



UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE  
CENTRO DE TECNOLOGIA  
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

Levy Gabriel da Silva Galvão

**Sistema embarcado portátil para aquisição de áudio  
ultrassônico em roedores**

Natal - RN  
Julho de 2022



UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE  
CENTRO DE TECNOLOGIA  
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

## **Sistema embarcado portátil para aquisição de áudio ultrassônico em roedores**

Trabalho de Conclusão de Curso submetido ao Departamento de Engenharia Elétrica do Centro de Tecnologia da Universidade Federal do Rio Grande do Norte, como requisito final à obtenção do título de Engenheiro Eletricista.

Orientador: Prof. Dr. Antônio Wallace Antunes Soares

Natal - RN  
Julho de 2022

À materialidade da nossa existência e à  
mundanidade dos nossos interesses.

## Resumo

As espécies animais utilizam vocalizações para comunicar estados emocionais importantes, seja na relação materna ou interações sociais entre grupos. Os ratos são um exemplo de animais que incorporaram a vocalização em sua vida social e, além disso, também são capazes de se comunicar na faixa ultrassônica por meio de vocalizações ultrassônicas (USV). Estudos utilizam o áudio das vocalizações para compreender o comportamento desses animais e, no caso do áudio de USVs, a aquisição deve ser feita com um sistema de aquisição com faixa de frequência elevada. Os sistemas utilizados para voz humana ou música são inadequados, pois operam unicamente na faixa audível devido os requisitos da própria aplicação, portanto sistemas especializados devem ser utilizados. Apesar disso as soluções atualmente disponíveis para aquisição de áudio ultrassônico em experimentos comportamentais são grandes e caras, necessitando de uma configuração experimental fixa e investimento elevado. Por isso, este trabalho se propõe a desenvolver um sistema de aquisição de áudio ultrassônico compacto, portátil e economicamente viável. Será utilizado um microfone micro-eletro-mecânico (MEMS) ultrassônico de pequenas dimensões para aquisição de áudio e um cartão microSD para armazenamento. O interfaceamento entre esses periféricos será feito pelo microcontrolador ESP32 que possui diversos periféricos e funcionalidades *Wi-Fi* e *Bluetooth* integrados. O sistema será desenvolvido em uma placa de circuito impressa (PCB) com dimensões na faixa de  $2 \times 2\text{cm}^2$  para ser embarcado individualmente em cada rato.

**Palavras-chaves:** áudio. ultrassom. rato. portabilidade. aquisição.

## Abstract

Animal species use vocalizations to communicate important emotional states, whether in maternal relationships or social interactions between groups. Rats are an example of animals that have incorporated vocalization into their social life and, in addition, are also able to communicate in the ultrasonic range through ultrasonic vocalizations (USV). Studies use the audio of vocalizations to understand the behavior of these animals and, in the case of audio from USVs, the acquisition must be done with an acquisition system with a high frequency range. The systems used for human voice or music are inadequate, as they operate only in the audible range due to the requirements of the application itself, therefore specialized systems must be used. Despite this, the solutions currently available for acquiring ultrasonic audio in behavioral experiments are large and expensive, requiring a fixed experimental setup and high investment. Therefore, this work proposes to develop a compact, portable and economically viable ultrasonic audio acquisition system. A small ultrasonic micro-electro-mechanical microphone (MEMS) will be used for audio acquisition and a microSD card for storage. The interfacing between these peripherals will be done by the ESP32 microcontroller that has several peripherals and integrated *Wi-Fi* and *Bluetooth* functionalities. The system will be developed on a printed circuit board (PCB) with dimensions in the range of  $2 \times 2\text{cm}^2$  to be embedded individually to each rat.

**Keywords:** audio. ultrasound. rat. portability. acquisition.

# **Lista de Figuras**

1	Encontro agressivo de ratos machos. (a) Boxe. (b) Luta livre. (c) Postura agressiva e submissa. (d) Agachado-submisso e noivo-agressivo. (e) Submissor-ereto e ameaçadora de lado. Fonte: Sales [2012] . . . . .	11
2	Sistemas de aquisição amplamente utilizados, com o microfone de captura e gravação (a) Ultramic 250K (130mm de comprimento e 20mm de diâmetro) Dodotronic [2016] e (b) o microfone condensador Avisoft-Bioacoustics CM16/CMPA (36 × 36 × 60mm <sup>3</sup> ) AvisoftBioacustics [2020]. . . . .	14
3	Sistemas alternativos de aquisição de áudio ambiental: (a) SOLO (65 × 56 × 12mm <sup>3</sup> , para o Raspberry Pi A+) Whytock and Christie [2017], (b) Aurita (350 × 230 × 86mm <sup>3</sup> , com base na caixa) Beason et al. [2019], (c) BatPi (85.6 × 56.5 × 17mm <sup>3</sup> , para o Raspberry Pi 3 B+) Aachen [2017] e (d) AudioMoth (58 × 48 × 17mm <sup>3</sup> ) Hill et al. [2019]. . . . .	16
4	Vista em corte da estrutura básica de um microfone MEMS. Adaptado de: CUI [2022] . . . . .	17
5	Vista do topo de um microfone MEMS analógico e digital sem o encapsulamento. Adaptado de: Lewis [2013] . . . . .	17
6	Resposta em frequência do microfone Knowles SPH0641LU4H-1 dividida na faixa de 100Hz-10kHz (esquerda) e na faixa de 10kHz-80kHz (direita). Fonte: Knowles [2015] . . . . .	18
7	Faixa dinâmica do microfone Knowles SPH0641LU4H-1 baseado na SNR e AOP. Inspirado em Infineon [2018]. . . . .	19
8	Funcionamento da interface I <sup>2</sup> S por meio do (a) diagrama de blocos e (b) diagrama temporal. Adaptado de NXP [2022]. . . . .	22
9	Pinagem de um cartão micro SD. . . . .	24
10	Arquitetura da cadeia de aquisição de áudio para microfone analógico e digital. . . . .	28
11	Visão do sistema de aquisição de áudio ao nível de microcontrolador e periféricos. . . . .	28
12	Noção de ocupação dos componentes na PCB no projeto final, indicado a face que possui a interface com o usuário (superior externa) e a face de contato com o animal (inferior externa). . . . .	30
13	PCB do protótipo 0 (14.59 × 9.8mm <sup>2</sup> ) com (a) desenho 2D da PCB no software EasyEDA e (b) versão física comparada com uma moeda de 20mm de diâmetro. . . . .	31

14	PCB do protótipo 1 ( $33.59 \times 57.28mm^2$ ) com (a) desenho 2D da PCB no <i>software</i> EasyEDA e (b) versão física comparada com uma moeda de 20mm de diâmetro. . . . .	32
15	PCB do protótipo 2 ( $44.26 \times 28.7mm^2$ ) com (a) desenho 2D da PCB no <i>software</i> EasyEDA e (b) versão física comparada com uma moeda de 20mm de diâmetro, mostrando as visões do topo (cima) e do fundo (baixo). . . . .	33
16	PCB do protótipo 2 ( $44.26 \times 28.7mm^2$ ) com os componentes soldados com a visão (a) inferior externa (área de contato com o animal) e visão (b) superior externa (interface com o usuário). . . . .	34
17	Fluxograma de atividades do <i>firmware</i> . . . . .	35
18	Espectro em magnitude de um sinal PDM codificando uma senoide de $20kHz$ a uma taxa de amostragem de $20.48MHz$ utilizando uma janela retangular. . . . .	38
19	Diagrama de blocos do sistema de filtragem do áudio PDM com um filtro CIC decimador como elemento de filtragem principal e um filtro FIR para compensar a resposta em frequência do CIC. Inspirado em Lyons [2020]. . . . .	38
20	Magnitude da resposta em frequência do filtro CIC para frequência de amostragem de entrada de $5MHz$ e OSR 16, variando a quantidade de estágios (N) para (esquerda) $M = 1$ e (direita) $M = 2$ . Inspirado em Lyons [2020]. . . . .	39
21	Magnitude da resposta em frequência (esquerda) do filtro em cascata (CIC+FIR) e filtro FIR compensador (frequência de corte de $90kHz$ de janela Chebyshev com atenuação de $300dB$ para o lobos secundários) para frequência de amostragem de $5MHz/OSR$ (OSR=16), variando a ordem do filtro para parâmetros fixos do filtro CIC ( $N = 2$ , $R = 16$ , $M = 1$ ). Inspirado em Lyons [2020]. . . . .	40
22	Típico programa de detecção e tratamento de USVs. (superior) espectrograma. (centro) entropia normalizada. (inferior) amplitude no tempo filtrada passa-faixa. Fonte: Bessa et al. [2019]. . . . .	43
23	Espectrogramas calculados com base no mesmo trecho de áudio ultrassônico com três USVs e re-quantizado para toda a faixa abaixo de 16 bits. . . . .	44
24	Espectrograma das múltiplas senoides com frequência variando em forma de degrau para (a) o sinal de teste e (b) o sinal adquirido no sistema de aquisição. . . . .	45
25	Magnitude do espectro de Fourier para o sinal adquirido ao longo da cadeia de processamento, i.e. (a) sinal PDM adquirido, (b) sinal PCM após a filtragem do CIC e (c) sinal PCM final após compensação do FIR. . . . .	47
26	Espectrograma do sinal de áudio pulsado de $40kHz$ obtido do sensor de proximidade avaliado na (a) saída do filtro CIC e na (b) saída do filtro FIR compensador. . . . .	47

27	Esquemático dos circuitos do protótipo 0 . . . . .	56
28	Esquemático dos circuitos do protótipo 1 . . . . .	57
29	Esquemático dos circuitos do protótipo 2 . . . . .	58

## **Lista de Tabelas**

1	Parâmetros importantes do microfone Knowles SPH0641LU4H-1. Outros parâmetros a se destacar são as dimensões de $3.5 \times 2.65 \times 1.1 \text{ mm}^3$ , diretividade omnidirecional, alimentação de $1.5V - 3.6V$ e corrente máxima de $1mA$ . [Knowles, 2015]	18
2	Faixa dinâmica ou SNR suportada para diversos valores de bits de resolução da codificação digital .	20
3	Configuração dos pinos do cartão microSD para operar em modo SPI. [Davis, 2015]	24
4	Configuração dos pinos do cartão microSD para operar em modo SD proprietário. [Davis, 2015]	24
5	Tecnologias utilizadas no desenvolvimento do projeto. Inspirado em Sarkar [2018].	27
6	Quantitativo dos componentes do sistema e balanço de carga mecânica e elétrica.	29

## Listas de abreviaturas

<b>USV</b>	Vocalização ultrassônica
<b>PCB</b>	Placa de circuito impressa
<b>SPL</b>	Nível de pressão sonora
<b>USB</b>	Barramento serial universal
<b>GPS</b>	Sistema de posicionamento global
<b>SMS</b>	Serviço de mensagem curta
<b>LED</b>	Diodo emissor de luz
<b>SD</b>	<i>Secure digital</i>
<b>MEMS</b>	Micro-eletro-mecânico
<b>ASIC</b>	Circuito integrado de aplicação específico
<b>ADC</b>	Conversor analógico-digital
<b>PDM</b>	Modulação por densidade de pulso
<b>PCM</b>	Modulação por codificação de pulso
<b>I<sup>2</sup>S</b>	<i>Inter-IC sound</i>
<b>THD</b>	Distorção harmônica total
<b>SNR</b>	Relação sinal-ruído
<b>AOP</b>	Ponto de sobrecarga acústica
<b>IoT</b>	Internet das coisas
<b>ULP</b>	Ultra-baixa potência
<b>eMMC</b>	Cartão multimídia embarcado
<b>SDIO</b>	<i>Secure digital input output</i>
<b>SPI</b>	Interface serial periférica
<b>PWM</b>	Modulação por largura de pulso
<b>RGB</b>	Vermelho, verde, azul
<b>DAC</b>	Conversor digital-analógico
<b>DMA</b>	Acesso direto à memória
<b>MSB</b>	Bit mais significativo
<b>LSB</b>	Bit menos significativo
<b>CPU</b>	Unidade central de processamento
<b>RAM</b>	Memória de acesso aleatório
<b>LEDC</b>	Controle LED
<b>GPIO</b>	Pinos de propósito geral
<b>PDM2PCM</b>	Conversão PDM para PCM
<b>CIC</b>	<i>Cascaded integrator-comb</i>
<b>FIR</b>	Finite impulse response
<b>OSR</b>	Taxa de sobreamostragem

# Sumário

<b>1</b>	<b>Introdução</b>	<b>10</b>
1.1	Motivação	10
1.2	Objetivos	10
1.2.1	Objetivos gerais	10
1.2.2	Objetivos específicos	11
<b>2</b>	<b>Aporte teórico</b>	<b>12</b>
2.1	Características físicas das USVs de roedores	12
2.2	Soluções correlatas	13
2.2.1	Sistemas amplamente utilizados	13
2.2.2	Sistemas alternativos	14
2.3	Microfone	15
2.4	Microcontrolador ESP32	20
2.5	Interface I <sup>2</sup> S/PDM com o microfone	21
2.6	Interface SPI com o cartão micro SD	23
<b>3</b>	<b>Desenvolvimento</b>	<b>26</b>
3.1	Requisitos da aplicação	26
3.2	<i>Hardware</i>	26
3.2.1	Arquitetura do sistema de aquisição	26
3.2.2	Protótipos	31
3.3	<i>Firmware</i>	33
3.3.1	Fluxograma de atividades	33
3.3.2	Adaptações na interface I <sup>2</sup> S para aquisição do fluxo de bits PDM	36
3.3.3	Sistema de filtragem digital	37
3.3.4	Prevenção de perda de dados	41
<b>4</b>	<b>Resultados</b>	<b>43</b>
4.1	Variação dos bits de resolução do áudio de USVs	43
4.2	Validação na faixa sonora	45
4.3	Validação na faixa ultrassônica	46
<b>5</b>	<b>Conclusões</b>	<b>48</b>
<b>A</b>	<b>Esquemáticos dos circuitos</b>	<b>56</b>

# 1 Introdução

## 1.1 Motivação

A vocalização das espécies animais constitui-se como um comportamento único, que do ponto de vista evolucionário, contribui para o desenvolvimento e sobrevivência do indivíduo. Esta permite comunicar estados emocionais importantes durante a relação materna e que, durante a vida adulta, permitem a interação social entre grupos ([Brudzynski \[2018\]](#)).

Os ratos são uma espécie extremamente comunicativa e incorporam a vocalização em sua vida social ([Brudzynski \[2013\]](#)). Além das vocalizações na faixa do som, os ratos e outros roedores desenvolveram a comunicação na faixa ultrassônica (20kHz - 100kHz) ([Boulanger-Bertolus et al. \[2017\]](#)) por motivos defensivos ([Brudzynski \[2018\]](#)). Estas são chamadas de vocalizações ultrassônicas, do inglês *ultrasonic vocalizations* (USVs).

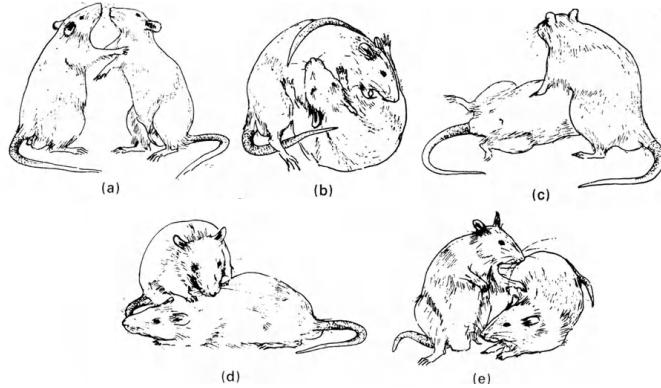
As características das vocalizações ultrassônicas são diversas, pois dependem da idade, espécie dos roedores ([Sales \[2012\]](#)), contexto social de emissão e podem variar desde 2kHz até 125kHz ([Brudzynski et al. \[1999\]](#)). Essas vocalizações podem ser analisadas com base na duração e na faixa de frequência da emissão. Por exemplo, ao considerar os ratos, USVs de longa duração (500-3000ms) na faixa de 22kHz (20-30kHz) podem ser utilizadas para comunicar estados emocionais negativos, como ansiedade, medo, situações de agressão entre machos, estado de derrota social, etc. USVs de curta duração (10-150ms) e emitidas na faixa de 50kHz (35-80kHz) podem indicar estados emocionais positivos, durante brincadeiras, uso de drogas ou medicamentos, interações sexuais, etc ([Boulanger-Bertolus et al. \[2017\]](#), [Wöhr and Schwarting \[2013\]](#), [Brudzynski \[2013\]](#)). Na Figura 1 estão representadas algumas situações de interação social agressiva entre roedores machos que podem eliciar USVs.

## 1.2 Objetivos

### 1.2.1 Objetivos gerais

Diante da importância da emissão das USVs, surge a necessidade da aquisição desse tipo de áudio para estudos aprofundados no comportamento dos animais. Os aparelhos expe-

Figura 1: Encontro agressivo de ratos machos. (a) Boxe. (b) Luta livre. (c) Postura agressiva e submissa. (d) Agachado-submisso e noivo-agressivo. (e) Submisso-ereto e ameaçadora de lado. Fonte: [Sales \[2012\]](#)



rimentais utilizados nestas situações comumente possuem uma robustez de um equipamento grande e estático ([Hofer et al., 1998, Roy et al., 2012, Zaytseva et al., 2019, Heckman et al., 2017, Saito et al., 2019, Wright et al., 2010, Brudzynski, 2009, Isobe and Kawaguchi, 2019]) e que necessita de alto investimento financeiro.

Nesse contexto, este trabalho tem como objetivo geral desenvolver uma solução de sistema de aquisição de áudio ultrassônico compacto, portátil e economicamente viável, que permite o registro local de USVs em roedores.

### 1.2.2 Objetivos específicos

Os objetivos específicos definem que o sistema desenvolvido deve ser capaz de:

- Escolha de um microfone adequado para aquisição de áudio na faixa audível e ultrassônica;
- Estabelecer uma interface de comunicação adequada entre microfone e controlador;
- Armazenamento local dos dados adquiridos em tempo real em memória do tipo cartão micro SD;
- Apresentar dimensões adequadas para ser embarcado em ratos filhotes, inicialmente com dimensões na ordem de  $2 \times 2\text{cm}^2$  para a placa de circuito impressa (PCB) projetada;
- Realizar testes e validação do protótipo.

## 2 Aporte teórico

### 2.1 Características físicas das USVs de roedores

As duas características físicas principais das vocalizações ultrassônicas que influenciam no projeto do sistema de aquisição de áudio são a faixa de frequência do sinal e o nível de pressão sonora. A faixa de frequência permite especificar a frequência de amostragem que influencia na velocidade de processamento e gravação dos dados, parâmetros de filtros e espaço para armazenamento. O nível de pressão sonora permite especificar a faixa dinâmica de operação do microfone e sistema de codificação digital.

As principais faixas de frequências discutidas na literatura para ratos e camundongos compreende frequências ao redor de 22kHz (20-30kHz) e ao redor de 50kHz (35-80kHz) ([Boulanger-Bertolus et al. \[2017\]](#), [Wöhr and Schwarting \[2013\]](#), [Brudzynski \[2013\]](#)). Apesar disso ainda existem emissões fora dessas faixas, como [Brudzynski et al. \[1999\]](#) apresenta em seus resultados, o menor pico de frequência foi obtido na faixa audível em 1.92kHz (para filhotes de 10 dias de idades) e o maior pico em 124.48kHz (para um filhote de 17 dias de idade). Em seus estudos [Sales \[2012\]](#) tabeliza as faixas de frequência de vocalizações emitidas por diversas espécies de roedores, o que permite observar que roedores neonatos emitem vocalizações em uma faixa ao redor de frequências menores (de 6kHz até 125kHz, levando em conta todas as espécies indicadas) e filhotes mais velhos possuem uma faixa centrada em frequências maiores (de 5kHz até 148kHz). Porém atender as faixas de frequência principais já é o suficiente para a maioria dos experimentos.

Em relação ao nível de pressão sonora da vocalização, esta também possui forte dependência da idade, espécie do roedor e tipo do experimento, mas possui aumento inversamente proporcional ao avançar da idade [Sales \[2012\]](#). Em seus estudos [Okon \[1970a,b\]](#) estabelece valores entre 80-90 dB SPL, [Bell et al. \[1972\]](#) indica entre 90-100 dB SPL, [Portfors \[2007\]](#) entre 65-85 dB SPL e [Branchi et al. \[2001\]](#) entre 60-100 dB SPL, sendo esta última a faixa mais abrangente entre todos os trabalhos avaliados.

## 2.2 Soluções correlatas

Os sistemas de aquisição de vocalizações ultrassônicas aqui avaliados que podem ser divididos em dois grupos: o primeiro é composto de equipamentos amplamente comercializados e validados por meio de diversos trabalhos acadêmicos; enquanto que o outro é composto por sistemas de caráter alternativo, que são mais acessíveis financeiramente e implementam soluções portáteis.

A compreensão das características desses sistemas é útil para conhecer boas práticas de projeto específicos da aplicação e entender o motivo que esses equipamentos são inviáveis para atender os objetivos propostos anteriormente.

### 2.2.1 Sistemas amplamente utilizados

Ao avaliar artigos com diferentes naturezas de experimentos notam-se alguns fabricantes de equipamentos e configurações de gravação mais recorrentes.

Todos os sistemas que especificam a resolução do áudio adquirido sempre utilizam 16 bits para as amostras. A taxa de amostragem costuma variar, mas quando é informada, varia entre 200kHz ([Wright et al. \[2010\]](#)), 250kHz ([Barker et al. \[2014\]](#), [Heckman et al. \[2017\]](#), [Saito et al. \[2019\]](#), [Roy et al. \[2012\]](#)) e, em casos especiais, até 384kHz ([Zaytseva et al. \[2019\]](#)).

É comum encontrar o uso de equipamentos produzidos pela Avisoft Bioacustics, incluindo microfones condensadores polarizados da família CM16/CMPA ([Avisoft \[2020\]](#)) de alta fidelidade de áudio ultrassom com resposta definida entre 2-200kHz, porém com grandes dimensões; e sistemas de gravação e *playback* como os equipamentos da família UltraSoundGate ([Bioacoustics \[2022\]](#)) que possuem várias configurações de taxa de amostragem e versões da conexão USB. [[Barker et al., 2014](#), [Heckman et al., 2017](#), [Roy et al., 2012](#), [Saito et al., 2019](#)]

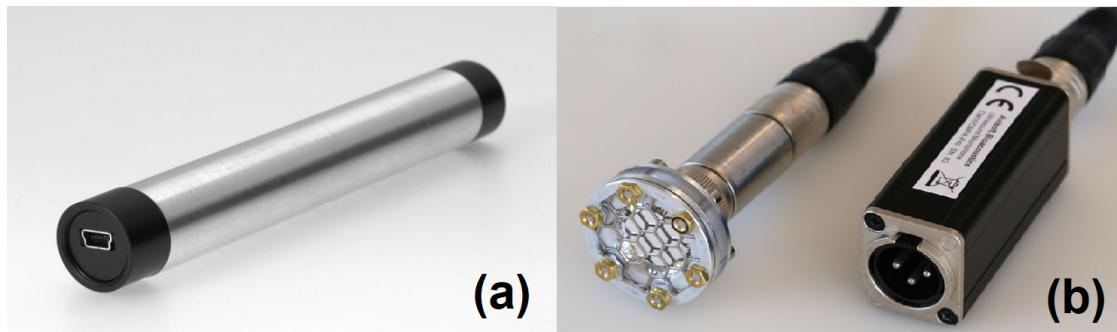
Também são encontrados dispositivos detectores ultrassônicos ou microfones da UltraSound Advice, gravadores da Pettersson com microfone embutido, microfones electret da Knowles e microfones Dodotronic Ultramic com sistema de gravação e conexão USB construídos no microfone. [[Brudzynski et al., 1999](#), [Hofer et al., 1998](#), [Zaytseva et al., 2019](#),

Wright et al., 2010, Bessa and Romcy-Pereira, 2019, Bessa et al., 2019]

Os microfones utilizados nesses experimentos possuem dimensões fora do escopo desse projeto e são fixados em posições estratégicas no topo da arena em que os roedores situam e direcionados os animais, mas sem interferir suas ações.

As dimensões físicas e alto preço desses sistemas inviabilizam seu uso no presente projeto. Por exemplo, o microfone Ultramic 250K ([Dodotronic \[2016\]](#)) da Dodotronics com capacidade de captura e gravação custa cerca de U\$240 (dólares americanos) e microfones mais robustos como o Avisoft-Bioacoustics CM16/CMPA ([AvisoftBioacustics \[2020\]](#)) que custa cerca de U\$2400, que além disso ainda necessitam uma interface de aquisição do áudio capturado. É possível observar na Figura 2 as fotos de alguns dos sistemas comentados.

Figura 2: Sistemas de aquisição amplamente utilizados, com o microfone de captura e gravação (a) Ultramic 250K (130mm de comprimento e 20mm de diâmetro) [Dodotronic \[2016\]](#) e (b) o microfone condensador Avisoft-Bioacoustics CM16/CMPA (36 × 36 × 60mm<sup>3</sup>) [AvisoftBioacustics \[2020\]](#).



### 2.2.2 Sistemas alternativos

A comunidade científica gera esforços no desenvolvimento de sistemas alternativos que usam microcontroladores ou computadores de placa única como o *Raspberry Pi* em sistemas de monitoração ambiental. [[Beason et al., 2019](#)]

O *Raspberry Pi Bat Project* ([Aachen \[2017\]](#)) é utiliza o *Raspberry Pi* em conjunto com o microfone Dodotronic Ultramic 250k e aceita uma taxa de amostragem de até 250kHz, além de possuir módulos adicionais com função GPS, SMS, *powerbank*, etc. O dispositivo Solo ([Whytock and Christie \[2017\]](#)) também utiliza o *Raspberry Pi* em um encapsulamento melhor

protegido e se vale do uso de uma placa de áudio de alta fidelidade que suporta microfones externos até uma taxa de 192kHz. O Solo não utiliza um microfone em específico, mas recomenda aqueles da série Knowles FG. Outra solução é o dispositivo AURITA ([Beason et al. \[2019\]](#)) que utiliza uma mesclagem de outros dispositivos como o Solo, microfones electret para a faixa audível e um sistema comercial para detectar e gravar emissões ultrassônicas.

Entre esses dispositivos, o mais compacto é o AudioMoth ([Hill et al. \[2019\]](#)) de tamanho  $58 \times 48 \times 15\text{mm}^3$  e desenvolvido em uma PCB customizada com conexão USB para configuração da rotina de gravação, LEDs, conexão micro SD, microfone microeletromecânico (MEMS) embutido e um microcontrolador de 32 bits baseado em arquitetura ARM Cortex-M4 48 MHz.

Apesar da proposta desses sistemas serem mais acessíveis, eles ainda possuem grandes dimensões que inviabilizam o uso no presente projeto. Além disso o *Raspberry Pi* é um computador de placa única e que possui propósitos gerais e para um projeto com poucas funcionalidades pode ser considerado exagerado, mas em projetos que realizam monitoração ambiental a robustez pode ser justificada. O AudioMoth é o único que utiliza um microcontrolador, apesar disso a complexidade da PCB é maior e com mais componentes passivos, o que dificulta a prototipagem e testes.

Em relação ao preço, as soluções que usam o Raspberry Pi possuem uma faixa variada, pois dependem da placa processadora e de acessórios externos. Mas uma noção do preço do sistema pode ser observado ao considerar o preço de uma Raspberry Pi 3 ou 4 na faixa de U\$30 (dólares americanos). O preço do AudioMoth varia conforme a versão e a faixa de preço inicia em U\$99. [[RaspberryPi, 2022](#), [OpenAcousticDevices, 2022](#)]

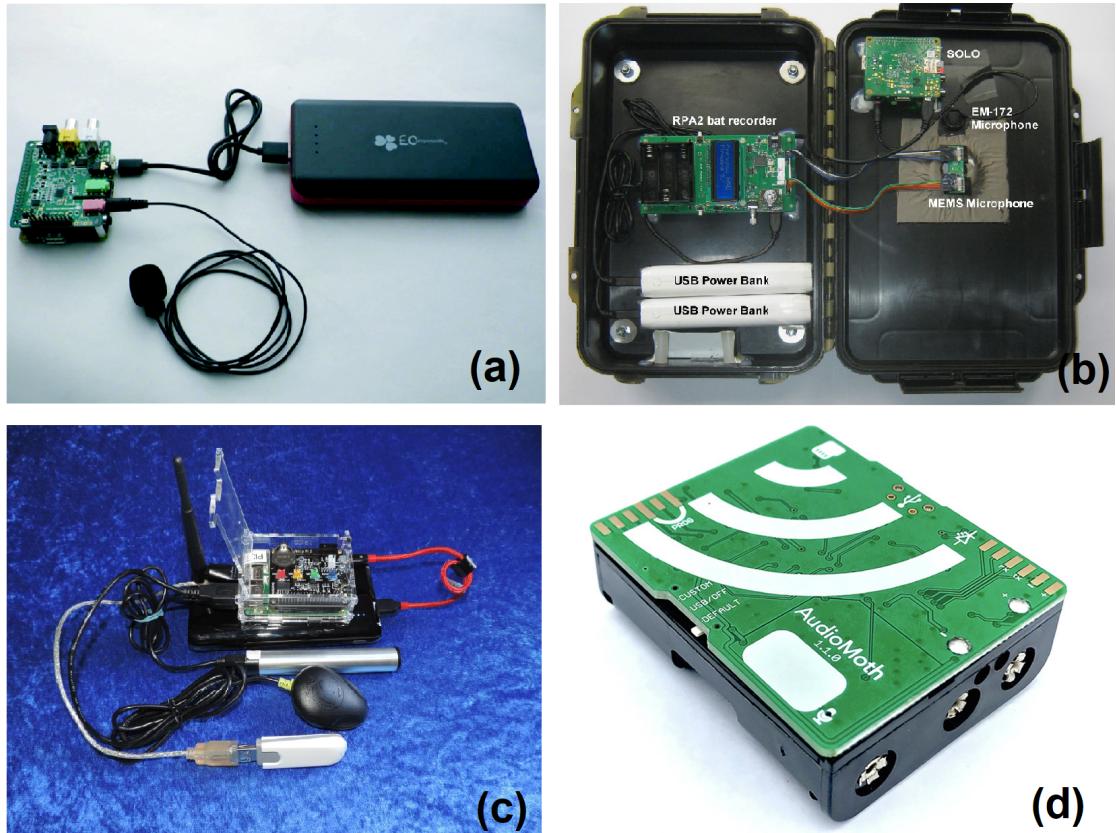
Na Figura 3 estão representadas fotos de cada um dos sistemas alternativos descritos acima.

## 2.3 Microfone

Os microfones MEMS são ideais para aplicações que exigem uma área de ocupação mínima na PCB. Eles são utilizados em *smartphones*, *tablets*, aplicações automotivas, industriais e médicas, mas não limitados a elas. [[Shah et al., 2019](#)]

É possível compreender a partir da Figura 4 a estrutura básica de um microfone

Figura 3: Sistemas alternativos de aquisição de áudio ambiental: (a) SOLO ( $65 \times 56 \times 12\text{mm}^3$ , para o Raspberry Pi A+) [Whytock and Christie \[2017\]](#), (b) Aurita ( $350 \times 230 \times 86\text{mm}^3$ , com base na caixa) [Beason et al. \[2019\]](#), (c) BatPi ( $85.6 \times 56.5 \times 17\text{mm}^3$ , para o Raspberry Pi 3 B+) [Aachen \[2017\]](#) e (d) AudioMoth ( $58 \times 48 \times 17\text{mm}^3$ ) [Hill et al. \[2019\]](#).

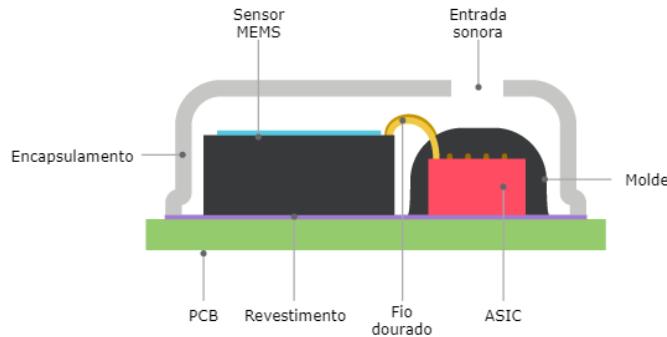


MEMS. Dentro de seu encapsulamento há dois *chips*, um *chip* sensor responsável pela captura das ondas sonoras e um circuito integrado de aplicação específica (ASIC, do inglês *application-specific integrated circuit*) para processamento do sinal sonoro e interfaceamento.

O sensor é um capacitor de placa com uma membrana e placas rígidas traseiras perfuradas. As ondas sonoras provocam deflexão na membrana, resultando em uma mudança na capacidade e, consequentemente, uma mudança na tensão que é monitorada pelo ASIC. [\[Loibl et al., 2016\]](#)

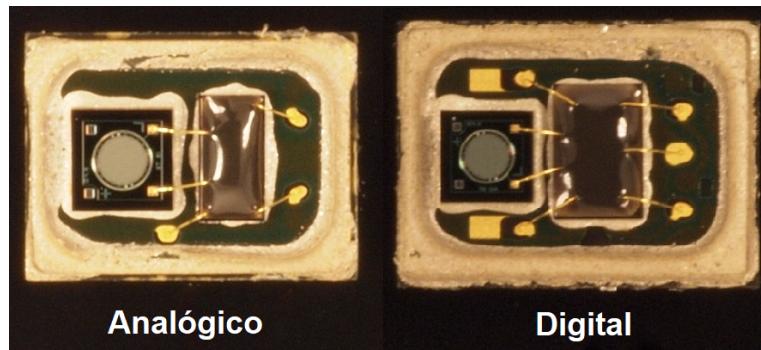
Esses microfones possuem versões com o sinal de saída analógico ou digital. Fisicamente se diferenciam pela pegada do componente eletrônico. Além do contatos de alimentação (VDD e GND), o microfone analógico possui apenas um contato de saída, enquanto que o microfone digital possui contatos adicionais de *clock* e de seleção do canal. Do ponto de

Figura 4: Vista em corte da estrutura básica de um microfone MEMS. Adaptado de: [CUI \[2022\]](#)



vista dos componentes internos apenas o ASIC possui diferenças [[Loibl et al., 2016](#)]. Na Figura 5 estão representadas as fotos de um microfone MEMS analógico e digital sem o encapsulamento, permitindo ter uma noção dos contatos externos e das conexões entre o sensor (esquerda) e ASIC (direita).

Figura 5: Vista do topo de um microfone MEMS analógico e digital sem o encapsulamento. Adaptado de: [Lewis \[2013\]](#)

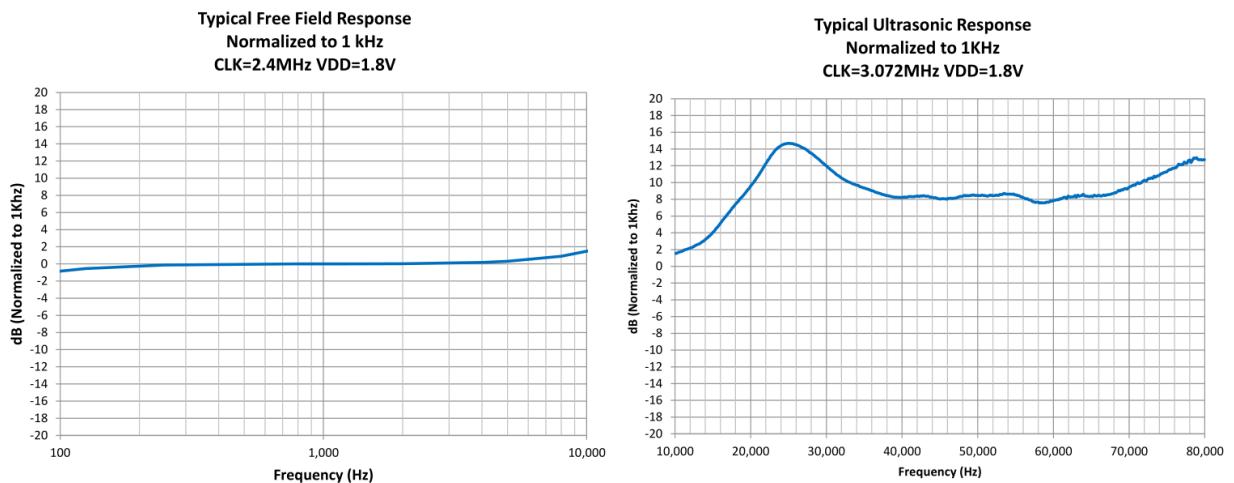


Em um microfone MEMS analógico o *chip* ASIC alimenta o sensor e amplifica o sinal sonoro e o distribui na saída. Adicionalmente o microfone digital possui uma etapa de conversão analógico-digital com um ADC do tipo Sigma-Delta que gera um sinal de saída PDM. Entre os microfones digitais, existem dois sub-tipos, o microfone digital com saída PDM e o com saída PCM. Aquele com saída PDM implementa até o estágio de conversão Sigma-Delta, enquanto que o microfone com saída PCM possui um estágio de filtragem e dizimação do sinal PDM para PCM e externaliza o sinal por uma interface I<sup>2</sup>S. [[Lewis, 2013](#), [Bruhns et al., 2012](#)]

Para este projeto a escolha de um microfone digital facilita a cadeia de processamento do sinal, pois dispensa estágios adicionais de amplificação e filtragem analógicas e a conversão analógico-digital. Apesar disso a resposta em frequência do microfone e a sensibilidade foram os parâmetros mais importantes para a escolha. Após a avaliação dos microfones MEMS disponíveis comercialmente, chegou-se a escolha do microfone SPH0641LU4H-1 da fabricante Knowles ([[Knowles, 2015](#)]).

É possível observar na Figura 6 a resposta em frequência do microfone Knowles SPH0641LU4H-1 oferecida no *datasheet*. Entre os microfones pesquisados, este possui a melhor resposta na faixa ultrassônica definida até 80kHz, com uma porção aproximadamente *flat* entre 32-70kHz, o que torna aceitável para trabalhar com USVs nas faixas principais de 22kHz e 50kHz como indicado em seções anteriores.

Figura 6: Resposta em frequência do microfone Knowles SPH0641LU4H-1 dividida na faixa de 100Hz-10kHz (esquerda) e na faixa de 10kHz-80kHz (direita). Fonte: [Knowles \[2015\]](#)



É possível observar na Tabela 1 algumas características eletromecânicas do microfone. Vale destacar que o parâmetro da faixa dinâmica e a quantidade de bits para codificação na tabela não é oferecido no *datasheet*, mas foi determinado como descrito a seguir.

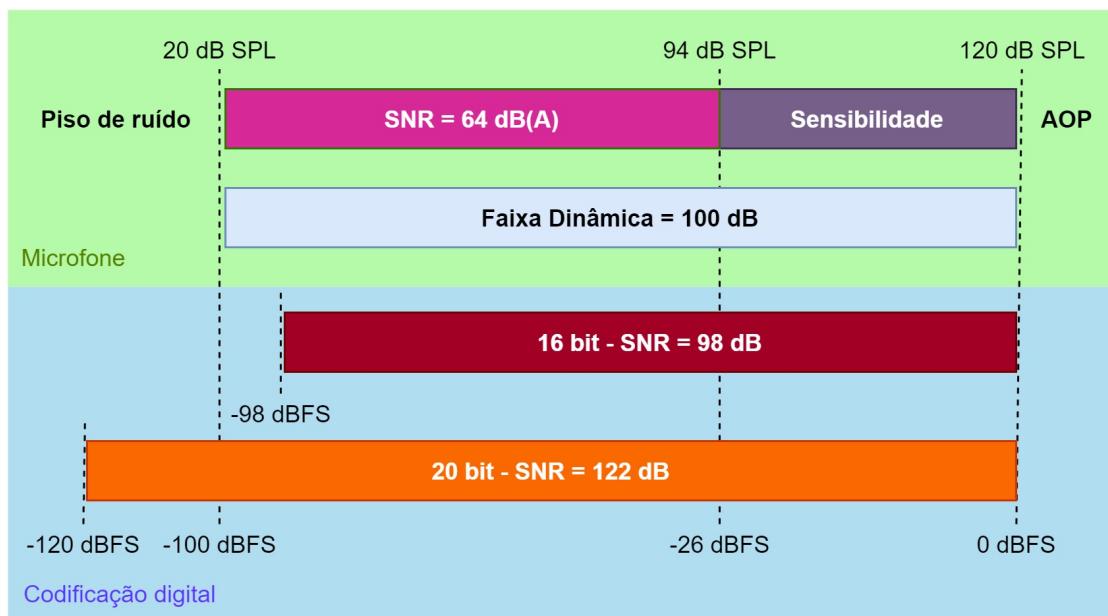
Tabela 1: Parâmetros importantes do microfone Knowles SPH0641LU4H-1. Outros parâmetros a se destacar são as dimensões de  $3.5 \times 2.65 \times 1.1 \text{ mm}^3$ , diretividade omnidirecional, alimentação de  $1.5V - 3.6V$  e corrente máxima de  $1mA$ . [[Knowles, 2015](#)]

Microfone	THD	Sensibilidade	SNR	AOP	DR	Codificação
SPH0641LU4H-1	0.2%	-26 dBFS	64.3 dB	120 dB SPL	100 dB	$\approx 16$ bits

A relação sinal-ruído (SNR) é a distância entre a potência do sinal de referência (a SNR é medida a partir de um sinal senoidal de potência sonora de  $94\text{dB SPL}$  e  $1\text{kHz}$ ) até o piso de ruído e permite encontrar o som menos intenso que o microfone suporta e o ponto de sobrecarga acústica (AOP) indica o som mais intenso, demarcado pelo ponto que a distorção harmônica total (THD) excede 10%, tornando estes dois parâmetros essenciais para entender a faixa dinâmica do microfone e a resolução que o sistema de codificação digital deve ter para representar o áudio em toda sua faixa de potência.

No caso do microfone utilizado o AOP é  $120\text{dB SPL}$  e a SNR de aproximadamente  $64\text{dB(A)}$  significa que a menor potência sonora está  $64\text{dB}$  abaixo de  $94\text{dB SPL}$ , ou seja, a menor potência sonora ou piso de ruído que o microfone suporta é  $20\text{dB SPL}$ . A distância entre é o AOP e o piso de ruído é a faixa dinâmica do microfone de  $100\text{dB}$ . Na Figura 7 está representado um diagrama que auxilia encontrar a faixa dinâmica do microfone.

Figura 7: Faixa dinâmica do microfone Knowles SPH0641LU4H-1 baseado na SNR e AOP. Inspirado em [Infineon \[2018\]](#).



Após encontrar a faixa dinâmica do microfone, a equação 1 ([Kester, 2005]) auxilia encontrar a faixa dinâmica (ou SNR) do sistema de codificação digital baseado na quantidade de *bits* de resolução que as amostras do áudio digital terá. Alguns valores de resolução típicos, como: 16, 20, 24 e 32 bits; são avaliados conforme representado na Tabela 2.

$$SNR/DR = 6.02N + 1.76 \text{ dB} \quad (1)$$

Tabela 2: Faixa dinâmica ou SNR suportada para diversos valores de bits de resolução da codificação digital .

Resolução	SNR/DR (dB)
16 bits	98.08
20 bits	122.16
24 bits	146.24
32 bits	194.40

De acordo com a tabela, apenas a resolução de 16 bits ( $SNR = 98dB$ ) não atenderia completamente a faixa dinâmica de  $100dB$  do microfone, pois não representaria o piso de ruído. Porém essa diferença não é significativa, pois tanto o microfone quanto a codificação digital com 16 bits são capazes de representar o sinal das USVs de acordo com as características especificadas, i.e. entre  $60dB SPL$  e  $100dB SPL$ .

## 2.4 Microcontrolador ESP32

O ESP32 é um *chip* microcontrolador da Espressif desenvolvido na tecnologia TSMC *ultra-low-power* de 40nm e possui conectividade sem fio e diversos periféricos integrados, habilitando seu uso para aplicações móveis, vestíveis e no campo da Internet das Coisas (IoT, do inglês *Internet of Things*) [Espressif, 2022a]. As características e periféricos do ESP32 pertinentes ao presente trabalho, são (Espressif [2022b], Espressif [2022a]):

- A versão ESP32-PICO-D4 baseada no *chip* do ESP32 possui 4MB de memória *flash* embutida no *chip*, por isso é uma boa escolha para o projeto, pois dispensa o uso de uma memória *flash* externa mandatória na versão mais simples do ESP32, consequentemente necessitando de menos componentes periféricos e minimizando a lista de materiais;
- O *chip* ESP32-PICO-D4 possui dimensões de  $7 \times 7 \times 0.94 \text{ mm}^3$ , permitindo economizar área na PCB;
- Conectividade *Wi-Fi* e *Bluetooth 2.4GHz* integrada no *chip*, necessitando de no mínimo uma antena de micro fita à parte (a antena é fornecida junto aos módulos que encapsulam o *chip*, porém ocupando mais espaço, e.g. o ESP32-WROOM-32 possui dimensões de  $18 \times 25.5 \times 3.1 \text{ mm}^3$  Espressif [2022c]), o que facilita a implementação

de uma transmissão sem fio em trabalhos futuros;

- Dois *cores* de 240 MHz, um *core ultra-low-power* (ULP) de 8MHz e 540kB de SRAM;
- Quatro periféricos SPI que podem ser utilizados para conexão com uma memória *flash* externa e/ou dispositivos de armazenamento do tipo cartão SD;
- *Host* SD/eMMC/SDIO para conexão com o cartão SD por meio de protocolo proprietário, porém a maior velocidade vem com o custo de um maior número de pinos que a conexão via interface SPI;
- Dois periféricos I<sup>2</sup>S que podem ser utilizados para capturar áudio digital no formato PCM, mas também pode receber áudio no formato PDM, pois possui um módulo de conversão PDM para PCM, porém suporta apenas duas taxas de redução de amostras: 64 e 128; o que inviabiliza a aquisição de áudio PDM com uma taxa de amostragem além de 98 kHz;
- Possui interface LED PWM e suporta até 16 canais e é particularmente útil para acionar LEDs RGB;
- Entre outros periféricos integrados, incluindo interfaces I<sup>2</sup>C que habilitam o uso de sensores, vários canais de conversão analógico-digital (ADC, do inglês *analog-to-digital converter*), dois conversores digital-analógico (DAC, do inglês *digital-to-analog converter*), trinta e quatro pinos de propósito geral (GPIO, do inglês *general-purpose input/output*, etc);

## 2.5 Interface I<sup>2</sup>S/PDM com o microfone

A interface *Inter-IC Sound* (I<sup>2</sup>S) é utilizada para comunicar áudio digital entre diversos dispositivos, e.g. microcontroladores, tocadores de som e microfones; considerando que estes possuem a interface implementada em seu *hardware*.

Um barramento I<sup>2</sup>S possui três linhas de sinal, i.e. *word select* (WS) para escolha do canal de áudio L/R, *clock* (SCK) para temporização dos dados e *data* (SD) para entrada ou saída de dados. O áudio digital transferido pode estar especificados em diversos padrões, e.g. Philips, alinhamento MSB, PCM. [Lewis, 2012]

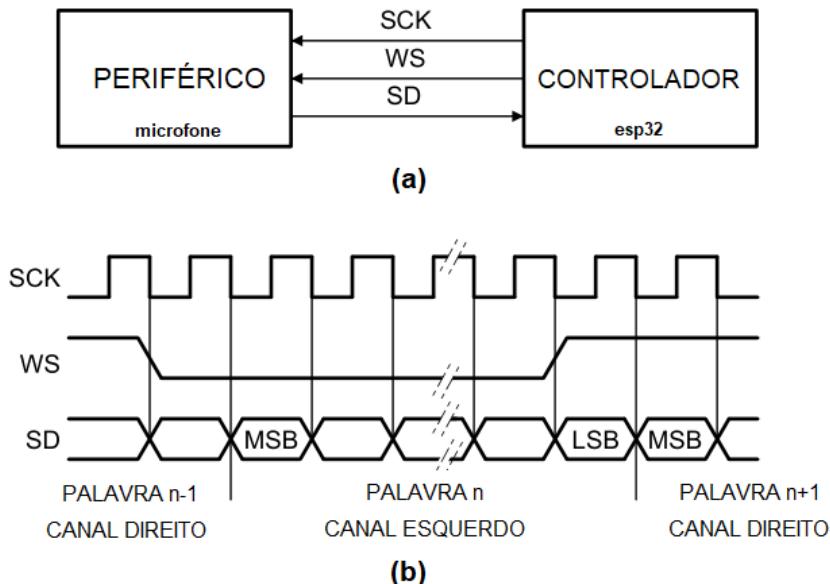
Na aplicação proposta para este trabalho, ao utilizar a interface I<sup>2</sup>S do ESP32, o módulo de conversão PDM para PCM do periférico I<sup>2</sup>S não possui taxas de sobre-amostragem (OSR) adequadas para trabalhar com o áudio ultrassônico do microfone utilizado. As taxas de sobre-amostragem possíveis de 64 e 128 combinadas ao *clock* entre 3.072-4.8 MHz

aceitas para o microfone SPH0641LU4H-1 trabalhar na faixa ultrassônica ([Knowles \[2015\]](#)) produzem um sinal PCM com uma taxa de amostragem que varia entre 48-75kHz para OSR de 64 e 24-37.5kHz para OSR de 128 o que não permite atender os requisitos da aplicação.

Com isso, o uso da conversão PDM para PCM em *hardware* disponibilizada pelo ESP32 se torna inviável para trabalhar na faixa ultrassônica. No entanto, é possível realizar a conversão via *software* mantendo todos os benefícios do periférico I<sup>2</sup>S como, por exemplo, o acesso direto a memória (DMA) e o ajuste da frequência do *clock* do microfone de acordo com a frequência de amostragem desejada. Para isso é importante ter uma noção básica do diagrama temporal do sinal PCM esperado no periférico I<sup>2</sup>S do microcontrolador, pois sem a conversão PDM para PCM em *hardware*, as palavras desse sinal digital representam um fluxo de bits PDM que pode ser pós-processado em *software*.

É possível observar na Figura 8 o diagrama de blocos e o diagrama temporal de uma transação I<sup>2</sup>S. Nota-se na Figura 8(a) o sentido dos sinais, com o *clock* e *word select* sendo produzidos pelo microcontrolador e os dados recebidos pelo microfone. No caso do microfone PDM o sinal de *word select* é dispensável, pois serve apenas para delimitar o início e fim dos bits de uma amostra de cada canal.

Figura 8: Funcionamento da interface I<sup>2</sup>S por meio do (a) diagrama de blocos e (b) diagrama temporal. Adaptado de [NXP \[2022\]](#).



Na condição de adaptar a interface para o recebimento de um sinal PDM bruto,

deve-se notar que é importante fazer com que a interface interprete as amostras como se fosse um áudio digital de dois canais, pois na condição de um canal a interface iria descartar os bits da amostra de um dos canais caso não houvessem dois microfones conectados a ela.

Outro detalhe é que os bits de extremidade devem ser devidamente identificados a depender do padrão de áudio digital que a interface está trabalhando, pois o correto ordenamento permite recuperar os bits do sinal PDM na ordem cronológica correta. No caso da Figura 8(b) o bit mais significativo (MSB) deve ser o bit mais antigo gerado pelo microfone e o bit menos significativo (LSB) o mais novo.

## 2.6 Interface SPI com o cartão micro SD

Um cartão *Secure Digital* (SD) é disponível em vários tamanhos, e.g. SD, miniSD e micro SD; sendo este último o mais adequado para usar em qualquer aplicação, pois tem as menores dimensões ( $11 \times 15\text{mm}^2$ ), capacidade de armazenamento de até 1TB e velocidade de escrita de até 90 MB/s (utilizando protocolo de comunicação proprietário). [Davis, 2015, [Kingston](#)]

A comunicação com o cartão micro SD pode ser feita por um protocolo SD proprietário ou por uma interface SPI. A diferença do ponto de vista é que a documentação do SPI é publicamente documentada e a do SD, não. Isso torna a interface SPI mais fácil de ser encontrada em microcontroladores.

É possível observar na Figura 9 a pinagem típica de um cartão micro SD com oito pinos. As Tabelas 3 e 4 representam, respectivamente as conexões dos pinos necessárias para o cartão operar no modo SPI e SD. Nota-se que o modo SPI utiliza dois pinos a menos que o modo SD e este possui várias linhas de dados e funções extras, separando a transmissão de dados e de comandos em diferentes pinos, enquanto que o modo SPI realiza essas duas funções na mesma linha.

Figura 9: Pinagem de um cartão micro SD.

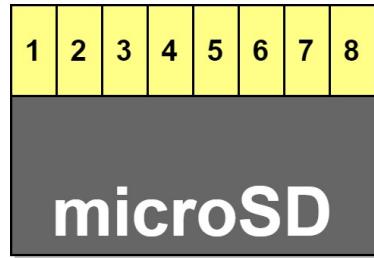


Tabela 3: Configuração dos pinos do cartão microSD para operar em modo SPI. [Davis, 2015]

	Modo SPI		
Pino	Nome	Tipo	Descrição
1	X	I	Reservado
2	CS	I	Chip select (ativo baixo)
3	DI	S	Dados e comandos host para cartão
4	VDD	S	Tensão de alimentação
5	SCLK	I	Clock
6	VSS	S	Tensão de referência
7	DO	O	Dados e status cartão para host
8	X	-	Reservado

Tabela 4: Configuração dos pinos do cartão microSD para operar em modo SD proprietário. [Davis, 2015]

	Modo SD		
Pino	Nome	Tipo	Descrição
1	DAT2	I,O	Linha de dados 2 (bit 2) / Ler esperar
2	CD/DAT3	I,O	Detecção de cartão / Linha de dados 3 (bit 3)
3	CMD	S	Comando / Resposta
4	VDD	S	Tensão de alimentação
5	CLK	I	Clock
6	VSS	S	Tensão de referência
7	DAT0	I,O	Linha de dados 0 (bit 0)
8	DAT1	I,O	Linha de dados 1 (bit 1) / Interrupção

Em relação à velocidade, o protocolo SD é mais rápido, pois atende os requisitos de classe de velocidade dos cartões SD, que variam desde a menor velocidade de 2 MB/s (classe

2) até 90 MB/s (classe V90). Em contra partida o modo SPI deve ser sempre considerado como classe 0, independente do que for informado no cartão, pois a classe de velocidade não pode ser garantida. A classe 0 não possui performance definida, pois deriva das especificações do sistema de classes, mas é dito que esta é mais lenta que a classe 2. [Specification, 2006]

## 3 Desenvolvimento

Ao decorrer do desenvolvimento do projeto foram utilizadas as tecnologias, linguagens de programação, bibliotecas e programas listados na Tabela 5. Todas as tecnologias listadas são gratuitas (gratuitas pelo menos durante o período de desenvolvimento deste trabalho) e, em alguns casos, de código aberto, permitindo que as etapas abordadas no desenvolvimento do sistema aqui proposto sejam reproduzidas facilmente, limitando-se apenas ao conhecimento dos conceitos de engenharia e do uso das tecnologias.

### 3.1 Requisitos da aplicação

As características físicas das USVs e os aparelhos experimentais que a literatura utiliza para sua aquisição foram extensamente discutidos em seções anteriores. Portanto, serão consideradas as seguintes especificações para o sistema de aquisição proposto:

- Nível de pressão sonora:  $60dB SPL$  até  $100dB SPL$ ;
- Banda do sinal:  $80kHz$ ;
- Frequência de amostragem mínima:  $160kHz$ ;
- Bits de resolução das amostras de áudio:  $16bits$ ;
- Dimensões aproximadas do dispositivo:  $2 \times 2cm^2$ ;
- Monitoração das condições do sistema por um LED RGB;
- Armazenamento de dados e configuração do sistema por um cartão microSD;
- Botões para operação do sistema;

### 3.2 Hardware

#### 3.2.1 Arquitetura do sistema de aquisição

A Figura 17 representa a proposta da arquitetura da cadeia do sistema de aquisição de áudio ultrassônico em duas versões, considerando o microfone analógico e digital. No caso do microfone analógico, a cadeia de processamento do sinal é mais extensa e difícil de projetar, pois necessita de um ou mais estágios de amplificação, filtragem passa-baixas para

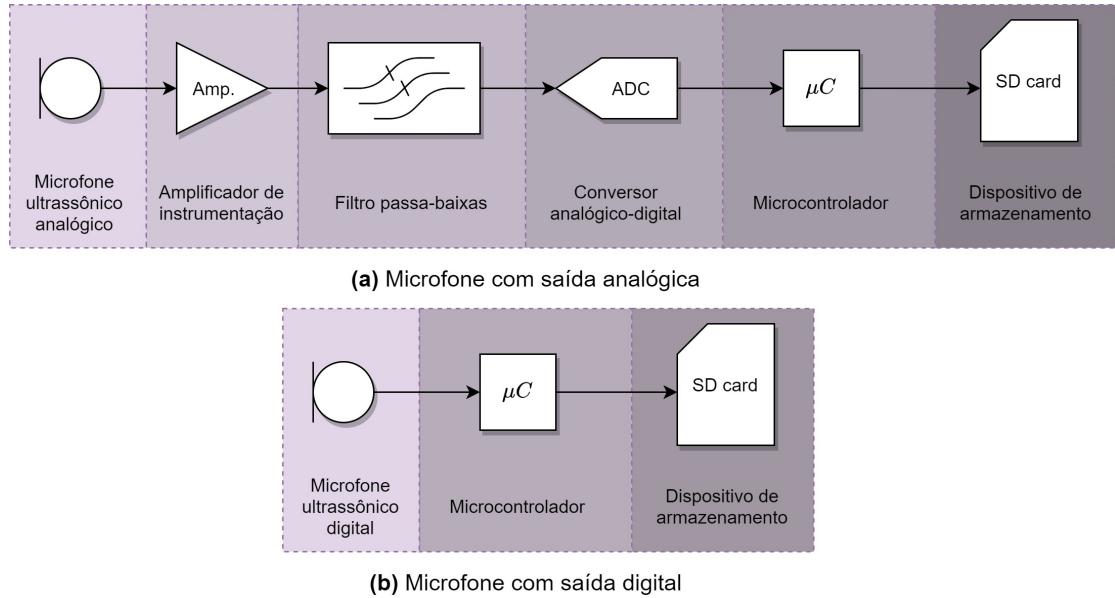
Tabela 5: Tecnologias utilizadas no desenvolvimento do projeto. Inspirado em [Sarkar \[2018\]](#).

Tarefas	Tecnologias	Bibliotecas
Ferramentas de colaboração em equipe	- Trello - Google Meets	
Documentação técnica e científica	- Doxygen - Mendeley	- Bibtex
Desenho de diagramas	- draw.io	- XML
Formatação de texto, planilhas e apresentações	- Overleaf - Google workspace	- LaTeX - Beamer - Docs, Sheets
Hospedagem e versionamento do código fonte	- Github	- Git
Projeto da PCB	- EasyEDA - Mouser Library Loader	- Componentes da Mouser - Componentes open source
Firmware do microcontrolador	- C	- GPIO - I2S - LEDC - SD/SPI - freeRTOS
Monitoramento do áudio	- Audacity - Python	- Chronux - Soundfile - Sounddevice - Pyaudio - Matplotlib - Plotly
Desenvolvimento e testes de algoritmos de processamento de áudio	- Python	- Numpy
Ambiente de desenvolvimento integrado	- VScode	- PlatformIO IDE - C/C++ IntelliSense - Pylance

evitar *aliasing* e a necessidade de um conversor analógico-digital. A solução para o presente trabalho é a arquitetura baseada no microfone digital, pois necessita apenas da interface com um microcontrolador, pois o processamento do sinal analógico é realizado no próprio microfone.

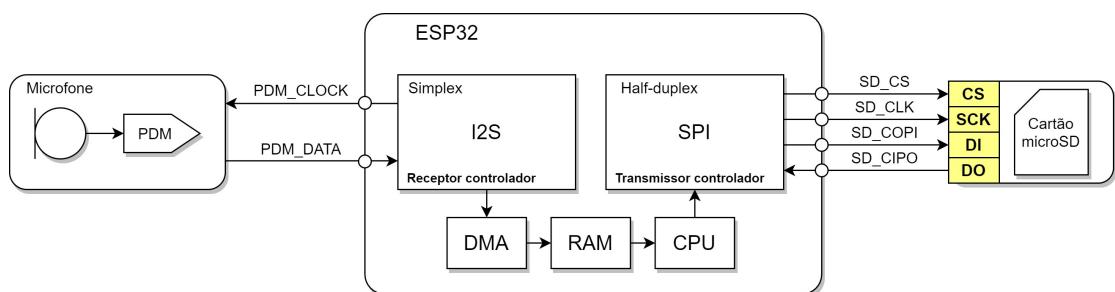
É possível observar na Figura 11 uma visão mais aprofundada do sistema com microfone digital escolhido. A figura estabelece o uso da interface de comunicação I<sup>2</sup>S entre o controlador

Figura 10: Arquitetura do cadeia de aquisição de áudio para microfone analógico e digital.



e o microfone PDM e a interface de comunicação SPI com o cartão micro SD. Além disso ela ilustra o uso do DMA para armazenar os dados obtidos do microfone direto na memória RAM (no caso do ESP32 a SRAM). Isso permite um envolvimento mínimo da CPU na coleta de dados, permitindo que um núcleo da CPU trabalhe no processamento dos dados e o outro na gravação do áudio processado no cartão SD.

Figura 11: Visão do sistema de aquisição de áudio ao nível de microcontrolador e periféricos.



Para atender a arquitetura do sistema de aquisição e os requisitos da aplicação, serão utilizados os seguintes componentes eletrônicos:

- 1 × Microfone digital Knowles SPH0641LU4H-1 ([Knowles, 2015]);
- 1 × Microcontrolador ESP32-PICO-D4 ([Espressif, 2022b]);

- 1 × Conector para cartão micro SD MEM2075-00-140-01-A ([[GCT, 2020](#)]);
- 1 × Regulador LDO de 3.3V TC1262-3.3VDB ([[Microchip, 2012](#)]);
- 1 × Conector para bateria formato moeda CR2450 (também funciona para LIR2450) ([[Keystone, 2010](#)]);
- 1 × LED RGB ASMB-MTB0-0B3A2 ([[Broadcom, 2015](#)]);
- 1 × bateria formato moeda LIR2450 ([[Stream, 2022](#)]);
- 2 × *Push-buttons*;

Na Tabela 6 estão representados os quantitativos de corrente consumida, peso e preço dos principais componentes utilizados. Do ponto de vista da potência consumida pelo sistema, a bateria LIR2450 possui uma capacidade nominal de  $100mAh$  (capacidade típica de  $120mAh$ ) e máxima corrente de descarga de  $200mA$ . Isso significa que a bateria é capaz de alimentar o sistema de carga  $105.7mA$  com duração de aproximadamente 55 minutos.

Tabela 6: Quantitativo dos componentes do sistema e balanço de carga mecânica e elétrica.

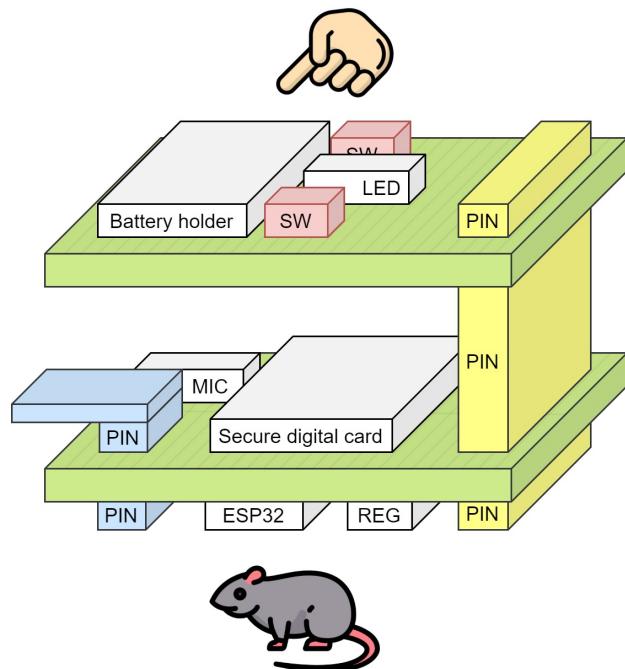
Função	Quantidade	Peso (mg)	Corrente (mA)	Preço (USD)
Microcontrolador	1	1479	50	\$4.95
Microfone	1	77	0.62	\$1.96
Conektor bateria	1	5663	-	\$1.48
Conektor microSD	1	720	30	\$1.96
Regulador	1	250	0.08	\$0.69
LED RGB	1	-	25	\$0.69
Botão	2	-	-	\$2.40
PCB	1	8000	-	\$0.40
<b>TOTAL</b>		16189	105.7	\$14.53

Após delimitar quais componentes serão utilizados e como estes interagem entre si, resta projetar a PCB para o protótipo final. Para isso foi desenvolvido um diagrama ilustrativo e fora de proporção da noção de ocupação dos componentes na PCB, como pode ser observado na Figura 12.

Inicialmente propõe-se que a placa final possua duas PCBs de duas camadas interligadas fisicamente por um barramento de pinos de transferência de dados e, se necessário, fixada mecanicamente pelas bordas. Essa organização permite o aproveitamento do espaço ocupado pela PCB ao expandi-la na dimensão da altura.

Na face superior externa foram dispostos os componentes que possuem direta interação com o usuário, aqui chamada de interface do usuário. A interface do usuário possui o

Figura 12: Noção de ocupação dos componentes na PCB no projeto final, indicado a face que possui a interface com o usuário (superior externa) e a face de contato com o animal (inferior externa).



conector da bateria, os botões para interação e o LED RGB para monitoramento. A partir da face superior sai um barramento de pinos contendo a alimentação da bateria, comando dos botões e do LED.

Na face inferior interna estão dispostos o microfone, o conector do cartão micro SD e um barramento de pinos para permitir a reprogramabilidade do microcontrolador. O conector do cartão na parte interna evita que o contato mecânico com o rato (que pode ser intenso a depender do experimento representado na Figura 1) desconecte o cartão micro SD, pois além da bateria este é o único componente removível do sistema. O microfone na posição descrita possui a entrada de som voltada para baixo em direção ao roedor, permitindo uma melhor captura das ondas sonoras.

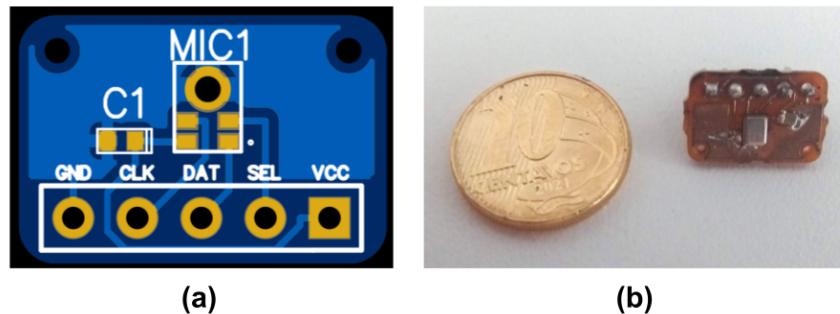
A face inferior externa possui contato direto com o animal. Nela estão localizados o microcontrolador e o regulador, assim como a maioria dos componentes passivos que auxiliam no funcionamento do sistema.

### 3.2.2 Protótipos

**Protótipo 0 - PCB do microfone** O primeiro protótipo se baseia no desenvolvimento de uma PCB de única face que comporta apenas o microfone para testes iniciais (por isso a numeração 0) antes de cometer à arquitetura final. Essa PCB conecta o microfone diretamente a um *kit* de desenvolvimento do ESP32 (DevKit ESP32) e com um conector de cartão SD, permitindo que o projetista realize testes no *firmware* sem se preocupar na alimentação por uma bateria ou depurar problemas que uma PCB mais complexa pode gerar.

É possível observar na Figura 27 a PCB do protótipo 0.

Figura 13: PCB do protótipo 0 ( $14.59 \times 9.8\text{mm}^2$ ) com (a) desenho 2D da PCB no *software* EasyEDA e (b) versão física comparada com uma moeda de 20mm de diâmetro.



**Protótipo 1 - PCB com DevKit do ESP32** Apesar do protótipo 0 ser útil para testes iniciais, ele não é válido para testes a longo prazo, pois a natureza da sua configuração é instável, visto que a conexão entre o módulo do microfone, microcontrolador e cartão micro SD é feita por fios que introduzem problemas de mau contato.

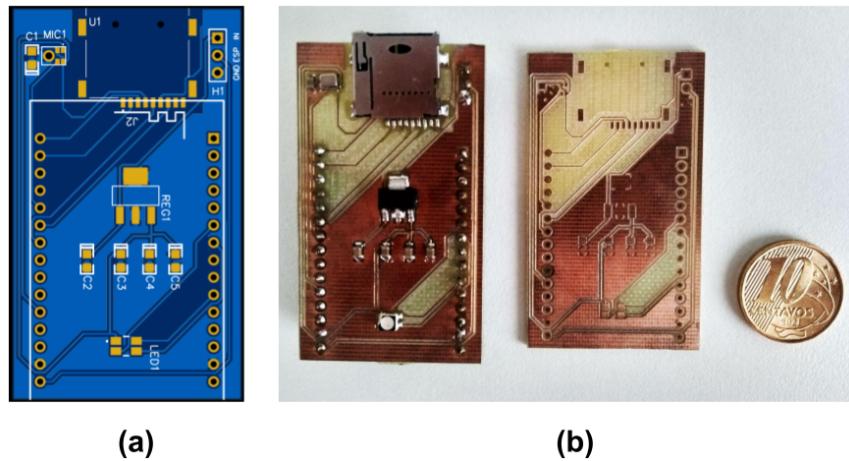
A PCB de única face implementou o conector do cartão micro SD, o LED, o regulador, o microfone e um DevKit ESP32 em uma PCB, permitindo uma conexão mais estável entre os componentes por meio de trilhas de cobre.

Nota-se a introdução do regulador neste protótipo, mas o DevKit ESP32 em si possui sua própria regulação. Portanto, para introduzir o uso do regulador proposto neste projeto, foi criada uma conexão de pinos que permite o usuário escolher alimentar o circuito por

uma alimentação externa que passa pelo regulador ou a alimentação do DevKit ESP32.

É possível observar na Figura 28 a PCB do protótipo 1.

Figura 14: PCB do protótipo 1 ( $33.59 \times 57.28\text{mm}^2$ ) com (a) desenho 2D da PCB no *software* EasyEDA e (b) versão física comparada com uma moeda de 20mm de diâmetro.



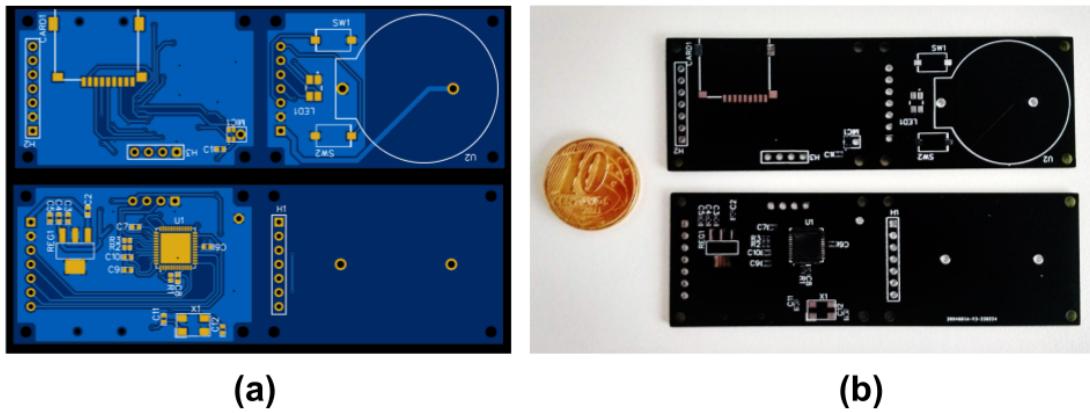
**Protótipo 2 - PCB final** O protótipo final busca implementar a organização dos componentes de acordo com a noção de ocupação representada na Figura 12.

É possível observar na Figura 29 a PCB do protótipo 2. Apesar dos esforços, as dimensões físicas do protótipo final ( $44.26 \times 28.7\text{mm}^2$ , considerando após o corte e junção das placas) superam àquelas especificadas de  $20 \times 20\text{mm}^2$ . Nota-se que a PCB possui espaços livres em todas as faces, mas que não foram aproveitados eficientemente devido à limitação imposta pelas grandes dimensões do conector da bateria e da necessidade dos botões e LED estarem na mesma face.

Ao soldar os componentes na placa, a massa total do sistema contabilizou cerca de 11 gramas, sem bateria, levando em conta os componentes representados na Tabela 6 (o conector da bateria foi cortado para ocupar menos espaço, portanto possuindo menor massa que a esperada na tabela).

É possível observar na Tabela 6 o preço total do sistema fica em torno de U\$15 (dólares americanos), um valor menor que as soluções tradicionais para aquisição de áudio e as soluções alternativas, como o AudioMoth e as soluções que usam o Raspberry Pi.

Figura 15: PCB do protótipo 2 ( $44.26 \times 28.7\text{mm}^2$ ) com (a) desenho 2D da PCB no *software* EasyEDA e (b) versão física comparada com uma moeda de 20mm de diâmetro, mostrando as visões do topo (cima) e do fundo (baixo).



O preço individual foi considerado para cada componente e tende a diminuir ao comprar em maior quantidade. No cálculo do preço total foram desconsiderados os componentes passivos

A foto do protótipo 2 com os componentes soldados e as duas placas unidas podem ser observadas na Figura 16.

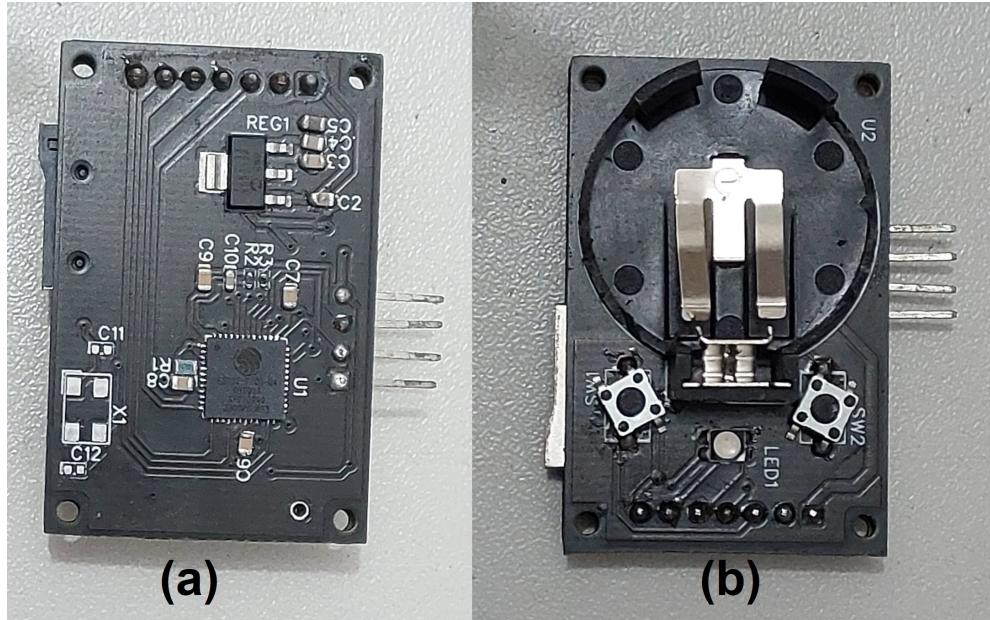
### 3.3 Firmware

#### 3.3.1 Fluxograma de atividades

O fluxograma das atividades executadas no código embarcado do microcontrolador é representado na Figura 17. A rotina é executada em três etapas distintas, i.e. inicialização do sistema, gravação do áudio e finalização da gravação.

De acordo com o fluxograma a rotina começa com as funcionalidades de inicialização. Nesta etapa do código serão inicializados todos os periféricos que serão utilizados, e.g. GPIO, I2S, SPI e LEDC; e, se não forem ser utilizados de imediato, serão forçados ao estado de suspensão até serem resumidos. Algumas variáveis globais também são inicializadas com base no conteúdo de um arquivo de configuração modificado pelo usuário e salvo no cartão

Figura 16: PCB do protótipo 2 ( $44.26 \times 28.7\text{mm}^2$ ) com os componentes soldados com a visão (a) inferior externa (área de contato com o animal) e visão (b) superior externa (interface com o usuário).



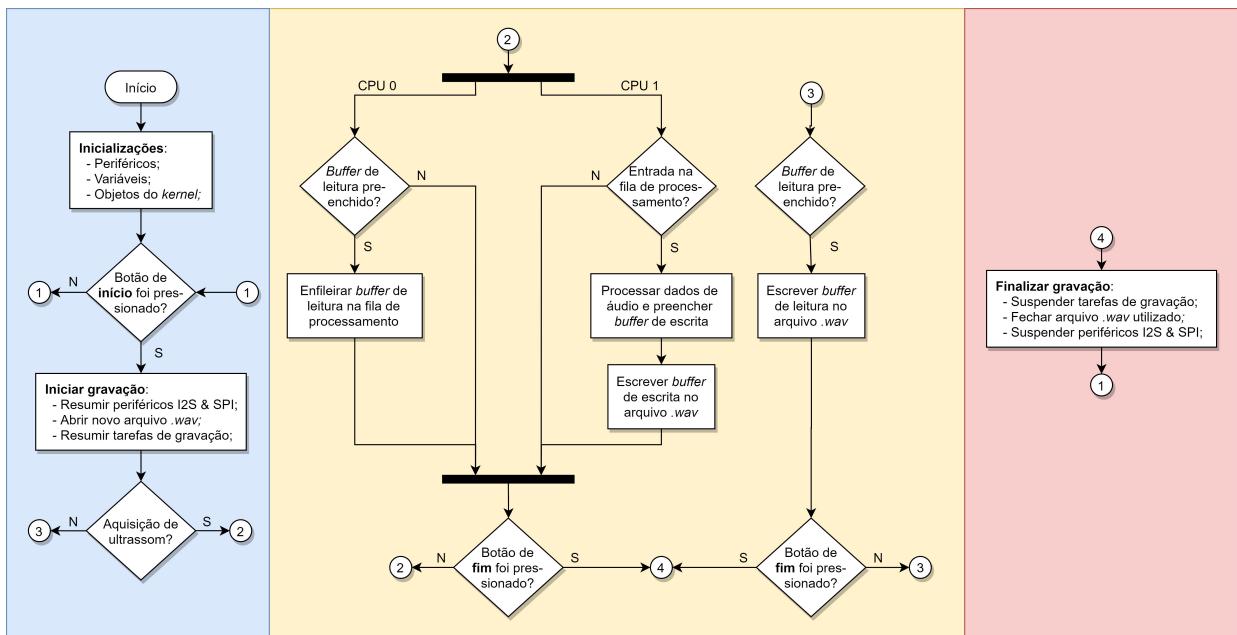
micro SD em uso.

Esse arquivo de configuração habilita mais um nível de customização do sistema de acordo com o interesse do usuário, também permitindo configurar: terminação dos nomes do arquivos gerados, cor do LED durante gravação, duração da gravação e duração dos intervalos entre gravações.

Os objetos do *kernel* do freeRTOS são inicializados conforme a disponibilidade de *tasks*, *mutexes*, *semaphores*, *queues* e *event groups*; e são fundamentais para habilitar todo o poder de processamento do ESP32 com o uso de seus dois núcleos.

Após a inicialização, o sistema entra em um estado inativo que aguarda por uma interrupção gerada ao pressionar o botão de início da gravação ou que a duração dos intervalos entre gravações seja atingido, caso especificada. Quando o sistema recebe o comando para iniciar a gravação, a tarefa de iniciação resume a operação dos periféricos I2S e SPI, que permitem a leitura e escrita dos dados, respectivamente. Em seguida abre-se um novo aquivo no sistema de arquivos do cartão micro SD e inicializa o cabeçalho *.wav*. Para finalizar esta etapa, as tarefas responsáveis pela gravação do áudio também são resumidas.

Figura 17: Fluxograma de atividades do *firmware*.



Nesse momento o sistema decide se vai realizar uma aquisição na faixa do ultrassom ou na faixa sonora.

Vale destacar que o *datasheet* do microfone SPH0641LU4H-1 prevê um diagrama de estados com a frequência de *clock* como variável de controle e que deve ser obedecido para acessar os diferentes modos de operação. Em resumo, o microfone pode ser inicializado com um *clock* de até 2.5MHz para operar nos diferentes modos de operação na faixa audível (modo de baixa potência, modo soneca ou modo padrão), porém não pode ser inicializado diretamente no modo ultrassônico que requer um *clock* além de 3MHz, devendo assim ser transicionado para o modo ultrassônico a partir dos modos anteriores, exceto o modo desligado (sem alimentação).

A gravação na faixa sonora é simples e realiza com apenas uma tarefa responsável por receber as transações do periférico I2S com os dados no formato PCM previamente convertidos no módulo de conversão PDM para PCM do periférico I2S e, em seguida, escrever os dados no cartão micro SD.

A gravação na faixa ultrassônica é realizada utilizando os dois núcleos do microcontrolador. Um núcleo é responsável por receber as transações do periférico I<sub>2</sub>S e colocar os dados em

uma fila para que a tarefa do outro núcleo realize a conversão do sinal PDM para PCM via *software* e escreva as amostras de áudio processadas no cartão micro SD.

Durante a execução das tarefas de gravação, caso ocorra mais uma interrupção por conta do botão ou caso o tempo especificado para gravação esgote, o sistema vai compreender o interesse do operador de finalizar a gravação e vai encaminhar a rotina para a etapa de finalização.

A tarefa de finalização apenas desfaz aquilo que a tarefa de iniciação fez e na ordem contrária. Primeiro as tarefas de gravação são suspensas para cessar qualquer leitura e gravação; o arquivo *.wav* é modificado mais uma vez para finalizar o cabeçalho e é fechado para evitar a corrupção do espaço de memória; por fim os periféricos I<sup>2</sup>S e SPI são suspensos novamente para economizar energia. A rotina retorna ao estado inativo e irá esperar que o botão seja pressionado novamente para dar continuidade à gravação ou que o tempo de intervalo entre gravações se esgote.

Apesar da atividade do LED não estar explícita no diagrama, este tem o papel fundamental de indicar a todo momento os estágios de execução do código por meio de um código de cores específico.

O código fonte do *firmware* e dos algoritmos utilizados podem ser consultados no repositório hospedado publicamente no Github [Galvão \[2022\]](#).

### 3.3.2 Adaptações na interface I<sup>2</sup>S para aquisição do fluxo de bits PDM

A operação de conversão do sinal PDM para o PCM será tratado neste trabalho como conversão *PDM2PCM* para facilitar a comunicação.

Como discutido anteriormente, a interface de conversão *PDM2PCM* embutida no periférico I<sup>2</sup>S do ESP32 é capaz de trabalhar apenas com taxas de sobre-amostragem que inviabilizam o uso do microfone na sua faixa de *clock* especificada para o modo ultrassônico (3.072-4.8 MHz, porém o microfone suporta frequências ligeiramente maiores, como 5MHz).

Não dispondo das ferramentas de *hardware* necessárias, a solução é recorrer para a conversão *PDM2PCM* por meio de *software*, mesmo que não seja ideal, pois tende a ser mais lenta e exige mais recursos computacionais da CPU.

Apesar do módulo de conversão *PDM2PCM* disponível no I<sup>2</sup>S ser dispensável para a aplicação, isso não exclui totalmente o uso do periférico na aquisição do áudio PDM, pois este ainda possui um controlador DMA e funções de fácil uso.

A interface I<sup>2</sup>S pode ser utilizada no seu modo de operação normal como se fosse coletar um sinal PCM a partir da linha PDM, porém utilizando apenas a linha de *clock* e dados, descartando a linha de *word select*. Os bits recebidos serão organizados em amostras de 16 ou 32 bits, que do ponto de vista da interface I<sup>2</sup>S seriam amostras de áudio padronizadas, mas que de fato codifica o fluxo de bits PDM e pode ser recuperado em *software* ao aplicar uma máscara de bits adequada.

A máscara de bits utilizada sobre as amostras obtidas deve ser cuidadosamente projetada para obedecer o padrão de áudio que a interface I<sup>2</sup>S usou (e.g. padrão Philips, ou padrão de alinhamento por MSB, padrão PCM, etc.), a ordenação dos bytes (*endianness*), a quantidade de bits por amostra (16 ou 32 bits), etc., permitindo extrair o fluxo de bits PDM original na ordem cronológica correta.

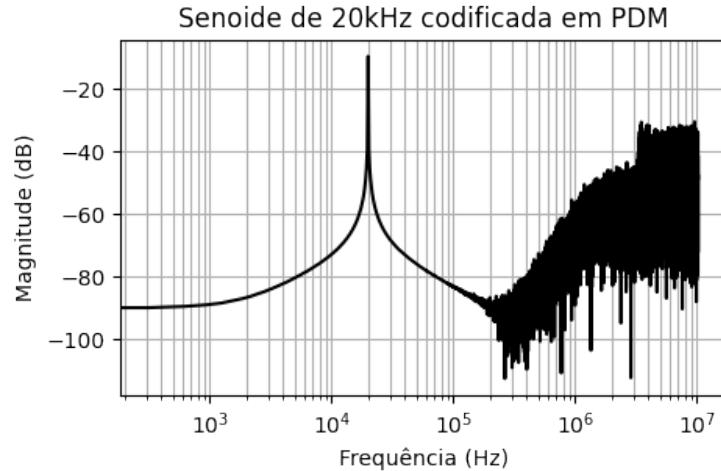
### 3.3.3 Sistema de filtragem digital

É possível observar na Figura 18 a natureza do espectro de magnitude de um sinal PDM. O sinal codifica uma senoide de 20kHz que pode ser identificada pelo pico de magnitude. Além disso, existe o ruído de quantização do sinal PDM deslocado para altas frequências, além da banda do sinal, devido a modelagem do ruído pelo conversor Sigma-Delta. [Loibl et al., 2016, Baker, 2011a]

Para obter o sinal senoidal original da Figura 18 é necessário realizar uma filtragem passa-baixas para restringir o espectro à banda passante do sinal de interesse, seguido de um filtro de dizimação para reduzir a taxa de amostragem e permitir um armazenamento mais eficiente das amostras resultantes. [Baker, 2011b]

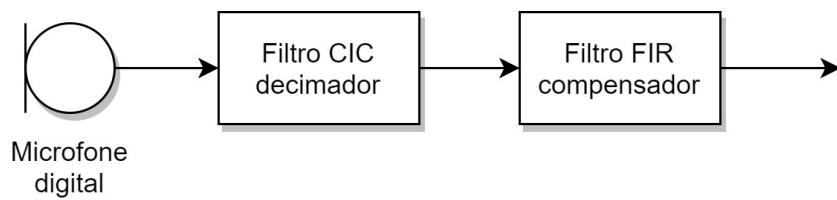
Como o sistema de filtragem vai ser implementado em *software*, sua arquitetura deve ser o mais eficiente possível. Uma forma de realizar a filtragem é com um filtro *Cascaded Integrator–Comb* (CIC) decimador. O filtro CIC decimador implementa uma filtragem passa-baixas seguida de uma dizimação em um único filtro e utiliza apenas operações de soma e subtração e requer o armazenamento de apenas três coeficientes: número de estágios

Figura 18: Espectro em magnitude de um sinal PDM codificando uma senoide de  $20kHz$  a uma taxa de amostragem de  $20.48MHz$  utilizando uma janela retangular.



(N), taxa de dizimação (R) e atraso diferencial (M). Porém sua magnitude da resposta em frequência não é suficientemente seletiva, pois segue o formato  $\text{sinc}(x)$ , necessitando que sua saída seja submetida por um filtro *Finite Impulse Response* (FIR) para compensar a banda passante não plana do filtro CIC. Na Figura 19 está representada a arquitetura do sistema de filtragem. [Lyons, 2020]

Figura 19: Diagrama de blocos do sistema de filtragem do áudio PDM com um filtro CIC decimador como elemento de filtragem principal e um filtro FIR para compensar a resposta em frequência do CIC. Inspirado em Lyons [2020].



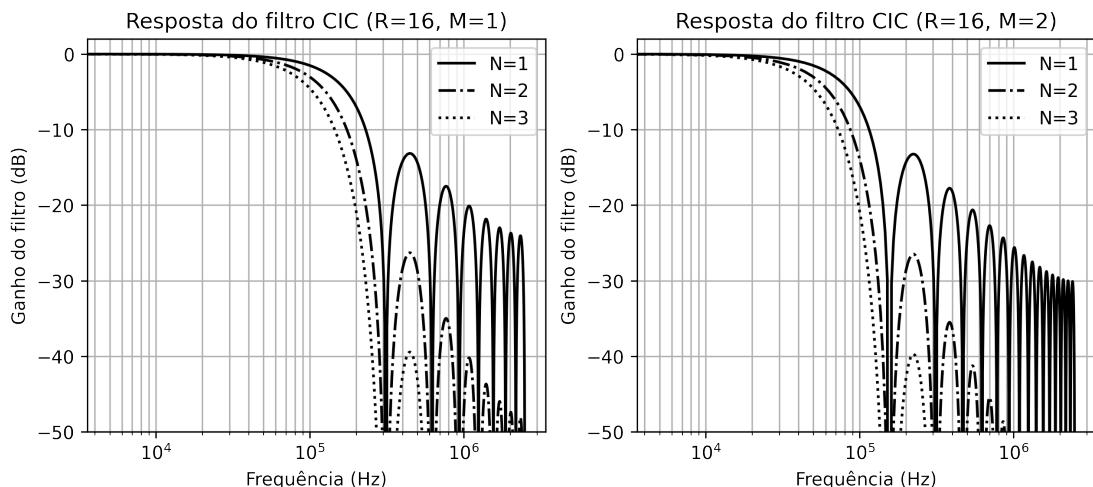
O microfone utilizado neste trabalho opera com um *clock* de até  $4.8MHz$  no modo ultrassônico, mas ele ainda retorna dados em frequências maiores. Portanto decidiu-se alimentar o *clock* de entrada do microfone com uma frequência de  $5MHz$  para o modo ultrassônico. Isso permite extrair a maior banda do sinal após a dizimação.

Em seguida define-se a taxa de sobreamostragem (OSR) ou taxa de dizimação (R),

em termos dos parâmetros do filtro CIC. Uma OSR de 32 e um *clock* de  $5MHz$  para o sinal PDM resulta em um sinal PCM com banda de  $78.125kHz$  que se torna próximo do limite definido para a banda do microfone e para a banda mais importante das USVs. Porém após a filtragem o sinal no limite da banda seria atenuado. Portanto a solução foi adotar uma OSR de 16 ( $R = 16$ ) que permite uma banda de  $156.25kHz$ , que apesar de ser bem superior à banda de interesse, ela permite maior flexibilidade nas respostas dos filtros.

Porém ainda resta definir o número de estágios e o atraso diferencial do filtro CIC. Para analisar seus efeitos no sinal foram avaliadas a magnitude das respostas em frequência de várias versões do filtro CIC conforme representado na Figura 20 com base na teoria proposta por [Lyons \[2020\]](#).

Figura 20: Magnitude da resposta em frequência do filtro CIC para frequência de amostragem de entrada de  $5MHz$  e OSR 16, variando a quantidade de estágios ( $N$ ) para (esquerda)  $M = 1$  e (direita)  $M = 2$ . Inspirado em [Lyons \[2020\]](#).



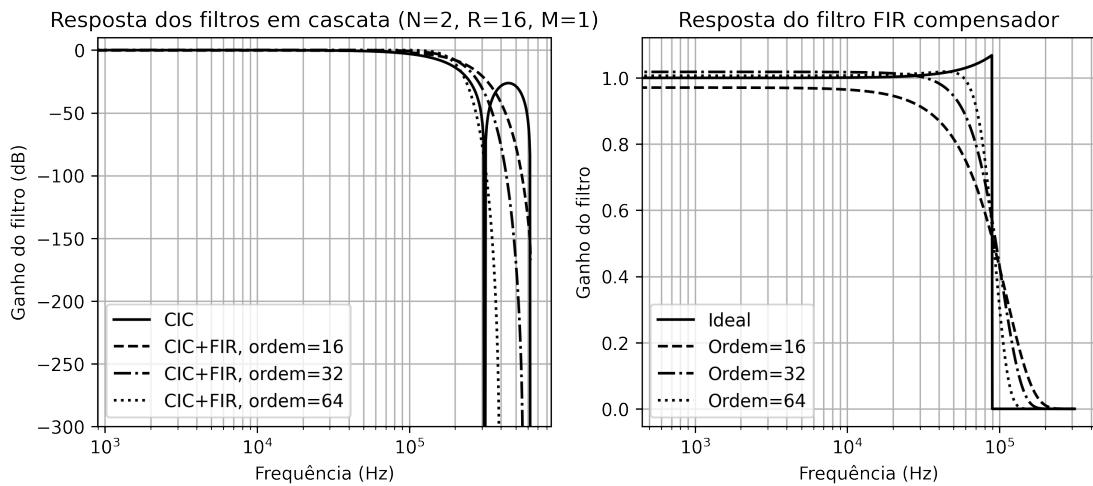
É possível observar na Figura 20 que a resposta torna-se mais seletiva com o aumento da quantidade de estágios do filtro ( $N$ ). Ao variar o atraso diferencial ( $M$ ), percebe-se que ao aumentar seu valor a resposta foi deslocada para a esquerda sem alteração da atenuação na resposta.

Para manter uma banda passante do sinal mínima de  $80kHz$  o atraso diferencial pode ser selecionado como 1 ( $M = 1$ ). O número de estágios escolhido como 2 ( $N = 2$ ) permite uma atenuação na faixa de rejeição que facilita a redução da ordem do filtro FIR compensador a ser utilizado e evita uma maior atenuação na banda passante do sinal

proporcionado por uma quantidade maior de estágios.

É possível observar na Figura 21 a magnitude da resposta em frequência do filtro FIR compensador e da sua associação com o filtro CIC anteriormente definido.

Figura 21: Magnitude da resposta em frequência (esquerda) do filtro em cascata (CIC+FIR) e filtro FIR compensador (frequência de corte de  $90kHz$  de janela Chebyshev com atenuação de  $300dB$  para os lobos secundários) para frequência de amostragem de  $5MHz/OSR$  ( $OSR=16$ ), variando a ordem do filtro para parâmetros fixos do filtro CIC ( $N = 2$ ,  $R = 16$ ,  $M = 1$ ). Inspirado em Lyons [2020].



Está representado na Figura 21(direita) apenas a resposta do filtro FIR e como ela se ajusta à resposta ideal. A Figura 21(esquerda) representa que, ao inserir o filtro FIR, o filtro em cascata proporciona melhor seletividade e atenua os lobos secundários do CIC, principalmente para uma ordem de 64 que alcança atenuação de  $-50dB$  antes de  $300kHz$ . Dessa forma, foi selecionado um filtro FIR de ordem 64.

Filtros FIR de ordem maiores poderiam possuiriam respostas mais seletivas, porém o consumo de recursos computacionais aumentaria, podendo ocasionar perda de dados durante a gravação.

Essa arquitetura do sistema de filtragem inicialmente foi implementado na linguagem Python para validação dos algoritmos. Após isso foi implementada de forma otimizada na linguagem C na máquina hospedeira para comparar os resultados com aqueles obtidos em Python. Por fim, o código na linguagem C foi adaptado para o *firmware* do microcontrolador para testes finais com fluxo de dados em tempo real.

### 3.3.4 Prevenção de perda de dados

Durante a execução do *firmware* deve ser garantido que não haja perda de dados entre cada ciclo de dados recebidos no I<sup>2</sup>S. Os principais parâmetros que controlam esse fenômeno é a quantidade de *buffers* (descritores) alocados para o DMA ( $DMA_{cnt}$ ) e a quantidade de amostras em cada um ( $DMA_{len}$ ). Para computar esses dois valores é necessário conhecer o menor intervalo que o I<sup>2</sup>S estará ocupado coletando dados ( $I2S_{time,min}$ ) e o maior tempo que o processador estará processando e armazenando os valores ( $CPU_{time,max}$ ). O ideal é que  $I2S_{time,min}$  seja o maior possível e  $CPU_{time,max}$  o menor possível.

O menor intervalo  $I2S_{time,min}$  é obtido no modo de operação na faixa audível, pois opera a uma taxa de amostragem maior que o modo ultrassônico ( $f_{s,std} = 98000Hz$ ). É importante esclarecer que a taxa de amostragem utilizada no I<sup>2</sup>S no modo ultrassônico é de  $f_{s,ult} = 78125Hz$ , pois ela equivale à adaptação necessária do I<sup>2</sup>S para que a interface emita um *clock* de 5MHz para o microfone, portanto  $f_{s,ult}$  não está relacionado à taxa de amostragem final após a filtragem e se relaciona com a frequência do *clock* PDM ( $PDM_{clock}$ ) pela equação 2 ( $m_{ult} = 32$  é a quantidade de bits utilizados na aquisição I<sup>2</sup>S).

$$PDM_{clock} = 2 \cdot m_{ult} \cdot f_{s,ult} = 5MHz \quad (2)$$

Arbitrando  $DMA_{len} = 1024$ , o intervalo  $I2S_{time,min}$  pode ser obtido de acordo com a equação 3:

$$I2S_{time,min} = \frac{DMA_{len}}{f_{s,std}} = 10.5ms \quad (3)$$

Porém o valor de  $DMA_{len}$  deve respeitar a relação 4, pois com base na quantidade de bits por amostras na aquisição I<sup>2</sup>S no modo padrão ( $m_{std} = 16$ ) e quantidade de canais ( $n = 1$ ), ela evita que o tamanho máximo do *bufferes* seja superado (4092).

$$\frac{1}{8} \cdot DMA_{len} \cdot m_{std} \cdot n < 4092 \Rightarrow 2048 < 4092 \quad (4)$$

O maior intervalo  $CPU_{time,max}$  é obtido no modo de operação na faixa ultrassônica

durante o tempo em que o processador está ocupado convertendo o fluxo de dados PDM para o áudio PCM e realizando a escrita no cartão micro SD. Ele pode ser computado temporizando a execução em tempo real do sistema e foi obtido como cerca de  $65ms$ .

Portanto a partir de  $I2S_{time,min}$  e  $CPU_{time,max}$  pode-se definir a quantidade de *bufferes* (descritores) que o DMA deve alocar para que não haja perda de dados e a relação é obtida pela equação 5:

$$DMA_{cnt} > \frac{CPU_{time,max}}{I2S_{time,min}} \Rightarrow DMA_{cnt} > 6 \quad (5)$$

Apesar disso foi escolhido o valor máximo suportado de  $DMA_{cnt} = 32$  para adequar-se ao pior caso e futuras expansões no sistema que podem causar atraso computacional.

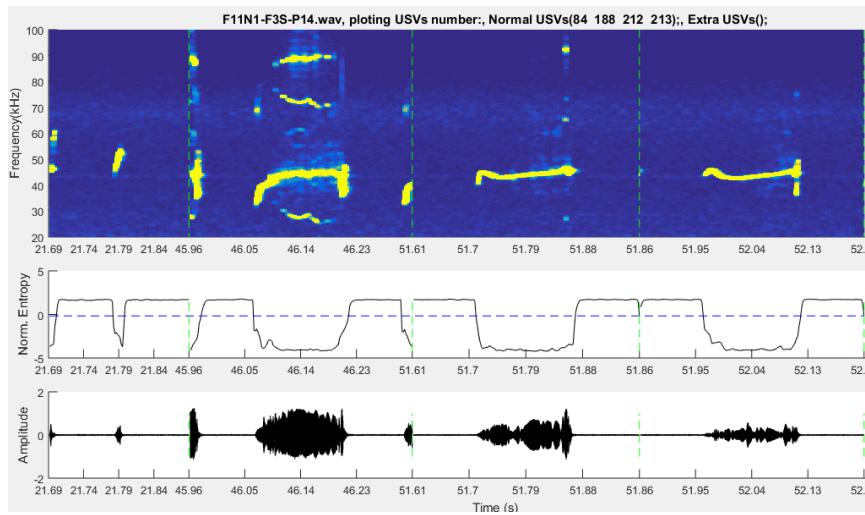
## 4 Resultados

Esta seção visa detalhar os resultados de alguns experimentos realizados para validar os protótipos desenvolvidos e entender a natureza do sinal das USVs.

### 4.1 Variação dos bits de resolução do áudio de USVs

Um programa típico de análise de USVs utiliza o espectrograma do sinal de áudio, como pode ser observado na Figura 22. O espectrograma indicado foi obtido a partir do áudio de vocalizações de um rato neonato com 14 dias de idades (P14) adquirido a uma taxa de amostragem de  $250\text{kHz}$ .

Figura 22: Típico programa de detecção e tratamento de USVs. (superior) espectrograma. (centro) entropia normalizada. (inferior) amplitude no tempo filtrada passa-faixa. Fonte: [Bessa et al. \[2019\]](#).



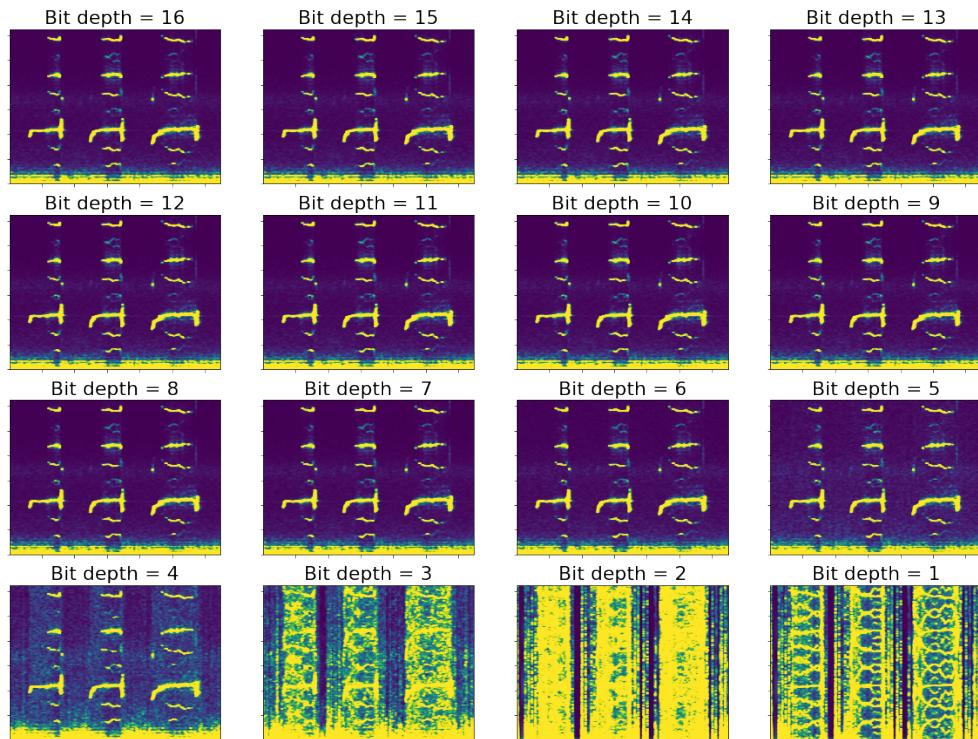
Apesar do sistema desenvolvido não ter realizado testes com a aquisição de USVs de roedores para observação do espectrograma como representado na Figura 22, arquivos de áudio previamente adquiridos em um sistema de aquisição amplamente utilizado podem ser utilizados para testes preliminares para compreender as características das USVs.

No caso deste teste foi utilizada uma amostra de áudio rica em USVs que foi adquirida com o sistema Ultramic 250k ([Dodotronic \[2016\]](#)) a uma taxa de amostragem de  $250\text{kHz}$  e

resolução de 16 bits, que de acordo com o que foi discutido sobre as características físicas das USVs é suficiente para um registro adequado.

O teste se baseou em quantizar o sinal de áudio de 16 bits para resoluções menores com o objetivo de observar, por meio do espectrograma, até que valor mínimo de resolução a informação da USV ainda é distinguível do ruído de quantização.

Figura 23: Espectrogramas calculados com base no mesmo trecho de áudio ultrassônico com três USVs e re-quantizado para toda a faixa abaixo de 16 bits.



Os resultados do teste estão representados na Figura 23 e, conforme a resolução cai, a perda da informação só é visualmente notada a partir de 4 bits. Esse tipo de resultado é fundamental para a arquitetura do sistema desenvolvido, pois ao indicar uma resolução do áudio menor que a esperada, a necessidade de espaço para armazenamento reduzirá e, em trabalhos futuros, ocasionará a redução na latência no envio de dados por comunicação sem fio devido a uma carga útil menor.

Apesar disso, esse áudio foi adquirido em um sistema de aquisição mais robusto que o desenvolvido neste trabalho. Portanto testes futuros devem ser feitos com relação à aquisição de USVs diretamente de roedores para conclusões mais confiáveis.

## 4.2 Validação na faixa sonora

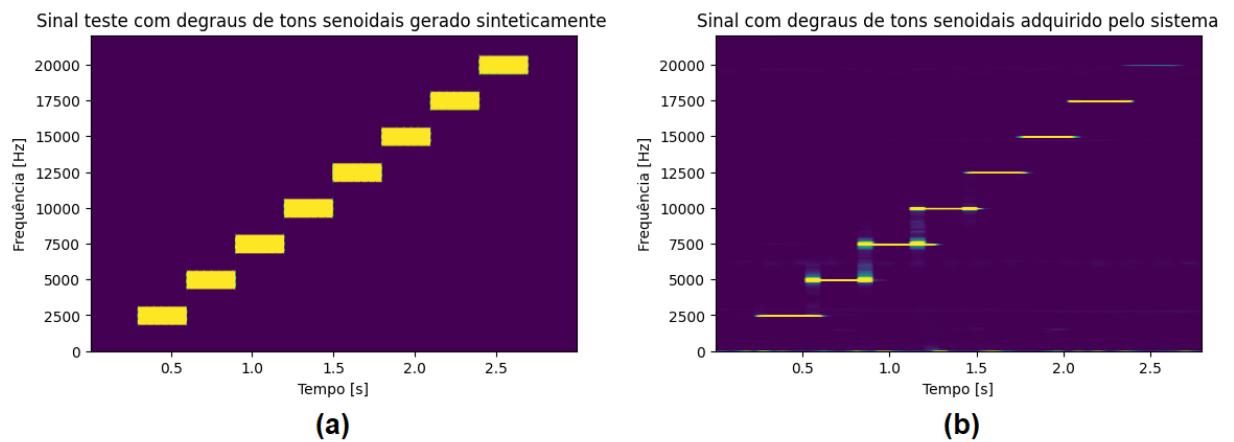
A validação na faixa sonora se baseia no uso da conversão *PDM2PCM* do *hardware* específico do periférico I2S do ESP32 a uma taxa de  $98ksp$ s.

Nas condições especificadas foi utilizado um áudio sintético de teste para excitar o microfone do sistema desenvolvido. O áudio é composto por tons senoidais cuja frequência aumenta em incrementos de  $2.5kHz$  em intervalos de  $0.3s$  até o limite da faixa sonora de  $20kHz$ .

O áudio de teste é emitido por um alto-falante Knup® KP-603 (resposta em frequência entre  $100-20kHz$ , impedância de  $4\Omega$  e sensibilidade  $70dB \pm 4dB$ ).

O espectrograma do sinal sintético de teste e do sinal adquirido pelo sistema de aquisição proposto pode ser observado na Figura 24.

Figura 24: Espectrograma das múltiplas senoides com frequência variando em forma de degrau para (a) o sinal de teste e (b) o sinal adquirido no sistema de aquisição.



O espectrograma resultante da aquisição indica a presença das 8 frequências exploradas no sinal de teste, i.e.  $2500$ ,  $5000$ ,  $7500$ ,  $10000$ ,  $12500$ ,  $15000$ ,  $17500$  e  $20000Hz$ . Apesar da intensidade da componente de frequência de  $20kHz$  ser menor que as demais, esta ainda está levemente presente no espectrograma. Como o sistema opera a uma taxa de amostragem de  $98kHz$  (banda passante de  $49kHz$ ), a frequência da senoide mais alta está compreendida na banda passante do sistema, então a razão para essa componente possuir menor intensidade

está fortemente relacionada à resposta em frequência limitada à  $20kHz$  do alto-falante utilizado.

Uma vez que a faixa audível é atendida pelo *hardware* específico de conversão *PDM2PCM* do ESP32, não há necessidade de recorrer ao módulo de conversão por *software*, já que a opção via *hardware* é mais eficiente computacionalmente.

### 4.3 Validação na faixa ultrassônica

A validação na faixa sonora se baseia no uso da conversão *PDM2PCM* via *software* a uma taxa de  $312.5ksp$ s (*clock* do microfone de  $5MHz$  e  $OSR = 16$ ).

Como o dispositivo alto-falante utilizado anteriormente não atende a faixa ultrassônica, faz-se necessário o uso de uma fonte confiável de áudio ultrassônico. Neste caso o sinal de testes utilizado tem a natureza de pulsos sonoros de  $40kHz$  (harmônicas em  $3 \cdot f = 120kHz$ ,  $5 \cdot f = 200kHz$ ,  $7 \cdot f = 280kHz$ , etc.) emitidos por um sensor de proximidade ultrassônico HC-SR04 ([Morgan \[2014\]](#)).

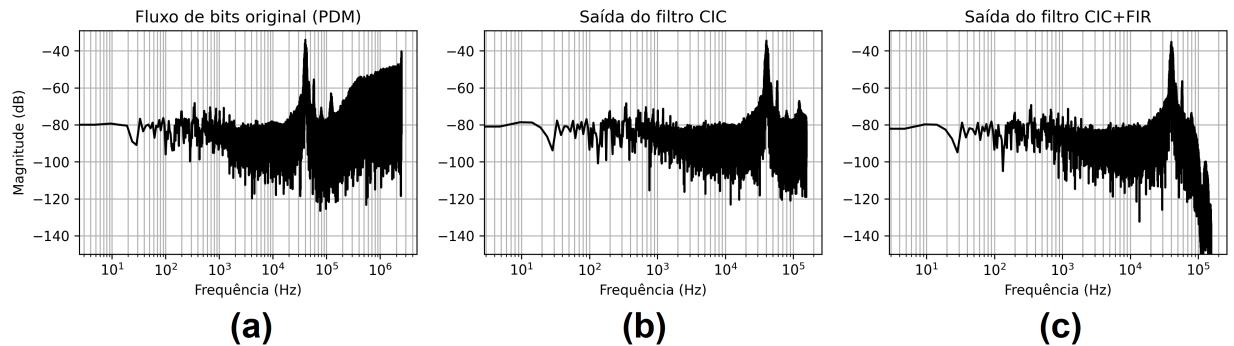
A partir dos testes feitos, avalia-se primeiro a magnitude do espectro de Fourier de um trecho de  $0.2s$  do sinal adquirido no sistema de aquisição conforme a observado na Figura 25. Nota-se na figura o espectro de magnitude para os três casos da cadeia de processamento de sinais, i.e. aquisição do sinal puro PDM em 25(a), sinal PCM na saída do filtro CIC 25(b) e sinal PCM na saída do filtro FIR compensador 25(c); e, em todos eles está presente o pico de magnitude referente à  $40kHz$ .

A presença do ruído típico em alta frequência como aquele representado na Figura 18, indica que o sinal representado na Figura 25(a) é, de fato, codificado em PDM. Imediatamente após a saída do CIC, o ruído de alta frequência é reduzido, mas a limitação da banda com maior atenuação só ocorre após a compensação pelo FIR.

Um detalhe sutil que pode ser notado é a atenuação das componentes harmônicas do sinal, que após a saída do CIC observa-se uma pequena atenuação e para o CIC mais o FIR, uma atenuação mais forte.

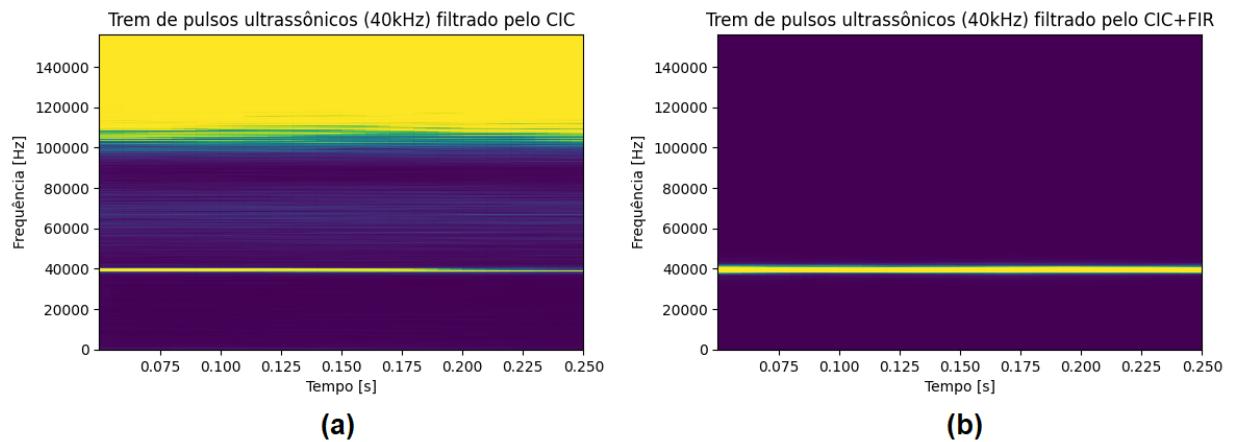
O sinal de teste também pode ser analisado por meio do spectrograma como pode ser observado na Figura 26. Este representa o sinal PCM convertido na saída do filtro CIC

Figura 25: Magnitude do espectro de Fourier para o sinal adquirido ao longo da cadeia de processamento, i.e. (a) sinal PDM adquirido, (b) sinal PCM após a filtragem do CIC e (c) sinal PCM final após compensação do FIR.



(Figura 26(a)) e na saída do filtro FIR compensador (Figura 26(b)).

Figura 26: Espectrograma do sinal de áudio pulsado de 40kHz obtido do sensor de proximidade avaliado na (a) saída do filtro CIC e na (b) saída do filtro FIR compensador.



Em ambos os estágios de filtragem a componente de  $40\text{kHz}$  é visível. A diferença que pode ser notada é a presença de ruído de alta frequência que contamina a saída do filtro CIC, enquanto que na saída do filtro FIR com frequência de corte de  $90\text{kHz}$  e ordem 64 (vide magnitude da resposta em frequência dos filtros na Figura 21) essas frequências são atenuadas.

## 5 Conclusões

Nas fases iniciais do projeto a existência da conversão PDM2PCM a partir do *hardware* do ESP32 foi essencial para validar o *firmware* programado e as conexões entre os periféricos. Porém, como essa interface está limitada a operar com áudio na faixa audível, a alternativa foi utilizar a interface PDM2PCM por *software* para processamento do áudio ultrassônico e que se mostrou eficaz de acordo com os testes realizados, permitindo que o sistema opere com uma banda de áudio ajustável na faixa ultrassônica e que atende o requisito mínimo de  $80\text{kHz}$ .

Os testes que resultaram nos espectrogramas dos sinais de áudio adquiridos foram essenciais para validar a resposta em frequência do sistema de aquisição. O sinal na faixa audível foi suficiente para validar a resposta audível, pois possui variações de frequência de acordo com o tempo. O sinal de teste na faixa ultrassônica possui apenas uma frequência e não permitiu emular o comportamento dinâmico das USVs. Porém o teste na faixa ultrassônica foi suficiente para compreender os limites da banda de áudio adquirida de acordo com os níveis de ruído presentes no espectrograma.

Também foi possível implementar eficientemente os acessórios de *software* e *hardware* projetados para facilitar a interação do sistema com o usuário, estes são: código de cores para indicar o *status* do sistema por um LED RGB; botões para controlar os estágios de operação; e uso de um arquivo texto de configuração no cartão micro SD que permite o usuário modificar parâmetros de funcionamento do sistema sem a necessidade de reprogramar o sistema.

As características mecânicas do protótipo desenvolvido garantiram um sistema final portátil e compacto, quando comparado às soluções correlatas, e possui dimensões  $44.26 \times 28.7\text{mm}^2$  e peso de 11 gramas. Apesar disso os requisitos físicos de dimensões de  $20 \times 20\text{mm}^2$  para o sistema não foram atendidos devido às limitações impostas pelos componentes utilizados e a tecnologia de projeto da PCB, porém as dimensões obtidas se mantiveram próximas das especificadas.

O estudo realizado permitiu a introdução à problemática de aquisição de áudio ultrassônico de roedores ao apresentar uma arquitetura funcional para o sistema com os seus elementos chaves bem definidos e que atendem as especificações físicas do sinal, i.e.

microfone que atende a faixa ultrassônico, microcontrolador com múltiplos periféricos que atendem as necessidades do projeto e um método viável de armazenamento dos dados.

A pesquisa também permitiu organizar as referências bibliográficas essenciais para orientação de futuros pesquisadores que necessitem se alinhar com os conceitos práticos e teóricos abordados. Os repositórios do trabalho estão documentados e versionados, permitindo monitorar a evolução do projeto ao longo do tempo. Nos repositórios encontram-se os códigos fontes do *firmware* em C, dos algoritmos de processamento de sinais em Python e C, dados de áudio obtidos que facilitam a replicação dos resultados expressos nesse trabalho, bem como diagramas, tabelas, figuras e textos desenvolvidos.

Os protótipos projetados abrem espaço para otimizações no sistema de aquisição para atender as especificações iniciais que não foram implementadas. Dessa forma, algumas propostas de trabalhos futuros podem ser citadas, como:

- Habilitar o envio dos dados de áudio por conexão sem-fio Wi-Fi ou Bluetooth;
- Inserir sensores ao sistema para gerenciamento da gravação por gatilhos biológicos, e.g. sensor de temperatura para identificar a respiração do animal que precede o início da gravação, etc.;
- Sintetizar a arquitetura do sistema de aquisição em uma FPGA (*field-programmable gate array*) e comparar o funcionamento com a alternativa desenvolvida com o microcontrolador;
- Desenvolver uma PCB mais compacta com melhor aproveitamento da área útil, principalmente focando no arranjo da bateria para diminuir o seu espaço consumido e utilizando uma placa de quatro camadas.

Também podem ser citados alguns testes futuros que seriam fundamentais para entender melhor o funcionamento dos protótipos atuais, como:

- Medir o consumo de potência do protótipo final;
- Testar a autonomia da bateria LIR2450 para alimentar o sistema;
- Traçar a característica de transferência do regulador de tensão para entender a tensão de entrada mínima que mantém a saída em valores aceitáveis para os componentes do circuito;
- Realizar aquisição de áudio de roedores em experimentos reais.

Finalmente, o desenvolvimento desse projeto permitiu a compreensão da cadeia de

processamento digital de áudio e como ela funciona ao nível de instrumentação eletrônica e de *software* embarcado, capacitando o projetista para que desenvolva habilidades de projeto de circuitos digitais, confecção de circuitos com tecnologia de componentes montados na superfície, programação de microcontroladores e processamento digital de sinais.

## Referências

- A. F. Aachen. Raspberry pi bat project, 2017. URL <http://www.bat-pi.eu/EN/index-EN.html>.
- Avisoft. Cm16/cmpa, Aug 2020. URL <https://www.avisoft.com/ultrasound-microphones/cm16-cmpa/>.
- AvisoSoftBioacustics. Cm16/cmpa, Aug 2020. URL <https://www.avisoft.com/ultrasound-microphones/cm16-cmpa/>.
- B. Baker. How delta-sigma adcs work, part 1. *Analog Applications*, 7:0000000, 2011a.
- B. Baker. How delta-sigma adcs work, part 2. *Analog Applications*, 7:0000000, 2011b.
- D. J. Barker, C. Herrera, and M. O. West. Automated detection of 50-khz ultrasonic vocalizations using template matching in xbat. *Journal of neuroscience methods*, 236: 68–75, 2014.
- R. D. Beason, R. Riesch, and J. Koricheva. Aurita: an affordable, autonomous recording device for acoustic monitoring of audible and ultrasonic frequencies. *Bioacoustics*, 28(4): 381–396, 2019.
- R. W. Bell, W. Nitschke, and T. A. Zachman. Ultra-sounds in three inbred strains of young mice. *Behavioral biology*, 7(6):805–814, 1972.
- R. d. S. d. Bessa and R. N. Romcy-Pereira. Early-life developmental profile of ultrasonic vocalizations in the valproic acid model of autism. 2019.
- R. S. d. Bessa, L. G. Galvão, T. M. A. Fialho, and R. N. Romcy-Pereira. Software development for the detection of ultrasonic vocalizations. 2019.
- A. Bioacoustics. AvisoSoft-ultrasoundgate, Jun 2022. URL <https://www.avisoft.com/ultrasoundgate/>.
- J. Boulanger-Bertolus, M. Rincón-Cortés, R. M. Sullivan, and A.-M. Mouly. Understanding pup affective state through ethologically significant ultrasonic vocalization frequency. *Scientific reports*, 7(1):1–12, 2017.

- I. Branchi, D. Santucci, and E. Alleva. Ultrasonic vocalisation emitted by infant rodents: a tool for assessment of neurobehavioural development. *Behavioural brain research*, 125 (1-2):49–56, 2001.
- Broadcom. *ASMB-MTB0-0A3A2 - PLCC-4 Tricolor Black Surface LED*. Avago Technologies, 2015.
- S. M. Brudzynski. Communication of adult rats by ultrasonic vocalization: biological, sociobiological, and neuroscience approaches. *Ilar Journal*, 50(1):43–50, 2009.
- S. M. Brudzynski. Ethotransmission: communication of emotional states through ultrasonic vocalization in rats. *Current opinion in neurobiology*, 23(3):310–317, 2013.
- S. M. Brudzynski. *Handbook of ultrasonic vocalization: A window into the emotional brain*. Academic Press, 2018.
- S. M. Brudzynski, P. Kehoe, and M. Callahan. Sonographic structure of isolation-induced ultrasonic calls of rat pups. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, 34(3):195–204, 1999.
- H. Bruhns, A. Marianovich, S. Rhein, and M. Wolff. P1.9.19 digital MEMS microphone with inter-IC sound interface for photoacoustic spectroscopy. In *Proceedings IMCS 2012*. AMA Service GmbH, Von-Münchhausen-Str. 49, 31515 Wunstorf, Germany, 2012. doi: 10.5162/imcs2012/p1.9.19. URL <https://doi.org/10.5162/imcs2012/p1.9.19>.
- CUI. Product spotlight:mems microphones, 2022. URL <https://www.cuidevices.com/product-spotlight/mems-microphones>.
- L. Davis. Microsd card pinout, microsd memory module pin out and description, 2015. URL [http://www.interfacebus.com/MicroSD\\_Card\\_Pinout.html](http://www.interfacebus.com/MicroSD_Card_Pinout.html).
- Dodotronic. Ultramic um250k, 2016. URL <https://www.dodotronic.com/product/ultramic-um250k/?v=2a47ad90f2ae>.
- Espressif. *ESP32 Series Datasheet*. Espressif Systems, 3 2022a. v3.9.
- Espressif. *ESP32-PICO-D4 Datasheet*. Espressif Systems, 3 2022b. v2.0.
- Espressif. *ESP32-WROOM-32 Datasheet*. Espressif Systems, 3 2022c. v3.3.
- L. G. d. S. Galvão. Mems2sd\_esp32, 2022. URL [https://github.com/Lwao/mems2sd\\_esp32](https://github.com/Lwao/mems2sd_esp32). Acessado em 16 de Julho de 2022.

GCT. *MEM2075-00-140-01-A*. Global Connector Technology, 9 2020. A1.

J. J. Heckman, R. Proville, G. J. Heckman, A. Azarfar, T. Celikel, and B. Englitz. High-precision spatial localization of mouse vocalizations during social interaction. *Scientific reports*, 7(1):1–16, 2017.

A. P. Hill, P. Prince, J. L. Snaddon, C. P. Doncaster, and A. Rogers. Audiomoth: A low-cost acoustic device for monitoring biodiversity and the environment. *HardwareX*, 6:e00073, 2019.

M. Hofer, J. Masmela, S. Brunelli, and H. Shair. The ontogeny of maternal potentiation of the infant rats' isolation call. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, 33(3):189–201, 1998.

Infineon. *AN558 - MEMS microphone electrical implementation*. Infineon Technologies, 1 2018. v1.0.

A. Isobe and M. Kawaguchi. Relationship between motor function and ultrasonic vocalizations induced by maternal separation in rat pups. *Journal of Veterinary Medical Science*, pages 18–0604, 2019.

W. Kester. Mt-001: Taking the mystery out of the infamous formula,"snr= 6.02 n + 1.76 db,"and why you should care. *REV. 0*, pages 10–03, 2005.

Keystone. *Battery holder for CR2450*. Avago Keystone Electronics Corp., 11 2010.

Kingston. A guide to speed classes for sd and microsd cards - kingston technology. URL <https://www.kingston.com/en/blog/personal-storage/memory-card-speed-classes>.

Knowles. *SPH0641LU4H-1 - Digital Zero-Height SiSonic™Microphone With Multi-Mode And Ultrasonic Support*. Knowles, 5 2015. B.

J. Lewis. *Technical Article MS-2275 - Common Inter-IC Digital Interfaces for Audio Data Transfer*. Analog Devices, 2012.

J. Lewis. Analog and digital mems microphone design considerations. *Technical Article MS-2472*. Analog Devices, 2013.

M. Loibl, S. Walser, J. Klugbauer, G. Feiertag, and C. Siegel. 4.3.2 - measurement of digital MEMS microphones. In *Tagungsband*. AMA Service GmbH, Von-Münchhausen-Str.

- 49, 31515 Wunstorf, Germany, 2016. doi: 10.5162/sensoren2016/4.3.2. URL <https://doi.org/10.5162/sensoren2016/4.3.2>.
- R. Lyons. A beginner's guide to cascaded integrator-comb (cic) filters, 2020.
- Microchip. *TC1262 - 500mA Fixed Output CMOS LDO*. Microchip, 2012.
- E. J. Morgan. Hc-sr04 ultrasonic sensor. *Nov*, 2014.
- NXP. *UM11732 - I2S bus specification*. NXP, 2 2022. v3.0.
- E. E. Okon. The effect of environmental temperature on the production of ultrasounds by isolated non-handled albino mouse pups. *Journal of Zoology*, 162(1):71–83, 1970a.
- E. E. Okon. The ultrasonic responses of albino mouse pups to tactile stimuli. *Journal of Zoology*, 162(4):485–492, 1970b.
- OpenAcousticDevices. Purchase: Open acoustic devices, 2022. URL <https://www.openacousticdevices.info/purchase>.
- C. V. Portfors. Types and functions of ultrasonic vocalizations in laboratory rats and mice. *Journal of the American Association for Laboratory Animal Science*, 46(1):28–34, 2007.
- RaspberryPi. Buy a raspberry pi 4 model b, 2022. URL <https://www.raspberrypi.com/products/raspberry-pi-4-model-b/>.
- S. Roy, N. Watkins, and D. Heck. Comprehensive analysis of ultrasonic vocalizations in a mouse model of fragile x syndrome reveals limited, call type specific deficits. *PLoS one*, 7 (9):e44816, 2012.
- Y. Saito, R. O. Tachibana, and K. Okanoya. Acoustical cues for perception of emotional vocalizations in rats. *Scientific reports*, 9(1):1–9, 2019.
- G. Sales. *Ultrasonic communication by animals*. Springer Science & Business Media, 2012.
- E. Sarkar. *Artificial Neural Networks: Kohonen Self-Organising Maps*. PhD thesis, University of Liverpool, 2018.
- M. A. Shah, I. A. Shah, D.-G. Lee, and S. Hur. Design approaches of mems microphones for enhanced performance. *Journal of sensors*, 2019, 2019.
- S. Specification. Sd specifications part 1 physical layer simplified specification. 2006.

- P. Stream. *PowerStream Li-ion Button Cell Lir2450 Data Sheet*. Power Stream, 2022.
- R. C. Whytock and J. Christie. Solo: an open source, customizable and inexpensive audio recorder for bioacoustic research. *Methods in Ecology and Evolution*, 8(3):308–312, 2017.
- M. Wöhr and R. K. Schwarting. Affective communication in rodents: ultrasonic vocalizations as a tool for research on emotion and motivation. *Cell and tissue research*, 354(1):81–97, 2013.
- J. M. Wright, J. C. Gourdon, and P. B. Clarke. Identification of multiple call categories within the rich repertoire of adult rat 50-khz ultrasonic vocalizations: effects of amphetamine and social context. *Psychopharmacology*, 211(1):1–13, 2010.
- A. S. Zaytseva, I. A. Volodin, O. G. Ilchenko, and E. V. Volodina. Ultrasonic vocalization of pup and adult fat-tailed gerbils (*pachyuromys duprasi*). *PloS one*, 14(7):e0219749, 2019.

## A Esquemáticos dos circuitos

Figura 27: Esquemático dos circuitos do protótipo 0.

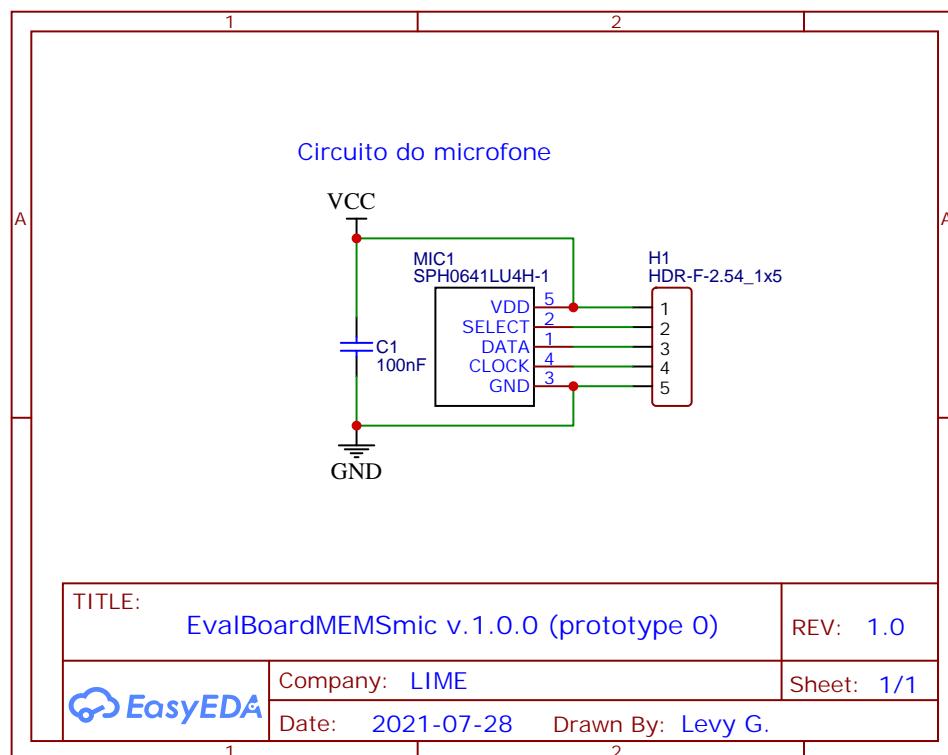


Figura 28: Esquemático dos circuitos do protótipo 1.

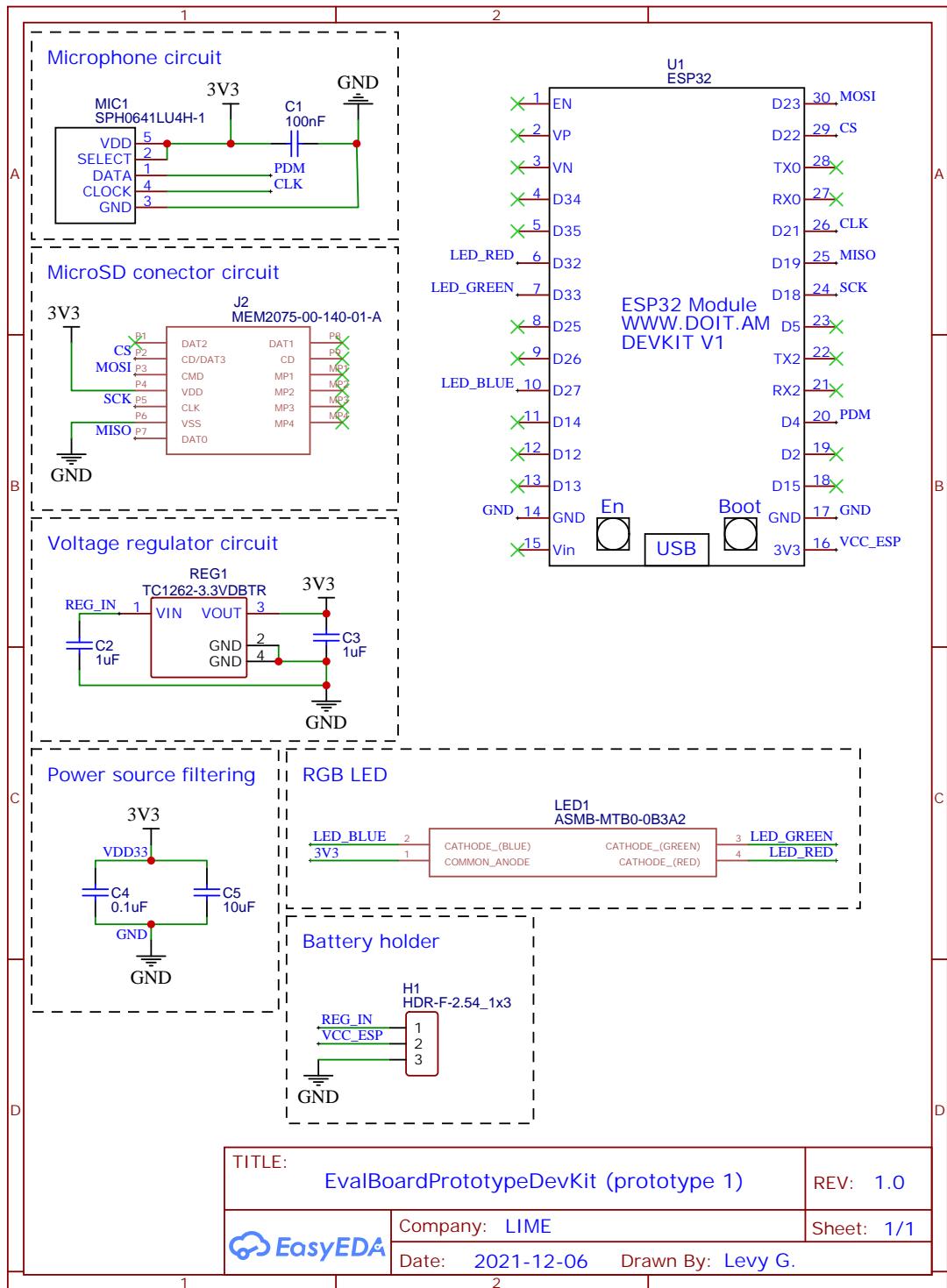


Figura 29: Esquemático dos circuitos do protótipo 2.

