

EchoLogger: sistema embarcado para cálculo do índice de riqueza acústica no monitoramento ambiental não invasivo

Pedro Lucas Bezerra Mendes¹, Rodolfo Stoll², Juan G. Colonna¹

¹Instituto de Computação (IComp) - Universidade Federal do Amazonas (UFAM)
Manaus, Amazonas, Brasil

²Facultad de Ingeniería - Universidad Nacional de Río Cuarto (UNRC)
Río Cuarto, Córdoba, Argentina.

{pedro.mendes, juancolonna}@icomp.ufam.edu.br
rstoll@ing.unrc.edu.ar

Abstract. This article presents the EchoLogger, a low-cost embedded system for environmental bioacoustic monitoring. Based on the ESP32, the device captures and processes audio in real time, calculating the Acoustic Richness Index (AR) to estimate sound diversity. The methodology includes a comparison of EchoLogger's calculations with the scikit-maad library and a field test to assess the feasibility of embedded processing. The results indicate high accuracy in the median and entropy calculations, with an MAE of 0.0429 for the median and 0.0776 for the entropy, demonstrating that real-time processing does not compromise data quality. The study confirms the efficiency and accessibility of EchoLogger for environmental monitoring, highlighting its potential in biodiversity and conservation.

Resumo. Este artigo apresenta o EchoLogger, um sistema embarcado de baixo custo para monitoramento bioacústico ambiental. Baseado no ESP32, o dispositivo captura e processa áudio em tempo real, calculando o Índice de Riqueza Acústica (AR) para estimar a diversidade sonora. A metodologia inclui a comparação dos cálculos do EchoLogger com a biblioteca scikit-maad e um teste de campo para avaliar a viabilidade do processamento embarcado. Os resultados indicam alta precisão nos cálculos de mediana e entropia, com um MAE de 0,0429 para a mediana e 0,0776 para a entropia, demonstrando que o processamento em tempo real não compromete a qualidade dos dados. O estudo confirma a eficiência e acessibilidade do EchoLogger para monitoramento ambiental, destacando seu potencial em biodiversidade e conservação.

1. Introdução

A biodiversidade, essencial para os ecossistemas terrestres, enfrenta crescentes desafios devido à pressões antropogênicas, como desmatamento e mudanças climáticas, que comprometem o equilíbrio ecológico. O monitoramento bioacústico surge como uma estratégia não invasiva para avaliar a saúde ambiental, registrando sons da fauna e do ambiente para inferir padrões de biodiversidade. No entanto, os métodos tradicionais exigem gravação contínua e processamento posterior em laboratório, gerando grandes volumes de dados que dificultam a escalabilidade e a análise manual, tornando essencial a busca por soluções mais eficientes [Colonna et al. 2020].

Característica	Song Meter	AudioMoth	AURITA	EchoLogger
Baixo Custo	✗	✓	✓	✓
Gravação de áudio	✓	✓	✓	✓
Variáveis ambientais	✗	✗	✗	✓
Código aberto	✗	✓	✓	✓
Índice acústico em Tempo Real	✗	✗	✗	✓

Tabela 1. Comparação entre diferentes dispositivos de gravação de áudio

Os índices acústicos constituem ferramentas essenciais para o monitoramento da biodiversidade, permitindo a estimativa automatizada da diversidade biológica a partir de paisagens sonoras. A integração entre a gravação de sons ambientais e o cálculo desses índices possibilita a coleta contínua de dados ecológicos em longo prazo [Bradfer-Lawrence et al. 2024]. Tradicionalmente, os registros sonoros são obtidos em campo com alta qualidade e posteriormente processados em laboratório, o que gera grandes volumes de dados brutos. Nesse contexto, a implementação do cálculo dos índices diretamente em dispositivos embarcados representa um avanço significativo, ao tornar a análise mais eficiente e facilitar a triagem automática de segmentos sonoros mais relevantes e reduzir a necessidade de armazenamento extensivo.

Dispositivos comerciais, como o nó sensor Song Meter (Wildlife Acoustics), são amplamente usados no monitoramento bioacústico [Mennill 2024], mas não permitem reprogramação ou integração com outras variáveis ambientais, como temperatura e umidade. Alternativas abertas e acessíveis, como os sensores AURITA [Beason et al. 2019] e AudioMoth [Hill et al. 2018], vêm ganhando espaço na comunidade científica. No entanto, ainda possuem limitações, como a falta de módulos sincronizados para registrar variáveis ambientais e a ausência de integração para o cálculo de índices acústicos.

A Tabela 1 apresenta uma comparação entre os dispositivos de gravação de áudio. Diante dessas limitações dos sensores comparados e considerando a necessidade de soluções de baixo custo, sistemas abertos e flexibilidade de dados coletados, o presente estudo fará uso de um dispositivo de hardware e software abertos, denominado EchoLogger¹, desenvolvido na Universidade Federal do Amazonas - UFAM. O EchoLogger, o qual utiliza o microcontrolador ESP32, foi concebido para atuar como um sensor inteligente com capacidade de processamento embarcado, possibilitando a realização de cálculos *in situ*, executar algoritmos de reconhecimento de espécies baseados em Aprendizado de Máquina e a obtenção de índices acústicos em tempo real. Além disso, o nó sensor foi otimizado para maximizar o armazenamento, reduzir o consumo de energia e simplificar a programação, uma vez que utiliza MicroPython.

Nesse contexto, o presente artigo tem como objetivo desenvolver e validar um *firmware* embarcado para o EchoLogger, capaz de calcular em tempo real o índice de riqueza acústica (AR). A proposta visa reduzir a dependência de processamento posterior, incorporando esse índice ao conjunto de variáveis ambientais que o nó sensor já registra. O AR é tratado como uma “variável de software”, pois difere de variáveis como umidade, pressão, luminosidade, temperatura, qualidade do ar e até mesmo do som bruto

¹<https://github.com/juancolonna/EchoLogger>

gravado — todas obtidas diretamente por sensores físicos, sem processamento adicional. Para avaliar a precisão numérica do índice calculado no dispositivo, utiliza-se como referência a biblioteca scikit-maad² [Bogdanov et al. 2013], comparando-se os resultados com os de um algoritmo não embarcado. Essa comparação permite verificar a fidelidade da implementação em MicroPython. Assim, o estudo propõe uma solução de baixo custo e alta escalabilidade voltada à conservação ambiental, tendo na ecoacústica sua principal aplicação.

2. Trabalhos Relacionados

Os dispositivos de monitoramento acústico passivo (*Passive Acoustic Monitoring*, PAM), como o AudioMoth [Hill et al. 2018] e o Aurita [Beason et al. 2019] têm impulsionado soluções não intrusivas para o monitoramento bioacústico. O AudioMoth destaca-se pela facilidade de implantação e personalização, enquanto o Aurita, baseado em Raspberry Pi, captura uma ampla faixa de frequências, incluindo ultrassons. No entanto, ambos não oferecem cálculo de índices acústicos em tempo real e apresentam consumo elevado de energia.

O BioAcoustic Index Tool (BAIT), desenvolvido com microcontrolador ARM Cortex-M4F, oferece cálculo de índices acústicos em tempo real, como o Índice de Diversidade Acústica (ADI) e o Índice de Complexidade Acústica (ACI), além de registrar dados ambientais. Apesar de seu bom desempenho, a dependência de hardware específico limita sua adoção em outros sistemas [Kadish and Stoy 2022].

O Soundscape Explorer Terrestrial (SET) [Farina 2018], desenvolvido pela Lu-niletronik³, utiliza computação embarcada para calcular o Índice de Complexidade Acústica (ACI) e registrar dados ambientais simultaneamente. O dispositivo possui dois microfones especializados: um para frequências audíveis e outro para ultrassônicas, permitindo gravações contínuas por até duas semanas em configuração padrão. No entanto, e apesar de ser um sistema fechado, este não está disponível para adquisição.

Embora o BAIT e o SET incorporem o cálculo de índices acústicos para caracterização de paisagens sonoras, nenhum deles está disponível para adoção por outros pesquisadores. Diante dessa limitação, optamos pela plataforma EchoLogger como alternativa.

3. Fundamentos Teóricos

3.1. Índice de Riqueza Acústica

O índice AR combina a entropia temporal H_t e a mediana M da amplitude $A(t)$ para inferir a diversidade acústica em ambientes naturais [Depraetere et al. 2012]. O estudo de [Towsey et al. 2014] mostra que este índice possui correlação positiva com a riqueza de espécies de aves em habitats temperados. Além disso, [Buxton et al. 2018] destaca que o AR contribui positivamente com a capacidade preditiva de modelos de Aprendizado de Máquina na estimativa da diversidade animal em gravações ambientais.

O AR foi desenvolvido para quantificar a riqueza acústica de um ambiente natural, considerando tanto a distribuição temporal de energia sonora quanto a intensidade do

²<https://scikit-maad.github.io/>

³http://www.iinsteco.org/soundscape_explorer/

sinal, sendo particularmente útil em estudos de bioacústica para estimar a presença das espécies de aves. Este índice é definido de acordo com:

$$AR = \frac{\text{rank}(H(t)) \times \text{rank}(M)}{n^2}, \quad (1)$$

onde $H(t)$ é a entropia temporal do sinal acústico, refletindo a variabilidade da energia sonora ao longo do tempo, M é a mediana da amplitude do envelope do sinal, representando a intensidade do som captado, a função $\text{rank}()$ atribui *rankings* (ordenações) aos elementos de um vetor em uma escala normalizada, e n é o número total de segmentos analisados.

O índice de AR (equação 1) varia de 0 a 1, onde valores altos indicam maior diversidade sonora e variação na energia acústica, enquanto valores baixos sugerem um ambiente acusticamente homogêneo. Sua implementação em dispositivos embarcados enfrenta três desafios computacionais: (i) Entropia: requer operações complexas, como logaritmos, que demandam alto processamento, (ii) Mediana: exige ordenação dos pontos do sinal, um processo iterativo e lento, consumindo memória adicional, (iii) Envelope: depende do valor máximo absoluto do sinal, necessitando do preenchimento completo do buffer do microfone, o que pode impactar a aquisição contínua dos dados.

3.2. Métricas de avaliação

O Erro Médio Absoluto (MAE) foi utilizado para avaliar a precisão numérica dos índices acústicos implementados no sistema embarcado em relação aos valores de referência gerados pela biblioteca scikit-maad. Essa métrica foi escolhida por expressar a média das diferenças absolutas entre os valores estimados e os de referência na mesma unidade de medida dos índices, o que facilita a interpretação do erro. Assim, o MAE permite quantificar diretamente o grau de aproximação numérica entre os índices calculados pelo *firmware* e aqueles fornecidos pela biblioteca de referência. O MAE é definido conforme:

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i|, \quad (2)$$

onde Y_i são os valores de referência gerados pela biblioteca scikit-maad, \hat{Y}_i são os valores gerados pelo *firmware*, e n é o número total de amostras. Valores de MAE próximos de zero indicam boa aproximação numérica. Por outro lado, valores elevados de MAE apontariam discrepâncias, indicando que o índice implementado no *firmware* não seria uma aproximação boa.

4. Materiais

4.1. Dispositivo EchoLogger

O EchoLogger, Figura 1, desenvolvido pelo Instituto de Computação (IComp) da Universidade Federal do Amazonas (UFAM), surge como uma solução inovadora para suprir a necessidade de ter uma plataforma nacional para monitoramento das florestas. Este dispositivo integra tecnologias de comunicação sem fio (bluetooth low-energy e wi-fi), sensores ambientais de variáveis abióticas e captação de áudio digital de alta resolução, visando otimizar o monitoramento bioacústico em ambientes naturais.

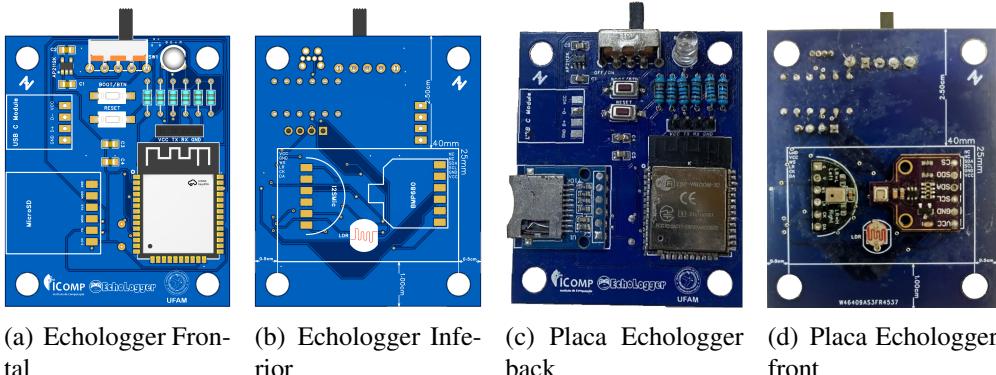


Figura 1. EchoLogger device.

O EchoLogger, baseado no ESP32-WROOM-32S, é ideal para operações prolongadas em campo. Equipado com um sensor BME680 e um LDR, monitora variáveis ambientais como temperatura, umidade, pressão, qualidade do ar e luminosidade, permitindo correlações com dados bioacústicos. Seu microfone omnidirecional do tipo MEMS capta áudio digital com alta fidelidade e se conecta diretamente ao barramento I2S, sem necessidade de conversores externos, garantindo registros precisos. A integração desses dados possibilita a análise de padrões entre condições ambientais e comportamentos acústicos de espécies.

4.2. Scikit-maad

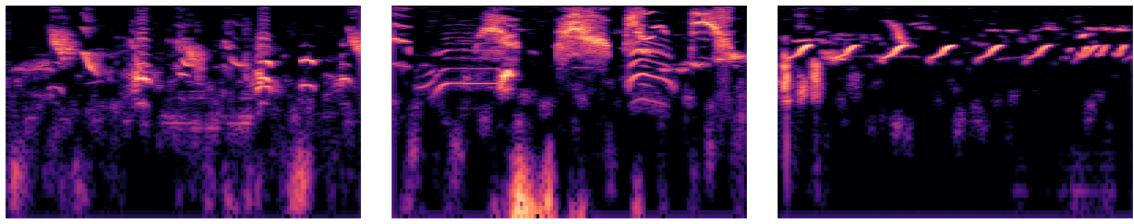
O scikit-maad é uma biblioteca Python de código aberto desenvolvida para simplificar a análise de dados de áudio ambiental em estudos de monitoramento ecológico e de biodiversidade [Ulloa et al. 2021]. Com foco em aplicações ecoacústicas, a biblioteca oferece ferramentas para pré-processamento, extração de características e caracterização de paisagens sonoras, permitindo que pesquisadores quantifiquem a diversidade acústica e os impactos antrópicos em ecossistemas. Entre suas funcionalidades principais estão a segmentação de áudio, análise espectral e cálculo de índices acústicos.

4.3. Dynamic Time Warping

O Dynamic Time Warping (DTW) mede a similaridade entre séries temporais, mesmo com variações de velocidade ou pequenas distorções [Berndt and Clifford 1994]. Ele ajusta dinamicamente a escala temporal dos sinais para minimizar diferenças, identificando padrões em momentos distintos. O cálculo usa uma matriz de custo, onde cada elemento representa a distância entre pontos das séries, e um caminho de deformação é determinado para minimizar essa soma. O DTW é amplamente aplicado em reconhecimento de fala, análise biológica e processamento de sinais, sendo útil para séries não perfeitamente alinhadas.

4.4. BirdCLEF

BirdClef-2022 [Kaggle et al. 2022] é um dataset desenvolvido para a identificação de vocalizações de aves em monitoramento acústico passivo. Coletado a partir de diversas fontes, incluindo o Xeno-canto e gravações de campo, o conjunto reúne milhares de amostras de áudio de diferentes regiões do mundo. As gravações apresentam variadas



(a) Espécies Houspa, Cangoo, Pibgre, Commyn e Gadwal (b) Espécies Laugul, Wessan, Hawcre e Bcnher (c) Espécies Warwhe e Sposan

Figura 2. Espectrogramas extraídos do dataset BirdClef-2022, os quais foram utilizados na primeira abordagem experimental.

taxas de amostragem e durações, refletindo a diversidade das condições de captura. Cada arquivo de áudio é acompanhado de metadados detalhados, incluindo local, data, autor da gravação e espécies identificadas. A Figura 2 apresenta espectrogramas gerados a partir de amostras do dataset BirdClef-2022, ilustrando a diversidade de padrões acústicos encontrados nas gravações.

5. Método proposto

O *firmware* do cálculo do AR foi implementado diretamente no buffer do microfone, onde o sinal é armazenado. Ao atingir a capacidade máxima, uma interrupção é acionada e o envelope do sinal é calculado e armazenado em um vetor. Após a aquisição, a lista é usada para calcular a entropia e a mediana, possibilitando a extração de informações em tempo real.

Este estudo adota duas abordagens para avaliar a eficiência do cálculo de métricas acústicas no EchoLogger. A primeira analisa a precisão e o tempo de processamento dos cálculos de entropia, mediana e envelope, usando a scikit-maad como referência e focando no tempo de execução do envelope, essencial para captura em tempo real. A segunda abordagem investiga o impacto do cálculo do AR na captura de áudio em tempo real por meio de dois experimentos: (1) verificando se há interferência na gravação e (2) avaliando a precisão do AR em um ambiente urbano com diferentes fontes sonoras.

5.1. Primeira abordagem - Avaliação de precisão e tempo de processamento

Nesta abordagem, foram analisados três cálculos fundamentais (entropia, mediana e envelope) com referência na biblioteca scikit-maad e implementação na biblioteca ulab⁴. A ulab possui aproximações numéricas e operações complexas, justificando a utilização.

Foram usados seis áudios do BirdCLEF (22,05 kHz, 16 bits), selecionados conforme o número de vocalizações de pássaros (0 a 5 indivíduos). Três espectrogramas destas amostras estão ilustrados na Figura 2. Os áudios foram processados no EchoLogger em blocos de 20.000 amostras, ignorando os primeiros 44 bytes do cabeçalho. A normalização foi feita para [-1,1], e o envelope foi calculado pelo valor máximo absoluto de cada bloco. A entropia foi extraída a partir da normalização do envelope em uma função de massa de probabilidade e aplicada a entropia de Shannon. Já a mediana foi calculada sobre os valores do envelope.

⁴<https://micropython-ulab.readthedocs.io/en/latest/>



Figura 3. Experimentos de verificação e coleta áudios.

		Valores							MAE
Mediana	EchoLogger	0,0004	0,5285	0,1524	0,5623	0,4342	0,6649		
	Scikit-maad	0,0006	0,6109	0,2879	0,5732	0,4612	0,6663		
	Erro Absoluto	0,0002	0,0824	0,1355	0,0109	0,0270	0,0014	0,0429	
Entropy	EchoLogger	0,8107	0,7088	0,7532	0,9481	0,8858	0,9746		
	Scikit-maad	0,8644	0,9939	0,7151	0,9727	0,8366	0,9898		
	Erro Absoluto	0,0537	0,2851	0,0381	0,0246	0,0492	0,0152	0,0776	

Tabela 2. Comparação de valores da mediana e da entropia usando o *firmware* embarcado e a biblioteca de referência. Cada coluna de valores corresponde a um áudio de exemplo da base de dados BirdCLEF.

O tempo de processamento foi medido em microssegundos via funções do MicroPython, com 100 repetições para média e desvio padrão, garantindo uma análise detalhada da viabilidade para processamento em tempo real.

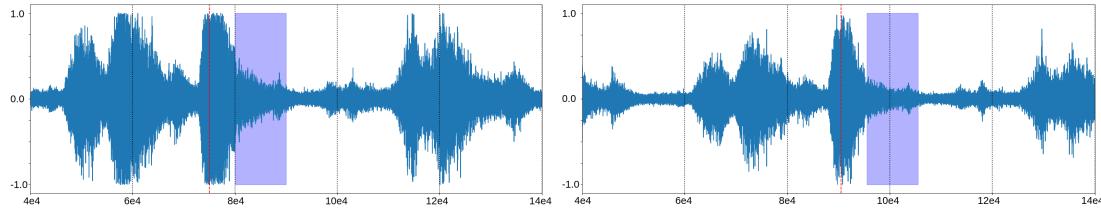
5.2. Segunda abordagem - Experimentos em campo

Os experimentos avaliaram o impacto do cálculo do AR na captura de áudio. No primeiro, foi examinada a interferência na gravação das vocalizações de saúim-de-coleira com dois EchoLoggers: um calculando o AR em tempo real e armazenando os áudios, outro apenas gravando, conforme ilustrado na Figura 3(a). Ambos operaram simultaneamente com 22,05 kHz, 16 bits e um buffer de 20.000 amostras. No segundo experimento, conduzido em ambiente urbano Figura 3(b), o dispositivo captou biofonia e antropofonia por 2h30m, seguindo um protocolo de 1 minuto de gravação e 2 minutos de sono Figura 3(c).

6. Experimentos e resultados

A Tabela 2 compara os valores calculados pelo EchoLogger e pela biblioteca scikit-maad para mediana e entropia. Para cada métrica, são exibidos os valores obtidos ao longo de diferentes gravações, além do Erro Absoluto e do MAE, que quantifica a discrepância entre as abordagens.

Os valores da mediana em ambos os sistemas foram similares, com pequenas variações. No EchoLogger, variaram de 0,0004 a 0,6649, enquanto na scikit-maad ficaram entre 0,0006 e 0,5732, resultando em um MAE de 0,0429, indicando boa precisão.



(a) Sinal do EchoLogger que realizou o processamento em tempo real. (b) Sinal do EchoLogger que apenas armazenou os dados brutos.

Figura 4. Análise de similaridade entre sinais de áudio utilizando DTW

A comparação da entropia também mostrou boa concordância. Os valores no EchoLogger variaram entre 0,7088 e 0,9746, e na scikit-maad, entre 0,7151 e 0,9898, com um MAE de 0,0776, reforçando a exatidão do EchoLogger.

Quanto ao tempo de execução, o cálculo do envelope foi o mais demorado ($13363,30 \mu\text{s} \pm 2409,81 \mu\text{s}$), seguido pela entropia ($2247,74 \mu\text{s} \pm 9,08 \mu\text{s}$) e pela mediana ($326,57 \mu\text{s} \pm 3,99 \mu\text{s}$). O envelope também apresentou maior variação temporal, sugerindo maior impacto no processamento.

A Figura 4 apresenta dois sinais de áudio sobrepostos, segmentados em blocos de 20.000 amostras. As linhas vermelhas indicam a sincronização temporal, e a região destacada em roxo foi analisada via DTW. O objetivo dessa análise é avaliar o impacto do cálculo do índice AR nos sinais de áudio. Para isso, comparamos o início do bloco do Sinal 4(a) com o sinal correspondente que é o sinal 4(b), com normalização no intervalo -1 a 1 e métrica de distância euclidiana. O cálculo resultou em uma distância DTW de 2346,62, indicando o grau de dissimilaridade entre os sinais. Essa diferença pode ser atribuída a três fatores principais: (i) latência introduzida pelo processamento em tempo real do AR, que pode ter causado pequenos deslocamentos temporais entre os sinais; (ii) variações na captura devido a ruídos externos dos dispositivos; (iii) diferenças na amplitude dos sinais, que podem influenciar a métrica de distância e impactar a comparação entre as formas de onda.

No segundo experimento, foram identificados diferentes tipos de áudio no ambiente urbano, incluindo vocalizações do bem-te-vi (Figura 5(a)), ruídos musicais (Figura 5(b)) e som de moto (Figura 5(c)). Esses registros demonstram que a gravação manteve sua integridade, permitindo uma identificação precisa dos diferentes eventos sonoros. Ao realizar a escuta dos áudios coletados, foi possível perceber que o cálculo do AR não impactou a qualidade das gravações. A reprodução dos áudios manteve-se fiel às características sonoras do ambiente, sem a introdução de ruídos ou distorções perceptíveis. A presença de vocalização de pássaros, permaneceu nítida e bem definida, garantindo a preservação das características acústicas essenciais para a análise e classificação das vocalizações.

7. Conclusões

Os resultados demonstram a alta precisão do EchoLogger no cálculo do índice AR em tempo real, comparado à biblioteca scikit-maad, com pequenas variações confirmadas pelo MAE. Isso evidencia a confiabilidade do dispositivo para análise acústica em tempo

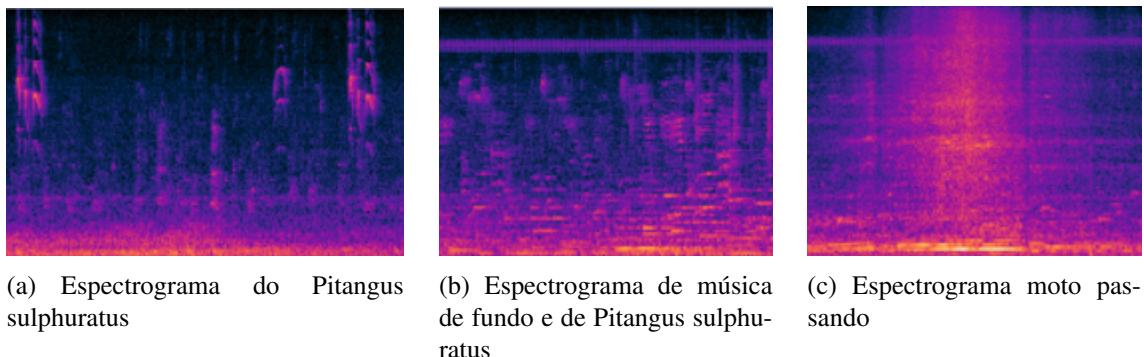


Figura 5. Espectrograma de espécie de pássaros e Influências antropogênica

real. Além disso, os tempos de execução indicam um bom equilíbrio entre desempenho e eficiência, reforçando sua adequação para aplicações de processamento *in situ*.

A análise dos sinais de áudio em ambiente urbano revelou vocalizações de aves e ruídos antropogênicos, destacando a eficiência do EchoLogger na detecção acústica. A distância DTW de 2346,62 indica uma diferença mensurável nos sinais, possivelmente devido à latência, ruídos externos e variações de amplitude. No entanto, a sincronização temporal e a preservação das características acústicas sugerem um impacto reduzido do cálculo do índice AR, demonstrando a viabilidade do método para análise bioacústica.

Embora os resultados sejam promissores, há espaço para otimizações, como o desenvolvimento de algoritmos mais eficientes para o cálculo do envelope e a implementação de outros índices bioacústicos, como o Bioacoustic Index (BI), o Índice de Complexidade Acústica (ACI) e o Índice de Diversidade Acústica (ADI), para expandir as capacidades de análise do sistema.

Agradecimentos

Este trabalho contou com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES-PROEX), Código de Financiamento 001, e da Fundação de Amparo à Pesquisa do Estado do Amazonas - FAPEAM, por meio dos projetos PDPG/CAPES e POSGRAD 2024. Agradecemos também à FAPEAM pelo suporte por meio do projeto intitulado “Diferentes Abordagens Computacionais para Monitoramento Ecoacústico Autônomo da Região Amazônica”, vinculado à CHAMADA PÚBLICA Nº 013/2022 - PRODUTIVIDADE CT&I. Agradecemos ainda ao Instituto Eldorado, unidade Manaus, pelo apoio concedido. Agradecemos à OpenAI pelo apoio por meio da parceria estabelecida com seu representante, Nicolas Robinson Andrade, como parte do projeto amazonLLM.

Referências

- [Beason et al. 2019] Beason, R. D., Riesch, R., and Koricheva, J. (2019). Aurita: an affordable, autonomous recording device for acoustic monitoring of audible and ultrasonic frequencies. *Bioacoustics*, 28(4):381–396.
- [Berndt and Clifford 1994] Berndt, D. J. and Clifford, J. (1994). Using dynamic time warping to find patterns in time series. In *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining*, AAAIWS’94, page 359–370. AAAI Press.

- [Bogdanov et al. 2013] Bogdanov, D., Wack, N., Gómez, E., Gulati, S., Herrera, P., Mayor, O., Roma, G., Salamon, J., Zapata, J., and Serra, X. (2013). Essentia: an open-source library for sound and music analysis. In *Proceedings of the 21st ACM international conference on Multimedia*, pages 855–858.
- [Bradfer-Lawrence et al. 2024] Bradfer-Lawrence, T., Duthie, B., Abrahams, C., Adam, M., Barnett, R. J., Beeston, A., Darby, J., Dell, B., Gardner, N., Gasc, A., Heath, B., Howells, N., Janson, M., Kyoseva, M.-V., Luypaert, T., Metcalf, O. C., Nousek-McGregor, A. E., Poznansky, F., Ross, S. R. P.-J., Sethi, S., Smyth, S., Waddell, E., and Froidevaux, J. S. P. (2024). The acoustic index user’s guide: A practical manual for defining, generating and understanding current and future acoustic indices. *Methods in Ecology and Evolution*.
- [Buxton et al. 2018] Buxton, R. T., McKenna, M. F., Clapp, M., Meyer, E., Stabenau, E., Angeloni, L. M., Crooks, K., and Wittemyer, G. (2018). Efficacy of extracting indices from large-scale acoustic recordings to monitor biodiversity. *Conservation Biology*, 32(5):1174–1184.
- [Colonna et al. 2020] Colonna, J. G., Carvalho, J. R. H., and Rosso, O. A. (2020). Estimating ecoacoustic activity in the amazon rainforest through information theory quantifiers. *PLOS ONE*, 15(7).
- [Depraetere et al. 2012] Depraetere, M., Pavoine, S., Jiguet, F., Gasc, A., Duvail, S., and Sueur, J. (2012). Monitoring animal diversity using acoustic indices: Implementation in a temperate woodland. *Ecological Indicators*, 13(1):46–54.
- [Farina 2018] Farina, A. (2018). Perspectives in ecoacoustics: A contribution to defining a discipline. *Journal of Ecoacoustics*, 2:TRZD5I.
- [Hill et al. 2018] Hill, A. P., Prince, P., Piña Covarrubias, E., Doncaster, C. P., Snaddon, J. L., and Rogers, A. (2018). Audiometer: Evaluation of a smart open acoustic device for monitoring biodiversity and the environment. *Methods in Ecology and Evolution*, 9(5):1199–1211.
- [Kadish and Stoy 2022] Kadish, D. and Stoy, K. (2022). Bioacoustic index tool: long-term biodiversity monitoring using on-sensor acoustic index calculations. *Bioacoustics*, 31(3):348–378.
- [Kaggle et al. 2022] Kaggle, Xeno-canto, and Contributors, V. (2022). Birdclef 2022 - bird species recognition in soundscape recordings. Kaggle Competition Dataset.
- [Mennill 2024] Mennill, D. J. (2024). Field tests of small autonomous recording units: an evaluation of in-person versus automated point counts and a comparison of recording quality. *Bioacoustics*, 33(2):157–177.
- [Towsey et al. 2014] Towsey, M., Wimmer, J., Williamson, I., and Roe, P. (2014). The use of acoustic indices to determine avian species richness in audio-recordings of the environment. *Ecological Informatics*, 21:110–119. Ecological Acoustics.
- [Ulloa et al. 2021] Ulloa, J. S., Haupert, S., Latorre, J. F., Aubin, T., and Sueur, J. (2021). scikit-maad: An open-source and modular toolbox for quantitative soundscape analysis in python. *Methods in Ecology and Evolution*, 12(12):2334–2340.