elétrica, que por fim é convertida para som por meio de alto-falantes.

⁴ Disponível em: <<http://blog.mundomax.com.br/wp-content/uploads/2010/02/Partes-de-Saxofone.jpg>>. Acesso em 27 de outubro de 2018.

4 TESTE QUALITATIVO

Com o objetivo de montar um teste prático para a comparação entre áudios com diferentes taxas de bits e frequências de amostragem, alguns itens são necessários para tirar conclusões. São eles: o dispositivo de reprodução desses áudios, o funcionamento da audição humana e como este processo ocorre no corpo humano. Com a compreensão dos parâmetros qualitativos envolvidos na dinâmica do teste, um modelo de testes pode então ser escolhido.

4.1 DISPOSITIVOS DE ALTA FIDELIDADE

Dentre os dispositivos de reprodução mais utilizados, o mais simples são os fones de ouvido, devido a sua mobilidade e individualidade, são muito utilizados em ambientes públicos. Em sua grande maioria são itens passivos, que recebem um sinal e o reproduzem sem nenhum tratamento do dispositivo. Entretanto, se observadas suas especificações apresentam resposta de frequência entre 20 Hz a 20000Hz, ou seja, dentro do espectro audível.

Outros dispositivos comuns são as caixas portáteis, aparelhos que geralmente tem conexão wireless, bons drivers de reprodução e em suas especificações as faixas de frequência são as mesmas já citadas, dentro da faixa audível do ser humano, e estas ainda contam com filtros passa faixa, para que frequências inaudíveis evitando vibrações julgadas desnecessárias dos alto falantes.

No ambiente residencial ainda precisam ser citados os aparelhos de televisão, onde a evolução na qualidade de reprodução de imagem tem crescido vertiginosamente nos últimos anos. Porém, essa evolução não é refletida nas propriedades de áudio, com um sistema de qualidade muito próximo às caixas de som portáteis, operam de forma semelhante, com filtros que limitam a reprodução apenas às faixas audíveis na maioria dos modelos. Isso se deve ao fato de a faixa de 20 Hz a 20000 Hz ser amplamente utilizada na gravação de CDs de músicas, DVDs de filmes e show, e principalmente ao fato da transmissão de canais de TV seja via satélite ou cabo acontecerem em 16 bit 44000 Hz. O máximo de evolução que acontece é a utilização de vários canais de áudio, gerando uma melhor

sensação de posição da origem do som, porém sempre utilizando a taxa de bits e frequência já difundida.

Para este projeto foi utilizado um dispositivo de alta fidelidade, diferente desses citados acima para que todo o espectro de frequência gravado seja reproduzido, e assim as pessoas escolhidas tenham como perceber uma melhora na qualidade de reprodução devido à gravação em alta fidelidade. Todavia essa percepção depende do costume com áudio do usuário e também de propriedades fisiológicas do ouvido humano.

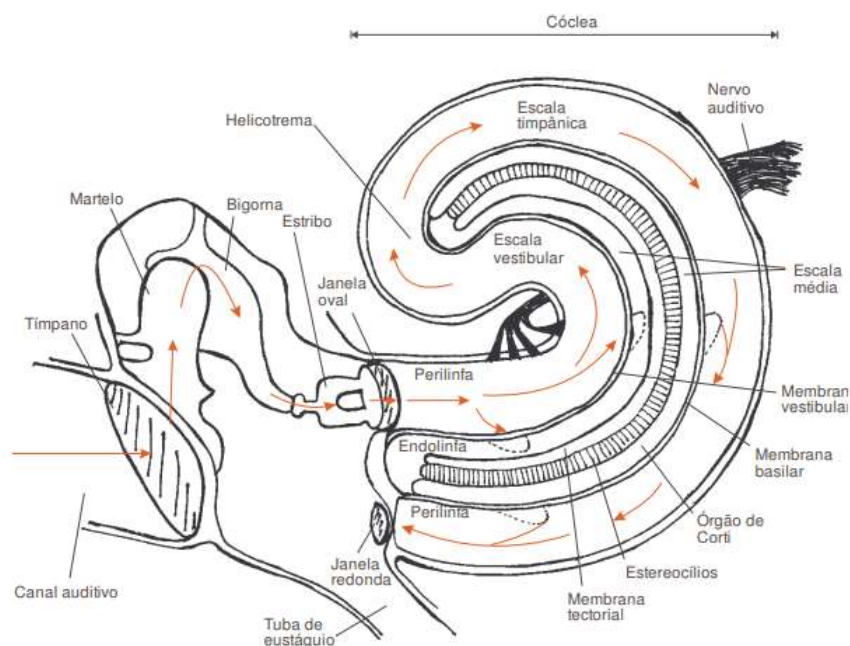
4.2 OUVIDO HUMANO

O processo de audição necessita de várias condições externas e internas ao ser humano. Externamente, um som audível deve ser gerado e propagado por algum meio, chegando assim ao aparelho auditivo.

Então, inicia-se o processo interno. O canal auditivo tem como função levar as informações até o nervo auditivo. Este, por fim, leva os sinais ao encéfalo que os traduz como o que é reconhecido como som (RITA, 2007).

A transmissão do som acontece pelo efeito da vibração no meio de propagação. Entretanto, nem toda vibração é audível para seres humanos. Dentre os diversos fatores que influenciam a audição, dois deles se destacam: intensidade e a frequência do som. Quando a intensidade do som é baixa, o som é fraco e gera dificuldade na audição, e essa dificuldade tende a aumentar com a idade. Com relação à frequência o ouvido humano, em geral, não é capaz de sentir frequências abaixo de 20 Hz, ou frequências acima dos 20 kHz, tendo seu limiar de audição influenciado pela idade também (RITA, 2007).

Figura 19 - Representação do aparelho auditivo



Fonte: RITA (2007).

A figura 19 é uma representação do aparelho auditivo humano sem a orelha. Como explanado em Rita (2007), o aparelho auditivo humano é composto por 3 partes principais: ouvidos externo, médio e interno. Cada uma dessas partes tem uma função específica na audição, sendo direcionamento e proteção no externo, recepção e amplificação no médio e conversão no interno.

O ouvido externo é composto pela orelha e pelo canal auditivo. A orelha tem um formato característico com curvas que formam um caminho de reflexão para que o som seja direcionado para o canal auditivo. Este tem aproximadamente 2,5 cm e serve como proteção para o tímpano. Dentro do canal há pelos e cera para exercer proteção contra sujeiras e, ao mesmo tempo, manter o caminho livre para que as ondas sonoras cheguem ao tímpano. Ainda, a diferença do tempo que um impulso sonoro chega quando ao tímpano diretamente e uma outra onda deste mesmo impulso chega pela reflexão causada pela orelha é o que permite a distinção da direção do som.

O ouvido médio é composto pelo tímpano e três pequenos ossos. O tímpano é uma membrana, com espessura de aproximadamente 0,1mm bem esticada, como se fosse a película de um tambor. Quando a oscilação do ar chega ao tímpano, ele vibra. Esta vibração é transportada pelos pequenos ossos do ouvido, o martelo, a bigorna e o estribo. O tímpano é conectado ao martelo, que por sua vez é conectado

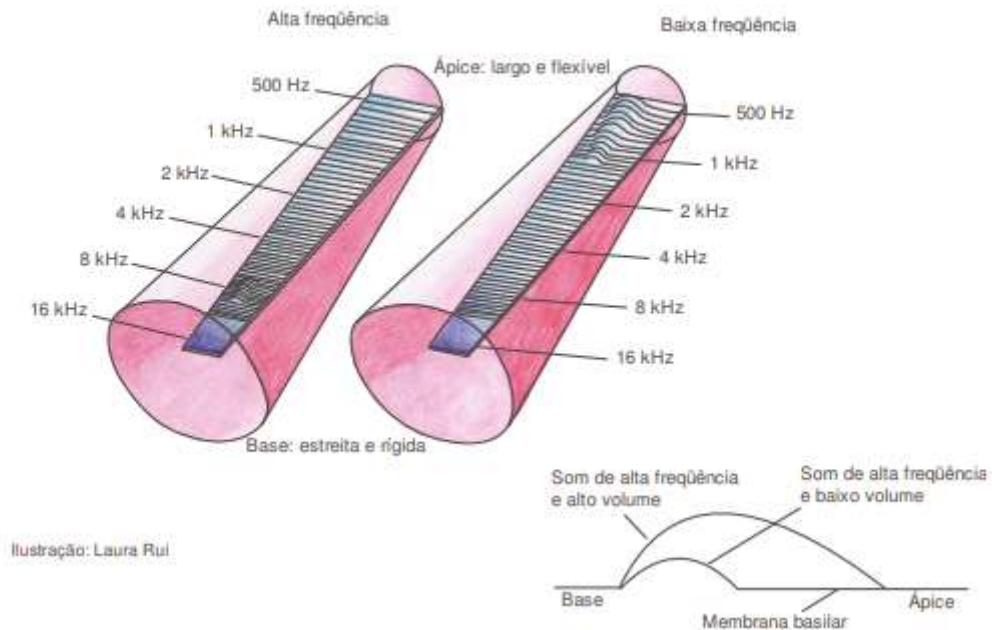
à bigorna, que por fim é conectada ao estribo. Esse conjunto de ossos tem a função de amplificar a vibração, pois no ouvido interno há um líquido e a pressão sonora necessária para fazer vibrar um líquido é maior do que a necessária para virar o ar. É dentro do ouvido médio também que ficam os músculos responsáveis pela atenuação de sons muito intensos, evitando danos ao ouvido. Entretanto, esses sistemas de proteção atuam na faixa de 50 e 100ms após a percepção de altas intensidades sonoras. Por este motivo, explosões e impulsos de som podem danificar a audição humana.

O ouvido interno é composto basicamente pela cóclea, vestíbulo e canais semicirculares. É nele que acontece a conversão da pressão sonora em impulsos elétricos para serem enviados ao cérebro.

A cóclea é um canal ósseo semelhante à uma concha cheia de água que se conecta ao ouvido médio pela janela oval. O estribo fica encaixado nessa janela evitando que o líquido que está dentro da cóclea saia, e faz esse líquido vibrar transmitindo as vibrações que perturbam o tímpano no ouvido externo. A vibração acontece porque o líquido que temos dentro do ouvido, chamado de perilinfa, é um líquido incompressível que transporta toda a vibração que chega a ele. O transporte é feito entre a escala timpânica e vestibular, que por sua vez transmite a vibração à membrana basilar, que são as paredes da escala média. Essa membrana tem 2 funções. A primeira é que sobre ela está o órgão de Corti, que é o órgão sensorial da audição e a segunda é que nessa membrana acontece a distinção da frequência dos sons.

Como ela tem espessura diferente na base e no ápice, há um comportamento igual para qualquer frequência. Contudo, para cada frequência existe um ponto de ressonância ao longo da membrana, que conseqüentemente é o ponto de maior amplitude de um som naquela frequência. Esse ponto excita o órgão de Corti, conforme observado na figura 20.

Figura 20 - Representação do órgão de Corti.



Fonte: RITA (2007).

O órgão de Corti possui fibras ciliadas sendo aproximadamente 3500 na parte interna distribuídas em uma única fileira e 20000 na parte externa distribuídas em 3 fileiras. Essas fibras ciliadas quando em movimento geram uma cadeia de reações eletroquímicas convertendo a vibração em impulso elétrico, que serão enviadas ao encéfalo pelo nervo auditivo.

4.3 ANÁLISES SENSORIAIS

Esse tipo de análise é aquele que visa os sentidos humanos como ferramenta para comparação de amostras. Existem vários tipos de análise estatística que podem ser feitas através de testes utilizando os sentidos humanos. É possível dividir essa categoria de análise em 3 classes:

- Discriminativo: Estabelecem diferenças quantitativas e qualitativas das amostras, são baseados em sensibilidade e medem os limites de percepção
- Descritivo: avaliação por escala de atributos e proporção, descrição de uma propriedade e sua intensidade.

- Afetivos: expressão opinião pessoal dos testados. Expressam confiabilidade de uma amostra frente a outras.

Dentro dos 3 grupos apontados, o utilizado no presente trabalho é o caso discriminativo, visto que inicialmente não há um parâmetro que defina se a diferença de qualidade obtida na gravação realmente é percebida por ouvidos humanos na reprodução. Logo, como o limiar de sensibilidade é desconhecido para as faixas de frequências não audíveis ao ouvido humano o teste discriminativo se faz necessário. Existem vários tipos de testes discriminativos que podem ser aplicados, sendo que os testes comparativos que mais se encaixam para este projeto são o teste triangular e o teste Duo-Trio.

O teste escolhido para este projeto foi o teste triangular. A NBR ISO 4120:2013 descreve o procedimento para determinar se existe uma diferença perceptível ou similaridade sensorial para alimentos e bebidas entre amostras de dois produtos utilizando o teste triangular. Ela não regulamenta para outros produtos, mas o procedimento metodológico é o mesmo.

Neste processo existe uma hipótese nula que, segundo Palermo (2015), é a hipótese, que afirma que não existe diferença significativa entre as amostras, essa é a hipótese a ser testada, e um teste é aplicado para validá-la. Em contrapartida existe a hipótese alternativa, que geralmente é a hipótese da pesquisa. Caso a hipótese nula seja rejeitada, tem-se que a hipótese alternativa é aceita.

Para a aplicação do teste triangular deve ser criada uma equipe julgadora com no mínimo 12 integrantes, sendo que o número ideal está entre 20 a 40 integrantes para a identificação de diferenças, como é o caso deste projeto. O teste é aplicado da seguinte forma: São apresentadas 3 amostras, e uma instrução de que 2 são iguais e uma diferente, sendo que o julgador deve indicar qual é o diferente. O teste não avalia o grau de diferença entre as amostras nem os agentes responsáveis pela diferença. Por isso, ele é o melhor para detectar pequenas diferenças.

Para o teste triangular existem duas possibilidades, o primeiro em que não existe obrigatoriedade de resposta, e um outro que o avaliador é obrigado a responder independente de realmente perceber a diferença nas amostras, este segundo é chamado de técnica da escolha forçada.

Neste projeto foi escolhida a técnica da escolha forçada, onde o avaliador deve dar uma resposta identificando uma das 3 amostras como a amostra diferente.

Isso permite aplicar a distribuição qui-quadrada para avaliar o resultado. Segundo Taylor (1997), a distribuição qui-quadrada avalia se uma amostra tem relação entre seus argumentos ou se o resultado é aleatório. Para tal, existe um certo número de amostras que deve ficar dentro da faixa esperada, conforme a figura 21. Quanto maior o número de respostas corretas menor o grau de liberdade e quanto menor o grau de liberdade, maior é a percepção de diferença.

Figura 21 - Níveis de Probabilidade teste qui-quadrado.

Nº total de julgamentos	Níveis de probabilidade (α)						
	5%	4%	3%	2%	1%	0,5%	0,1%
5	4	5	5	5	5	5	-
6	5	5	5	5	6	6	-
7	5	6	6	6	6	7	7
8	6	6	6	6	7	7	8
9	6	7	7	7	7	8	8
10	7	7	7	7	8	8	9
11	7	7	8	8	8	9	10
12	8	8	8	8	9	9	10
13	8	8	9	9	9	10	11
14	9	9	9	9	10	10	11
15	9	9	10	10	10	11	12
16	9	10	10	10	11	11	12
17	10	10	10	11	11	12	13
18	10	11	11	11	12	12	13
19	11	11	11	12	12	13	14
20	11	11	12	12	13	13	14
21	12	12	12	13	13	14	15
22	12	12	13	13	14	14	15
23	12	13	13	13	14	15	16
24	13	13	13	14	15	15	16
25	13	14	14	14	15	16	17
26	14	14	14	15	15	16	17
27	14	14	15	15	16	17	18
28	15	15	15	16	16	17	18
29	15	15	16	16	17	17	19
30	15	16	16	16	17	18	19
31	16	16	16	17	18	18	20
32	16	16	17	17	18	19	20
33	17	17	17	18	18	19	21
34	17	17	18	18	19	20	21
35	17	18	18	19	19	20	22
36	18	18	18	19	20	20	22
37	18	18	19	19	20	21	22
38	19	19	19	20	21	21	23
39	19	19	20	20	21	22	23
40	19	20	20	21	21	22	24
41	20	20	20	21	22	23	24
42	20	20	21	21	22	23	25
43	20	21	21	22	23	24	25
44	21	21	22	22	23	24	26
45	21	22	22	23	24	24	26
46	22	22	22	23	24	25	27
47	22	22	23	23	24	25	27
48	22	23	23	24	25	26	27
49	23	23	24	24	25	26	28
50	23	24	24	25	26	26	28
60	27	27	28	29	30	31	33
70	31	31	32	33	34	35	37
80	35	35	36	36	38	39	41
90	38	39	40	40	42	43	45
100	42	43	43	44	45	47	49

Fonte: ABNT, NBR 12995 (1993)

Na figura 22 é mostrado um modelo de formulário para ser entregue aos julgadores, considerando que eles receberam amostras aleatórias dentre os modelos a seguir: AAB, BAA, ABA, ABB, BBA e BAB, onde essas são as sequências possíveis de apresentar as amostras para os avaliadores, sendo que neste projeto umas das amostras foi o áudio de alta fidelidade e a outra o áudio padrão *RedBook*.

Figura 22 - Formulário de julgamento

Amostra:	Julgador:	Data:
Você está recebendo três amostras codificadas, sendo duas iguais e uma diferente. Identifique com um círculo a amostra diferente.		
_____	_____	_____
Comentários:		

Fonte: ABNT, NBR 12995 (1993)

Como o teste exige uma resposta do avaliador, para evitar que respostas aleatórias corretas sejam consideradas, um segundo teste, de favorabilidade, é aplicado ao mesmo tempo. Nele é verificado se o avaliador deu a resposta correta com consciência ou se ele apenas escolheu uma das 3 amostras ao acaso e acabou acertando. O formulário da figura 22 é entregue em conjunto com o da figura 23, onde ele deve responder se o elemento analisado em questão foi perceptivelmente diferente, respondendo entre 5 níveis sendo eles:

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

Figura 23 - Formulário de favorabilidade

Eu Consigo Perceber a diferença entre os Audios reproduzidos



Discordo totalmente ()



Discordo Parcialmente ()



Nem Concordo Nem Discordo ()



Concordo Parcialmente ()



Concordo Totalmente ()

Fonte: Os autores

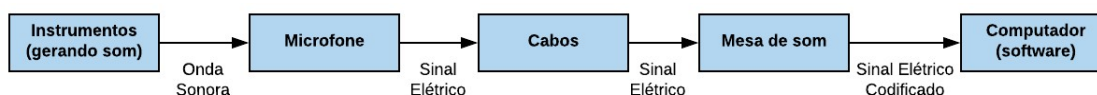
Analisando em conjunto ambos os testes, há a definição que se a resposta, mesmo que correta, estiver juntamente com uma das 3 primeiras respostas citadas, ela é considerada errada, visto que se o avaliador não consegue perceber diferença alguma nos áudios, a chance de um acerto ao acaso é maior. O foco é verificar os que realmente conseguiram distinguir alguma diferença do objeto de teste em questão.

5 GRAVAÇÃO

Foi definido que a gravação seria realizada com três instrumentos musicais diferentes, cada um sendo gravado individualmente.

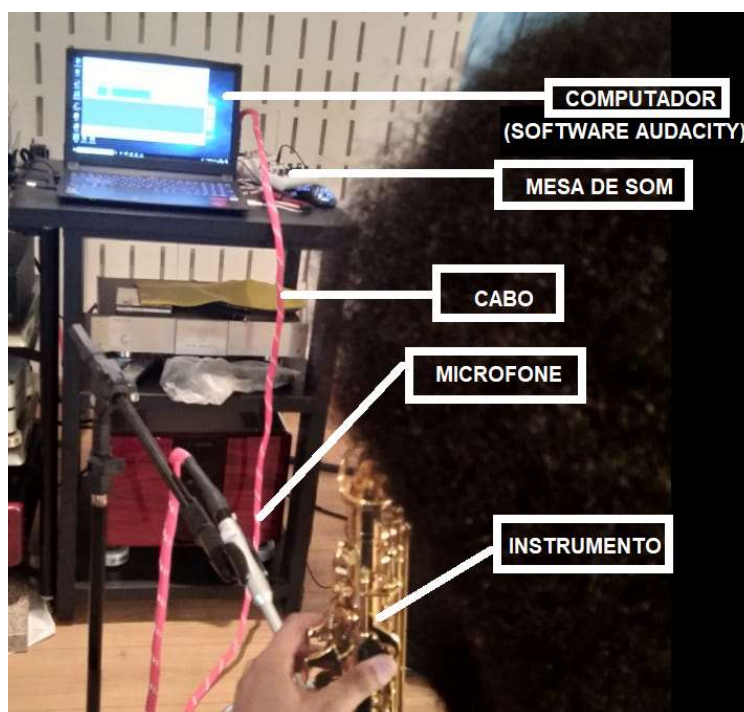
Assim, dois instrumentos de sopro e um de corda foram escolhidos. Respectivamente, saxofone, escaleta e violão. Os equipamentos necessários, além dos instrumentos, para realizar a gravação foram microfone, mesa de gravação, cabos e um computador. As etapas do processo de gravação estão descritas na figura 24. Já na figura 25 uma fotografia do processo de gravação realizado, com os elementos indicados é mostrada.

Figura 24 – Diagrama de blocos das etapas da gravação



Fonte: Os autores

Figura 25 – Fotografia das etapas de gravação



Fonte: Os autores

Foram gravados áudios dos instrumentos individuais, buscando a maior gama de tipos de som possíveis. Com o saxofone, foram tocadas seqüências de notas, percorrendo algumas escalas, alternando entre sons graves e agudos em seqüência.

No violão foram feitos dedilhados, para obter som referente a notas, alguns acordes e diferentes ritmos. Já na escaleta, assim como no saxofone, foram tocadas sequências de notas, percorrendo algumas escalas, alternando entre sons graves e agudos em sequência, também foram executados alguns acordes.

Durante a gravação foi utilizado um microfone, uma mesa de som e um computador para gravar o áudio em 192kHz / 24 bits. Para se obter o áudio em padrão *Redbook* foi utilizado o *software* Audacity© para *downscaling* (redução da frequência de amostragem e quantização). Dessa forma foi possível fazer apenas uma gravação em alta qualidade e utilizar o *software* para gerar, a partir desta, um novo áudio em menor qualidade.

5.1 MICROFONE

Para as gravações em 24bits/192kHz foi utilizado o microfone Behringer ECM8000® (figura 26). Segundo seu manual, é um microfone onidirecional, condensador (*phantom power* de +15V até +48V). Sua curva de resposta em frequência vai de 15 Hz até 20kHz.

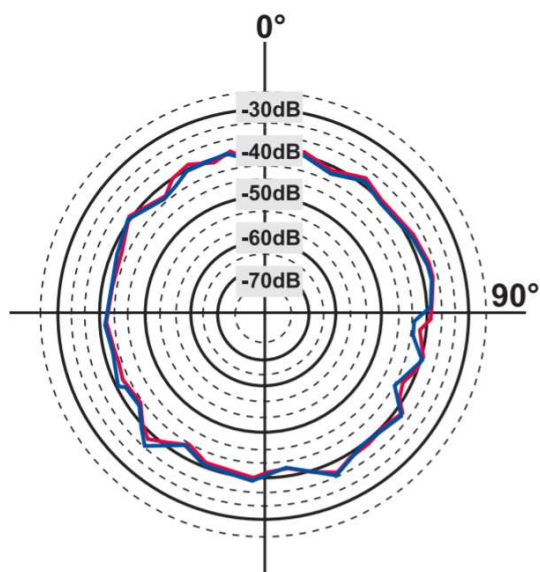
Figura 26 - Microfone Behringer ECM8000®



Fonte: Página de venda do microfone no site Audiodriver⁵

⁵ Disponível em: <<http://www.audiodriver.com.br/produto/microfone-condensador-ecm8000-behringer/22639>>. Acesso em 27 de outubro de 2018.

Figura 27 - Diagrama direcional do microfone Behringer ECM8000®.



Fonte: Datasheet Behringer ECM8000®.

5.2 MESA DE GRAVAÇÃO

Para as gravações efetuadas em 24bits/192kHz, foi utilizada a mesa de som Yamaha® AG06 (figura 28).

Figura 28 - Mesa de Som Yamaha® AG06



Fonte: Os autores

5.3 INSTRUMENTOS

Levando em consideração a disponibilidade dos instrumentos e dos instrumentistas, foram escolhidos três instrumentos, sendo estes escaleta (popularmente conhecido como melódica, clavieta ou pianica), saxofone e violão.

Para a gravação da escaleta foi utilizada uma escaleta de três oitavas. Ela é um aerofone de palhetas livres, similar a uma gaita de boca (harmônica) ou a um acordeão. As notas são produzidas através de uma coluna de ar, que é gerada por um orifício numa extremidade do instrumento, e das teclas escolhidas pelo instrumentista. O instrumento possui um teclado análogo a um piano, o qual permite a execução de notas e acordes.

Figura 29 - Escaleta



Fonte: Imagem disponível na página sobre Escaleta da Wikipédia⁶

⁶ Disponível em: <<https://pt.wikipedia.org/wiki/Escaleta>>. Acesso em 04/11/2018

Figura 30 – Gravação do áudio da escaleta



Fonte: Os autores

A gravação do saxofone utilizou um saxofone alto, afinação natural em mi bemol (Eb). O saxofone é um aerofone e instrumento de sopro, onde o som é gerado a partir de uma coluna de ar e uma palheta de bambu. Nesse instrumento as notas devem ser geradas separadamente, impossibilitando a presença de acordes (conjuntos de notas que quando combinadas lembram uma nota específica).

Figura 31 - Saxofone Alto



Fonte: Imagem disponível na página sobre Saxofone alto da Wikipédia⁷

⁷ Disponível em: <https://pt.wikipedia.org/wiki/Saxofone_alto>. Acesso em 30/10/2018

Figura 32 - Gravação do áudio do saxofone

Fonte: Os autores

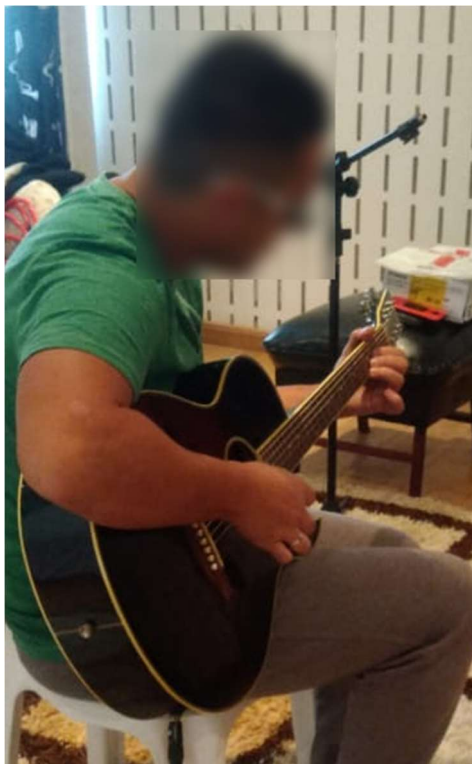
Para gravação do violão foi utilizado uma guitarra acústica (popularmente chamada de violão), com saída de som, em cordas de aço, com afinação padrão (E, A, D, G, B, E). O violão é um cordofone e seu som é gerado a partir das 6 cordas tensionadas. O som é manipulado através do braço do violão, onde cada posição reduz ou aumenta o comprimento da corda, produzindo uma frequência de onda estacionária diferente. A presença das 6 cordas permite a reprodução de até 6 notas diferentes ao mesmo tempo, logo, possibilitando a reprodução de acordes.

Figura 33 – Guitarra acústica clássica de 6 cordas

Fonte: Imagem disponível na página sobre Guitarra acústica da Wikipédia⁸

⁸ Disponível em: < https://pt.wikipedia.org/wiki/Guitarra_ac%C3%BAstica>. Acesso em 30/10/2018

Figura 34 - Gravação do áudio do violão



Fonte: Os autores

5.4 AMBIENTE DE GRAVAÇÃO

As paredes e teto do ambiente de gravação são revestidos com painéis de gesso destinados à absorção sonora. Esse revestimento, similar ao *drywall*, fornece uma maior absorção sonora quando comparado com outros revestimentos (alvenaria de tijolos por exemplo). Na figura 35. tem-se algumas imagens do ambiente de gravação, focadas nos detalhes das paredes e nas placas do teto.

Figura 35 – Imagens do ambiente de gravação



Fonte: Os autores

6 ANÁLISE QUANTITATIVA

6.1 MATLAB

O *software* Matlab permite produzir *scripts* de código para a realização de tarefas. Aproveitando essa característica e as bibliotecas já embutidas nele para a análise de sinais, foi possível gerar os *scripts* do apêndice C. Com eles, os áudios após serem gravados, foram analisados.

O código foi dividido nas funções da tabela 1 para aumentar a organização e dividir as etapas do processo de análise. As funções já prontas do Matlab foram encapsuladas em funções para facilitar seu uso e permitir, caso fosse necessário, a troca do algoritmo com mais facilidade.

Tabela 1 – Lista de *scripts* criados e sua respectiva função

Arquivo	Função
analiseAudio With.m	Rotina que gera os parâmetros e gráficos sobre os áudios.
getAudio.m	Rotina que retorna o arquivo de áudio dados o seu caminho e o seu tipo.
getAudioPower.m	Rotina que calcula a potência média do áudio.
getMSE.m	Rotina que calcula a MSE entre 2 áudios.
getSNR.m	Rotina que calcula a SNR entre os áudios.
mainTCC.m	Rotina que inicia a análise dos áudios.
plotAudio.m	Rotina que gera o gráfico do áudio passado como parâmetro.
plotSpectrogram.m	Rotina que gera o espectrograma do áudio passado como parâmetro.

Fonte: Os autores

Com os arquivos gravados com o auxílio do *software* Audacity foi possível exportar a gravação no formato WAV, que é um formato sem compressão, com diferentes taxas de amostragem e quantização. Essa transformação é necessária para que o *software* Matlab possa os reconhecer e analisar.

Como existem duas variáveis que podem ser modificadas, a quantização e a frequência de amostragem, a análise foi dividida em 3 partes. Na primeira, os áudios

foram exportados com a mesma frequência de amostragem e uma taxa de bits diferente. Na segunda a quantização foi mantida constante e a frequência de amostragem foi modificada. E por fim, ambos os parâmetros foram modificados.

Os áudios gerados foram analisados e comparados com as seguintes características:

- Formato da onda
- Espectrograma de frequências do sinal
- Potência do sinal
- MSE e SMSE
- SNR

A MSE e a SNR necessitam que os áudios tenham o mesmo número de amostras, pois fazem uma comparação ponto a ponto. Assim, para os casos em que as frequências de amostragem são diferentes, esses parâmetros não foram calculados, visto que uma interpolação poderia introduzir mais um erro e atrapalhar a análise.

Dentre as gravações realizadas, 3 foram escolhidas para uma análise matemática. Uma de cada um dos instrumentos (escaleta, violão e saxofone). Cada uma delas foi exportada com diferentes taxas de amostragem e quantização. A relação das combinações que foram exportadas está na tabela 2.

Tabela 2 – Relação dos áudios exportados dos instrumentos.

Instrumento	Duração (s)	Frequência (kHz)	Quantização (bits)	Tamanho (kB)
Escaleta	34	48	16	3195
	34	48	24	4792
	34	192	16	12778
	34	192	24	19167
Violão	30	48	16	2903
	30	48	24	4354
	30	192	16	11609
	30	192	24	17414
Saxofone	39	48	16	3727
	39	48	24	5590

Instrumento	Duração (s)	Frequência (kHz)	Quantização (bits)	Tamanho (kB)
	39	192	16	14906
	39	192	24	22359

Fonte: Os autores.

Para comparar a diferença de tamanho dos arquivos no padrão *RedBook* com o de alta fidelidade em 24bit 192kHz utiliza-se a fórmula 1 e como mostrado abaixo o áudio *high-res* é da ordem de seis vezes maior.

Áudio Redbook: 96 kbps,

$$S_{RedBook} = \frac{48000 \cdot 16 \cdot 1 \cdot t}{8} = 96000 \cdot t \text{ e,}$$

high-res em 24bit 192kHz: 576 kbps,

$$S_{hi-fi} = \frac{192000 \cdot 24 \cdot 1 \cdot t}{8} = 576000 \cdot t.$$

6.2 MESMA FREQUÊNCIA DE AMOSTRAGEM E QUANTIZAÇÃO DIFERENTE

Como os áudios foram gravados em 24 bits e 192 kHz e o processo de transformação consiste em abaixar os parâmetros, a frequência de amostragem de 192 kHz foi mantida e uma versão de cada áudio com 16 bits foi gerada. Após esse processo, os *scripts* do Matlab foram utilizados para gerar as informações sobre os áudios individualmente e comparativamente. Como os áudios possuem a mesma frequência de amostragem, foi possível subtrair a onda de 16 bits da onda da onda de 24 bits e obter a onda resultante das diferenças.

Nesse ponto foram analisadas as potências dos sinais, suas diferenças matemáticas e visualizados os espectrogramas.

6.2.1 Escaleta

Nessa análise foram utilizados os áudios da escaleta exportados ambos com 192 kHz. Sua diferença está na quantização, um deles foi quantizado com 24 bits e outro com 16 bits. Considerando o de 24 bits como o original, foi possível calcular a diferença direta e os parâmetros da tabela 4.

Tabela 3 – Potência do áudio da escaleta.

Áudio	Potência
Escaleta 24 bits/192 kHz	1.48027093e-03
Escaleta 16 bits/192 kHz	1.48027487e-03

Fonte: Os autores.

Tabela 4 – Parâmetros da análise da escaleta.

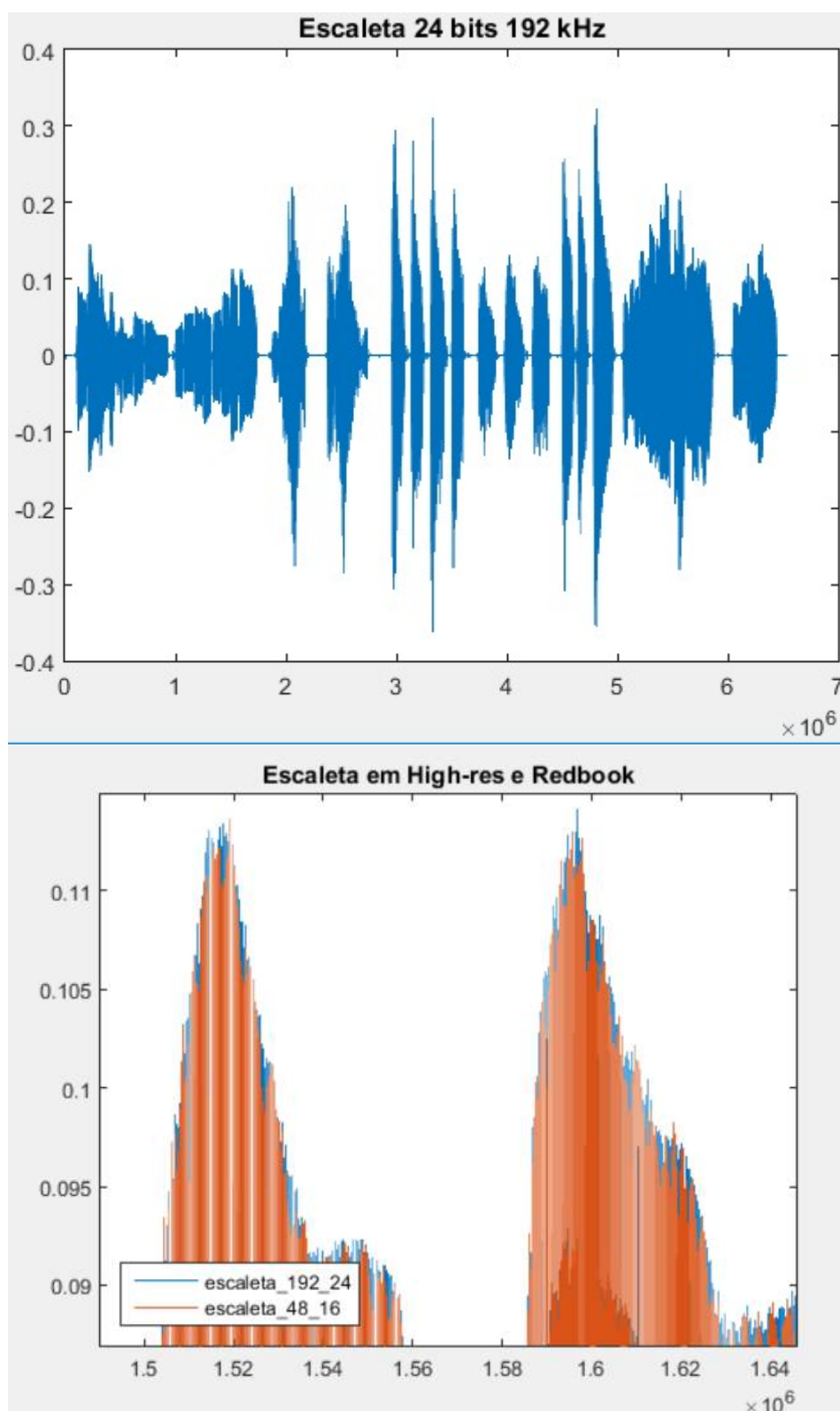
Diferença	SNR	MSE	SMSE
-3.9423e-09	55.8371	3.8603e-09	6.2131e-05

Fonte: Os autores.

As figuras a seguir mostram os formatos de onda e os espectrogramas de ambos os áudios. Ainda há a forma de onda da diferença entre eles.

Na figura 36 tem-se o formato de onda do áudio da escaleta em 24 bits/192 kHz, obtido pelo *script* do Matlab, e um *zoom* feito com as duas ondas sobrepostas, de forma a verificar sua diferença.

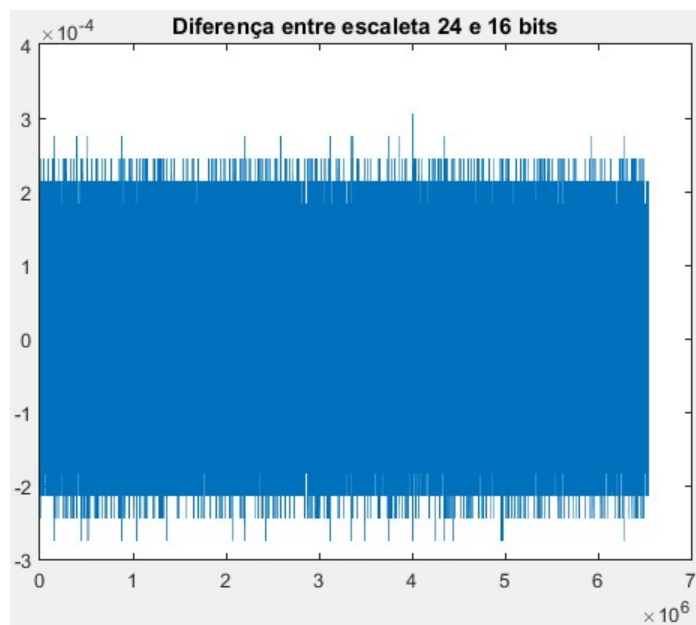
Figura 36 – Formato de onda do áudio da escaleta em 24 bits e 192 kHz e diferença entre as ondas sobrepostas (zoom).



Fonte: Os autores.

Como os áudios possuem a mesma frequência de amostragem, foi possível subtrair a onda de 16 bits da onda da onda de 24 bits. A onda de diferença é mostrada na figura 37.

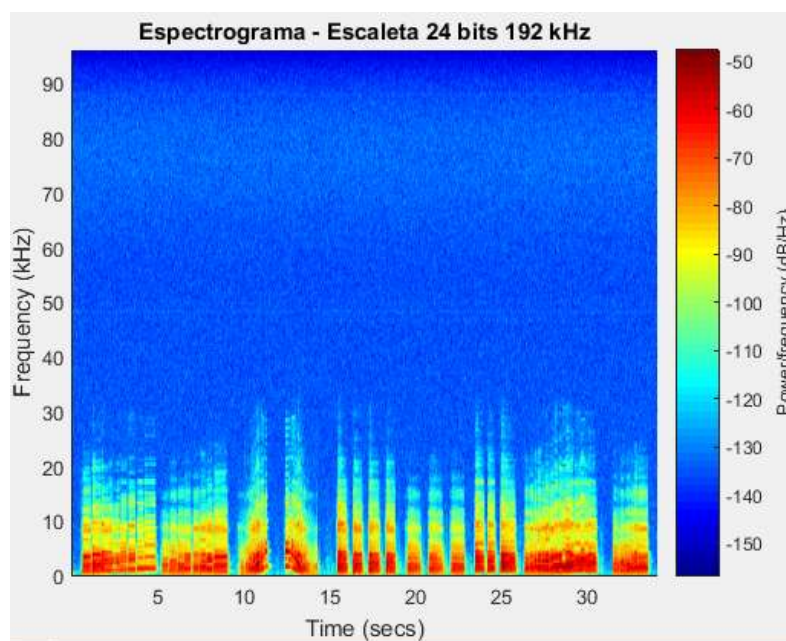
Figura 37 – Formato de onda da diferença entre o áudio em 24 bits e o em 16 bits.



Fonte: Os autores.

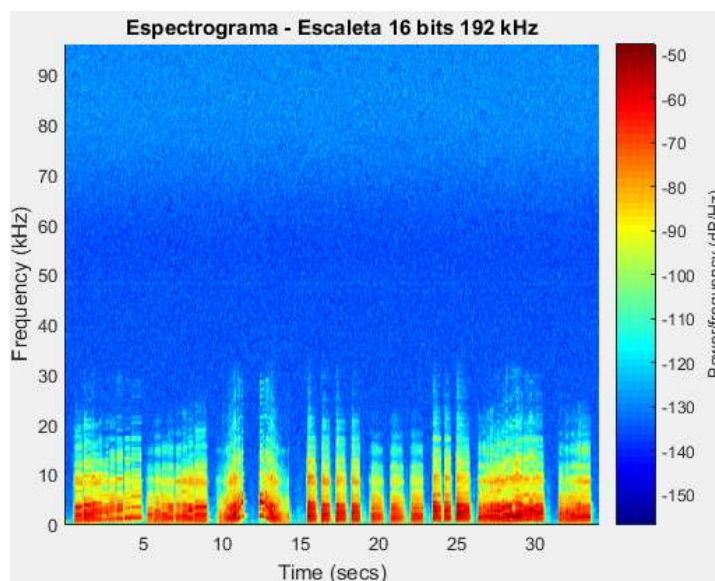
Nas figura 38 e figura 39 têm-se os espectrogramas dos áudios da escaleta em 24 bits/192 kHz e 16 bits/192 kHz.

Figura 38 – Espectrograma do áudio da escaleta em 24 bits e 192 kHz



Fonte: Os autores.

Figura 39 – Espectrograma do áudio da escaleta em 16 bits e 192 kHz



Fonte: Os autores.

6.2.2 Violão

O mesmo procedimento realizado com o áudio da escaleta, foi utilizado com o áudio do violão. Sua diferença está na quantização, um deles foi quantizado com 24 bits e outro com 16 bits. Considerando o de 24 bits como o original, foi possível calcular a diferença direta e os parâmetros da tabela 6.

Tabela 5 – Potência dos áudios do violão

Áudio	Potência
Violão 24 bits/192 kHz	2.6590717e-03
Violão 16 bits/ 192 kHz	2.6590757e-03

Fonte: Os autores.

Tabela 6 – Parâmetros da análise do violão

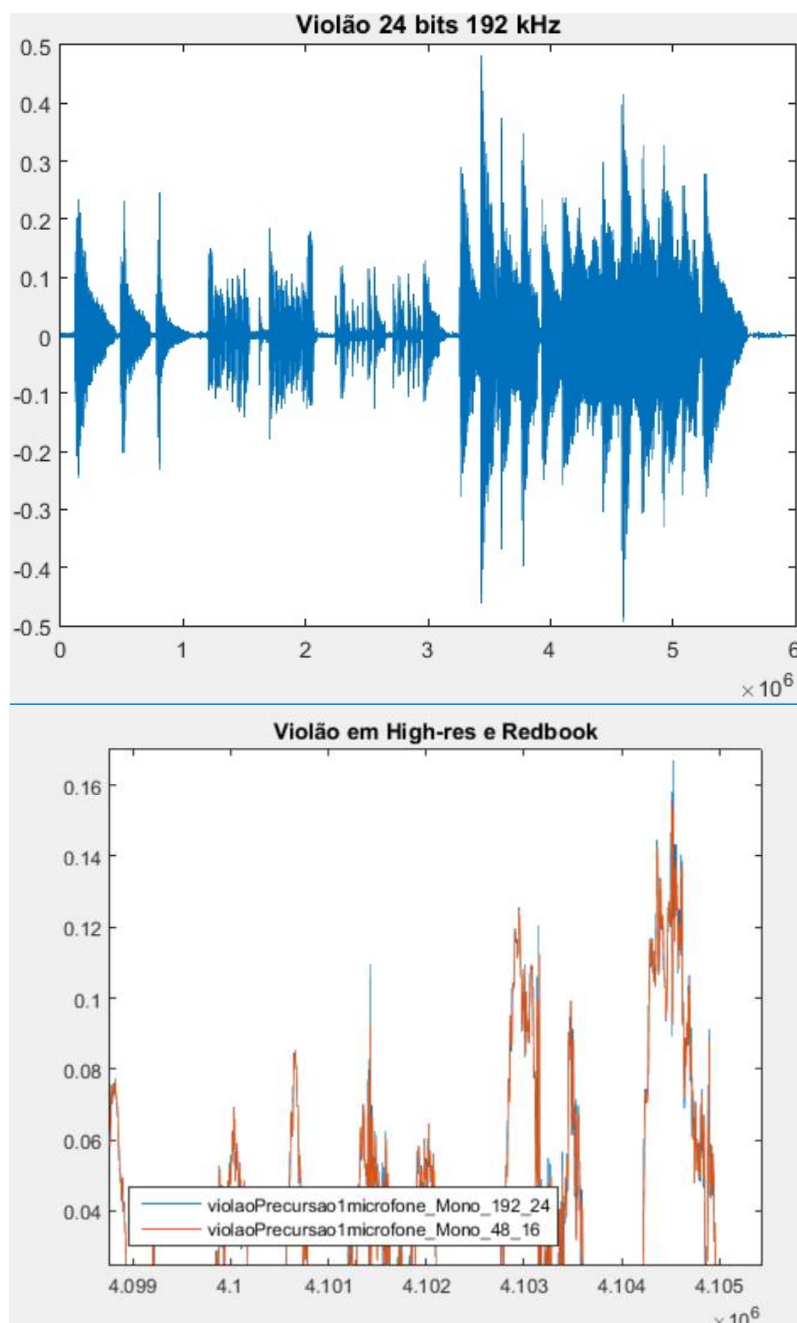
Diferença na Potência	SNR	MSE	SMSE (%)
-3.9412e-09	58.3784	3.8626e-09	6.2150e-03

Fonte: Os autores.

As figuras a seguir mostram os formatos de onda e os espectrogramas de ambos os áudios. Ainda há a forma de onda da diferença entre eles.

Na figura 40. tem-se o formato de onda do áudio do violão em 24 bits/192 kHz, obtido pelo *script* do Matlab, e um *zoom* feito com as duas ondas sobrepostas, de forma a verificar sua diferença.

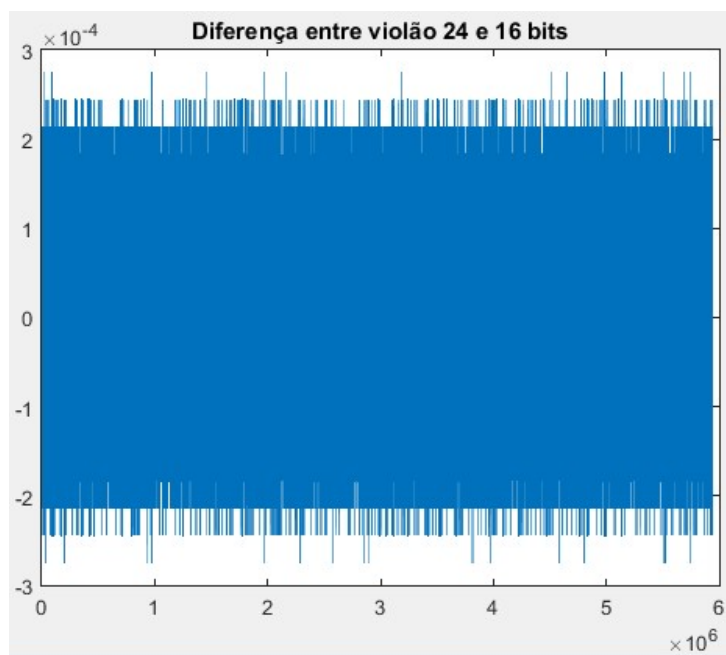
Figura 40 – Formato de onda do áudio do violão em 24 bits e 192 kHz e diferença entre as ondas sobrepostas (zoom).



Fonte: Os autores.

Como os áudios possuem a mesma frequência de amostragem, foi possível subtrair a onda de 16 bits da onda da onda de 24 bits. A onda de diferença é mostrada na figura 41.

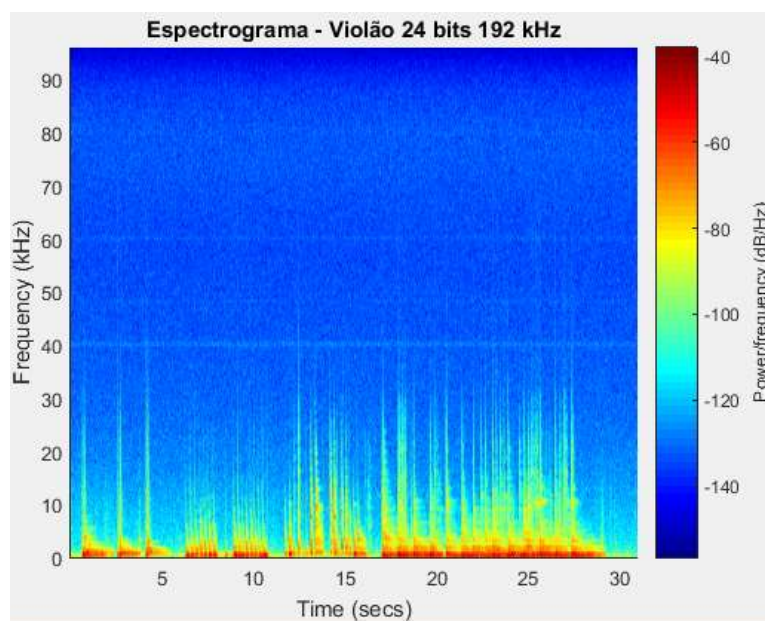
Figura 41 – Formato de onda da diferença entre o áudio em 24 bits e o em 16 bits.



Fonte: Os autores.

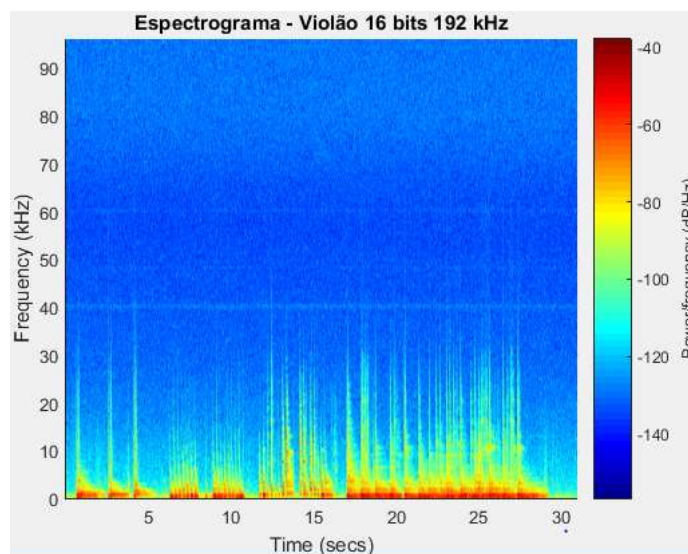
Nas figura 42 e figura 43 têm-se os espectrogramas dos áudio do violão em 24 bits/192 kHz e 16 bits/192 kHz.

Figura 42 – Espectrograma do áudio do violão em 24 bits e 192 kHz



Fonte: Os autores.

Figura 43 – Espectrograma do áudio do violão em 16 bits e 192 kHz



Fonte: Os autores.

6.2.3 Saxofone

O mesmo procedimento realizado com o áudio da escaleta e do violão, foi utilizado com o áudio do saxofone. Sua diferença está na quantização, um deles foi quantizado com 24 bits e outro com 16 bits. Considerando o de 24 bits como o original, foi possível calcular a diferença direta e os parâmetros da tabela 8.

Tabela 7 – Potência dos áudios do saxofone.

Áudio	Potência
Saxofone 24 bits/192 kHz	3.495154e-03
Saxofone 16 bits/192 kHz	3.495158e-03

Fonte: Os autores.

Tabela 8 – Parâmetros da análise dos áudios do saxofone.

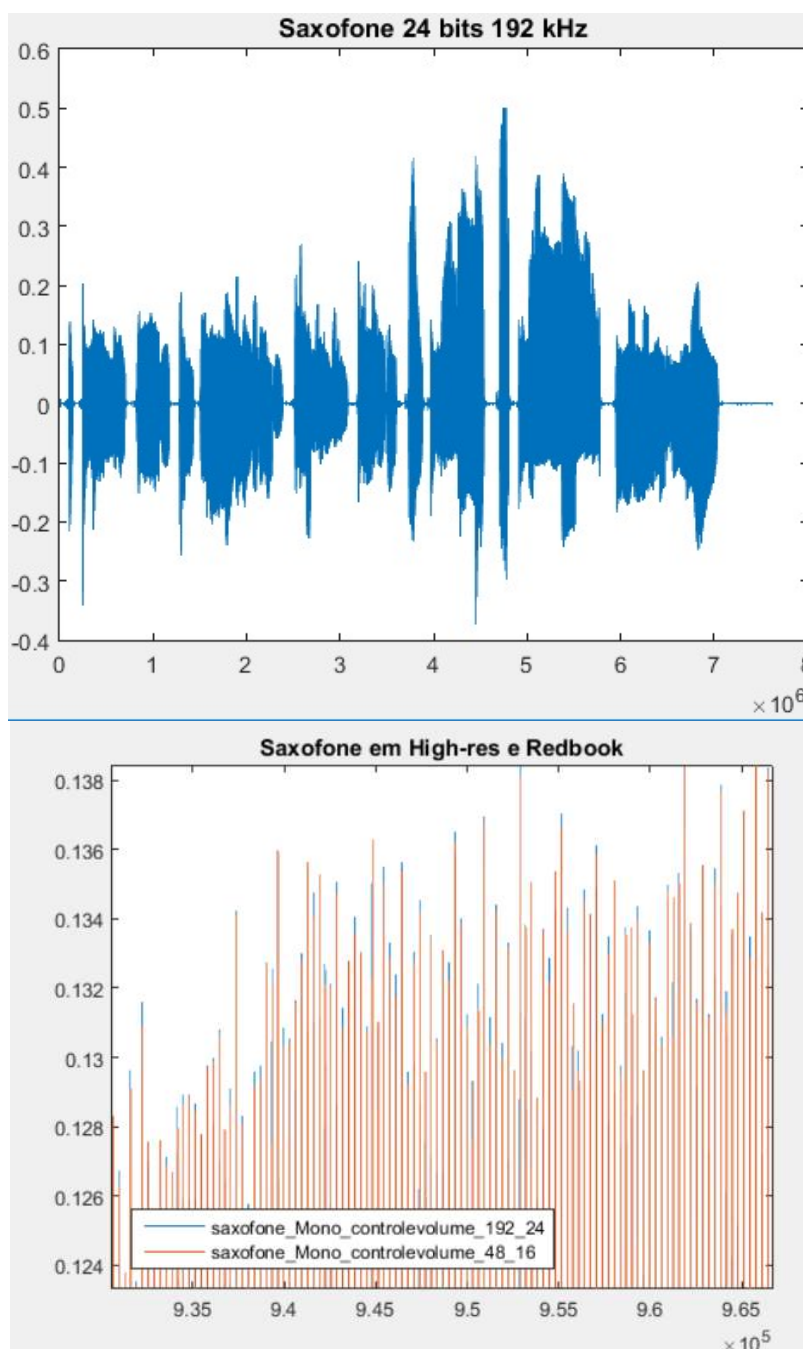
Diferença na Potência	SNR	MSE	SMSE (%)
-3.9295e-09	59.5777	3.8520e-09	6.2065e-03

Fonte: Os autores.

As figuras a seguir mostram os formatos de onda e os espectrogramas de ambos os áudios. Ainda há a forma de onda da diferença entre eles.

Na figura 44 tem-se o formato de onda do áudio do saxofone em 24 bits/192 kHz, obtido pelo *script* do Matlab, e um *zoom* feito com as duas ondas sobrepostas, de forma a verificar sua diferença.

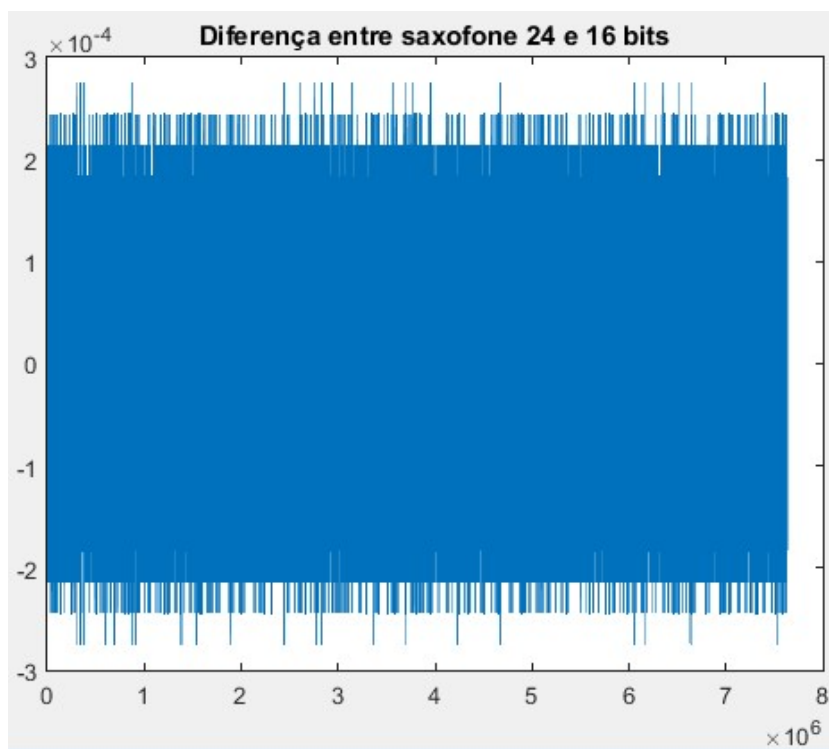
Figura 44 – Formato de onda do áudio do saxofone em 24 bits e 192 kHz e diferença entre as ondas sobrepostas (*zoom*).



Fonte: Os autores.

Como os áudios possuem a mesma frequência de amostragem, foi possível subtrair a onda de 16 bits da onda da onda de 24 bits. A onda de diferença é mostrada na figura 45.

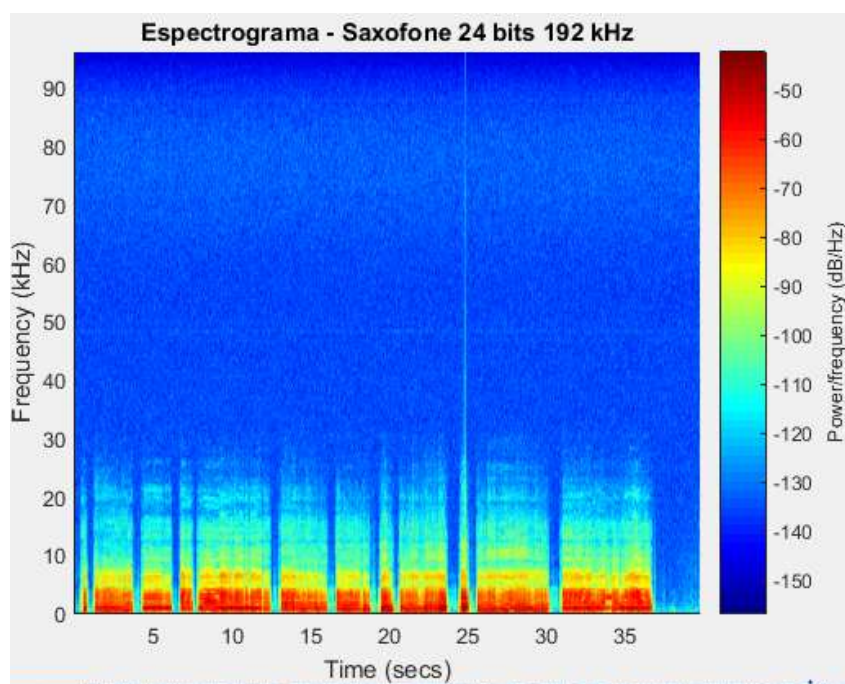
Figura 45 – Formato de onda da diferença entre o áudio em 24 bits e o em 16 bits.



Fonte: Os autores.

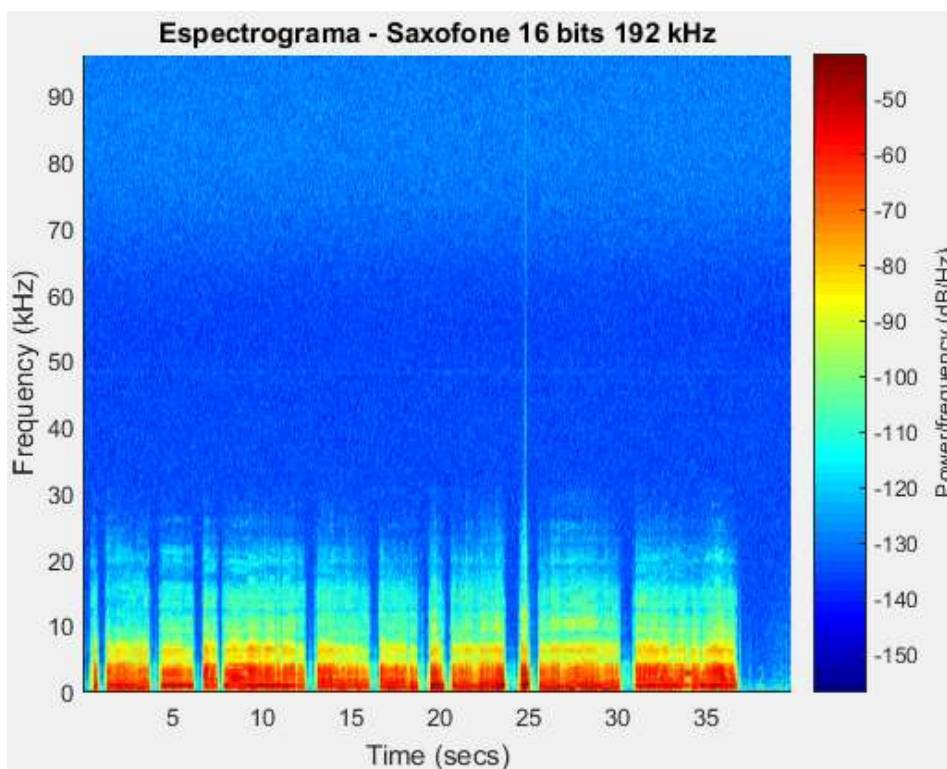
Nas figura 46 e figura 47 têm-se os espectrogramas dos áudio do saxofone em 24 bits/192 kHz e 16 bits/192 kHz.

Figura 46 – Espectrograma do áudio do saxofone em 24 bits e 192 kHz.



Fonte: Os autores.

Figura 47 – Espectrograma do áudio do saxofone em 16 bits e 192 kHz.



Fonte: Os autores.

6.3 DIFERENTES FREQUÊNCIAS E MESMA QUANTIZAÇÃO

Como os áudios foram gravados em 24 bits e 192 kHz e o processo de transformação consiste em abaixar os parâmetros, a frequência de amostragem alterada de 192 kHz para 48 kHz. Após esse processo, os *scripts* do Matlab foram utilizados para gerar as informações sobre os áudios individualmente e comparativamente. Nesse ponto foram analisadas as potências dos sinais, suas diferenças matemáticas e visualizados os espectrogramas.

6.3.1 Escaleta

Nessa análise foram utilizados os áudios da escaleta exportados ambos com 24 bits. Sua diferença está na frequência de amostragem, um deles foi amostrado com 192 kHz e outro com 48 kHz. Calculando a potência para ambos sinais, foi possível calcular a diferença entre elas, e os valores encontrados estão na tabela 9.

Tabela 9 – Potência dos áudios da escaleta com a mesma quantização.

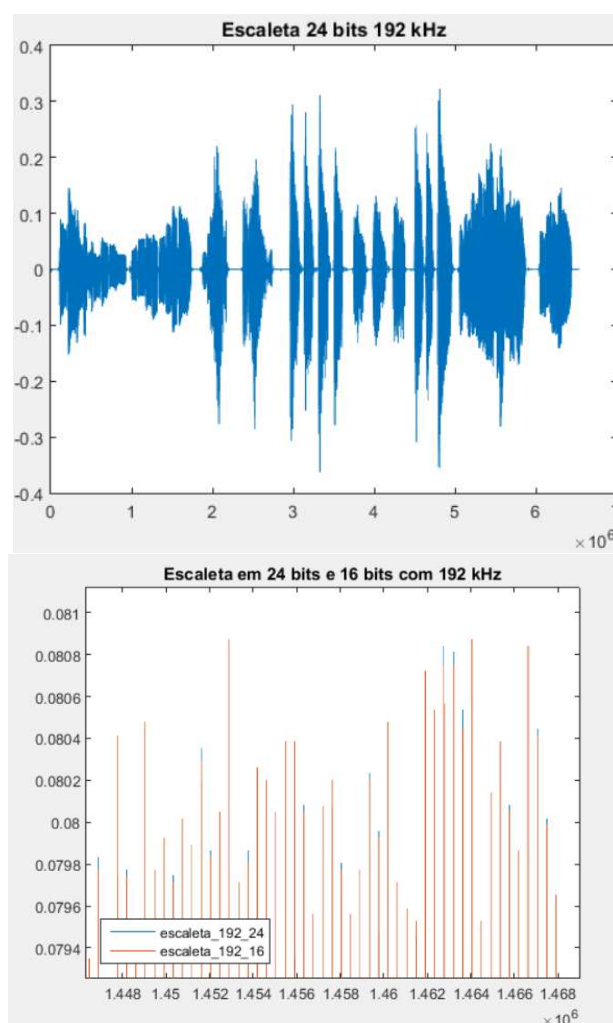
Áudio	Potência
Escaleta 24 bits/192 kHz	1.480271E-3
Escaleta 24 bits/48 kHz	1.480266E-3
DIFERENÇA DE POTÊNCIA	4.566E-09 (3.0846E-4%)

Fonte: Os autores.

As figuras a seguir mostram os formatos de onda e os espectrogramas de ambos os áudios. Ainda há a forma de onda da diferença entre eles.

Na figura 48 tem-se o formato de onda do áudio da escaleta em 24 bits/192 kHz, obtido pelo *script* do Matlab, e um *zoom* feito com as duas ondas sobrepostas, de forma a verificar sua diferença.

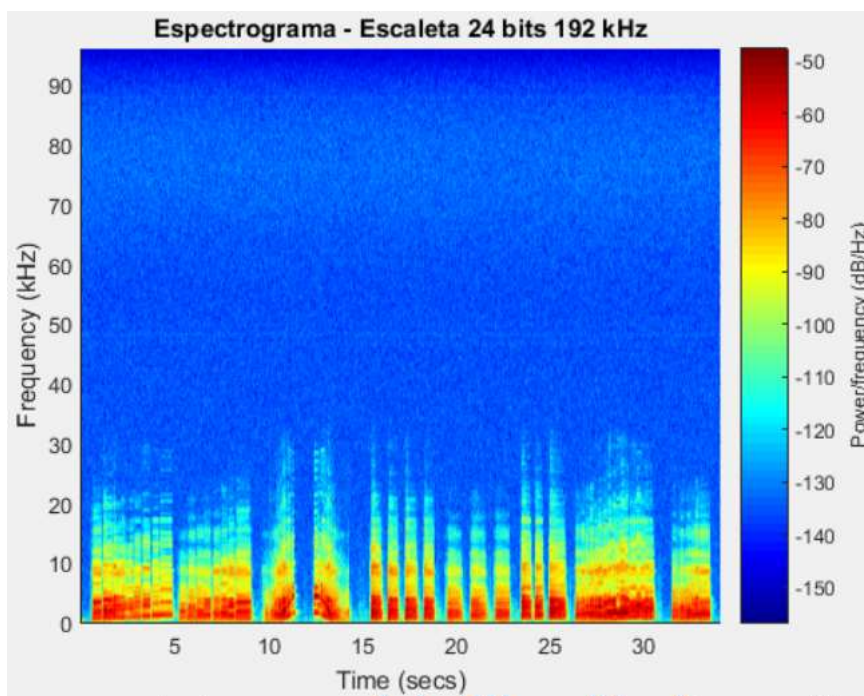
Figura 48 – Formato de onda do áudio da escaleta em 24 bits e 192 kHz e diferença entre as ondas sobrepostas (zoom).



Fonte: Os autores.

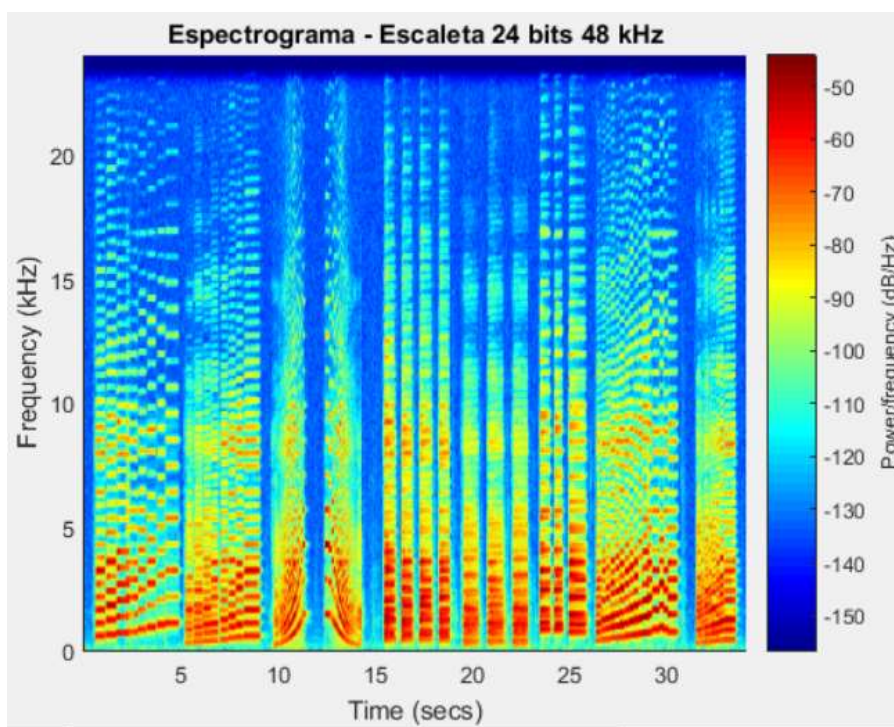
Nas figura 49 e figura 50 têm-se os espectrogramas dos áudio da escaleta em 24 bits/192 kHz e 24 bits/48 kHz.

Figura 49 – Espectrograma do áudio da escaleta em 24 bits e 192 kHz.



Fonte: Os autores.

Figura 50 – Espectrograma do áudio da escaleta em 24 bits e 48 kHz.



Fonte: Os autores.

6.3.2 Violão

O mesmo procedimento realizado com o áudio da escaleta, foi utilizado com o áudio do violão. Sua diferença está na frequência de amostragem, um deles foi amostrado com 192 kHz e outro com 48 kHz. Calculando a potência para ambos sinais, foi possível calcular a diferença entre elas, e os valores encontrados estão na Tabela 10.

Tabela 10 – Potência dos áudios do violão com mesma quantização.

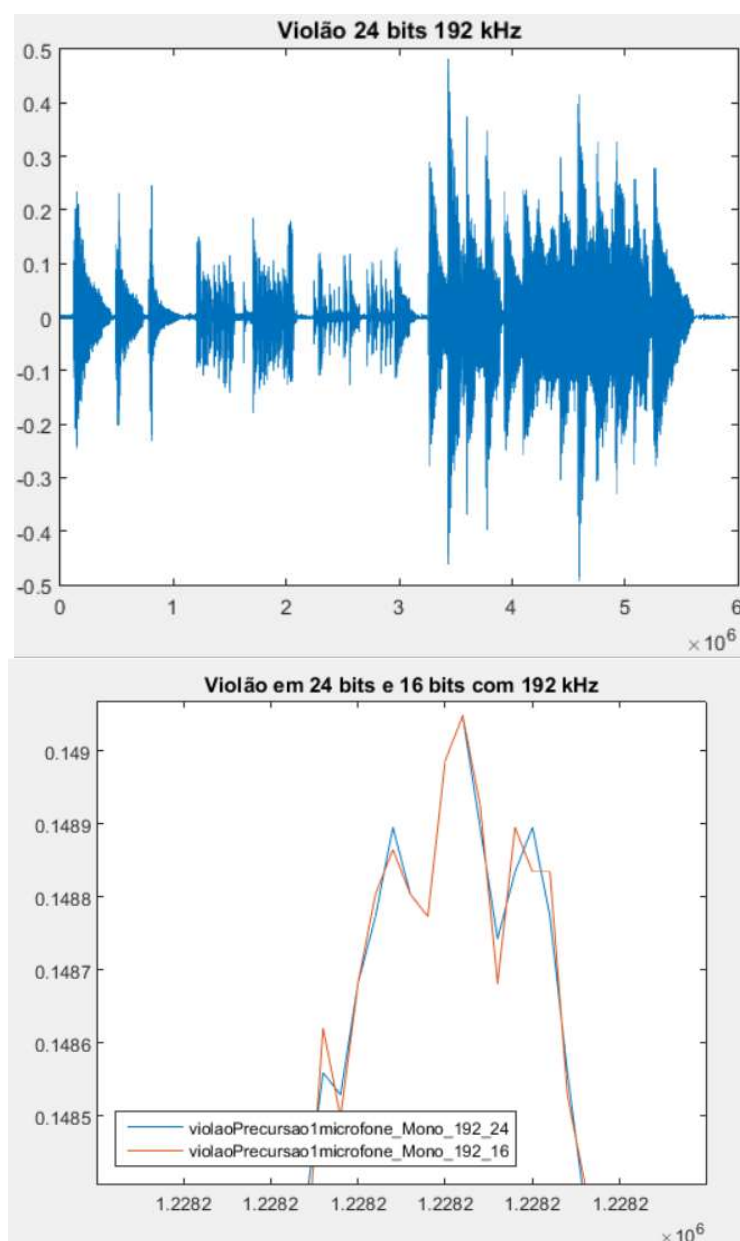
Áudio	Potência
Violão 24 bits/192 kHz	2.659072E-3
Violão 24 bits/48 kHz	2.659026E-3
DIFERENÇA DE POTÊNCIA	4.5287E-08 (1.7031E-3%)

Fonte: Os autores.

As figuras a seguir mostram os formatos de onda e os espectrogramas de ambos os áudios. Ainda há a forma de onda da diferença entre eles.

Na figura 51 tem-se o formato de onda do áudio do violão em 24 bits/192 kHz, obtido pelo *script* do Matlab, e um *zoom* feito com as duas ondas sobrepostas, de forma a verificar sua diferença.

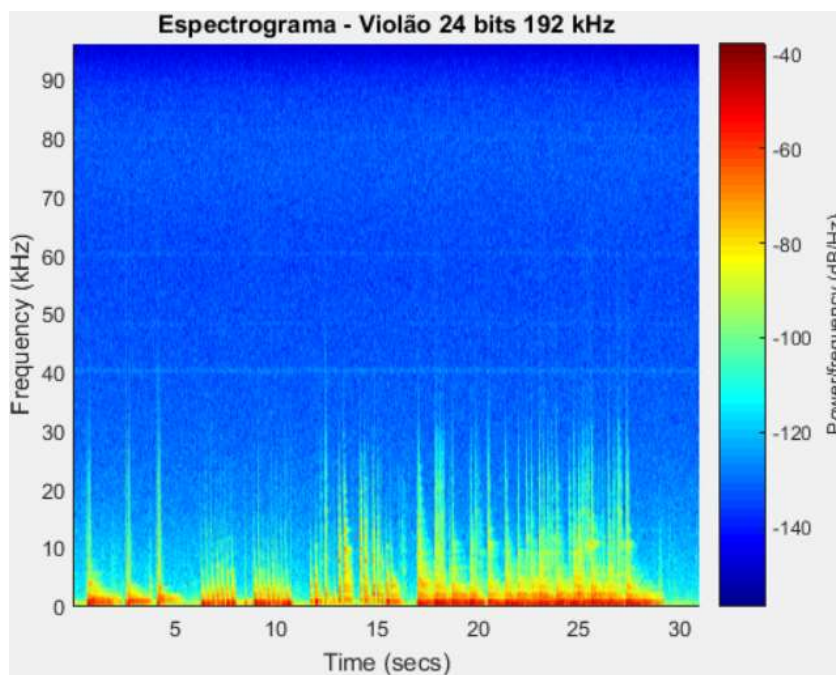
Figura 51 – Formato de onda do áudio do violão em 24 bits e 192 kHz e diferença entre as ondas sobrepostas (zoom).



Fonte: Os autores.

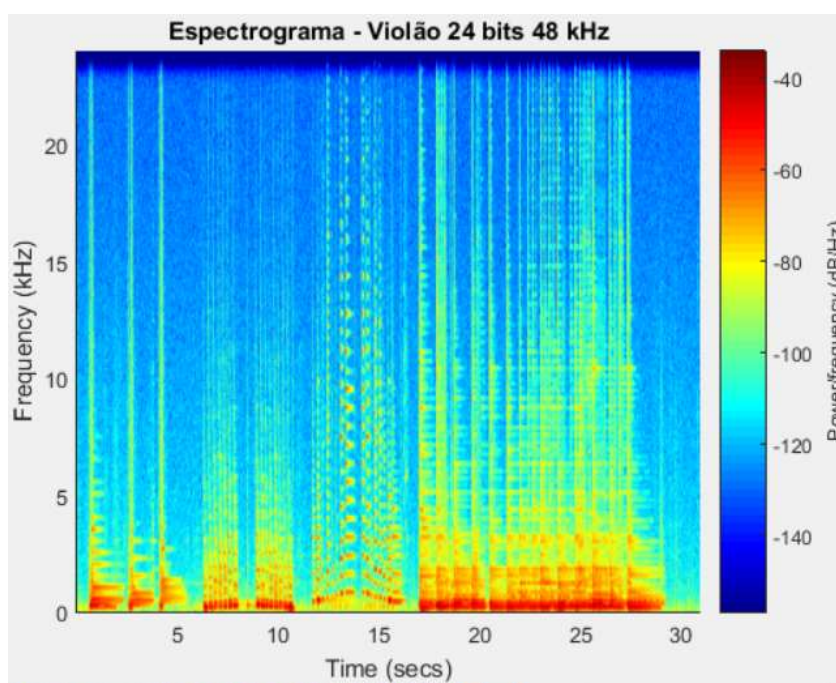
Nas figura 52 e figura 53 têm-se os espectrogramas dos áudio do violão em 24 bits/192 kHz e 24 bits/48 kHz.

Figura 52 – Espectrograma do áudio do violão em 24 bits e 192 kHz.



Fonte: Os autores.

Figura 53 – Espectrograma do áudio do violão em 24 bits e 48 kHz.



Fonte: Os autores.

6.3.3 Saxofone

O mesmo procedimento realizado com o áudio da escaleta e do violão, foi utilizado com o áudio do saxofone. Sua diferença está na frequência de amostragem, um deles foi amostrado com 192 kHz e outro com 48 kHz. Calculando a potência para ambos sinais, foi possível calcular a diferença entre elas, e os valores encontrados estão na tabela 11.

Tabela 11 – Potência dos áudios do saxofone com mesma quantização.

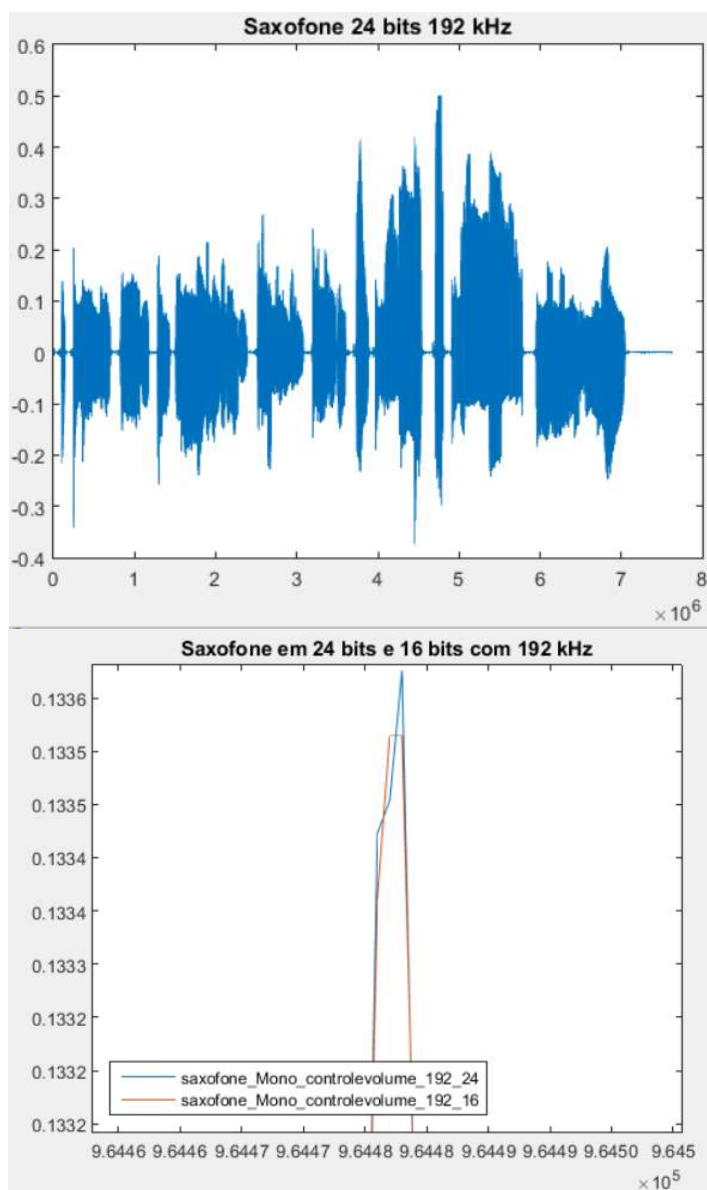
Áudio	Potência
Saxofone 24 bits/192 kHz	3.495154E-3
Saxofone 24 bits/ 48 kHz	3.495149E-3
DIFERENÇA DE POTÊNCIA	5.3475E-09 (1.5299E-4%)

Fonte: Os autores.

As figuras a seguir mostram os formatos de onda e os espectrogramas de ambos os áudios. Ainda há a forma de onda da diferença entre eles.

Na figura 54 tem-se o formato de onda do áudio do saxofone em 24 bits/192 kHz, obtido pelo *script* do Matlab, e um *zoom* feito com as duas ondas sobrepostas, de forma a verificar sua diferença.

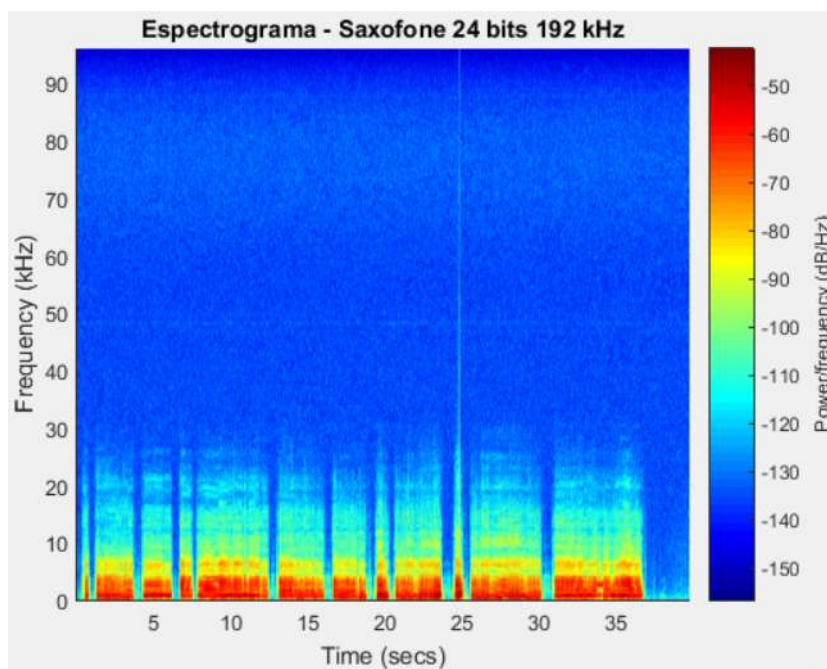
Figura 54 – Forma de onda do áudio do saxofone em 24 bits e 192 kHz e diferença entre as ondas sobrepostas (zoom).



Fonte: Os autores.

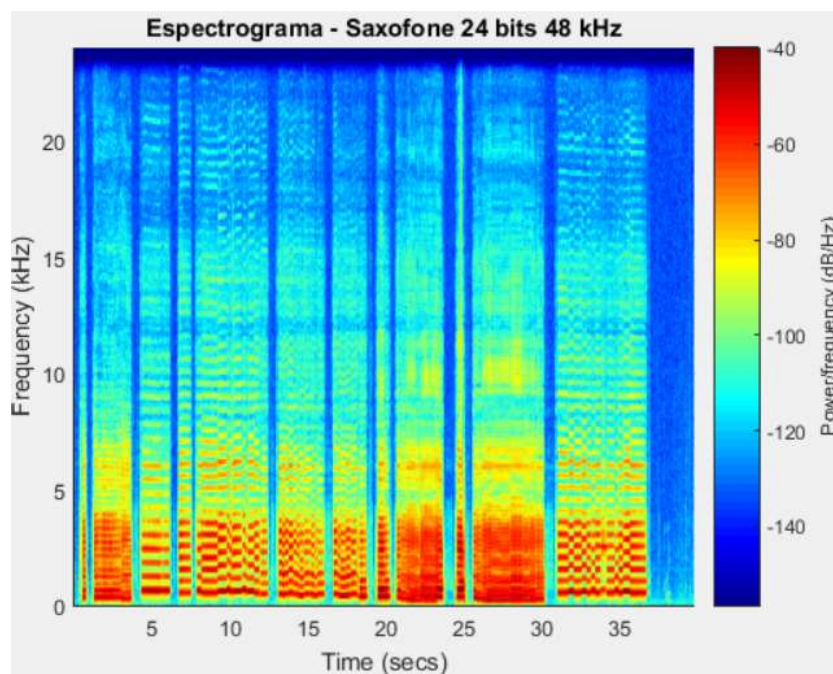
Nas figura 55 e figura 56 têm-se os espectrogramas dos áudio do saxofone em 24 bits/192 kHz e 24 bits/48 kHz.

Figura 55 – Espectrograma do áudio do saxofone em 24 bits e 192 kHz.



Fonte: Os autores.

Figura 56 – Espectrograma do áudio do saxofone em 24 bits e 48 kHz.



Fonte: Os autores.

6.4 DIFERENTES FAIXAS DE FREQUÊNCIA E QUANTIZAÇÃO

Como os áudios foram gravados em 24 bits e 192 kHz e o processo de transformação consiste em abaixar os parâmetros, a quantização foi alterada de 24 para 16 bits e a frequência de amostragem alterada de 192 kHz para 48 kHz no áudio de 16 bits. Após esse processo, os *scripts* do Matlab foram utilizados para gerar as informações sobre os áudios individualmente e comparativamente. Nesse ponto foram analisadas as potências dos sinais, suas diferenças matemáticas e visualizados os espectrogramas.

6.4.1 Escaleta

Nessa análise foram utilizados os áudios da escaleta exportados em 24 bits/192 kHz e 16 bits/48 kHz. Nem a frequência de amostragem nem a quantização são similares entre os arquivos. Calculando a potência para ambos sinais, foi possível calcular a diferença entre elas, e os valores encontrados estão na Tabela 12.

Tabela 12 - Potência dos áudios da escaleta com quantização e frequência de amostragem diferentes.

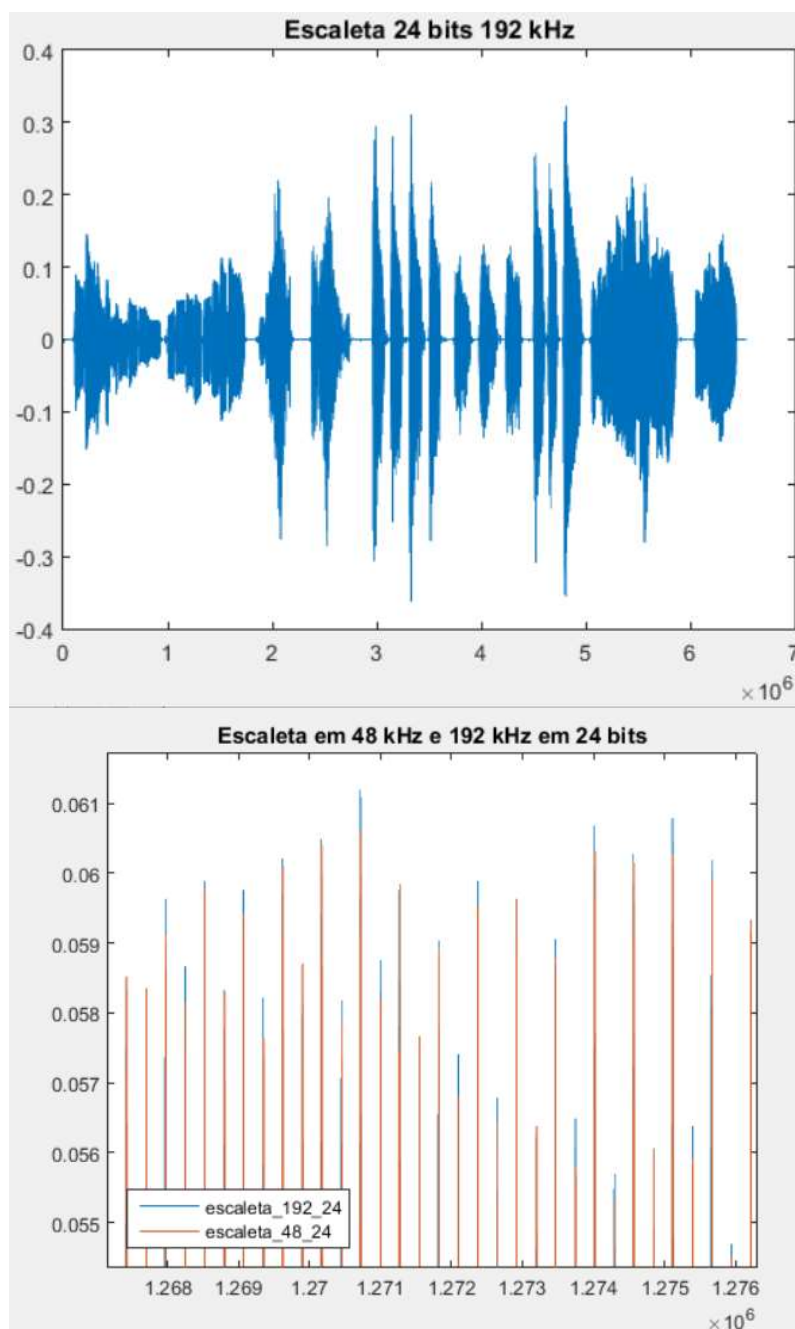
Áudio	Potência
Escaleta 24 bits/192 kHz	1.48027093E-3
Escaleta 16 bits/48 kHz	1.48027025E-3
DIFERENÇA DE POTÊNCIA	6.7936E-10 (4.5894E-5 %)

Fonte: Os autores.

As figuras a seguir mostram os formatos de onda e os espectrogramas de ambos os áudios.

Na figura 57 tem-se o formato de onda do áudio da escaleta em 24 bits/192 kHz, obtido pelo *script* do Matlab, e um *zoom* feito com as duas ondas sobrepostas, de forma a verificar sua diferença.

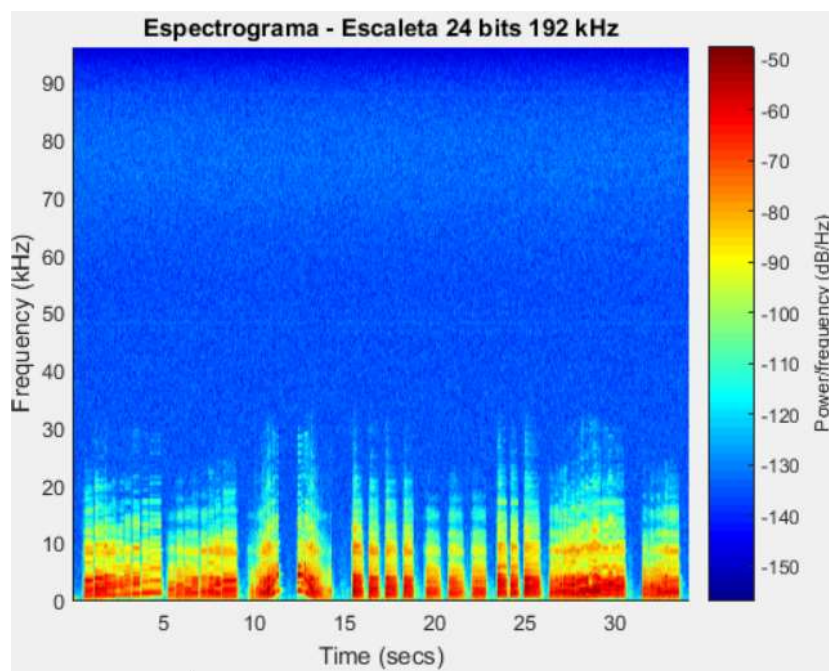
Figura 57 – Forma de onda do áudio da escaleta em 24 bits e 192 kHz e diferença entre as ondas sobrepostas (zoom).



Fonte: Os autores.

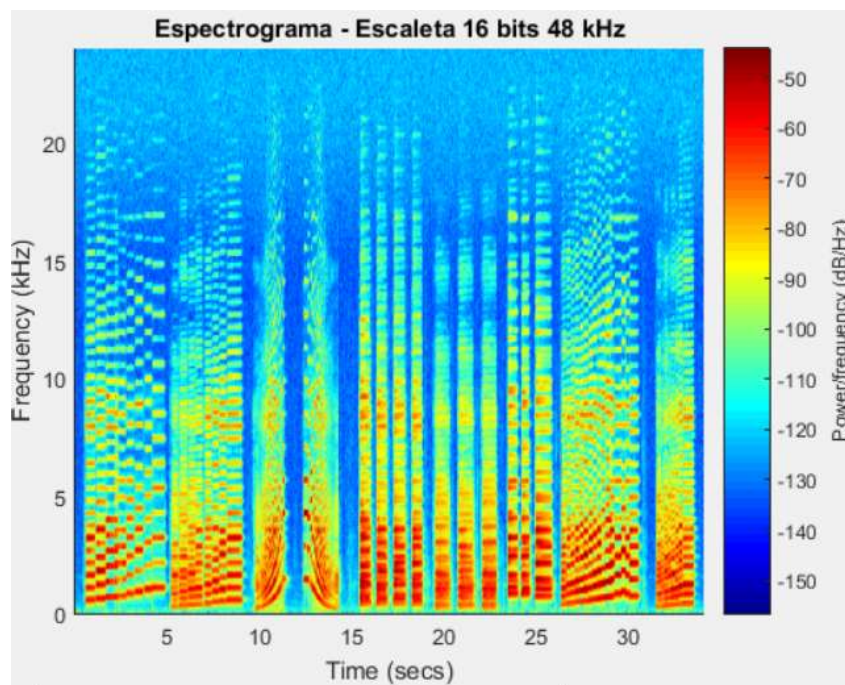
Nas figura 58 e figura 59 têm-se os espectrogramas dos áudios da escaleta em 24 bits/192 kHz e 16 bits/48 kHz.

Figura 58 – Espectrograma do áudio da escaleta em 24 bits e 192 kHz.



Fonte: Os autores.

Figura 59 – Espectrograma do áudio da escaleta em 16 bits e 48 kHz.



Fonte: Os autores.

6.4.2 Violão

O mesmo procedimento realizado com o áudio da escaleta, foi utilizado com o áudio do violão. Nem a frequência de amostragem nem a quantização são similares entre os arquivos. Calculando a potência para ambos sinais, foi possível calcular a diferença entre elas, e os valores encontrados estão na tabela 13.

Tabela 13 – Potência dos áudios do violão com mesma quantização.

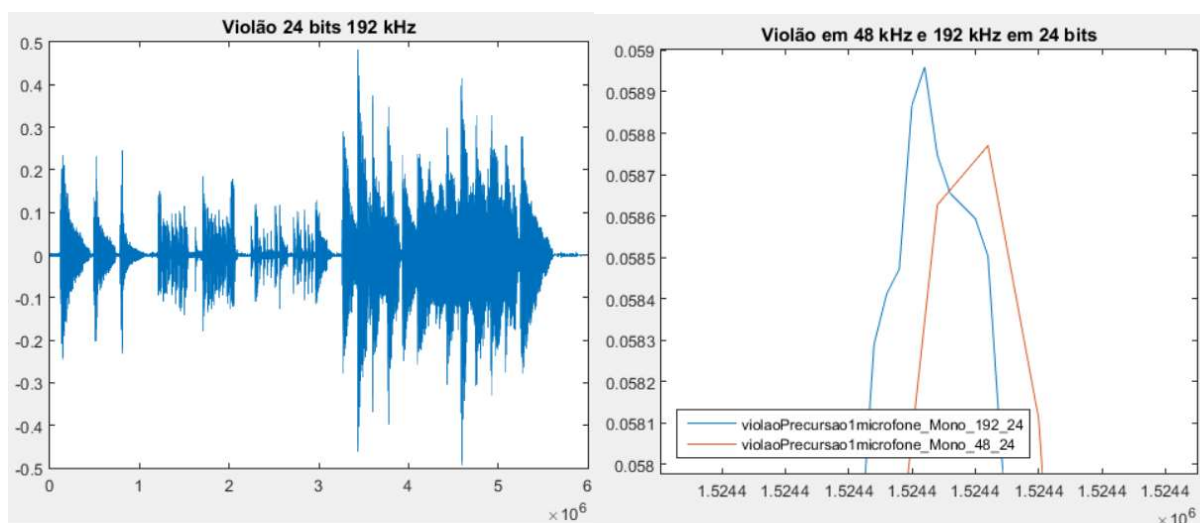
Áudio	Potência
Violão 24 bits/192 kHz	2.659072E-3
Violão 16 bits/48 kHz	2.659030E-3
DIFERENÇA DE POTÊNCIA	4,1588E-08 (1.5640E-3%)

Fonte: Os autores.

As figuras a seguir mostram os formatos de onda e os espectrogramas de ambos os áudios. Ainda há a forma de onda da diferença entre eles.

Na figura 60 tem-se o formato de onda do áudio do violão em 24 bits/192 kHz, obtido pelo *script* do Matlab, e um *zoom* feito com as duas ondas sobrepostas, de forma a verificar sua diferença.

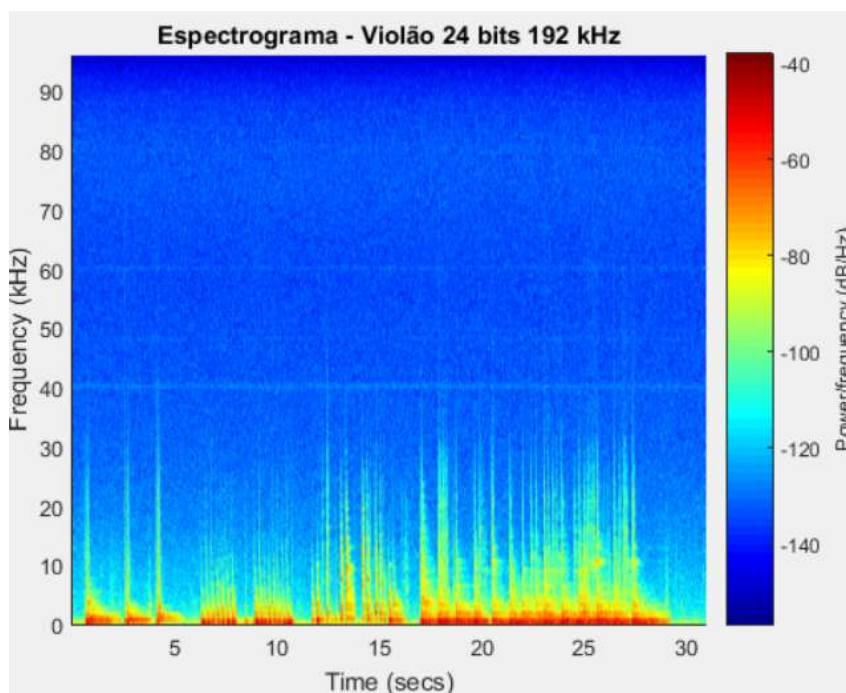
Figura 60 – Forma de onda do áudio do violão em 24 bits e 192 kHz e diferença entre as ondas sobrepostas (zoom).



Fonte: Os autores.

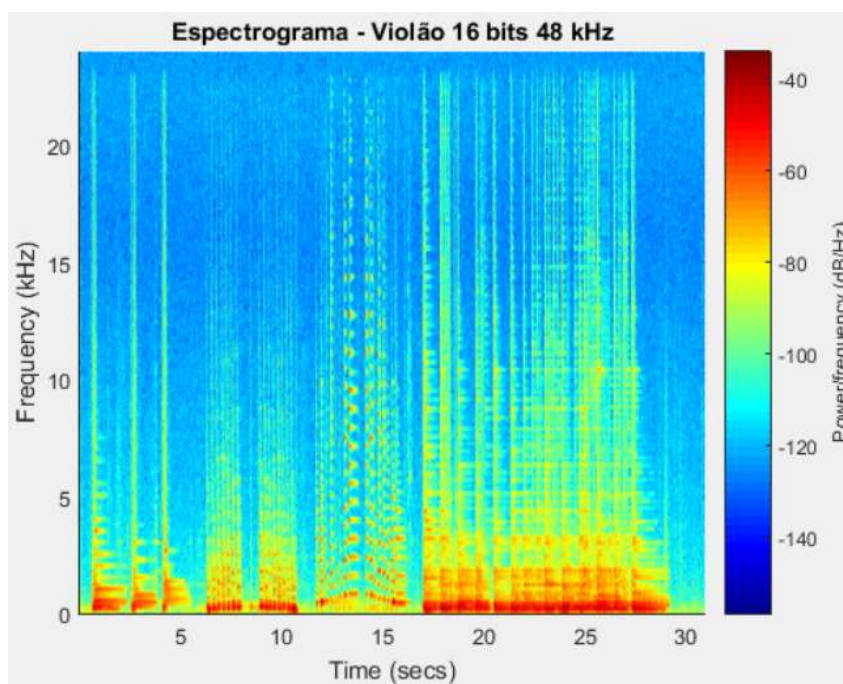
Nas figura 61 e figura 62 têm-se os espectrogramas dos áudio do violão em 24 bits/192 kHz e 16 bits/48 kHz.

Figura 61 – Espectrograma do áudio do violão em 24 bits e 192 kHz.



Fonte: Os autores.

Figura 62 – Espectrograma do áudio do violão em 16 bits e 48 kHz.



Fonte: Os autores.

6.4.3 Saxofone

O mesmo procedimento realizado com o áudio da escaleta e do violão, foi utilizado com o áudio do saxofone. Nem a frequência de amostragem nem a quantização são similares entre os arquivos. Calculando a potência para ambos sinais, foi possível calcular a diferença entre elas, e os valores encontrados estão na Tabela 14.

Tabela 14 – Potência dos áudios do saxofone com frequência e quantização diferentes.

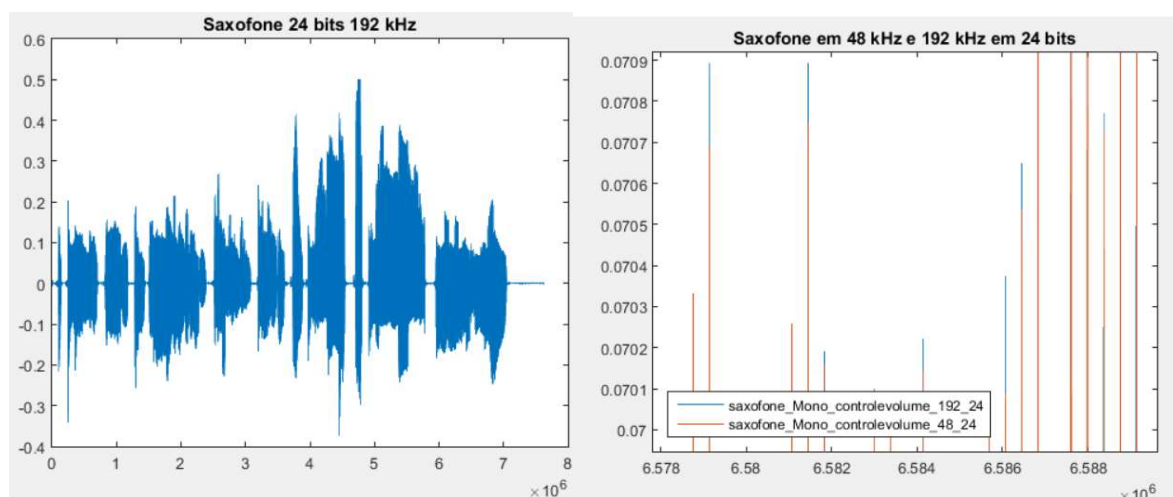
Áudio	Potência
Saxofone 24 bits/192 kHz	3.4951543E-3
Saxofone 16 bits/ 48 kHz	3.4951524E-3
DIFERENÇA DE POTÊNCIA	1.9398E-09 (5.5501E-5%)

Fonte: Os autores.

As figuras a seguir mostram os formatos de onda e os espectrogramas de ambos os áudios. Ainda há a forma de onda da diferença entre eles.

Na figura 63 tem-se o formato de onda do áudio do saxofone em 24 bits/192 kHz, obtido pelo *script* do Matlab, e um *zoom* feito com as duas ondas sobrepostas, de forma a verificar sua diferença.

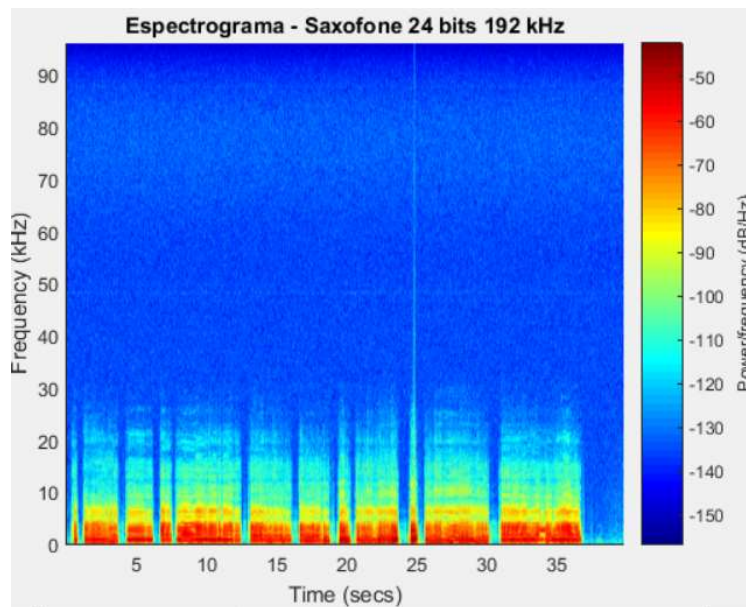
Figura 63 – Forma de onda do áudio do saxofone em 24 bits e 192 kHz e diferença entre as ondas sobrepostas (zoom).



Fonte: Os autores.

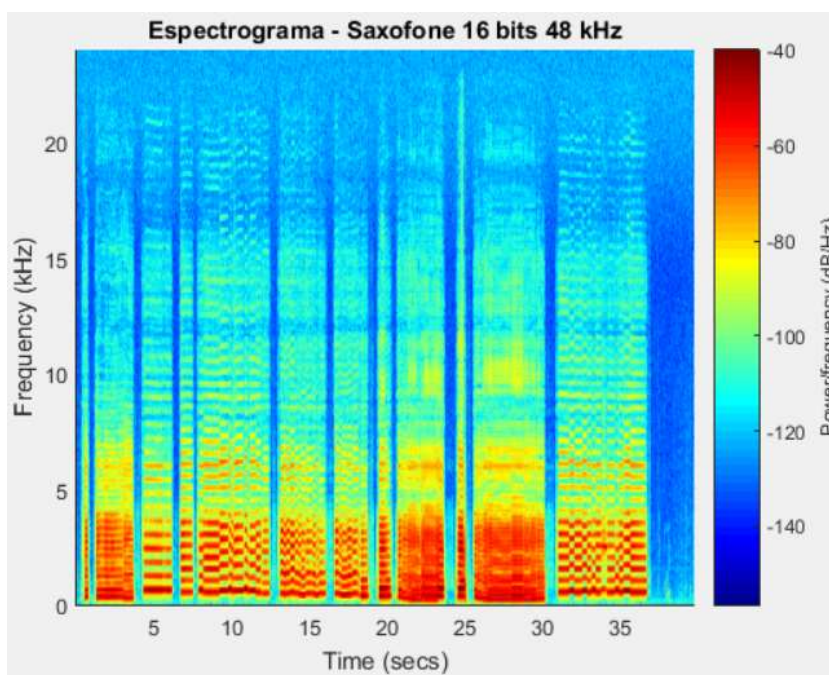
Nas figura 64 e figura 65 têm-se os espectrogramas dos áudio do saxofone em 24 bits/192 kHz e 16 bits/48 kHz.

Figura 64 – Espectrograma do áudio do saxofone em 24 bits e 192 kHz.



Fonte: Os autores.

Figura 65 – Espectrograma do áudio do saxofone em 16 bits e 192 kHz.



Fonte: Os autores.

7 ANÁLISE SENSORIAL

7.1 HIPÓTESE NULA

Com o objetivo de comparar os áudios com frequências e quantizações diferentes um teste sensorial foi construído e aplicado. Para compor este teste, foi assumida a hipótese nula de que os áudios gerados nos diferentes métodos de gravação não geram diferença de percepção ao usuário.

Em contrapartida, a hipótese alternativa é para que caso a hipótese nula seja refutada é aceita. Assim, a hipótese alternativa é que os áudios geram uma percepção diferente ao usuário.

Logo, caso a hipótese nula seja aceita, a conclusão direta é que não há diferença significativa entre os áudios gerados na gravação. Já se ela for rejeitada, pode-se concluir que há diferença entre os áudios de alta fidelidade e os não.

7.2 AUDIÇÕES

7.2.1 Objetos de teste

Tendo realizado a sessão de gravações foi optado por realizar uma bateria de testes composta por um áudio de violão, um áudio de saxofone e um áudio de escaleta, em 2 qualidades cada. Para isso foram preparados 6 áudios, sendo eles nomeados conforme abaixo

1. Áudio do violão gravado em 48hz e 16bits – Violão Amostra A
2. Áudio do violão gravado em 192hz e 24bits – Violão Amostra B
3. Áudio do saxofone gravado em 48hz e 16bits – Sax Amostra A
4. Áudio do saxofone gravado em 192hz e 24bits – Sax Amostra B
5. Áudio da escaleta gravado em 48hz e 16bits – Escaleta Amostra A
6. Áudio da escaleta gravado em 192hz e 24bits – Escaleta Amostra B

Como citado no capítulo 5, os áudios de diferentes qualidades foram gravados e posteriormente convertidos para as amostragens e quantizações desejadas, na hora de montar os áudios para a audição, foram escolhidos trechos

com uma duração aproximada de 30 segundos de uma das gravações feitas de cada instrumento.

7.2.2 Avaliadores

Para a escolha dos avaliadores, foram chamadas pessoas de diferentes idades, sexo, profissão e convivência com áudio, com o objetivo de encontrar possíveis diferenças de percepção do áudio em função dos fatores citados.

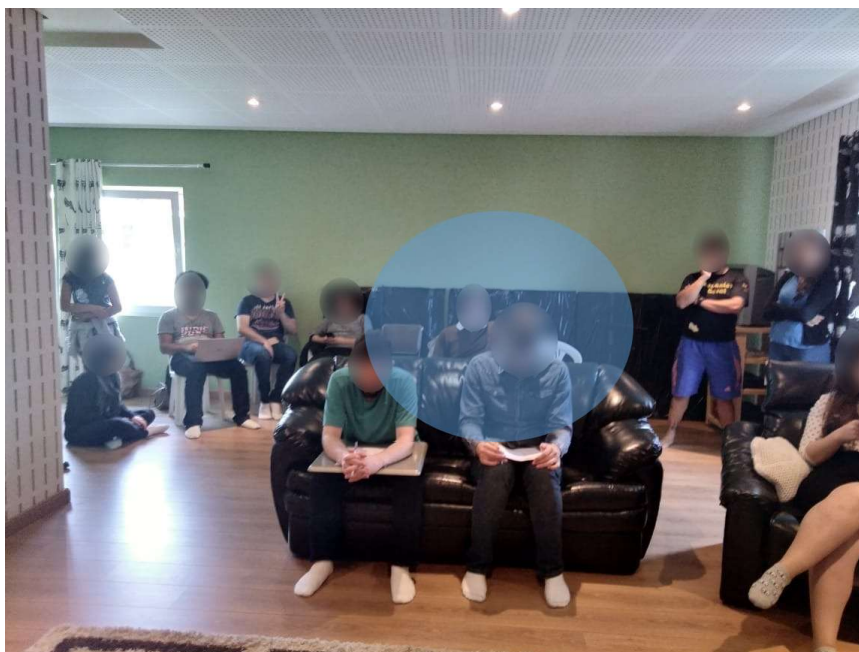
Foi entregue a cada participante, quando chegou ao local da audição, uma ficha de identificação, mostrada no apêndice A. Essa ficha continha perguntas para levantar informações sobre a quais grupos ele pertenceria. Após preencherem-na os avaliadores foram para a fase de audição.

Para responder a cada audição o avaliador recebeu uma ficha de avaliação igual à do apêndice B, e deveria responder as duas questões obrigatoriamente, após realizar as três baterias de audições, os avaliadores eram dispensados.

7.2.3 Ambiente de teste

As audições foram realizadas no mesmo local das gravações, pois essa sala tem boas propriedades acústicas para reprodução de áudio. Nesta sala também se encontravam os equipamentos para realização da audição. Um computador da marca Avell, enviava para um conversor Digital Analógico SOLO da Etalon Acoustics os arquivos salvos em .wav. Este foi conectado à um amplificador integrado CurrentA, também da Etalon Acoustics, responsável pela amplificação do sinal de tensão e transformação em corrente, estando conectado as caixas acústicas Special 25 da Dynaudio. Por fim, a caixa converte de um sinal de corrente para um sonoro.

A fim de buscar a melhor performance, os avaliadores foram divididos em grupos de no máximo seis avaliadores por vez, assim garantiu-se que estivessem no melhor ponto de áudio da sala. Conforme apresentado na figura 66, um grupo de 3 avaliadores, marcados com o círculo azul, estavam sentados ao centro da sala em frente às caixas de som da forma mais equidistante possível.

Figura 66 – Sala de Audição.

Fonte: Os autores.

7.2.4 Realização do teste

Para o teste sensorial, conforme o capítulo 4, foram preparadas 3 baterias de testes. Cada bateria consiste em tocar 3 áudios do mesmo instrumento, sendo que dois deles são iguais e um diferente. Cabe ao participante identificar qual a amostra tocada era diferente. Então, existe apenas uma resposta correta e o avaliador tinha 33% de chance de acertar aleatoriamente qual era o áudio diferente, conforme prevê este modelo de teste.

O teste de favorabilidade não alterou o resultado da avaliação da hipótese nula, e não tem resposta correta, servindo apenas como termômetro de percepção. Foi considerado como favorável, os casos em que o avaliador respondeu que concorda parcialmente ou totalmente que existe diferença entre os áudios, qualquer resposta diferente dessas foi considerada como não favorável.

A audição foi realizada com um grupo de quatorze pessoas, sendo 5 mulheres e 9 homens de idades entre 14 e 53 anos de idade, de várias profissões e intimidades com áudios e música.

Esses foram divididos em três grupos, sendo o primeiro com seis integrantes o segundo com cinco e o terceiro com três, isso para que cada grupo tivesse ao menos doze avaliações para ter resultados significativos entre a ordem de audição.

Os testes foram aplicados na mesma sequência sendo inicialmente os áudios de violão, seguidos do saxofone e por último a escaleta. Para evitar que as pessoas associassem um número de amostra à alguma resposta, foram criados números aleatórios de amostra bem como a ordem de reprodução dos áudios. Estes números e ordens aleatórios foram criados com uma fórmula no *software* Microsoft Excel, que gera números aleatoriamente.

A reprodução dos áudios foi feita através de *software* de áudio em um computador com sistema operacional Windows 10, e utilizando um sistema de som de alta fidelidade.

7.3 RESULTADOS OBTIDOS

7.3.1 Teste Sensorial

Realizados os testes, chegou-se a um número total de 42 audições, sendo 14 para cada instrumento. O levantamento de dados mostrou que houve um total de 16 respostas corretas para 42 avaliações às cegas. Conforme a figura 21 mostrada no capítulo 4, conclui-se que este número não é suficiente para rejeitar a hipótese nula, e sendo assim, assume-se que não há diferença significativa entre as amostras de 192Hz com 24 bits e o áudio padrão *RedBook*.

Na figura 67 abaixo é mostrada a tabela com os números obtidos para todos os grupos criados, e em todos os casos a hipótese nula não foi rejeitada, endossando nos agrupamentos o que ocorreu no teste global.

Os grupos apresentados são: tipo de instrumento, gênero do avaliador, faixa de idade, utilização de fones de ouvido e intimidade com fontes de áudio e música.

Figura 67 – Tabela de Resultados por Grupo

Grupo	Total	Corretas	Mínimo Necessario	Hipótese Nula
Instrumentos	-	-	-	-
Violão	14	4	9	Aceita
Saxofone	14	6	9	Aceita
Escaleta	14	6	9	Aceita
Genero	-	-	-	-
Feminino	15	5	9	Aceita
Masculino	27	11	14	Aceita
Idade	-	-	-	-
Até 25 Anos	24	10	13	Aceita
Mais de 25 Anos	18	6	10	Aceita
Uso de Fones	-	-	-	-
Diário	24	10	13	Aceita
Ocasional	18	6	10	Aceita
Intimidade	-	-	-	-
Possui Conhecimento	18	7	10	Aceita
Apenas Usuario	24	9	13	Aceita

Fonte: Os autores.

O número de acertos por grupo e no teste global foi próximo, conforme apresentado na figura 68, mostrando que o índice de percepção não se altera com idade, sexo, ou conhecimento e convivência com áudio. Todos foram próximos dos 38% de acerto, que foi a média de todos os testes, e esta média é ligeiramente acima da média de acertos aleatórios que era de 33,3%, medidas estas que ficam dentro do desvio padrão, logo o resultado obtido era o esperado para o caso da hipótese nula não ser rejeitada.

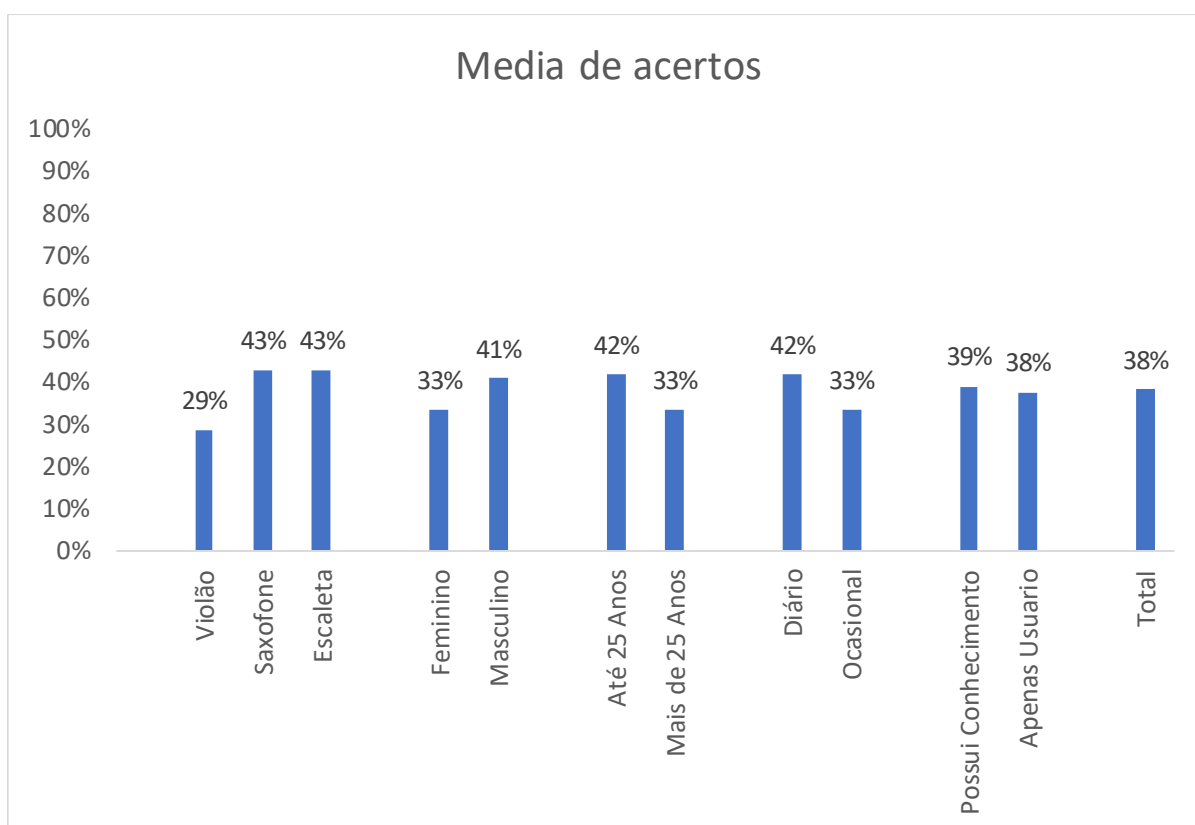
Tabela 15 – Tabela de Perfil dos avaliadores

Avaliador	Idade	Sexo	Profissão	Intimidade com Audio	Conhecimento de audio	Uso de fones de ouvido
1	23	M	Metalurgico	Baixo	Baixo	Ocasional
2	29	F	Tatuadora	Baixo	Baixo	Ocasional
3	14	F	Estudante	Baixo	Baixo	Diario
4	32	M	Engenheiro	Médio	Médio	Ocasional
5	29	M	Analista de Rede	Baixo	Médio	Raro
6	24	M	Estudante	Baixo	Baixo	Diario
7	20	F	Luthier	Alto	Médio	Diario
8	23	M	Tecnico produto	Médio	Médio	Diario

9	24	F	Relações Publicas	Baixo	Médio	Diario
10	23	F	Tecnico Radiologia	Baixo	Baixo	Diario
11	26	M	Empresario	Médio	Baixo	Raro
12	53	M	Professor	Alto	Alto	Ocasional
13	24	M	Desenvolvedo r	Baixo	Médio	Diario
14	35	M	Programador	Médio	Médio	Diario

Fonte: Os autores.

Figura 68 – Gráfico de média por Grupo



Fonte: Os autores.

7.3.2 Teste de favorabilidade

Como já citado, o teste de favorabilidade tem apenas 2 níveis de resposta, mesmo tendo 5 opções no formulário, sendo elas favorável e não favorável.

Na realização deste teste houve um resultado de 50% de favorabilidade, ou seja, em 21 das 42 audições o avaliador respondeu que concorda parcialmente ou

concorda totalmente que há diferença entre os áudios, mesmo com os números da audição às cegas informando que não há diferença significativa entre os áudios.

Mesmo entre os avaliadores que concordaram totalmente que existia diferença entre os áudios houveram erros no formulário sensorial, indicando um possível efeito placebo.

8 RESULTADOS E CONCLUSÕES

Realizando o processo de gravar e reproduzir um áudio em alta fidelidade permitiu ver os desafios práticos dessa modalidade de reprodução de áudio.

Como é possível observar nos espectrogramas dos áudios que foram exportados com uma frequência de 48 kHz, mesmo que tenham taxas de quantização maiores, a sua frequência máxima é de 24 kHz. Há uma clara faixa de corte nas frequências no espectrograma. Isso está de acordo com o Teorema de Nyquist, visto que a maior frequência que o sinal poderia conter seria metade da frequência de amostragem.

O mesmo é válido para os sinais com a frequência de amostragem de 192 kHz. A frequência máxima possível é de 96 kHz, como apontado nos espectrogramas. Entretanto, as intensidades de frequências próximas ao limite de Nyquist para esse caso são quase imperceptíveis. Mesmo assim, em todas as gravações se observam frequências maiores que 24 kHz, ou seja, estas foram ceifadas nos áudios com amostragem menor.

Como esperado, os áudios em 192 kHz tiveram uma potência média maior do que os em 48 kHz, devido as componentes superiores a 24 kHz. Contudo, para os áudios gravados no presente trabalho, essas componentes têm uma participação muito pequena na composição total do áudio, o que é visto em todos os cálculos de potência do capítulo 6, em que a diferença de potência média de todos os casos não ultrapassou 0,0017%.

Analisando os dois parâmetros principais que foram variados no capítulo 6, é possível notar que a forma de onda não se alterou significativamente para mudanças na quantização, na frequência de amostragem, ou em ambas ao mesmo tempo. Cabe frisar que essa afirmação é válida apenas para alterações acima do padrão *RedBook*, visto que se fosse diminuída a quantização e/ou a frequência de amostragem, o áudio apresentaria distorções maiores, devido à baixa amplitude dinâmica e o Teorema de Nyquist, respectivamente.

Como visto no capítulo 4, o limiar teórico da audição é de 20 kHz, assim, o fato de o padrão comercial ser de 44.1 kHz, aqui representado por 48 kHz, seria mais que suficiente para conter todos os sons perceptíveis pelo ser humano.

Entretanto, as altas frequências ainda poderiam ter algum efeito físico que alterasse a percepção do som como um todo.

Utilizando equipamentos capazes de reproduzir além dos 20 kHz, as audições foram realizadas e com elas foi possível observar que a hipótese nula foi aceita, chegando à conclusão que nesse projeto não foram gerados áudios com diferenças significativas de percepção.

Assim, não foram encontrados dados que corroborem que um áudio *high-res* é melhor que um *Redbook*. Mesmo possuindo algumas frequências a mais, estas não geraram impacto na percepção dos ouvintes.

Utilizar frequências de amostragem maiores não se mostrou relevante para os áudios gravados. O mesmo fato ocorre com a quantização. Aumentar de 16 para 24 bits aumentou a variabilidade dinâmica que o arquivo é capaz de armazenar. Todavia, para as gravações realizadas, a precisão de 16 bits foi equivalente à de 24 bits, não afetando a percepção dos ouvintes.

Logo, a frequência de 44.1 kHz e a quantização de 16 bits é suficiente para a gravação e reprodução de sons de maneira satisfatória, e aumentar esses parâmetros nesse trabalho não trouxe melhorias para a experiência do usuário. Ainda cabe ressaltar que cada parâmetro que é aumentado, torna o arquivo maior, sendo necessário mais espaço e poder de processamento para reproduzir o som. Então, esse aumento de espaço, que pode chegar a 6 vezes nesse caso, como mostrado no capítulo 6, pelas técnicas realizadas nesse trabalho, não foi justificado.

Para trabalhos futuros, há a possibilidade de realizar gravações simultâneas, com os mesmos tipos de equipamentos em configurações diferentes, para que qualquer efeito de erro induzido pelo *software* ao mudar a frequência e a quantização não afete o processo. Ainda, existem amplas possibilidades de estudo no campo da alta fidelidade, como os aparelhos de captação, transmissão e reprodução, que foram brevemente citados nesse trabalho.

Na parte dos testes práticos, a bibliografia sugere que realizar treinamentos pode aumentar a chance de percepção da diferença entre áudios *high-res* dos áudios nos padrões *RedBook*.

Os áudios ainda podem ser analisados em combinação com filtros, isolando faixas de frequência e testando elas uma a uma para ver se aumentar os parâmetros de informação do arquivo de áudio alteram a percepção para diferentes faixas de frequência.

REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 12995: Teste triangular em análise sensorial de alimentos e bebidas**. Rio de Janeiro, 1993.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 4120: Teste triangular em análise sensorial dos alimentos e bebidas**. Rio de Janeiro, 2013. 5p.

AMERICAN MUSICAL. **SM58 user guide**. Disponível em: <https://www.americanmusical.com/itemfiles/manual/shu_sm58_userguide.pdf>. Acesso em: 22 out. 2018

ANJOS, Adilson. **SENSOMETRIA Teste Triangular: Adilson dos Anjos Departamento de Estatística Universidade Federal do Paraná** <<https://docs.ufpr.br/~aanjos/SENSOMETRIA/slides/triangular.pdf>> Acesso em 28 set. 2018

AUDIODRIVER. **Microfone condensador ecm8000 - behringer**. Disponível em: <<http://www.audiodriver.com.br/produto/microfone-condensador-ecm8000-behringer/22639>>. Acesso em: 22 out. 2018.

AVID.COM. **Fast track c400™ user guide**. Disponível em: <https://verleih.fhstp.ac.at/upload/m-audio_fast_track_c400_usb_interface_1452382965.pdf>. Acesso em: 14 out. 2018.

BEHRINGER MUSIC GROUP. **ECM800 specs user guide**. Disponível em: <https://downloads.music-group.com/software/behringer/ecm8000/ecm8000_c_specs.pdf>. Acesso em: 22 out. 2018.

BURKERT, Janaina. **Métodos sensoriais**: Disponível em: <<http://www.sabercom.furg.br/bitstream/1/1569/1/m%c3%89todos%20sensoriais%20discriminativos1.pdf>>. Acesso em: 20 set. 2018.

GOMES, Daniel Alexandre Azevedo. **Caracterização da resposta em atraso de grupo de altifalantes e microfones**. Faculdade de engenharia da universidade do porto, 2012.

HALLIDAY, RESNICK, WALKER. **Fundamentos de Física: Vol. 2**. 8 ed. Editora LTC, 2009

HARLEY, Robert. **The complete guide to high-end audio**. 5. Carlsbad, California: Acapella Publishing, 2015

HENRIQUE, Luis L. **Acústica musical**. 3ª ed. Lisboa: Fundação Calouste Gulbenkian, 2011. 1130 p.

IAZZETA, Fernando. **Tutoriais De Áudio E Acústica - Áudio digital**. Disponível em: <http://www2.eca.usp.br/prof/iazzetta/tutor/audio/a_digital/a_digital.html>. Acesso em: 22 set. 2018.

INTERNATIONAL ELECTROTECHNICAL COMMISSION. **IEC 60908** – Audio recording – Compact disc digital audio system: General – Recorded parameters. Geneva, 1999.

KIESER, Robert; REYNISSON, Pall; MULLIGAN, Timothy. Definition of signal-to-noise ratio and its critical role in split-beam measurements. **Ices journal of marine science**, Cidade, v. 62, n. 1, p. 123-130, jan. 2005.

MARCHI, Leonardo De. **A angústia do formato: uma história dos formatos fonográficos**. Revista da associação nacional dos programas de pós-graduação em comunicação, Rio de Janeiro, v.2, 2015.

MARTINS, Tiago de Almeida. **Análise e tratamento acústico para estúdios musicais**. Trabalho de conclusão de curso. SÃO PAULO, 2010.

MELO, D. H. C. T. B. **Uso de dados Ikonos II na análise urbana: testes operacionais na zona leste de São Paulo**. INPE, 2002. 146p.

OPPENHEIM, Alan V.; WILLSKY, Alan S.; NAWAB, Syed Hamid. **Signals & systems**. 2nd ed. Upper Saddle River, NJ: Prentice-Hall, c1997. xxx, 957 p. (Prentice Hall signal processing series). ISBN 0138147574.

OWEN, F.F.E. **PCM and Digital Transmission Systems**. McGraw-Hill, New York, 1982.

PALERMO, Jane Rizzo. **Análise Sensorial Fundamentos e Métodos**. Atheneu, 2015.

PICCINO, Evaldo. **Um breve histórico dos suportes sonoros analógicos**. Campinas, 2005. Disponível em: <http://www.sonora.iar.unicamp.br/sonora1/artigos_pdf/02suportesanalogicos.pdf>. Acesso em: 22 set. 2018.

RUI, Laura Rita. **A física na audição humana**. Porto Alegre: UFRGS, Instituto de Física, Programa de Pós-Graduação em Ensino de Física, 2007.

SHURE.COM. **Sm58**. Disponível em: <<https://br.shure.com/produtos/microfones/sm/microfone-para-voz-principal-e-backing-vocal-sm58>>. Acesso em: 22 out. 2018.

SICA, Carlos. **Sistemas Automáticos com Microcontroladores 8031/8051**. Editora Novatec, 2006. VELLOSO. 186p.

SILVA, R. **High-res Audio: The Basics**. Lifewire, 2018. Disponível em: <<https://www.lifewire.com/hi-res-audio-4134630>>. Acesso em: 07 out. 2018

STUART, Bob. **High-resolution audio a perspective**. Audio eng. soc., v. 63, n. 10, p. 831-832, out. 2015.

TAYLOR, John R., **An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements**, 1997

WIKIPÉDIA. **Guitarra acústica**. Disponível em: <https://pt.wikipedia.org/wiki/Guitarra_ac%C3%A9stica>. Acesso em: 30 out. 2018.

WIKIPÉDIA. **Escaleta**. Disponível em: <<https://pt.wikipedia.org/wiki/Escaleta>>. Acesso em: 04 nov. 2018.

WIKIPÉDIA. **Saxofone alto**. Disponível em: <https://pt.wikipedia.org/wiki/Saxofone_alto>. Acesso em: 30 out. 2018.

YAMAHA PRO AUDIO. **Ag series manual**. Disponível em: <http://www.yamahaproaudio.com/global/en/products/mixers/ag/manual/index_pt.jsp>. Acesso em: 14 out. 2018.

ZUBEN, P. **Música e Tecnologia O Som e Seus Novos Instrumentos**. Rio de Janeiro: Irmãos Vitale.

APÊNDICE A – FICHA DE IDENTIFICAÇÃO

Você será o avaliador número: ____

Nome: _____

Idade: ____ Sexo: () Masculino () Feminino () Outro: _____

Profissão: _____

1. Intimidade com áudio
 - a. Trabalho com música ou tratamento de áudio.
 - b. Sou músico ou tenho hobbies relacionados à áudio ou música.
 - c. Sou apenas um consumidor final de áudios em geral.

2. Com relação à sons em geral
 - a. Trabalho com música ou tratamento de áudio.
 - b. Sou músico ou tenho hobbies relacionados à áudio ou música.
 - c. Sou apenas um consumidor final de áudios em geral.

3. Conhecimento de áudio
 - a. Domino o assunto.
 - b. Conheço algumas propriedades do som e sobre qualidade de áudio
 - c. Não sei o que pode diferenciar um áudio comum de um com alta fidelidade

4. Uso de fones de ouvido
 - a. Utilizo diariamente.
 - b. Utilizo algumas vezes por semana.
 - c. Utilizo apenas para chamadas ou situações atípicas.
 - d. Não utilizo, pois, acho desconfortável.

5. Quando vai ouvir música, qual a principal origem de consumo?
 - a. Aplicativos de streaming: (ex: deezer, Spotify, youtube music)
 - b. Arquivos de áudio em meu dispositivo: (ex: MP3, FLAC, OGG, WAV)
 - c. CDs, DVDs, ou LPs de música.
 - d. Não costumo ouvir música

APÊNDICE B – FICHA DE AVALIAÇÃO

Amostra: ____ - Solicite o numero da amostra ao examinador

Julgador: ____ - Informe o numero que estava na sua ficha de identificação

Você ouviu 3 amostras de áudio codificadas, sendo 2 iguais e uma diferente.

Circule abaixo qual amostra da sequência que julga ser a diferente.

1

2

3

Preencha a tabela abaixo conforme sua percepção da diferença entre os áudios.

Eu Consigo Perceber a diferença entre os Audios reproduzidos



Discordo totalmente ()



Discordo Parcialmente ()



Nem Concordo Nem Discordo ()



Concordo Parcialmente ()



Concordo Totalmente ()

APÊNDICE C – CÓDIGOS MATLAB

mainTCC.m

```
clc, close all, clear all

analiseDiferentesBits()

analiseDiferentesFrequencias()

analiseDiferentesBitseFrequencias()
```

analiseDiferentesBits.m

```
function analiseDiferentesBits()

%% escaleta
escaletaPath = 'Audios/TCC_02_11_18/';
escaleta24bitsName = 'escaleta_192_24';
escaleta16bitsName = 'escaleta_192_16';

escaleta24bitsTitle = 'Escaleta 24 bits 192 kHz';
escaleta16bitsTitle = 'Escaleta 16 bits 192 kHz';
escaletaDifferenceAudiosTitle = 'Diferença entre escaleta 24 e 16 bits'

escaleta24bitsSpectrogramTitle = 'Espectrograma - Escaleta 24 bits 192 kHz';

escaleta16bitsSpectrogramTitle = 'Espectrograma - Escaleta 16 bits 192 kHz';

[powerFirst, powerSecond, powerDifference, mse, smse, snr] =
analiseAudioWith(escaletaPath, escaleta24bitsName, escaleta16bitsName,
escaleta24bitsTitle, escaleta16bitsTitle, escaletaDifferenceAudiosTitle,
escaleta24bitsSpectrogramTitle, escaleta16bitsSpectrogramTitle);

%% Violão
violaoPath = 'Audios/TCC_02_11_18/';
violao24bitsName = 'violaoPrecursaoImicrofone_Mono_192_24';
violao16bitsName = 'violaoPrecursaoImicrofone_Mono_192_16';

violao24bitsTitle = 'Violão 24 bits 192 kHz';
violao16bitsTitle = 'Violão 16 bits 192 kHz';
violaoDifferenceAudiosTitle = 'Diferença entre violão 24 e 16 bits'

violao32bitsSpectrogramTitle = 'Espectrograma - Violão 24 bits 192 kHz';

violao16bitsSpectrogramTitle = 'Espectrograma - Violão 16 bits 192 kHz';
```

```
[powerFirst, powerSecond, powerDifference, mse, smse, snr] =
analiseAudioWith(violaoPath, violao24bitsName, violao16bitsName,
violao24bitsTitle, violao16bitsTitle, violaoDifferenceAudiosTitle,
violao32bitsSpectrogramTitle, violao16bitsSpectrogramTitle);
```

```
%% Saxofone
```

```
saxofonePath = 'Audios/TCC_02_11_18/';
saxofone24bitsName = 'saxofone_Mono_controlevolume_192_24';
saxofone16bitsName = 'saxofone_Mono_controlevolume_192_16';
```

```
saxofone24bitsTitle = 'Saxofone 24 bits 192 kHz';
saxofone16bitsTitle = 'Saxofone 16 bits 192 kHz';
saxofoneDifferenceAudiosTitle = 'Diferença entre saxofone 24 e 16 bits'
```

```
saxofone24bitsSpectrogramTitle = 'Espectrograma - Saxofone 24 bits 192
kHz';
```

```
saxofone16bitsSpectrogramTitle = 'Espectrograma - Saxofone 16 bits 192
kHz';
```

```
[powerFirst, powerSecond, powerDifference, mse, smse, snr] =
analiseAudioWith(saxofonePath, saxofone24bitsName, saxofone16bitsName,
saxofone24bitsTitle, saxofone16bitsTitle, saxofoneDifferenceAudiosTitle,
saxofone24bitsSpectrogramTitle, saxofone16bitsSpectrogramTitle);
end
```

analiseDiferentesFrequencias.m

```
function analiseDiferentesFrequencias()
```

```
%% escaleta
```

```
escaletaPath = 'Audios/TCC_02_11_18/';
escaleta192kHzName = 'escaleta_192_24';
escaleta48kHzName = 'escaleta_48_24';
```

```
escaleta192kHzTitle = 'Escaleta 24 bits 192 kHz';
escaleta48kHzTitle = 'Escaleta 24 bits 48 kHz';
escaletaDifferenceAudiosTitle = 'Diferença entre escaleta 192 kHz e 48
kHz';
```

```
escaleta24bitsSpectrogramTitle = 'Espectrograma - Escaleta 24 bits 192
kHz';
```

```
escaleta16bitsSpectrogramTitle = 'Espectrograma - Escaleta 24 bits 48 kHz';
```

```
[powerFirst, powerSecond, powerDifference, mse, smse, snr] =
analiseAudioWith(escaletaPath, escaleta192kHzName, escaleta48kHzName,
escaleta192kHzTitle, escaleta48kHzTitle, escaletaDifferenceAudiosTitle,
escaleta24bitsSpectrogramTitle, escaleta16bitsSpectrogramTitle);
```

```
%% violão
```

```
clc, close all, clear all
violaoPath = 'Audios/TCC_02_11_18/';
violao192kHzName = 'violaoPrecursoraolmicrofone_Mono_192_24';
```

```

violao48kHzName = 'violaoPrecursaoImicrofone_Mono_48_24';

violao192kHzTitle = 'Violão 24 bits 192 kHz';
violao48kHzTitle = 'Violão 24 bits 48 kHz';
violaoDifferenceAudiosTitle = 'Diferença entre violão 48kHz e 192 kHz'

violao192kHzSpectrogramTitle = 'Espectrograma - Violão 24 bits 192 kHz';

violao48kHzSpectrogramTitle = 'Espectrograma - Violão 24 bits 48 kHz';

[powerFirst, powerSecond, powerDifference, mse, smse, snr] =
analiseAudioWith(violaoPath, violao192kHzName, violao48kHzName,
violao192kHzTitle, violao48kHzTitle, violaoDifferenceAudiosTitle,
violao192kHzSpectrogramTitle, violao48kHzSpectrogramTitle);

%% Saxofone

saxofonePath = 'Audios/TCC_02_11_18/';
saxofone192kHzName = 'saxofone_Mono_controlevolume_192_24';
saxofone148kHzName = 'saxofone_Mono_controlevolume_48_24';

saxofone192kHzTitle = 'Saxofone 24 bits 192 kHz';
saxofone48kHzTitle = 'Saxofone 24 bits 48 kHz';
saxofoneDifferenceAudiosTitle = 'Diferença entre saxofone 24 e 21 bits'

saxofone192kHzSpectrogramTitle = 'Espectrograma - Saxofone 24 bits 192
kHz';

saxofone48kHzSpectrogramTitle = 'Espectrograma - Saxofone 24 bits 48 kHz';

[powerFirst, powerSecond, powerDifference, mse, smse, snr] =
analiseAudioWith(saxofonePath, saxofone192kHzName, saxofone148kHzName,
saxofone192kHzTitle, saxofone48kHzTitle, saxofoneDifferenceAudiosTitle,
saxofone192kHzSpectrogramTitle, saxofone48kHzSpectrogramTitle);

end

```

analiseDiferentesBitseFrequencias.m

```

function analiseDiferentesBitseFrequencias

%% escaleta
escaletaPath = 'Audios/TCC_02_11_18/';
escaleta192kHzName = 'escaleta_192_24';
escaleta48kHzName = 'escaleta_48_16';

escaleta192kHz24bitsTitle = 'Escaleta 24 bits 192 kHz';
escaleta48kHz16bitsTitle = 'Escaleta 16 bits 48 kHz';
escaletaDifferenceAudiosTitle = 'Diferença entre escaleta 48kHz/192 kHz e
16/24bits';

escaleta192kHz24bitsSpectrogramTitle = 'Espectrograma - Escaleta 24 bits
192 kHz';

escaleta24kHz16bitsSpectrogramTitle = 'Espectrograma - Escaleta 16 bits 48
kHz';

```

```

[powerFirst, powerSecond, powerDifference, mse, smse, snr] =
analiseAudioWith(escaletaPath,          escaleta192kHzName, escaleta48kHzName,
escaleta192kHz24bitsTitle,              escaleta48kHz16bitsTitle,
escaletaDifferenceAudiosTitle,
escaleta192kHz24bitsSpectrogramTitle, escaleta24kHz16bitsSpectrogramTitle);

%% violão
violaoPath = 'Audios/TCC_02_11_18/';
violao192kHz24bitsName = 'violaoPrecursaoImicrofone_Mono_192_24';
violao48kHz16bitsName = 'violaoPrecursaoImicrofone_Mono_48_16';

violao192kHz24bitsTitle = 'Violão 24 bits 192 kHz';
violao48kHz16bitsTitle = 'Violão 16 bits 48 kHz';
violaoDifferenceAudiosTitle = 'Diferença entre violão 48kHz/192 kHz e
16/24bits'

violao192kHz24bitsSpectrogramTitle = 'Espectrograma - Violão 24 bits 192
kHz';

violao48kHz16bitsSpectrogramTitle = 'Espectrograma - Violão 16 bits 48
kHz';

[powerFirst, powerSecond, powerDifference, mse, smse, snr] =
analiseAudioWith(violaoPath, violao192kHz24bitsName, violao48kHz16bitsName,
violao192kHz24bitsTitle,              violao48kHz16bitsTitle,
violaoDifferenceAudiosTitle,
violao192kHz24bitsSpectrogramTitle, violao48kHz16bitsSpectrogramTitle);

%% Saxofone
saxofonePath = 'Audios/TCC_02_11_18/';
saxofone192kHzName = 'saxofone_Mono_controlevolume_192_24';
saxofone148kHzName = 'saxofone_Mono_controlevolume_48_16';

saxofone192kHzTitle = 'Saxofone 24 bits 192 kHz';
saxofone48kHzTitle = 'Saxofone 16 bits 48 kHz';
saxofoneDifferenceAudiosTitle = 'Diferença entre saxofone 48kHz/192 kHz e
16/24bits'

saxofone192kHzSpectrogramTitle = 'Espectrograma - Saxofone 24 bits 192
kHz';

saxofone48kHzSpectrogramTitle = 'Espectrograma - Saxofone 16 bits 48 kHz';

[powerFirst, powerSecond, powerDifference, mse, smse, snr] =
analiseAudioWith(saxofonePath,          saxofone192kHzName, saxofone148kHzName,
saxofone192kHzTitle,  saxofone48kHzTitle,  saxofoneDifferenceAudiosTitle,
saxofone192kHzSpectrogramTitle, saxofone48kHzSpectrogramTitle);

end

```

analiseAudioWith.m

```

function [powerFirst, powerSecond, powerDifference, mse, smse, snr] =
analiseAudioWith(path, nameFirstAudio, nameSecondAudio, title1, title2,
title3, title4, title5)

firstAudioPath = strcat(path, nameFirstAudio);

```

```

secondAudioPath = strcat(path,nameSecondAudio);

firstAudio = getAudio(firstAudioPath, 'wav');

secondAudio = getAudio(secondAudioPath, 'wav');

%% Plot Files and Spectrograms
figure
plotAudio(firstAudio(:,1), title1)
figure
plotAudio(secondAudio(:,1), title2)
if length(secondAudio) == length(firstAudio)
difference = secondAudio - firstAudio;
figure
plotAudio(difference(:,1), title3)
mse = getMSE(firstAudio, secondAudio);
smse = getSMSE(firstAudio, secondAudio);
snr = getSNR(firstAudio, secondAudio);
else
mse = 0;
smse = 0;
snr = 0;
end
figure
plotSpectrogram(firstAudio(:,1), 192000,title4)
figure
plotSpectrogram(secondAudio(:,1), 48000,title5)

%% Return parameters
powerFirst = getAudioPower(firstAudio);
powerSecond = getAudioPower(secondAudio);
powerDifference = powerFirst-powerSecond;

```

```
end
```

getAudio.m

```

function audio = getAudio(name,type)

audioName = strcat(name, '.');
audioName = strcat(audioName,type);
audio = audioread(audioName);
end

```

getAudioPower.m

```

function power = getAudioPower(audio)

powerArray = bandpower(audio);

if min(length(powerArray)) > 1
power = (powerArray(1)+powerArray(2))/2;
else
power = powerArray;
end
end

```

getMSE.m

```
function mse = getMSE(originalSound, compressedSound)
mse = immse(originalSound, compressedSound);
end
```

```
function smse = getSMSE(originalSound, compressedSound)
smse = sqrt(immse(originalSound, compressedSound));
end
```

getSMSE.m

```
function smse = getSMSE(originalSound, compressedSound)
smse = sqrt(immse(originalSound, compressedSound));
end
```

getSNR.m

```
function SNR = getSNR(originalSound, compressedSound)
noise = originalSound - compressedSound;
SNR = snr(originalSound, noise);
end
```

plotAudio.m

```
function plotAudio(audio, titleText)

plot(audio);
title(titleText);

end
```

plotSpectrogram.m

```
function plotSpectrogram(audio, frequency, titleText)

spectrogram(audio, hanning(512), 128, 512, frequency, 'yaxis')
colormap jet
title(titleText)
end
```