

UNIVERSIDADE ESTADUAL DO PARANÁ
DEPARTAMENTO DE MÚSICA
CURSO DE ESPECIALIZAÇÃO EM MÚSICA ELETROACÚSTICA

FILIPE BORATO DE CASTRO

**MEMORIAL DE COMPOSIÇÃO: PEÇA ELETROACÚSTICA *UM
AMONTOADO DE GENTE***

CURITIBA
2023

FILIPE BORATO DE CASTRO

**MEMORIAL DE COMPOSIÇÃO: PEÇA ELETROACÚSTICA *UM
AMONTÓADO DE GENTE***

Trabalho de Conclusão de Curso apresentado ao
Curso de Especialização em Música Eletroacús-
tica da Universidade Estadual do Paraná.
Orientador: Prof. Dr. Alvaro Henrique Borges

CURITIBA
2023

Ficha catalográfica elaborada pelo Sistema de Bibliotecas da UNESPAR e Núcleo de Tecnologia de Informação da UNESPAR, com Créditos para o ICMC/USP e dados fornecidos pelo(a) autor(a).

Borato de Castro, Filipe
MEMORIAL DE COMPOSIÇÃO: PEÇA ELETROACÚSTICA "UM AMONTOADO DE GENTE" / Filipe Borato de Castro. -- Curitiba-PR, 2023.
36 f.: il.

Orientador: Alvaro Henrique Borges.
Especialização em Música Eletroacústica - Universidade Estadual do Paraná, 2023.

1. Inteligência Artificial. 2. Computação para música. 3. DDSP. 4. NSynth. 5. Música Eletroacústica. I - Borges, Alvaro Henrique (orient). II - Título.

RESUMO

Este trabalho tem como objetivo a criação e relatos composicionais da peça eletroacústica *Um Amontado de Gente* que utiliza de ferramentas digitais em fase pesquisa e inovação na área de Inteligência Artificial para música e/ou algoritmos digitais já consolidados. Narrando-se as etapas do processo composicional com suas ferramentas, técnicas e motivações.

Palavras-chaves: Música Eletroacústica. Composição. IA Voltado a Composição Musical. DDSF. NSynth.

ABSTRACT

The objective of this work is the creation and compositional reports of the electroacoustic piece *Um Amontado de Gente* that uses digital tools in the research and innovation phase in the area of Artificial Intelligence for music and/or already consolidated digital algorithms. Narrating the stages of the compositional process with their tools, techniques and motivations.

Key-words: Electroacoustic Music. Composition. AI Focused on Music Composition. DDSP. NSynth.

LISTA DE FIGURAS

Figura 1	– Cabeçalho do notebook Nsynth	15
Figura 2	– Cabeçalho do notebook DDSP	16
Figura 3	– Reconhecimento de fundamentais	18
Figura 4	– Processo de parametrização de fundamentais e ajuste de <i>Loudness</i>	20
Figura 5	– Espectrograma de Frequência x Tempo da Ressíntese	20
Figura 6	– <i>Patch Automatonism 3.1</i> implementando Granulador	21
Figura 7	– Software <i>Emission Control 2</i>	22
Figura 8	– <i>Plugins PitchAccum, Evolution, ++pitchdelay e SpaceGrain</i>	24
Figura 9	– O projeto no <i>Reaper</i> da peça eletroacústica	25
Figura 10	– Trecho da melodia da flauta criada pelo <i>DDSP</i>	29
Figura 11	– Espectrograma Melódico de um trecho de objetos graves	30
Figura 12	– Espectrograma de Picos de frequência de uma faixa textural	31
Figura 13	– Espectrograma de picos de frequência de uma faixa textural	32
Figura 14	– Espectrograma de picos de frequência e a forma da onda no tempo de trechos graves e médios	32
Figura 15	– Espectrograma de picos de frequência e a forma da onda domínio do tempo de trechos de texturas e movimentações	33
Figura 16	– Análise Geral da Peça	34
Figura 17	– Análise de Temas e Variações	35

LISTA DE ABREVIATURAS E SIGLAS

AI	<i>Artificial Intelligence</i>
AM	<i>Amplitude Modulation</i>
CPU	<i>Central Process Unit</i>
DAW	<i>Digital Audio Workstation</i>
DNN	<i>Deep Neural Networks</i>
DDSP	<i>Differentiable Digital Signal Processing</i>
DSP	<i>Digital Signal Processing</i>
FFT	<i>Fast Fourier Transform</i>
FM	<i>Frequency Modulation</i>
GPU	<i>Graphics Processing Unit</i>
GRM	<i>Groupe de recherches musicales</i>
Hz	Hertz
IA	Inteligência Artificial
VST	<i>Virtual Studio Technology</i>

SUMÁRIO

1	INTRODUÇÃO	8
2	TÉCNICA E POÉTICA	10
2.1	TÉCNICAS DE COMPOSIÇÃO MUSICAL ASSISTIDA POR COMPUTADOR ..	10
2.2	POÉTICA E CAMINHOS COMPOSICIONAIS	12
3	PROCESSO COMPOSICIONAL E TÉCNICAS APLICADAS	14
3.1	NOVAS TÉCNICAS DE PROCESSAMENTO DIGITAL NO ÁUDIO	14
3.2	TÉCNICAS DIFUNDIDAS DE PROCESSAMENTO DIGITAL NO ÁUDIO	21
3.3	EDIÇÃO, MIXAGEM E COMPOSIÇÃO	24
4	ANÁLISE MUSICAL DA PEÇA ELETROACÚSTICA <i>UM AMONTOADO DE GENTE</i>	27
4.1	MICROELEMENTOS, <i>MEZZOELEMENTOS</i> E MACROELEMENTOS	28
4.2	MICROELEMENTOS - FRAGMENTOS	29
4.3	<i>MEZZOELEMENTOS</i> - PARTES DE UM TODO	31
4.4	MACROELEMENTOS - ESTRUTURA TOTAL	34
5	CONCLUSÃO	36
	REFERÊNCIAS	37

1 INTRODUÇÃO

A utilização de ferramentas para composição por meio computacional e o auxílio de ferramentas de inteligência artificial é aplicado no meio musical há algumas décadas (MANTARAS; ARCOS, 2002, pag. 43-44). Os computadores pessoais são usados como ferramenta de editoração musical, registro sonoro, processamento sonoro, síntese sonora e entre outros. A composição musical assistida por computador que obedece um conjunto de regras, que podemos chamar de algoritmos, são como receitas de bolo muito sofisticadas. Quanto mais detalhadas forem essas regras, que o computador executa em sequência, mais a composição se assemelha a uma obra humana. Nesse caso, em uma composição com um conjunto aprimorado de regras executadas sequencialmente por uma máquina, a criação da composição musical está totalmente sujeita ao desenvolvedor destes códigos para que a máquina os execute.

O estudo da inteligência artificial (IA) tem tido um grande avanço em várias áreas do conhecimento que envolvem a tentativa de implementação de sistemas inteligentes auto-regulados e que não dependem do homem para decidir que escolha tomar (KOVACS, 2006, pag. 3-4). É provável que a tentativa de implementação de uma IA autônoma que colabore com o compositor na criação do material composicional não esteja tão distante de ser criada, ou talvez já tenha sido criada com determinadas limitações de projeto. Pode-se definir a inteligência artificial em duas categorias: aprendizado de máquina supervisionado e não supervisionado. O aprendizado supervisionado é como uma alegoria da receita de bolo, depende do criador da receita para criar as instruções a serem seguidas (HAYKIN, 2001, pag. 27-29). Já o aprendizado não supervisionado, utilizando outra alegoria, a própria máquina cria a sua receita de bolo.

Definindo que as novas ferramentas digitais são apenas meios e que o objetivo final é a criação composicional, pode-se inferir que as técnicas envolvidas serão usadas como recursos sonoros para a criação de uma obra musical. A finalidade deste trabalho é composicional e trata-se de uma descrição desse processo. Este trabalho não tem a pretensão de definir fórmulas composicionais criadas pela máquina, mas sim de utilizar e incrementar diferentes e novos algoritmos de processamento de áudio que têm surgido em novas pesquisas. O objetivo é encontrar ferramentas que possibilitam novos sons ou novos materiais para compor.

Desta feita, os processos artísticos estão geralmente interligados ao momento que a sociedade é o que ela é e a tentativa do artista em expressar o momento da história aonde ele esteve inserido. Ignorar a tecnologia pode ser um ato de tentar preservar a música tradicional como uma conservação de uma época que não existe mais, talvez seja uma tentativa desesperada da sobrevivência social do músico contemporâneo, mas ousar na utilização de novos recursos e ferramentas composicionais é provável que gere inovação na criação musical.

A possibilidade de criação de novos timbres desenvolvidos por aprendizado de máquina, inova no alcance de novas técnicas de composição para música eletroacústica. Tendo em vista que a composição do timbre é formada por frequências que variam no tempo (ROEDE-

RER, 2002, pag. 11) e que por meio de algoritmos de aprendizado que analisam e categorizam dados de áudio, há a possibilidade de extrair e criar timbres ricos e complexos objetos sonoras.

Devido a quantidade de materiais, pesquisas e plataformas que envolvem a criação de algoritmos para aprendizado de máquinas, uma das plataformas mais populares sendo o Google Colab (COLAB, 2022), a utilização operacional é facilitada, não necessitando de um alto nível de conhecimento computacional e matemático, mas um conhecimento mais básico sobre o fluxo da programação e alteração dos seus parâmetros para utilização de suas funcionalidades principais.

O objetivo principal deste Memorial é apresentar os fundamentos e o processo da composição musical de uma peça eletroacústica feita por meio do auxílio computacional e o auxílio da inteligência artificial para a criação do material composicional. Neste processo foram contemplados os seguintes passos: registros de material sonoro utilizados como "matéria prima" para peça eletroacústica; a criação de timbres por meio de síntese sonora e utilizando novas técnicas de síntese baseadas em aprendizado de máquina e processamento digital de áudio (MAGENTA, 2021); a criação de material composicional por aprendizado de máquina supervisionado e não supervisionado e, por fim, a composição de uma peça eletroacústica para mais de dois canais estereofônicos.

Como o trabalho teórico faz parte da proposta do curso, haverá um esforço em demonstrar e relatar verbalmente os processos composicionais e as técnicas para o desenvolvimento da composição eletroacústica. A dificuldade em relatar ou comunicar uma narrativa composicional eletroacústica gera a necessidade de utilizar termos e ideias de teóricos da área. É emprestado o termo "objeto sonoro" e "objeto musical" de Pierre Schaeffer para definir trechos e faixas sonoras que foram criados ou capturados pelo compositor, com a intenção de criar um afastamento do sentido semântico e uma valorização das características do som (SCHAEFFER, 2003).

2 TÉCNICA E POÉTICA

Este capítulo será definido questões pré-compositivas e técnicas aprendidas pelo compositor ao longo de sua trajetória.

2.1 TÉCNICAS DE COMPOSIÇÃO MUSICAL ASSISTIDA POR COMPUTADOR

A primeira consideração nesta seção é a definição do que é composição musical assistida por computador. Na definição de Jean-Claude Risset, no artigo *Musique, calcul secret?*, a música assistida por computador é um meio, ou ferramenta, para um fim composicional. Na visão de Risset, na década de setenta, o músico deve ser capaz de dominar o computador para se organizar e criar ideias, possibilitando o seu próprio universo musical (RISSET, 1977, pag. 3). Segundo Otto Laske, o computador, por meio de algoritmos, regula a criação com total controle e rigor, com o qual o compositor-programador compõe sua música. No entanto, existe uma zona intermediária que flexibiliza as regras computacionais e é aproveitada pela teoria composicional (LASKE, 1981, pag. 45).

Uma das questões que envolvem o questionamento da composição por auxílio computacional é o quanto a máquina interfere nas escolhas da criação e pensamento humano sobre a música composta e o quanto de abertura e controle é possível ser dado a uma máquina. Quando se fala em algoritmos de aprendizagem de máquina, ocorre uma certa ilusão de que o controle das escolhas não está mais nas mãos do programador que utilizou determinada técnica, mas a escolha pela utilização é humana, independentemente do tipo de processo envolvido. Neste trabalho, a escolha é utilizar o computador como ferramenta, baseando-se nos conceitos citados acima.

Em 2022, ano em que este trabalho está em desenvolvimento, o músico tem em suas mãos um leque de ferramentas digitais para o processamento e criação sonora facilitado por meio de computadores pessoais. A acessibilidade dessas ferramentas é facilitada, mas o desafio reside no conhecimento técnico necessário para a criação composicional de música eletroacústica, bem como no desejo e ação para compor. Existem vários caminhos para se compor música eletroacústica, e neste trabalho foi escolhido utilizar um material gravado em áudio digital de uma execução vocal. Esse material será utilizado como uma *pedra bruta*, que será transformada em uma obra.

A maior parte da teoria matemática envolvida em softwares de processamento de áudio foi formulada até o final do século XIX e aplicada na tecnologia ao longo do século XX. A formulação mais utilizada são os princípios de Fourier sobre a representação funcional de fenômenos e oscilações periódicas (OPPENHEIM; SCHAFER, 1999, pag. 2-7). Os processos digitais utilizados no áudio, como a manipulação do tempo, amplitude, frequência e modulações

do tipo AM e FM, utilizam as descobertas de Fourier sobre o sinal, discretizadas e otimizadas para a computação.

Hoje, em 2022, as pesquisas de algoritmos que possuem multiprocessamento e cálculos avançados em GPU, priorizando a operação digital de multiplicações de pontos flutuantes, fomentam pesquisas na área denominada de Inteligência Artificial. Na área de produção de capital, a IA tem sido amplamente estudada e implementada em várias corporações privadas e estatais. Grande parte deste estudo é muito mais antigo do que a sua aplicação e implementam pesquisas acadêmicas de diversos momentos históricos. Os novos estudos na área do áudio são um reflexo desse momento. Tenho encontrado com facilidade *papers* e códigos publicados em diversas plataformas de computação colaborativa que disponibilizam estudos com esse viés.

A primeira parte do estudo de novas tecnologias que envolvem aprendizado de máquina foi realizada a partir de um levantamento de estudos recentes nessa área e da sua acessibilidade para execução. Encontrei publicações da comunidade *Magenta*, que disponibilizam estudos e códigos-fonte preparados para computação em nuvem para qualquer pessoa com acesso à internet (MAGENTA, 2021). Os códigos-fonte são disponibilizados em uma plataforma utilizada para pesquisas nesse ramo chamada *Google Colab*, aparentemente gratuita e com um determinado interesse nesta área de pesquisa.

Pesquisando a comunidade *Magenta* encontrei um algoritmo de ressíntese que implementa um dos conceitos do aprendizado de máquina para recriar sinteticamente um instrumento acústico a partir de gravações previamente coletadas, a partir de uma entrada de áudio ou midi. Um dos projetos dessa comunidade é o *NSynth*. Ao contrário de um sintetizador tradicional que gera áudio a partir de componentes projetados à mão, como osciladores e wavetables, o *NSynth* usa redes neurais profundas para gerar sons no nível de amostras individuais. Aprendendo diretamente com os dados, o *NSynth* fornece aos artistas controle intuitivo sobre timbre e dinâmica e a capacidade de explorar novos sons que seriam difíceis ou impossíveis de produzir com um sintetizador ajustado à mão (MAGENTA, 2021).

Outro algoritmo de aprendizado de máquina foi desenvolvida pela equipe da *Magenta* é o *DDSP: Differentiable Digital Signal Processing*. A ideia-chave do projeto é usar elementos *DSP* simples e interpretáveis para criar sinais realistas complexos, controlando com precisão seus parâmetros. Por exemplo, uma coleção de filtros lineares e osciladores senoidais (elementos *DSP*) pode criar o som de um violino realista se as frequências e respostas forem ajustadas da maneira certa. No entanto, é difícil controlar dinamicamente todos esses parâmetros manualmente, e é por isso que sintetizadores com controles simples geralmente soam artificiais e “sintéticos”. Com o *DDSP*, usa-se uma rede neural para converter a entrada de um usuário em controles *DSP* complexos que podem produzir sinais mais realistas. Essa entrada pode ser qualquer forma de sinal de controle, incluindo recursos extraídos do próprio áudio. Como as unidades *DDSP* são diferenciáveis (portanto, o D extra), pode-se treinar a rede neural para se adaptar a um conjunto de dados por meio de retro-propagação padrão (ENGEL; HANTRAKUL; GU, 2020).

No *DDSP* é implementado a técnica de Síntese de Modelagem Espectral que corrobora

para um sinal mais realista devido a quantidade de harmônicos e ruídos inerentes a instrumentos acústicos. A Rede Neural Recorrente que extrai as informações de áudios entrada é acoplada a essa técnica que gera instrumentos virtuais complexos. Por exemplo, nos 4 segundos de áudio de 16kHz nos conjuntos de dados considerados no estudo da construção do *DDSP*, os coeficientes do sintetizador realmente têm 2,5 vezes mais dimensões do que a própria forma de onda de áudio ((1 amplitude + 100 harmônicos + 65 banda de ruído magnitudes) * 1.000 intervalos de tempo = 165.000 dimensões, contra 64.000 amostras de áudio). Isso os torna passíveis de controle por uma rede neural, pois seria difícil especificar realisticamente todos esses parâmetros à mão (ENGEL; HANTRAKUL; GU, 2020).

Neste trabalho, são utilizados algoritmos já consolidados e amplamente difundidos, criados a partir do princípio da transformada rápida de Fourier, para realizar alterações nas características sonoras, como mudanças de *pitch*, *stretch* e utilização de filtros simples ou combinatórios. Como sistema de processamento sonoro, optei pelo software *Pure Data* e como editor e mixer, escolhi o software *Reaper*. No *Pure Data* (PD, 2022), utilizei *patches*, que são estruturas de objetos modulares criados neste tipo de software, como por exemplo, o *Automatizism 3.1*, que conta com diversos tipos de processamento digital já criados e de fácil utilização (ERIKSSON, 2022). A busca pela utilização desses softwares tem como objetivo a reconstrução e criação sonora de novos timbres e entidades sonoras que sejam adequados à peça composta.

2.2 POÉTICA E CAMINHOS COMPOSICIONAIS

O desejo por compor esta peça eletroacústica está interligado ao caminho que percorri na minha trajetória musical. Devido ao meu interesse por tecnologia em música e áudio, eu encontrei na música eletroacústica fatores que me fazem querer compor neste estilo. Também tenho interesse em compor para instrumento e eletrônica, mas por questões técnicas e logística escolhi que esta peça será acusmática. Tenho o interesse em manipular as ferramentas digitais de maneira que se aproxima execução de um instrumentista. Acredito no uso da tecnologia para compor, trata-se de uma ferramenta para alcançar seus objetivos previamente elaborados. A experimentação faz parte do processo composicional, mas não é o meu intuito final.

A poética pensada para essa composição tem a ver com alguns estudos livres em sociologia e filosofia que tenho feito ao longo dos anos, especificamente a tentativa da construção da ideia de povo brasileiro. A utilização desse tema é de caráter artístico e não pretende definir os diversos conceitos ou teorizações a cerca dele. Parto dessa ideia como ímpeto inicial para a composição musical da peça *Um Amontoado de Gente*. O desejo em criar esta peça está relacionado à vontade em expressar diversos sentimentos e disposições políticas na minha trajetória e trata-se, também, de um tributo há um grande amigo meu assassinado brutalmente na cidade de Curitiba em 2022. É da minha prática refletir, conversar, ler e ouvir sobre assuntos filosóficos que moldam as atitudes diárias e para mim estes conhecimentos são transformadores e ao

mesmo tempo libertadores de diversas formas de controle e dominação social. Imprimir tais posicionamentos de forma musical me motiva a compor.

O título da obra "*Um Amontoado de Gente*" é uma crítica ou uma reflexão sobre falta de unicidade nacional no Brasil. Entendo que não é um assunto de poucas linhas e este trabalho não pretende definir a construção da identidade do povo brasileiro, mas expor de maneira artística na obra algumas impressões e sentimentos sobre o tema. Tais reflexões estão centradas em leituras de alguns pensadores brasileiros como Florestan Fernandes, Celso Furtado, Sergio Buarque de Holanda, Darcy Ribeiro, Fábio Comparato e Jessé Souza. Tais teóricos que dedicaram parte do seu trabalho pensando, refletindo, construindo e debatendo questões referentes a sociedade brasileira. Acredito que seja um assunto de determinada relevância cotidiana para quem reflete sobre situações sociais no Brasil e recorrentes em debates formais ou informais.

Como material primário, decidi utilizar vozes e peças vocais com o objetivo de proporcionar maior expressividade e articulação em cada objeto sonoro criado (SCHAEFFER, 2003, pag. 26-27). Manipulo os objetos sonoros utilizando técnicas de variações e contraponto para compor a peça. A técnica de variação na eletroacústica pode ser usada em diversos aspectos do som, como timbre, modulação, tempo e reverberação. O canto possui grande articulação e movimento interno tímbrico, e a manipulação e alteração do som desses registros criam diferentes timbres e trechos momentâneos com grande intencionalidade humana. A variação e manipulação do som pode remover a inteligibilidade da origem da fonte sonora, tornando a fonte geradora do som irreconhecível ao ouvinte e formando novos objetos musicais com um potencial a ser utilizado na música eletroacústica (SCHAEFFER, 2003).

3 PROCESSO COMPOSICIONAL E TÉCNICAS APLICADAS

Este capítulo apresenta uma visão geral das técnicas e processos que foram utilizados para a composição da peça *Um Amontoado de Gente*. Descreve-se os métodos tecnológicos utilizados e as variações sonoras obtidas a partir do processamento digital do áudio. Serão relatados os processos de ressíntese sonora obtidos a partir de aprendizado de máquina, filtros digitais, granuladores, *reverbs*, mixagem e espacialização.

3.1 NOVAS TÉCNICAS DE PROCESSAMENTO DIGITAL NO ÁUDIO

Ao escolher o *Notebook NSynth*, comecei a manipular registros de áudio vocais de trechos da peça *Sequenza III* de Luciano Berio. O *Notebook* é um arquivo que contém textos, figuras e códigos em Python, disponibilizados pelo *Framework Jupyter*, e é atualmente a ferramenta mais difundida para o desenvolvimento de estudos em aprendizado de máquina. O *Google Colab* é um dos fornecedores da plataforma *Jupyter*. O objetivo do algoritmo *NSynth*¹ é coletar informações sobre vários instrumentos musicais a partir de uma rede neural e transpor os aspectos tímbricos para o registro de áudio de entrada. Neste algoritmo, a impressão é de que o resultado sonoro é uma interpolação da voz humana com instrumentos do grupo das Madeiras de uma orquestra, pois o algoritmo foi pré-treinado utilizando uma grande base de trechos gravados de instrumentos musicais. Além disso, é possível utilizar dois ou mais áudios de entrada com timbres distintos, que serão interpolados e re-sintetizados para gerar um único timbre.

No código em Python² abaixo, temos um exemplo dos parâmetros de entrada disponíveis para alteração pelo programador/compositor. Esses parâmetros tratam do tempo e do período de interpolação entre o áudio de entrada e a síntese por rede neural. Ao alterar o parâmetro *SampleLength* para uma duração mais longa, percebe-se um “esticamento” da faixa no novo trecho de áudio, como se um trecho musical alterasse seu andamento para uma cadência mais lenta. Aumentar o *SampleLength* torna o algoritmo mais custoso em termos de duração do processamento. Com o parâmetro definido como 441000, o tempo de execução do algoritmo é em torno de 3 horas. Como é chamada uma função de ressíntese que é um módulo Python externo ao *Notebook*, é difícil analisar por que o tempo de processamento para ressíntese e interpolação dos parâmetros extraídos por uma rede neural é tão custoso.

```
#@title Set Sound Length (in Seconds) { vertical-output: true, run: "auto" }
Length = 10 #@param {type:"number"}
SR = 16000
SAMPLE_LENGTH = int(SR * Length)
```

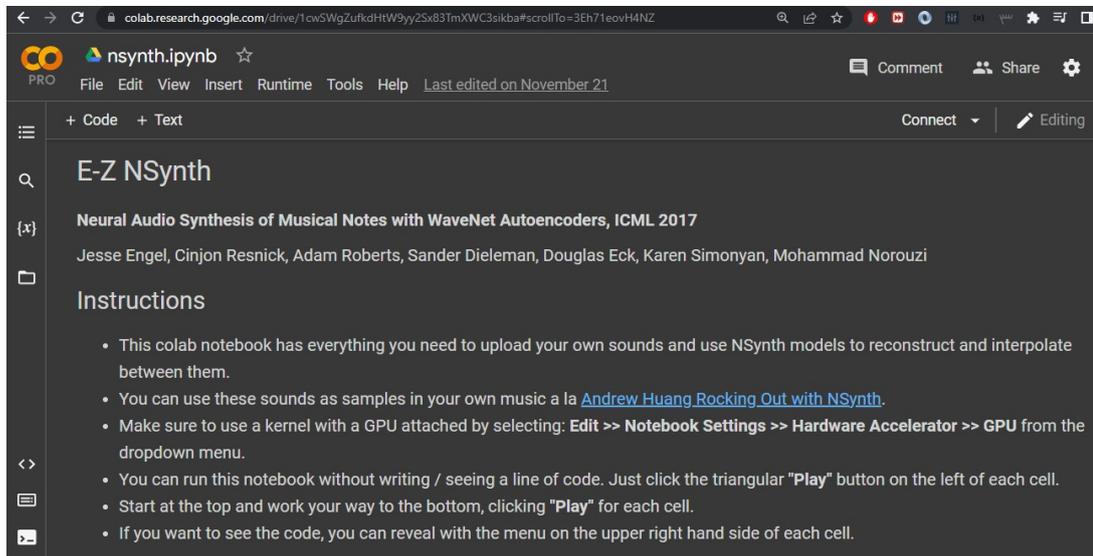
¹ Notebook encontrado em: <<https://colab.research.google.com/notebooks/magenta/nsynth/nsynth.ipynb>>

² Documentação da linguagem Python: <<https://docs.python.org/pt-br/3/tutorial/>>

Fonte: (MAGENTA, 2021)

Na figura 1 temos um exemplo do *Notebook NSynth* (MAGENTA, 2021):

Figura 1 – Cabeçalho do notebook Nsynth



Fonte: O autor(2022)

Após a execução deste algoritmo, temos como entrada um áudio a ser interpolado pela rede neural. Para a entrada, pode-se escolher um ou mais trechos de áudio, porém faixas com duração superior a 10 segundos tendem a gerar um tempo de execução longo devido aos cálculos de multiplicação em pontos flutuantes. Em seguida, é realizado o processo de interpolação entre dois áudios e, posteriormente, a Síntese por Interpolação, que é o algoritmo mais custoso. A seguir, temos um exemplo do código que executa o algoritmo de síntese:

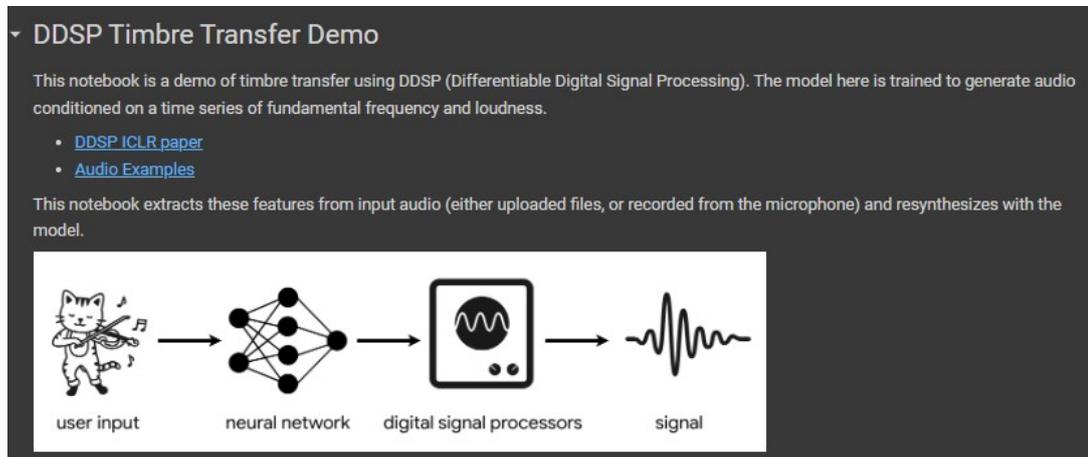
```
#@title Synthesize Interpolations. On the GPU, this should take about 4 minutes
      per 1 second of audio per a batch.
print('Total Iterations to Complete: %d\n' % SAMPLE_LENGTH)
encodings = np.array(z_list)
save_paths = ['/content/' + name + '.wav' for name in name_list]
fastgen.synthesize(encodings,
                   save_paths=save_paths,
                   checkpoint_path=ckpt_path,
                   samples_per_save=int(SAMPLE_LENGTH / 10))
```

Fonte: (MAGENTA, 2021)

resultado final não apresenta um espectro de frequência rico em harmônicos e não é tão surpreendente em comparação com a tecnologia envolvida, mas é audivelmente diferente de outros processos de convolução. Os objetos musicais mais ricos ocorrem quando envolvem mais de uma interpolação de entrada. Utilizei vários trechos da peça do Berio já citada para criar novos objetos musicais que farão parte da peça.

A outra rede neural utilizada é a *DDSP*³ com instrumentos pré-treinados criando parâmetros matemáticos que simulam instrumentos específicos. O algoritmo do *DDSP* não pode ser confundido com a técnica de disparo de *samplers* e sim uma síntese complexa a partir de informações e atributos do timbre modelados matematicamente pela rede neural. Utilizei as mesmas amostras de áudio da peça do Luciano Berio e criei diferentes tipos de objetos musicais com estes trechos. Ao utilizar e executar os códigos do *Notebook DDSP* fornecido pelo site da comunidade Magenta (MAGENTA, 2021) em ambiente do *Google Colab* alguns erros de execução ocorreram. O primeiro erro e o mais crítico foi a instalação de uma biblioteca python que entrava em fluxo interminável e não finalizava sua instalação no ambiente virtual do *Google Colab*. Toda vez que executamos um algoritmo no ambiente do *Google Colab* é disponibilizada uma máquina virtual temporária que apenas possui pacotes padrões de python e linux pré-instalados. Para utilizar outros pacotes é necessário rodar comandos como `"! pip install <xpacote>"` e `"! apt get <xpacote>"` para sua instalação temporária e execução em tempo de processamento do ambiente virtual. Depois de alguns dias de análise, encontrei uma versão antiga de um dos pacotes da versão do *DDSP* e fiz a alteração na primeira célula de código e o *Notebook DDSP* voltou a funcionar. Exemplos da importação de bibliotecas em ambiente virtual python corrigida e a figura 2 do *Notebook DDSP*:

Figura 2 – Cabeçalho do notebook DDSP



Fonte: O autor(2022)

```
#@title #Install and Import
```

```
#@markdown Install ddsp, define some helper functions, and download the model.
```

```
    This transfers a lot of data and _should take a minute or two_.
```

```
%tensorflow_version 2.x
```

```
print('Installing from pip package...')
```

```
!pip install -qU ddsp==0.14.0
```

³ Notebook encontrado em: <<https://github.com/magenta/ddsp/tree/main/ddsp/colab/tutorials>>

```

# Ignore a bunch of deprecation warnings
import warnings
warnings.filterwarnings("ignore")

import copy
import os
import time
import crepe
import ddsp
import ddsp.training
from ddsp.colab import colab_utils
from ddsp.colab.colab_utils import (
    auto_tune, detect_notes, fit_quantile_transform,
    get_tuning_factor, download, play, record,
    specplot, upload, DEFAULT_SAMPLE_RATE)
import gin
from google.colab import files
import librosa
import matplotlib.pyplot as plt
import numpy as np
import pickle
import tensorflow.compat.v2 as tf
import tensorflow_datasets as tfds

# Helper Functions
sample_rate = DEFAULT_SAMPLE_RATE # 16000

print('Done!')
```

Fonte: (MAGENTA, 2021)

Após a execução do pedaço de código acima, é requisitado na célula posterior o tipo de entrada de áudio, *upload* de arquivo de áudio ou gravação por microfone padrão do sistema operacional. Aqui ocorre mais um erro no código e possivelmente uma incompatibilidade de versões, apenas entrada de áudio por microfone funciona. Esses erros são relatados para demonstrar a dificuldade que um usuário ou músico encontraria na busca de novas ferramentas em desenvolvimento e que não têm o caráter de produto finalizado para utilização do músico ou técnico de som. Gravei sons capturados pelo microfone *AKG perception 120* e *Focusrite Scarlett Solo* em tempo do execução do algoritmo e utilizei um sistema com caixas e amplificadores para reproduzir trechos da peça do Berio que foram gravadas pelo algoritmo em Python. O tamanho da janela de tempo pode ser pré-definido pelo usuário. Segue abaixo um exemplo de código para entrada de áudio e um comentário em código sobre a análise espectral executada nesta célula:

```

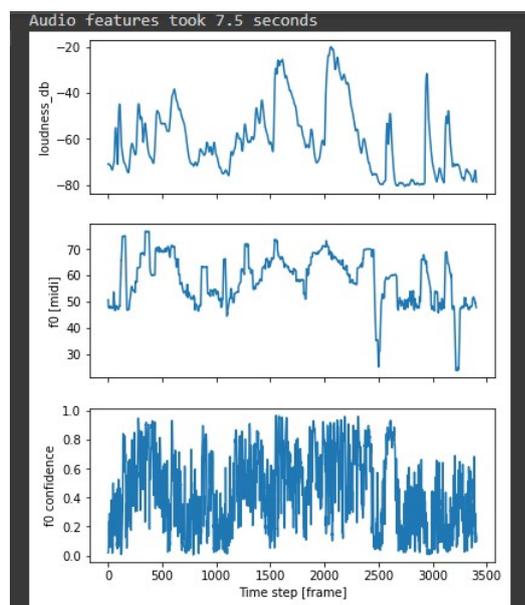
#@title Record or Upload Audio
#@markdown * Either record audio from microphone or upload audio from file
            (.mp3 or .wav)
#@markdown * Audio should be monophonic (single instrument / voice)
#@markdown * Extracts fundmanetal frequency (f0) and loudness features.
record_or_upload = "Record" #@param ["Record", "Upload (.mp3 or .wav)"]
record_seconds = 30#@param {type:"number", min:1, max:10, step:1}
if record_or_upload == "Record":
    audio = record(seconds=record_seconds)
else:
    # Load audio sample here (.mp3 or .wav3 file)
    # Just use the first file.
    filenames, audios = upload()
    audio = audios[0]
audio = audio[np.newaxis, :]
print('\nExtracting audio features...')

```

Fonte: (MAGENTA, 2021)

Na célula de código executada é feita uma análise de reconhecimento de fundamentais extraídas do áudio de entrada inicial. Esse reconhecimento é utilizado para a ressíntese do instrumento, como se o algoritmo tentasse “tocar” a extração de informação do áudio de entrada. Abaixo a figura 3 de um dos trechos captados e amostrados nesta célula de código que foi executada gerando a análise espectral:

Figura 3 – Reconhecimento de fundamentais



Fonte: O autor(2022)

O programador ou músico define qual o modelo pré-treinado por uma rede neural irá utilizar para a ressíntese e nesse notebook existem 5 modelos de instrumentos: violino, flauta 1 e 2, trompete e sax tenor. Em outro projeto há possibilidade de utilizar uma quantidade de samples para gerar um modelo a partir de um algoritmo de treino por redes neurais profundas que gerará um modelo matemático para um novo instrumento. Quanto maior é quantidade de amostras de determinado instrumento, melhor é o resultado do modelo que algoritmo construirá. Utilizei apenas os instrumentos pré-treinados para este trabalho que já está configurado neste *Notebook*. O código de definição do *dataset* utilizado:

```

if model in ('Violin', 'Flute', 'Flute2', 'Trumpet', 'Tenor_Saxophone'):
    # Pretrained models.
    PRETRAINED_DIR = '/content/pretrained'
    # Copy over from gs:// for faster loading.
    !rm -r $PRETRAINED_DIR &> /dev/null
    !mkdir $PRETRAINED_DIR &> /dev/null
    GCS_CHKPT_DIR = 'gs://ddsp/models/timbre_transfer_colab/2020-11-06'
    model_dir = os.path.join(GCS_CHKPT_DIR, 'solo_%s_ckpt' % model.lower())

    !gsutil cp $model_dir/* $PRETRAINED_DIR &> /dev/null
    model_dir = PRETRAINED_DIR
    gin_file = os.path.join(model_dir, 'operative_config-0.gin')

else:
    # User models.
    UPLOAD_DIR = '/content/uploaded'
    !mkdir $UPLOAD_DIR
    uploaded_files = files.upload()

    for fnames in uploaded_files.keys():
        print("Unzipping... {}".format(fnames))
        !unzip -o "/content/$fnames" -d $UPLOAD_DIR &> /dev/null
    model_dir = find_model_dir(UPLOAD_DIR)
    gin_file = os.path.join(model_dir, 'operative_config-0.gin')

```

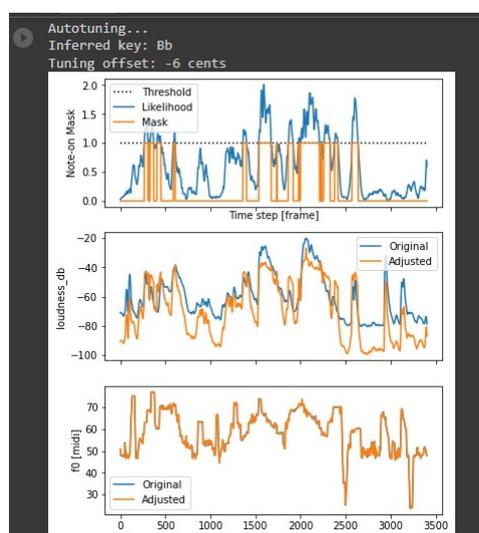
Outros parâmetros que podem ser definidos são: *Note Detection Threshold*, *Note Detection Automatic* com os parâmetros *quiet* e *autotune*, *Manual Pitch Shift* e *Loudness Shift*. Conforme os resultados eram executados, analiso cada faixa de ressíntese criada e ajusto cada parâmetro. Os parâmetros de detecção de nota podem gerar frequências mais assertivas com o áudio de entrada, mas a utilização do *autotune* as vezes gera trinados e mudanças de notas desnecessárias. A variável *Loudness Shift* e *Pitch Shift* alteram respectivamente a sensação de volume e a mudança de oitava para cada faixa de síntese criada. Os resultados gerados criaram interessantes objetos sonoros que serão inseridos para a composição da peça. Algumas faixas de

ressíntese tiveram "*desafinações*" em relação aos trechos vocais de entrada, mas serão tratados ou esteticamente manipulados na peça eletroacústica.

O único fato que não corrobora para ressíntese no *DDSP* é a inexistência de frequências a cima de 8kHz e talvez seja uma limitação de projeto para que não gere falhas no processamento de sinal, devida ao fator múltiplo de 2.5 ao tamanho matrizes de um arquivo de áudio. Há uma pré seleção de 16.000 *samples per second*, porém alteração deste parâmetro não afeta o espectro de frequências altas. 16k é duas vezes 8k, lembrando da regra de Nyquist, a taxa de amostragem deve ter no mínimo duas vezes o valor da frequência, devido a faixa negativa que é gerada nos cálculos transformação de sinal.

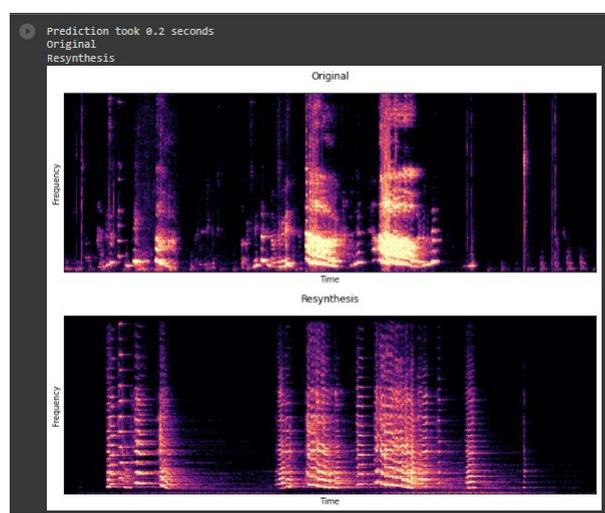
Nas figuras 4 e 5 estão a representação de frequência no tempo ajustadas com os parâmetros e espectrograma da ressíntese:

Figura 4 – Processo de parametrização de fundamentais e ajuste de *Loudness*



Fonte: O autor(2022)

Figura 5 – Espectrograma de Frequência x Tempo da Ressíntese



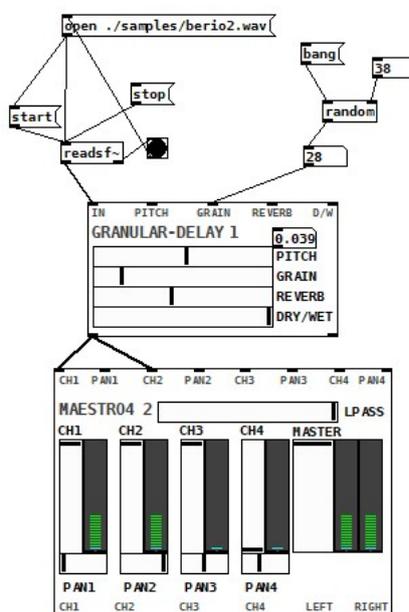
Fonte: O autor(2022)

3.2 TÉCNICAS DIFUNDIDAS DE PROCESSAMENTO DIGITAL NO ÁUDIO

O que difere esta seção da anterior é a tecnologia envolvida na criação de algoritmos e produtos de processamento digital para o áudio. Na seção anterior os algoritmos implementados ainda não são produtos finalizados ou difundidos. Nesta seção é descrito técnicas aplicadas amplamente difundidas, embora muitos produtos e processos tem um real caráter de diferenciação e divergência a produtos comerciais para manipulação de faixas de áudio. A escolha destas ferramentas é motivada pela transformação da fonte sonora inicial deslocando-a do reconhecimento semântico na tentativa de embaralhar e criar variações com objetos sonoros e musicais.

Com recortes da faixa de áudio da peça *Sequenza III* do Luciano Berio, organizo e penso em possíveis transformações e resultados sonoros. Elaboro a execução de algumas técnicas já implementadas e usadas em trabalhos anteriores. Uma dessas técnicas é a granulação de faixas de áudio. Nesta técnica a faixa de áudio é subdividida em pedaços, ou grãos, que utilizam a técnica de repetição deste pedaços de maneira ordenada ou aleatória, gerando um novo resultado sonoro que diverge da fonte sonora inicial (FRITSCH, 2008, pag. 186-192). No ambiente de desenvolvimento do *Pure Data* (PD, 2022) e com *patch Automatonism 3.1*⁴ (ERIKSSON, 2022), há a possibilidade de inserir um módulo de granulação e conectar um áudio de entrada e uma conversão de saída. Exemplo na figura 6 abaixo:

Figura 6 – *Patch Automatonism 3.1* implementando Granulador



Fonte: O autor(2022)

Outro granulador utilizado neste trabalho é o *Emission Control 2* software *opensource* desenvolvido pela supervisão do professor Curtis Roads e dois desenvolvedores, Jack Kilgore

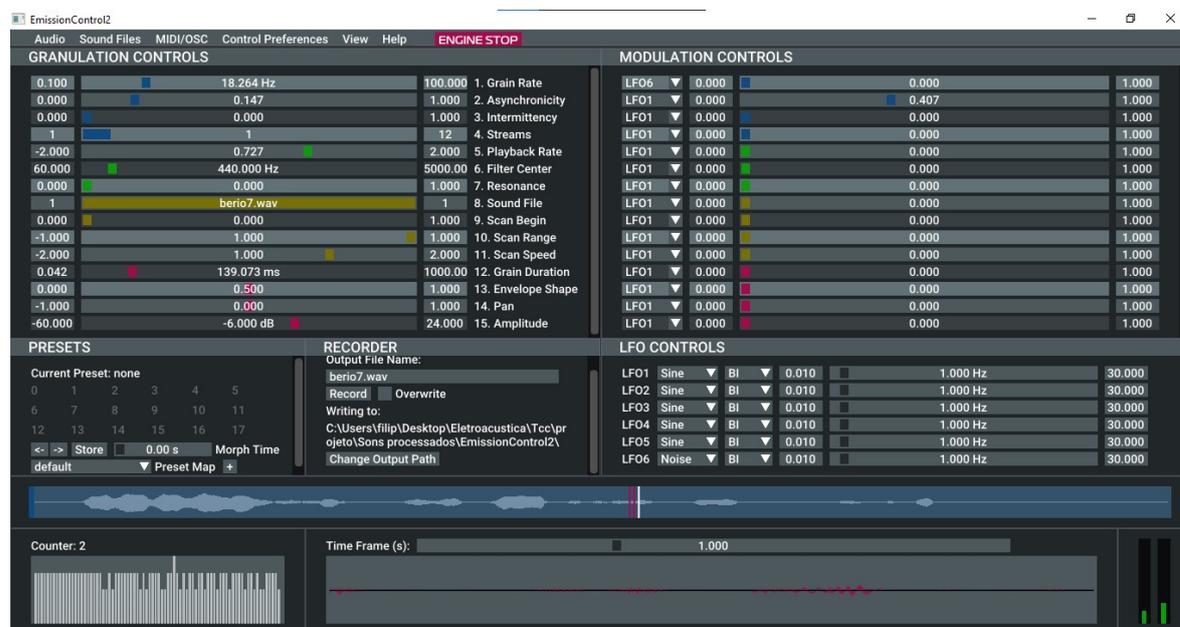
⁴ Patch disponível no site: <<https://www.automatonism.com/>>

e Rodney Duplessi⁵. A característica desse software é criar pequenos trechos de áudio, grãos, e utilizar o conceito de *delay*, repetições espaçadas no tempo com durações manipuláveis, para repetir estes trechos de áudio. Há também neste software vários moduladores de frequências baixas para cada parâmetro de entrada disponíveis em seus processadores.

O Emission Control 2 (EC2) é um aplicativo interativo em tempo real para síntese granular e granulação de arquivos de som. O software enfatiza o controle de grão fino, modulação poderosa e processamento de sinal por grão. O EC2 foi um esforço colaborativo com o Prof. Curtis Roads (Gerente de Projeto, Designer e Mentor) e Rodney DuPlessis (Designer, GUI). Observe que o EC2 é vagamente baseado no Emission Control de David Thall para o Supercollider⁶.

Na foto abaixo temos uma representação da interface ao usuário do software *Emission control 2* com todos os seus aspectos aqui comentados:

Figura 7 – Software *Emission Control 2*



Fonte: O autor(2022)

Utilizo estes granuladores para criar novos objetos sonoros a partir de registros que poderiam ter ou não algum processo de modificação sonora em relação a fonte. Na estrutura de processamento de áudio pode haver mais de um processo em cadeia, os processos de modificação de áudio podem ser acumulados e dispostos de várias formas e com diferentes parâmetros. Aplico processos como alteração de *pitch* para registros de frequências graves e agudas e *stretch* para prolongamento da duração das faixas que cria um efeito de “esticamento” no som, podendo ser combinados filtros de subtração de frequência com o intuito artístico. O que é interessante ressaltar que os trechos extraídos da peça Luciano Berio tem um caráter de extrapolação do canto

⁵ Software disponível no site: <<https://www.curtisroads.net/software>>

⁶ Citação direta e traduzida pelo autor para o resumo disponível no site: <<https://jackilgore.com/pages/software.html>>

e que a performance da peça tem uma certa similaridade aos processos de modificação sonoros utilizados na música eletroacústica.

Neste trabalho operei os *plugins VST GRM Tools Complete II*⁷ construídos pelo *Groupe de recherches musicales (GRM)*, os primeiros processadores foram desenvolvidos em 15 anos de pesquisa para um antigo sistema computacional personalizado pelo próprio grupo chamado de *Syter* (DAMBORG; HARLEY, 2003). Esta versão *Complete II* conta com 19 *plugins* com diferentes tipos de processadores. O *plugin SpaceGrain*, um dos VST que fazem parte deste pacote, é um granulador com a função de espacialização dos grãos. Crio sons graves com efeito de espacialização dos grãos e insiro um processador de reverberação para gerar sensação de “massa” sonora. O granulador espacializador conta com parâmetros de dispersão de grãos e a possibilidade de não repetição dos grãos em valores múltiplos, criando um aspecto não linearidade e arritmia.

O *Evolution*, *plugin VST* do pacote do *GRM Tools*, reinventa o processo de encurtamento e prolongamento da faixa de áudio, processo digital conhecido na área de áudio digital como *stretch*. A maneira como o *Evolution* foi projetado é diferente dos processos comerciais que utilizam desta técnica. Quando softwares como os Digital Audio Workstation (DAW) implementam o *stretch* o requisito da funcionalidade é o uso da FFT⁸ para criar ferramentas de edição e correção das faixas captadas, ajustando gravações e notas musicais no tempo do metrônomo. A proposta do *plugin Evolution* é diferente dos “estritadores” comerciais, cria-se a sensação de prolongamento do som sem necessariamente ter a fidelidade do sinal de áudio no tempo, em outras palavras, a finalidade é criar sons que mantém o harmônicos sem alteração das frequências fundamentais e utilizar a técnica de maneira criativa e artística. Em um dos trechos recortados para processamento da peça do Bérió, que tem bastante energia e a sustentação vocal é feita entorno de uma nota só, processo essa *track* com o *Evolution* e é gerado um efeito de prolongamento do próprio trecho e uma variação daquele momento musical que já tinha sido apresentado na peça de maneira mais próxima da fonte sonora original.

A possibilidade de mudança da nota fundamental de um objeto musical e a tentativa de fazer essa modulação também para os harmônicos que fazem parte da nota musical, este processo é chamado de *pitch shifter*. Sampleadores que importam uma amostra de áudio inicial e fazem uso do processo de *pitch shifter*, proporcionam a funcionalidade de “tocar” notas musicais em diferentes registros de frequências e intensidades sonoras. Com um controlador midi e um sampleador importo um objeto musical, ou uma faixa de áudio, gerada pela rede neural *DDSP* e “toco” esta faixa com diferentes alturas e faço a gravação desta performance no *Digital Audio Workstation (DAW) Reaper*. O sampleador faz parte do pacote disponível no *DAW*. Outros processadores de *pitch shifter* como *PitchAccum* que faz parte do pacote *GRM Tools* e o *++pitchdelay* que é da empresa *SoundHackReaper*⁹ foram empregados para criação de registros

⁷ software encontrado em: <<https://inagr.com/en/store>>

⁸ *Fast Fourier Transform*, técnica matemática que otimiza o tempo computacional da Transformada de Fourier para sinais digitais, amplamente utilizada em algoritmos de áudio.

⁹ software encontrado em: <<https://www.soundhack.com/>>

graves e agudos em objetos sonoros que já tinham sido manipulados e afastados da fonte sonora.

O *SpaceGrain* é um granulador com espacialização de grãos que também faz parte do pacote *GRM Tools Complete II*. Com a função de espacialização os grãos são manipulados em mais de canais para criar a sensação de cada grão em uma localização no espaço. Alguns dos seus parâmetros de modificação incluem dispersão, quantidade de grãos, frente/trás/direita/esquerda, duração dos grãos, densidade e randomização do duração dos grãos. Criei alguns objetos musicais graves que geram uma sensação de movimentação do som grave pela sala.

Figura 8 – Plugins *PitchAccum*, *Evolution*, *++pitchdelay* e *SpaceGrain*



Fonte: O autor(2022)

3.3 EDIÇÃO, MIXAGEM E COMPOSIÇÃO

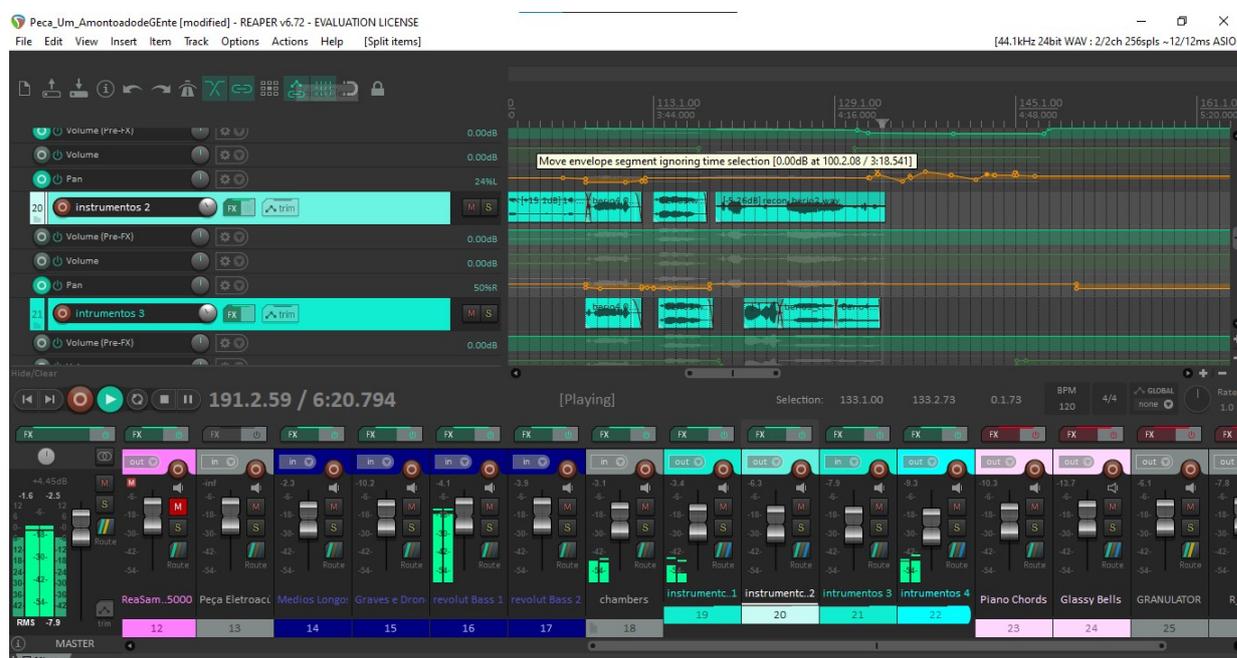
São diversos os softwares de edição para áudio e alguns softwares são mais amplamente usados devido a sua aceitação no âmbito de produção comercial e estúdios de música em geral. Nos dias de hoje há o conceito de software de áudio como estação de trabalho, os chamados de *Digital Audio Workstation (DAW)* que podem ser empregados em todo o processo de produção de áudio. Neste trabalho a DAW operada foi o *Reaper*¹⁰ e a escolha por esse software envolve o tempo de experiência que tenho com ferramenta, horas gastas para aprender as funcionalidades, a compatibilidade com hardwares de menor desempenho e a aceitação de diferentes tipos de *plugins*.

Tenho a prática de importar todas os registros de áudios já processados em faixas organizadas no *Reaper*, para facilitar a visualização e escolha de cada objeto. Quando ainda quero

¹⁰ software encontrado em: <<https://www.reaper.fm/>>

criar mais um processamento com um *plugin*, renderizo a faixa de áudio com o processamento em tempo real para outra faixa criando um novo arquivo de áudio ".wav". A prática de renderização libera processamento de máquina e possibilita a utilização do recurso de máquina apenas para mixagem. Neste momento de edição faço recortes e ajusto a intensidade sonora de inicialização e finalização das faixas chamados de *FadeIn/FadeOut* que proporcionam uma “suavização de volume” nos recortes gerados. Para correções e ajustes, em alguns momentos, recorro ao software *RX* da *Izotop* que reescreve o arquivo de áudio sem a necessidade de renderização, trazendo uma agilidade no processo e facilidade em criar novos objetos sonoros. Até aqui a edição não tem o intuito composicional, pode ser criativa, mas é mais um recurso de ajuste e exploração ou criação de novos objetos musicais.

Figura 9 – O projeto no *Reaper* da peça eletroacústica



Fonte: O autor(2022)

Quando separo estes momentos de edição corretiva e edição composicional, tenho em mente que há um momento que a edição é composicional, porque é organizado e combinado os materiais sonoros criados para formar a minha peça de maneira estrutural ao decorrer do tempo. Os materiais a serem escolhidos pelo compositor para a composição podem ser de um banco de sons já catalogado e organizado com o intuito de se criar uma composição eletroacústica, mas nem todo o bancos tem o dever de ser utilizado na totalidade para construção de peça. Este banco de sons pode contar com materiais “crus” que não foram processados, materiais processados e materiais sintetizados. No momento da edição composicional, eu continuo aplicando outros processos de modificação sonora e espacialização.

Na mixagem e finalização, ajusto os volumes e decido como os sons são apresentados no espaço. Faço o uso do recurso *Pan* que define se o som será apresentado na faixa direita ou esquerda, automatizando o pan com *faders* que variam ao longo do tempo, tendo em mente que

essa primeira versão é em estéreo. Outro recurso é *plugin Space3D* do pacote do *GRM Tools* que implementa funções matemáticas que criam a sensação de movimentação do som no espaço. Se algumas frequências são indesejadas ou se sobrepõe desnecessariamente, com o equalizador as atenuo e com o compressor reduz os picos de transiente. A última tarefa é o processo *Bounce*, que gera uma faixa *wave* em estéreo do projeto, no *Reaper* esse processo é chamado de *Render to File* e prevê a renderização para 8 canais. Depois de todas estas etapas temos a peça finalizada tecnicamente.

4 ANÁLISE MUSICAL DA PEÇA ELETROACÚSTICA *UM AMONTOADO DE GENTE*

Começo com uma frase que cada vez faz mais sentido para mim: “ninguém sai do nada”. Quando penso em composição, logo penso em instituições e trajetórias. Nossas escolhas estão entrelaçadas a nossa trajetória e a música eletroacústica escancara as nossas escolhas no momento da composição. Quando decido por uma material sonoro ao invés de outro, conscientemente e inconscientemente estou fazendo escolhas que são carregadas da construção do meu “Eu”. Pensando em minha trajetória, lembro das técnicas aprendidas e aperfeiçoadas, das estruturas composicionais estudadas e da estética ensinada tolhendo ou pronunciando o que não deve ser feito. As escolhas e regras composicionais fazem parte do jogo social que cada um presenciou. Uma das regras escolhidas neste trabalho é a junção de sons que foram criados separadamente para fazer parte de um todo e acredito que essa escolha tem a ver com que aprendi ao longo do tempo. Um fluxo de trabalho cadenciado e organizado que fui desenvolvendo.

Deixando de lado os motivos do porquê fiz o que fiz, vamos nos dedicar as escolhas e no desenvolvimento que criei. Diferentemente do trabalho de análise musical de peças de outros compositores, a auto-análise composicional pode ser um relato da estrutura musical elaborada antes e no momento do ato composicional. Se encontro fatores que não foram pensados, evidencia-se que tive um certo nível de tentativas e experimentalismo que não foram planejados *a priori* e a auto-análise musical, nesse caso, pode ter o mesmo nível de dificuldade para alguém que está analisando aquela peça pela primeira vez. A singularidade na análise e auto-análise composicional ocorre na constatação de fatos que são percebidos indistintamente à intencionalidade do compositor, como que em uma equação matemática ou em uma prova lógica que se compara dois fatores por sua igualdade ou desigualdade, a exemplo uma repetição exata de um trecho musical já previamente apresentado. Não que não haja abertura para diferentes análises, focos e pontos de vista divergentes, mas não é difícil de evidenciar fatos em uma estrutura musical.

Definindo estes pontos para que tanto seja exposta a minha visão e para a formação do pensamento analítico composicional do próprio autor, parto desses pressupostos para minha análise de forma consciente da tarefa e a racionalização sobre a peça. Nas próximas seções será relatado como a estrutura foi definida por momentos composicionais e quais foram as decisões para gerar o resultado que gerou, demonstrando por meio de imagens e apontamentos gráficos cada detalhe da peça. Descreverei aspectos como timbres, texturas, alturas, ritmos, envelopes de som, melodias, massas, *drones*, ostinatos, variações, cadências, formas, estruturas e elementos musicais que constituem a peça eletroacústica *Um Amontoado de Gente*.

4.1 MICROELEMENTOS, *MEZZOELEMENTOS* E MACROELEMENTOS

Deixando a definição da área química de lado, em outro trabalho defino estes termos do título da secção como:

Microelementos musicais são os fragmentos, que podem ser ordenados ou não, que representam o que há de primário em música, por exemplo: Uma nota no piano, um ritmo de pequena duração e intervalos musicais. Os macroelementos são aqueles que possuem sentido e completude, é o significado de todo numa obra. Existem outros tipos de materiais que podemos chamá-los de *mezzoelementos*, são estruturas que têm uma ideia que não é completa, está inacabada, mas são perceptíveis as suas funções, por exemplo: frases melódicas, ritmos longos e gestos musicais. Para que uma peça tenha sentido e coerência esses elementos precisam ser usados tal qual ao engenheiro que projeta uma casa, tem que ser elaborado para que cumpra a sua função. Formulei alguns conceitos básicos estruturais para deixar claro cada terminologia utilizada neste trabalho (CASTRO, 2016, pag. 5-6).

Quando separo em três subcapítulos a análise da peça, tenho a intenção em demonstrar ou dar foco em estruturas que fazem sentido no momento que ocorre o meu processo composicional. A escolha em separar os tópicos desta forma não é uma convenção, mas apenas uma forma encontrada de relatar de forma textual a análise da obra. Faço, também, a escolha por organizar e estruturar a análise musical com critérios que fazem sentido a minha escuta, sendo esta a própria escuta que organizou o material escolhido.

No artigo *Metodologias de Análise Musical para Música Eletroacústica* é exposto diferentes métodos de análise e teorias que narram a discussão de diferentes teóricos sobre a possibilidade de uma metodologia para análise musical (GUBERNIKOFF, 2007). O que converge neste debate sobre a criação de uma ferramenta de análise de música eletroacústica é a falta ou a necessidade de expansão do léxico em descrever as diferentes formas do som, a percepção pela audição e a criação descritiva desta percepção da obra musical como um todo. A quantidade de possibilidades técnicas não convergem para um único estilo ou uma única técnica composicional, mas uma expansão de novas técnicas composicionais (GUBERNIKOFF, 2007).

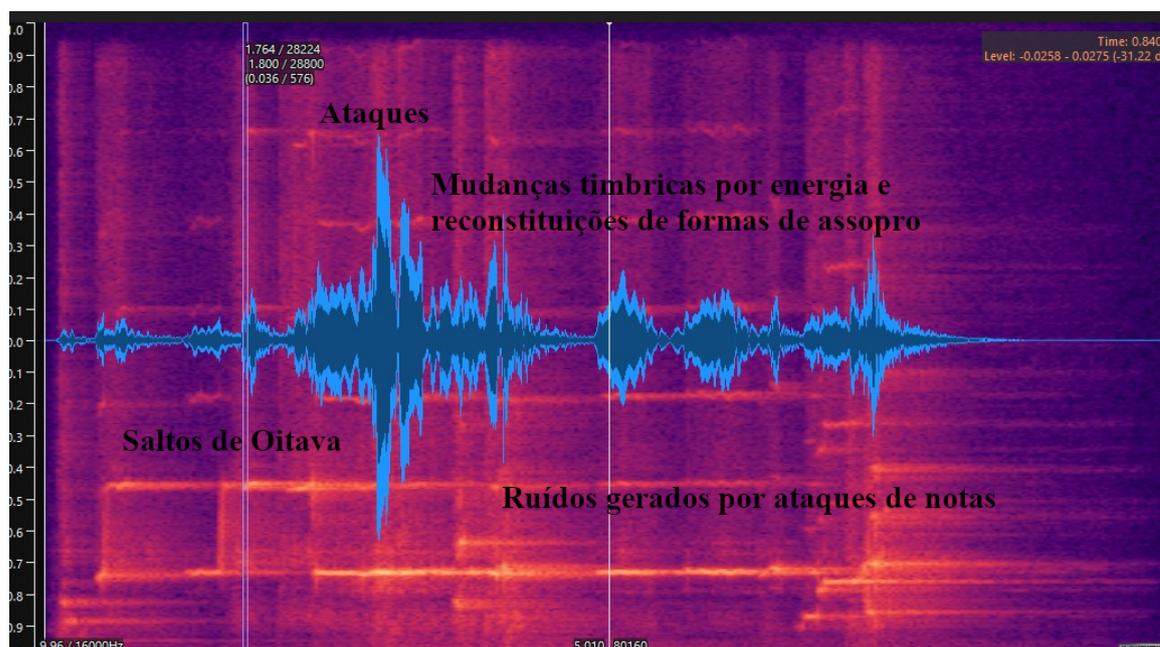
A intenção de distinguir, delimitar e denominar diferentes momentos e características do som é uma tentativa de dar foco a determinados elementos e desenvolver um pensamento analítico próprio que será empregado na análise deste capítulo. A tentativa de descrever os fragmentos sonoros criados que incorporam a construção da peça exemplificam os detalhes e características que só são reconhecidas pela audição, é um detalhamento em palavras e imagens na esperança que a imaginação do leitor reconstrua o som por meio de pistas semânticas fornecidas. A invenção de termos, analogias, derivações e invocações de similaridades fornecem um vocabulário próprio que é criado para gerar recurso para uma análise descritiva da peça. Na área da linguística este fenômeno é chamado de neologismo, ato de criar novas palavras ou mudar o sentido original geralmente por recursos lexicais de justaposição, prefixação, aglutinação, verbalização e sufixação (GUBERNIKOFF, 2007).

4.2 MICROELEMENTOS - FRAGMENTOS

Nesta seção será detalhado os diferentes trechos sonoros e pequenos elementos que fazem parte da peça eletroacústica. Detalhes musicais como ritmos, alturas, timbres, pequenos desenhos melódicos, rítmica, grãos e texturas fazem parte da análise desta seção.

Na criação de novos timbres e frases melódicas geradas por ressíntese pela rede neural *DDSP*, percebe-se a tentativa do algoritmo em recriar a execução de um instrumento musical específico, utilizando técnicas e ornamentos musicais humanas. A figura 10 é uma representação gráfica do espectrograma no tempo e na frequência de uma das melodias fabricadas pelo *DDSP*, o som de uma flauta sintética criada pelo objetivo do algoritmo na transdução do timbre vocal para este instrumento. O áudio de entrada é uma melodia cantada que por meio de reconhecimento de frequências fundamentais e análise da intensidade de energia de cada nota é escrita em um arquivo *midi* que recria a execução vocal. O resultado sonoro é uma melodia de flauta com stacatos, ataques, *frullatos*, mordentes, microtons e alguns gestos e técnicas impossíveis de serem executadas em um instrumento real.

Figura 10 – Trecho da melodia da flauta criada pelo *DDSP*

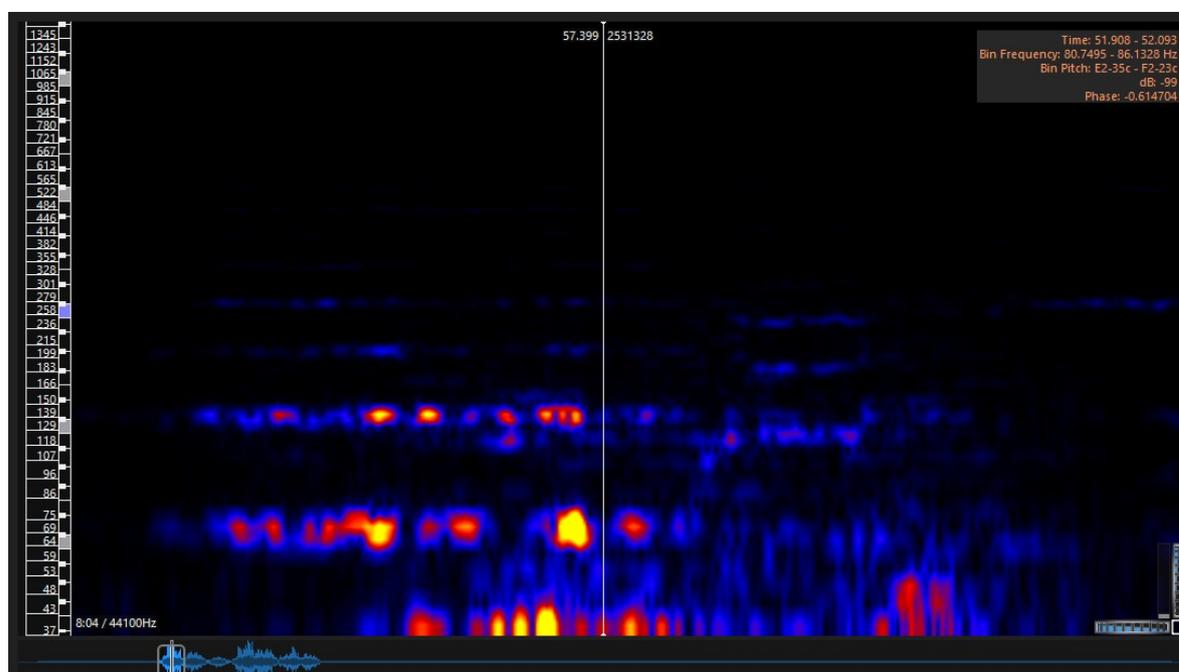


Fonte: O autor(2023)

Nos parâmetros de entrada do algoritmo pode ser feita a escolha de outro modelo matemático de instrumento musical, quando se escolhe o timbre violino, percebe-se que as técnicas aplicadas no instrumento virtual são técnicas do repertório clássico e não foram elaboradas, ainda, com bancos de sons de técnicas estendidas ou técnicas novas para execução instrumental. Já no Instrumento Virtual Flauta a algumas técnicas estendidas já contempladas e categorizadas no *dataset* do *DDSP* (ENGEL; HANTRAKUL; GU, 2020).

Iremos analisar os objetos sonoros criados a partir das técnicas de granulação, alteração do espectro de frequência e convolução, tantos de trechos da voz humana, como de instrumentos musicais. O primeiro objeto musical a ser analisado é um trecho de frequências grave, criado pela manipulação do *DDSP*, *SpaceGrain* e *++pitchdelay*. O objeto grave é gerado a partir de instrumentos melódicos que possuem diferentes características tímbricas quando executados, ou nesse caso, programados e estas características do som se matem ao ocorrer a modulação para a frequência grave, gerando harmônicos e movimentações espaciais. A representação do espectrograma melódico da figura 11 foi feita a partir de um trecho aonde os objetos graves predominam e conseguimos perceber as diferentes frequências e seus harmônicos que compõem esses timbres.

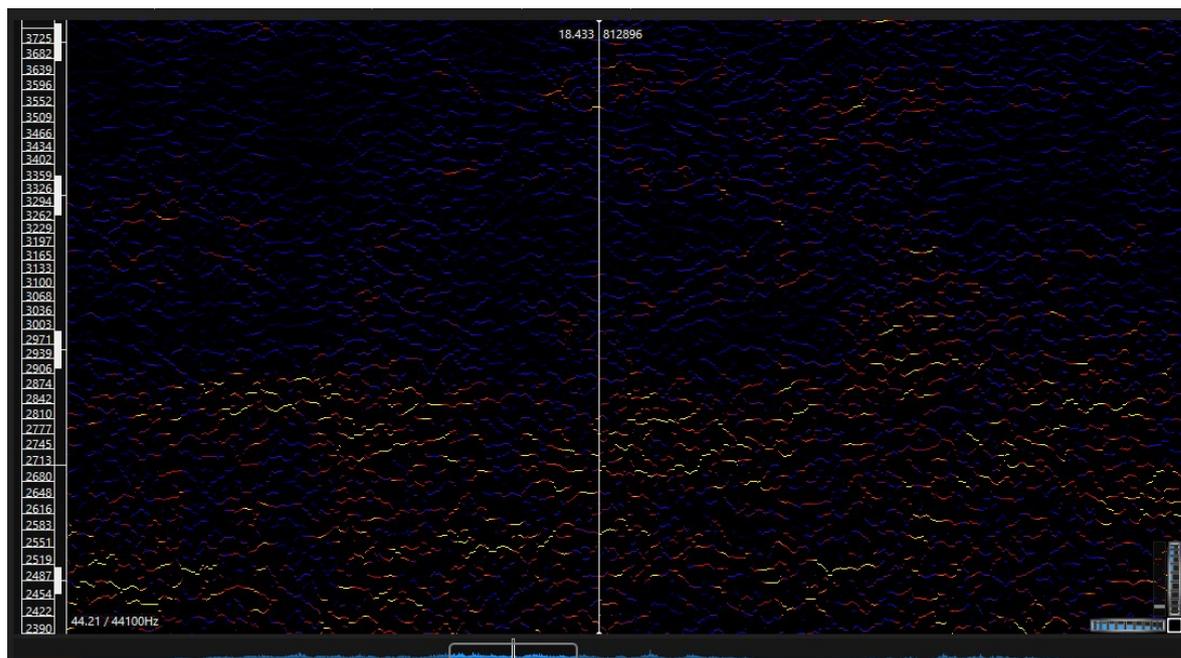
Figura 11 – Espectrograma Melódico de um trecho de objetos graves



Fonte: O autor(2023)

A criação de texturas surge da intenção em criar movimentações espaciais e o aproveitamento de toda a banda de frequência disponível ao som. Como já descrito, os timbres gerados pelas redes neurais não ultrapassavam os 8kHz. Foi adotada a técnica de granulação e esticamento do tempo para a criação destes objetos que abrangiam a maior parte do espectro de frequência. A partir de um trecho de canto que se entona a melodia ao redor de uma nota só, é aplicado os *plugins SpaceGrain* e *Evolutiam* criando som texturais que reforçam certas frequências ao invéns de outras. Na representação gráfica da figura 12, é representado um trecho textural que geram frequências de 30Hz a 1.3kHz, embora o foco da imagem seja por volta de 30Hz a 350Hz para a melhor visualização da movimentação de frequências em *glissando* e quantidade de notas.

Figura 12 – Espectrograma de Picos de frequência de uma faixa textural



Fonte: O autor(2023)

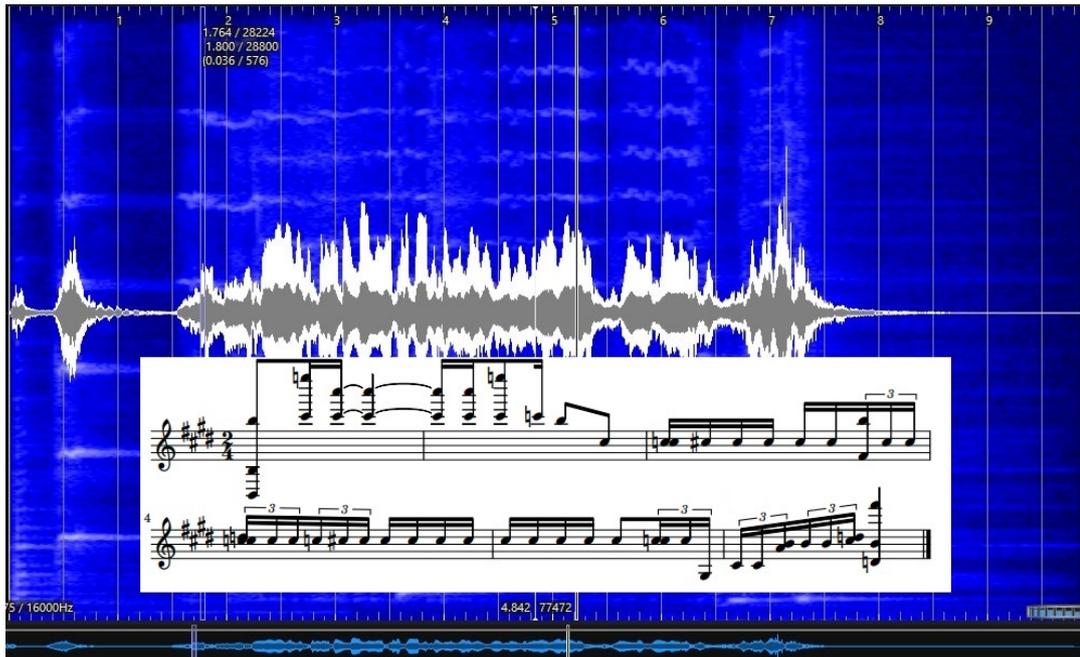
4.3 *MEZZOELEMENTOS* - PARTES DE UM TODO

Nesta seção são tratados assuntos musicais como: melodias, padrões rítmicos, repetições, variações, contraponto, harmonias, drones, camadas de textura, amalgamas graves, seções, espacialidade, granulação e momentos da peça.

O primeiro trecho musical tem a mesma melodia, agora no violino, da primeira análise na seção de *microelementos*, mas dessa vez o intuito é entender a melodia deste trecho que é reutilizada como variação em várias partes da peça. Com um software de transcrição para partitura é feita a notação musical deste trecho melódico. Como no código de reconhecimento de fundamentais, no momento da transcrição o algoritmo *confunde* notas quando 1º harmônico é provido de muita energia, gerando saltos de oitava no momento da transcrição ou notas oitavadas. Na figura 13 é apresentada a transcrição deste trecho da melodia no violino e o espectrograma e a forma da onda pelo tempo.

Estas melodias são trespostas para o grave, são sobrepostas e *tocadas* por disparadores de *samplers* em diferentes modulações de frequência e são apresentadas da forma inicial, a execução vocal do trecho da peça do Berio, e granuladas e variadas de diversas formas. Criando a forma musical de temas e variações de um jeito não tão explícito. Em relação a escolha das melodias, levando em consideração que elas já tinham sido criadas pelo Berio, são pela escuta e entendimento da estrutura da peça.

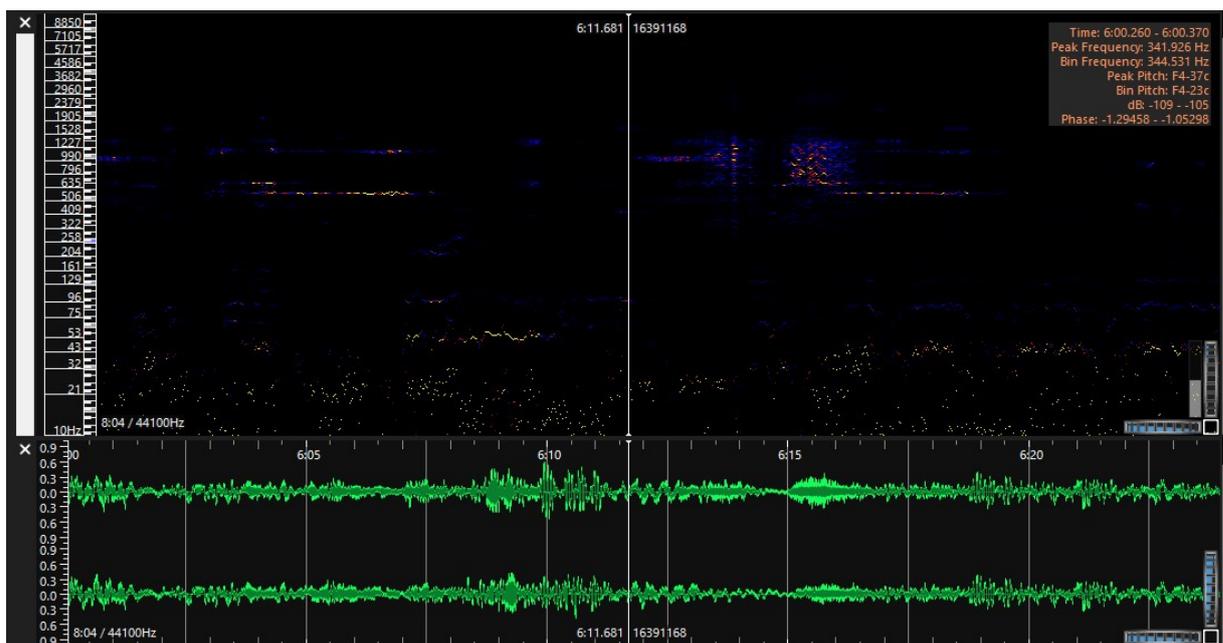
Figura 13 – Espectrograma de picos de frequência de uma faixa textural



Fonte: O autor(2023)

Em alguns momentos trechos graves são sobreposto por trechos médio-agudo e geram a sensação de um baixo que se sobrepõem a trechos melódicos e cantados. A figura 14 exemplifica este trecho. No momento da mixagem eu não tenho como referência os sons graves, o que no resultado geral pode ocasionar mais ainda a sensação de sobreposição dependendo dos sistemas de auto-falantes.

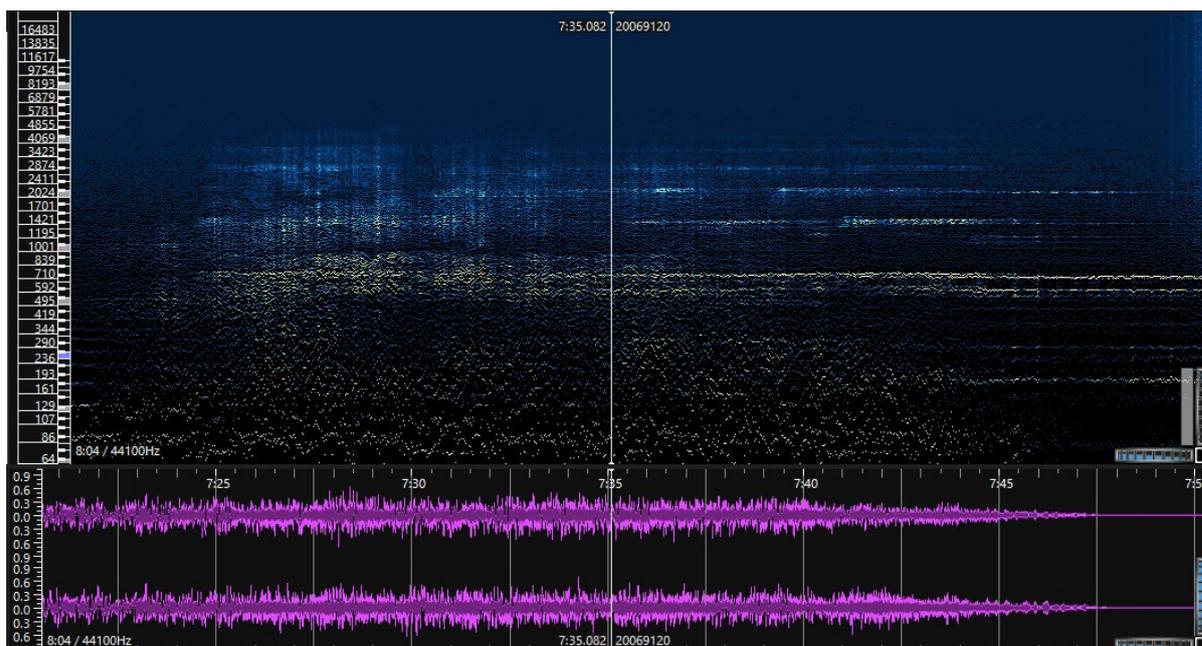
Figura 14 – Espectrograma de picos de frequência e a forma da onda no tempo de trechos graves e médios



Fonte: O autor(2023)

Na seção anterior menciono sobre a criação de texturas e sons agudos para gerar movimentação no espaço e preenchimento do espectro de frequência. Algumas tramas, texturas, grãos e *drones* foram criados com este intuito de gerar novos materiais que pudessem criar a percepção de movimento no espaço. Analisaremos um trecho onde este efeito de movimentação e texturas ocorrem na peça.

Figura 15 – Espectrograma de picos de frequência e a forma da onda domínio do tempo de trechos de texturas e movimentações



Fonte: O autor(2023)

A figura 15 é uma representação gráfica de texturas que são criadas no registro agudo e médio e são manipuladas para criar movimentos elípticos ao redor do ouvinte. Com o *plugin Space3D* do pacote do *GRM Tools* consegue-se provocar a sensação de movimentação por conceitos psicofísicos e matemáticos, que para esta mixagem está em dois canais. A sensação gerada é de que a fonte sonora se movimenta no espaço de forma determinística, devida as funções matemáticas de movimentação implementadas pelo *plugin Space3D*. Tais trechos criados com o intuito de engendrar movimentação no espaço são tentativas, ora alcançadas ou não, para aproveitar e fazer uso de técnicas e tecnologias de aproveitamento do espaço físico.

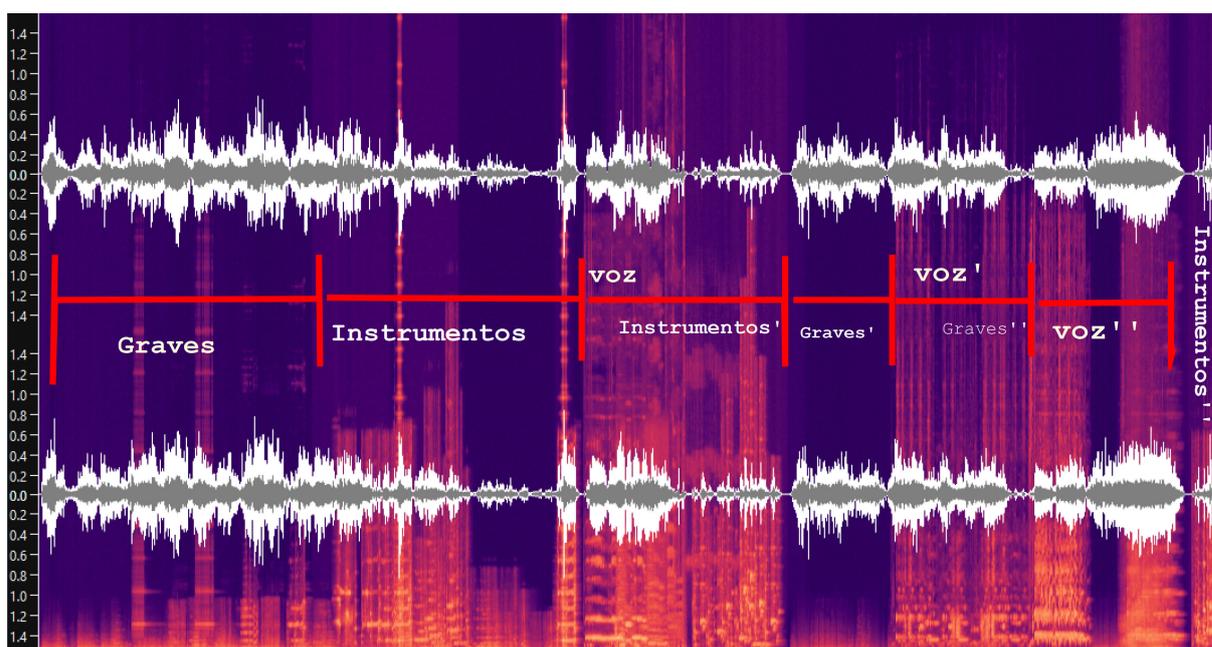
Na mesma ideia de criar elementos que fazem um complemento ou produzem materiais para um acompanhamento ou apresentações no espectro de frequência e no espaço, foram gerados objetos musicais que complementariam a seção do meio da peça. Esta seção é muito diferente dos trechos iniciais, embora os elementos primários utilizados para variação sejam os mesmos. Devido ao interesse em promover o resultado sonoro do algoritmo *NSynth* e inseri-lo na peça, percebi a necessidade de estabelecer uma coesão entre elementos texturais e o trecho com pouca riqueza de frequências agudas e assim o fiz.

4.4 MACROELEMENTOS - ESTRUTURA TOTAL

A estrutura geral ou total da peça é baseada em variações melódicas e tímbricas de faixas de áudio que foram “extraídas” ou “recortadas” da peça *Sequenza III* de Luciano Berio por meio do ferramental de áudio digital, sendo que a fonte geradora de todos os sons são trechos vocais dessa peça, embora exista um acorde, a a granulação deste acorde, feito por síntese FM e um instrumento virtual sampleado de um piano *Steinway*. A característica do material gera poucos objetos sonoros na região aguda e esse fato é evidenciado na peça inteira. Pode ser feita uma análise da estrutura da peça como um todo e, também, de suas subestruturas particulares. Outro tipo de análise possível seriam relações sobre a estrutura da peça original do Berio e a correlação com a minha peça. Devido alguns recortes estarem organizados em uma mesma cronologia na peça, seria possível criar pontos de relação e semelhança.

Neste momento da análise será priorizada a forma geral da obra, delimitando a suas partes e as nomeando. Pode-se dizer que há 3 seções de graves que se apresentam na peça, 3 seções de momentos onde foco principal são execução de instrumentos virtuais e 3 seções vocais que estão em primeiro plano. Todas as seções são variações ou temas novos que serão variados em algum momento. Este efeito é criado devido ao extenso trabalho de manipulação digital da fonte original, a voz. Lembrando que a peça *Sequenza III* de Luciano Berio já é uma tentativa de se manipular a voz de todas as formas possíveis, até posso dizer que a estética remete a manipulação de processos eletrônicos na música eletroacústica. A manipulação é tão grande na minha obra que há trechos que distinguem-se totalmente de outros. Representação analítica geral da peça na figura 16.

Figura 16 – Análise Geral da Peça

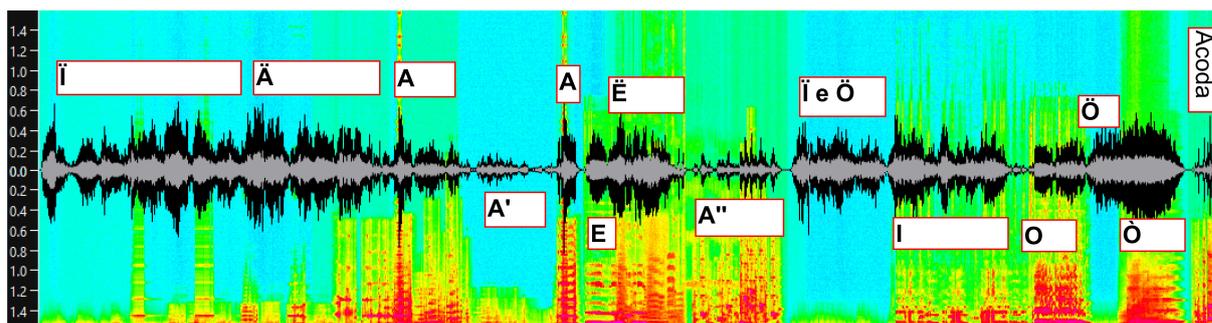


Fonte: O autor(2023)

Nas suas subestruturas pode ser feita uma análise da variação de figuras ou objetos musicais que apresentam-se e reapresentam-se de diversas formas. Na primeira seção intitulada como *Graves* todos os materiais melódicos são apresentados e ao mesmo tempo *escondidos*. Apresentados porque os aspectos físicos, ou morfológicos, estão lá e *escondidos* ou *embaçados*, porque não fica clara a percepção humana, redução fenomenológica, quais notas são ou quais figuras rítmicas estão presentes. Neste trecho podemos chamar uma de suas figuras de \ddot{A} , figura logo é reapresentada com intensidade no trecho posterior, chamemos está de A . O evento da não percepção melódica se dá por fatores psicoacústicos em relação a notas graves e é apresentado nos momentos *Graves*. Essa mesma técnica de ofuscamento é utilizada em variações por meio de granuladores e esticamento de objetos musicais no tempo, em toda a peça esses elementos de super-variação, é utilizado símbolo trema para exemplificar estes trechos.

O primeiro elemento ou objeto musical apresentado como tema é a flauta da figura 10 que é variada de diversas formas durante a peça, é chamado trecho A , o mesmo trecho citado no parágrafo acima. A segunda figura que denomino de E , é um trecho vocal que faz um intervalo ascendente de sétima maior e depois um intervalo descendente décima primeira. O objeto E é variado poucas vezes. O que podemos chamar de objeto I não é uma figura melódica, mas sim variações de elementos que remetem a fonte sonora voz já variada na peça do Berio, reapresentada de outra forma e com um grave que seria a própria variação \ddot{I} , elemento este apresentado no início da peça. Outro momento é O , \ddot{O} e o \ddot{O} , tendo o \ddot{O} apresentado antes do I . O trecho O e \ddot{O} é o momento de maior intensidade da peça e formada por um trecho de um poema cantado ao final da peça do Berio. A representação gráfica da \ddot{O} na figura 15 do subcapítulo anterior.

Figura 17 – Análise de Temas e Variações



Fonte: O autor(2023)

Na figura 17 esta representada a análise dos temas e variações mencionados no parágrafo anterior. O *Acoda* é o trecho conclusivo que tem um gestual de bastante expressão e gerado a partir do som da flauta.

5 CONCLUSÃO

A peça eletroacústica *Um Amontoado de Gente* foi finalizada seguindo a proposta de pesquisas relacionadas a inteligência artificial e novos métodos ou ferramentas para a composição de música eletroacústica e a composição musical incluindo estas novas técnicas. O trabalho contribui demonstrando novas técnicas de manipulação de áudio, tendo em vista o uso prático. É discutido no trabalho temáticas que envolvem o ato composicional da música eletroacústica e a apresentação de uma análise musical de técnica própria. A peça eletroacústica deste trabalho foi finalizada para ser reproduzida em três formatos de sistemas de falantes distintos: stereo, em 6 canais de áudio e para 8 canais de áudio e está disponível na nota de rodapé ¹.

¹ A peça finalizada em estéreo, 6 canais e em 8 canais nos formatos de arquivos de áudio *wave* e *mp3* e as algumas amostras e exemplos de arquivos de áudio manipulados que integram a peça eletroacústica estão neste link: <https://drive.google.com/drive/folders/1AgUt1Hh0Y3zdXJoeqKsZhyc2Zu6sE9v?usp=share_link>

REFERÊNCIAS

- CASTRO, F. B. de. **Memorial de Composição: Peça Eletroacústica Prataria Nº 3**. 127 f. Monografia (Graduação) — Departamento de Música, Universidade Federal do Paraná, Paraná, 2016.
- COLAB, G. Ambiente computacional compartilhado para desenvolvimento de algoritmos em python. 2022. Disponível em: <<https://colab.research.google.com/>>.
- DAMBORG, D.; HARLEY, J. Grm tools. **Computer Music Journal**, The MIT Press, v. 27, n. 4, p. 103–108, 2003.
- ENGEL, J.; HANTRAKUL, L.; GU, C. Ddsp: Differentiable digital signal processing. In: **International Conference on Learning Representations**. [s.n.], 2020. Disponível em: <<https://openreview.net/forum?id=B1x1ma4tDr>>.
- ERIKSSON, J. Site oficial do software automatonism 3.1. 2022. Disponível em: <<https://www.automatonism.com/>>.
- FRITSCH, E. F. **Música Eletrônica Uma Introdução Ilustrada, 1ª Edição**. [S.l.]: EDITORA DA UFRGS, 2008. ISBN 9788570259998.
- GUBERNIKOFF, C. Metodologias de análise musical para música eletroacústica. **Revista eletrônica de musicologia**, v. 11, p. 17, 2007.
- HAYKIN, S. **Redes Neurais: Princípios e prática, Segunda Edição**. [S.l.]: BOOKMAN COMPANHIA ED, 2001. ISBN 9788573077186.
- KOVACS, L. K. **Redes neurais artificiais, fundamentos e aplicações, Segunda Edição**. [S.l.]: Livrarias da física, 2006. ISBN 8588325144.
- LASKE, O. Composition theory in koenig’s project one and project two. **Computer Music Journal**, 1981.
- MAGENTA, C. Documentação dos algoritmos da comunidade magenta de aprendizado de máquina. 2021. Disponível em: <<https://magenta.tensorflow.org/>>.
- MANTARAS, R. L. de; ARCOS, J. L. Ai and music: From composition to expressive performance. **AI Magazine Volume 23 Number 3**, 2002. Disponível em: <<https://ojs.aaai.org//index.php/aimagazine/article/view/1656>>.
- OPPENHEIM, A.; SCHAFER, R. **Discrete-Time Signal Processing, Second Edition**. [S.l.]: Prentice-Hall, Inc., 1999. ISBN 0137549202.
- PD, C. of. Documentação do software *Pure Data*. 2022. Disponível em: <<https://puredata.info/docs>>.
- RISSET, J.-C. Musique, calcul secret ? **Critique Edition 395 p.414-429**, 1977.
- ROEDERER, J. G. **Introdução à física e psicofísica da música, First Edition**. [S.l.]: Editora da Universidade de São Paulo, Edusp, 2002. ISBN 8531404576.
- SCHAEFFER, P. . **O Tratado de Los Objetos Musicales, Second Edition**. [S.l.]: Alianza Editorial, 2003. ISBN 8120685402.