

RELATÓRIO FINAL DE ATIVIDADES DE INICIAÇÃO CIENTÍFICA

Localização Espacial de Fontes Sonoras

vinculado ao projeto

**Técnicas de redução de ruído, dereverberação e preservação do cenário
acústico para aparelhos auditivos biauriculares**

Yuri Gabriel dos Reis Souza

Voluntário

Engenharia de Computação

Data de ingresso no programa: 11/2024

Prof. Dr. Johnny Werner

Área do Conhecimento: Processamento de Sinais Biológico

CAMPUS Pato Branco, 2025

YURI GABRIEL DOS REIS SOUZA

PROF. DR. JOHNNY WERNER

LOCALIZAÇÃO ESPACIAL DE FONTES SONORAS

Relatório de Pesquisa do Programa
de Iniciação Científica ou Programa
de Iniciação Tecnológica da
Universidade Tecnológica Federal do
Paraná.

PATO BRANCO, 2025

SUMÁRIO

RESUMO	1
INTRODUÇÃO	2
METODOLOGIA	3
RESULTADOS E DISCUSSÕES	4
CONCLUSÕES	5
REFERÊNCIAS	6
ANEXOS	7

1. RESUMO

Este trabalho apresenta o desenvolvimento de um sistema embarcado de baixo custo para localização espacial de fontes sonoras, voltado a aplicações em robótica, automação e tecnologias assistivas. A proposta busca responder à crescente demanda por sistemas autônomos mais inteligentes e acessíveis, capazes de identificar a direção de sons no ambiente. O projeto foi estruturado em três etapas: concepção e definição da técnica, implementação em hardware e validação experimental. Para a estimativa da direção, adotou-se a técnica de Diferença de Tempo de Chegada (TDOA), com cálculo do atraso por meio do algoritmo de Correlação Cruzada Generalizada com Transformada de Fase (GCC-PHAT), reconhecido por sua robustez frente a ruído e reverberação. A implementação foi realizada em um microcontrolador ESP32, associado inicialmente a dois microfones digitais INMP441, permitindo captura e processamento em tempo real. Os testes em ambiente fechado e de baixa reverberação demonstraram que o protótipo foi capaz de identificar corretamente a direção da fonte sonora, apresentando resultados consistentes com trabalhos da literatura em sistemas leves. Os resultados comprovam a viabilidade da solução, abrindo caminho para sua expansão a cenários mais complexos, como *arrays* com mais microfones e aplicações práticas.

Palavras-chave: localização-sonora, sistemas-embarcados, processamento-de-sinais, TDOA.

2. INTRODUÇÃO

A localização de fontes sonoras é um tema de grande relevância em áreas como robótica, automação e tecnologias assistivas. A capacidade de estimar a direção de chegada de um som é fundamental para tornar sistemas autônomos mais inteligentes e inclusivos, com aplicações que abrangem desde a navegação de robôs até dispositivos de apoio auditivo.

Apesar de sua importância, a implementação de sistemas de localização enfrenta dois grandes obstáculos: custo e complexidade. As soluções de alta precisão costumam usar muitos microfones e algoritmos pesados, necessitando de processadores potentes e de custo elevado, o que as torna inviáveis para projetos simples ou com orçamento limitado. Existe, portanto, uma necessidade clara de desenvolver sistemas de localização sonora^[3] que sejam ao mesmo tempo eficazes e acessíveis, podendo ser implementados em hardware simples e de baixo custo.

Dessa maneira, este trabalho insere-se neste contexto com o objetivo principal de desenvolver e validar um protótipo de baixo custo para localização espacial de fontes sonoras em tempo real. A hipótese investigada foi a de que, mesmo com recursos limitados, como os oferecidos pelo microcontrolador ESP32, seria possível implementar algoritmos robustos de processamento de sinais para estimar a direção de chegada do som.

O foco do projeto foi a implementação da técnica de Diferença de Tempo de Chegada (TDOA)^[1], através do algoritmo de Correlação-Cruzada Generalizada com Transformada de Fase (GCC-PHAT)^[2], uma metodologia reconhecida pela sua eficiência e robustez. Desta forma, a pesquisa visa demonstrar a viabilidade prática da localização sonora em sistemas embarcados, contribuindo com uma solução que torna esta tecnologia mais acessível para inovação e desenvolvimento, favorecendo a disseminação dessa tecnologia em contextos acadêmicos e práticos.

3. METODOLOGIA

O desenvolvimento do projeto seguiu as etapas de definição da metodologia de localização, seleção de hardware e implementação do algoritmo.

3.1. TDOA E GCC-PHAT

A primeira fase do trabalho consistiu na revisão de literatura para investigar as principais técnicas de localização de fontes sonoras. Foram estudadas abordagens como *beamforming*^[3] e métodos baseados em intensidade sonora, mas a técnica de Diferença de Tempo de Chegada (TDOA - *Time Difference of*

Arrival)^[1] foi selecionada como a mais promissora para este projeto. O princípio do TDOA baseia-se na medição do pequeno atraso de tempo (Δt) que ocorre quando uma onda sonora atinge os diferentes microfones de um *array*. Este atraso é diretamente relacionado à direção da fonte sonora. Dessa forma, os dois microfones capturam o mesmo sinal sonoro em instantes ligeiramente diferentes, devido à diferença de caminho percorrido pelo som.

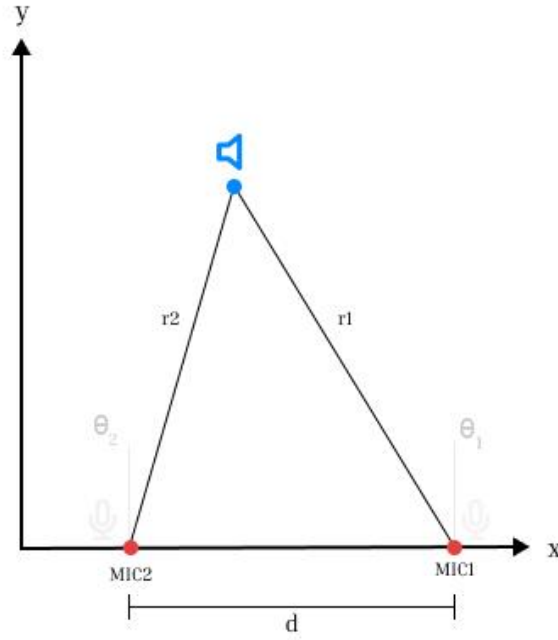


Figura 1. Par de microfones em relação à fonte sonora.

Para formalizar o problema, consideramos apenas um *array* com dois microfones, M_1 e M_2 , localizados nas posições conhecidas (x_1, y_1) e (x_2, y_2) , e uma fonte sonora P em uma posição desconhecida (x_p, y_p) . A onda emitida por P percorre as distâncias r_1 e r_2 . A Diferença de Tempo de Chegada, Δt , é a diferença no tempo que a onda leva para percorrer essas duas distâncias. Sendo v a velocidade do som, o TDOA é dado pela relação (1):

$$\Delta t = \frac{r_2 - r_1}{v} = \left| \frac{\|P - M_2\| - \|P - M_1\|}{v} \right| \quad (1)$$

onde v é a velocidade do som. $\|\cdot\|$ é a norma euclidiana.

Para estimar o Δt de forma robusta, foi escolhido o algoritmo de Correlação-Cruzada Generalizada com Transformada de Fase (GCC-PHAT)^[2]. A Correlação-Cruzada é uma medida matemática da semelhança entre dois sinais. A sua forma generalizada, quando calculada no domínio da frequência através da Transformada Rápida de Fourier (FFT)^[2].

Em vez de utilizar informações de fase, é empregada a diferença de tempo de chegada do sinal entre os elementos. O algoritmo de Correlação-Cruzada^[2] é usado para calcular essa diferença de tempo. A partir disso, o ângulo é calculado com base na diferença de tempo de chegada. O algoritmo de triangulação é baseado em uma fórmula trigonométrica simples. Conforme mostrado na Figura 1, assume-se que o arranjo de microfones está separado, e o ângulo da fonte sonora em relação aos dois arranjos de microfones é dado por θ_1 e θ_2 .

Uma vez que o Δt é estimado pelo algoritmo, o passo seguinte é convertê-lo em uma informação geométrica: o Ângulo de Chegada. Assumindo que a fonte sonora está suficientemente distante, a relação entre o TDOA, a geometria do array e o ângulo é descrita pela seguinte equação trigonométrica (2):

$$\cos \theta = \frac{\Delta d}{d} = \frac{\Delta t \cdot v}{d} \quad (2)$$

Onde:

- θ é o ângulo de chegada da fonte sonora em relação ao eixo perpendicular do array.
- Δd é a diferença de distância percorrida pela onda sonora até os dois microfones.
- Δt é o TDOA estimado pelo GCC-PHAT (em segundos).
- v é a velocidade do som no meio (aproximadamente 343 m/s no ar).
- d é a distância física entre os dois microfones.

Isolando o ângulo θ , obtemos a equação para determinar a direção da fonte, que foi implementada no código (3):

$$\theta = \cos^{-1}\left(\frac{\Delta t \cdot v}{d}\right) \quad (3)$$

Com um *array* de dois microfones, esta equação permite determinar a direção da fonte em um plano de 180°. É importante notar que métodos mais complexos, utilizam múltiplos *arrays* para obter diferentes ângulos e, a partir deles, realizar uma triangulação para encontrar as coordenadas (x, y) exatas da fonte.

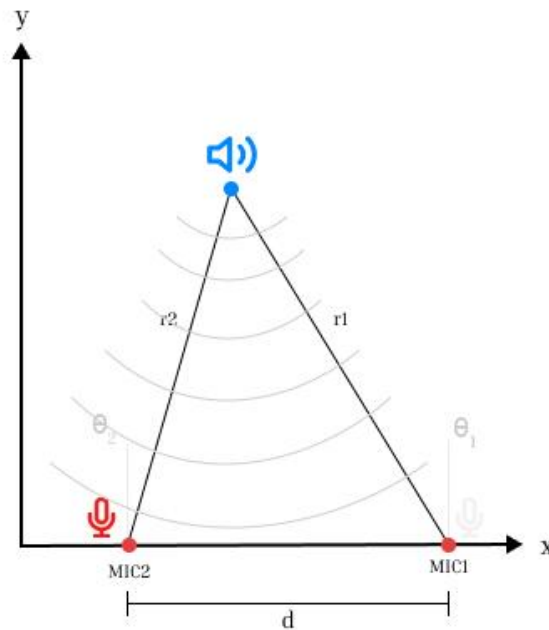


Figura 2. Exemplo do funcionamento do sistema.

3.2. SELEÇÃO DE HARDWARE

O hardware foi selecionado com base nos critérios de baixo custo, acessibilidade e capacidade de processamento. O protótipo foi construído com os seguintes componentes:

- **Microcontrolador:** ESP32 (WROOM-32D). Sua capacidade de processamento e suporte nativo à comunicação I2S o tornam uma escolha mais que suficiente para a aplicação, além da facilidade de implementação de software.
- **Microfones:** Para a validação inicial, inicialmente foi montado um *array* linear com 2 microfones digitais I2S do modelo INMP441, posicionados a uma distância fixa de 15 cm.

- **Feedback Visual:** Foram utilizados 2 LEDs para fornecer uma indicação visual imediata da direção detectada (esquerda/direita/centro).

3.3. IMPLEMENTAÇÃO DO ALGORITMO

O algoritmo GCC-PHAT^[2] foi implementado em C++ no ambiente Arduino IDE. O software foi estruturado para operar em um ciclo contínuo de processamento em tempo real, que executa os seguintes passos: captura de um bloco de amostras de áudio via I2S, pré-processamento, aplicação da FFT em cada canal, cálculo da correlação GCC-PHAT no domínio da frequência, aplicação da Transformada de Fourier Inversa (IFFT) para retornar ao domínio do tempo e, finalmente, a localização do pico de correlação para determinar o TDOA^[1].

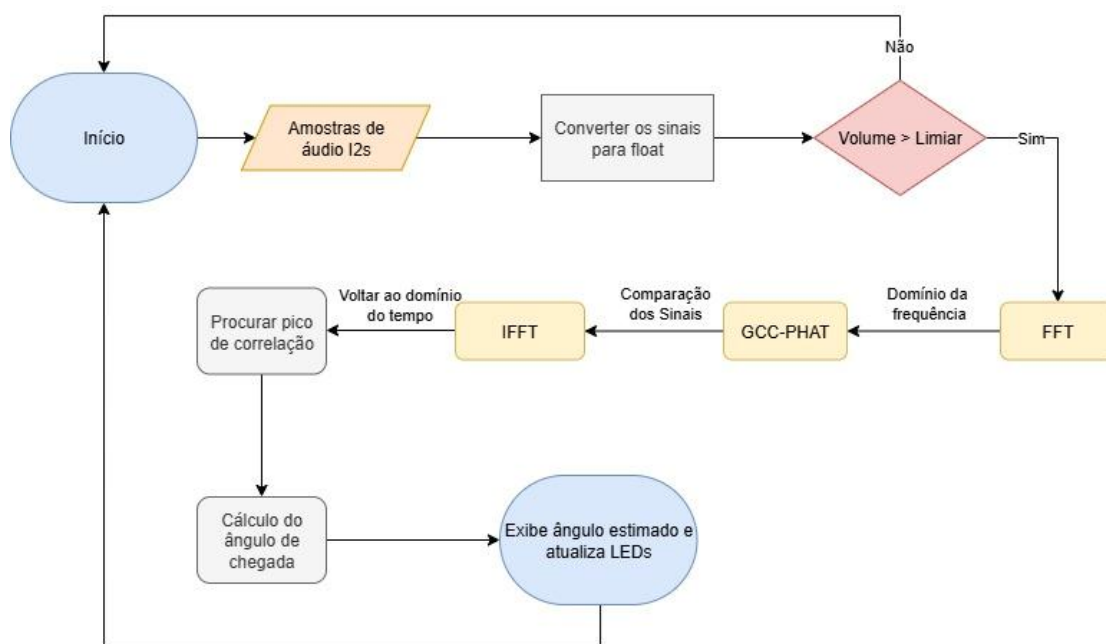


Figura 3. Fluxograma seguido pelo código para determinação do ângulo de chegada do som.

4. RESULTADOS E DISCUSSÕES

Os testes iniciais em ambiente fechado e de baixa reverberação mostraram que o protótipo conseguiu estimar a direção da fonte com certa precisão, em consonância com estudos recentes em sistemas leves^[4].

Durante o desenvolvimento, o protótipo identificou corretamente se a fonte sonora estava posicionada à esquerda, à direita ou ao centro do conjunto de microfones, utilizando os *LEDs* para indicação visual. Além disso o algoritmo foi capaz de calcular e exibir o ângulo de chegada do som na tela, com valores coerentes ao movimento da fonte sonora. Conforme esperado pela geometria do sistema, a faixa de ângulos estimado entre 0° e 180° , sendo 90° o centro do *array*. Essa abordagem, entretanto, torna uma ambiguidade quanto a determinação da origem do som na parte da frente ou atrás do protótipo.

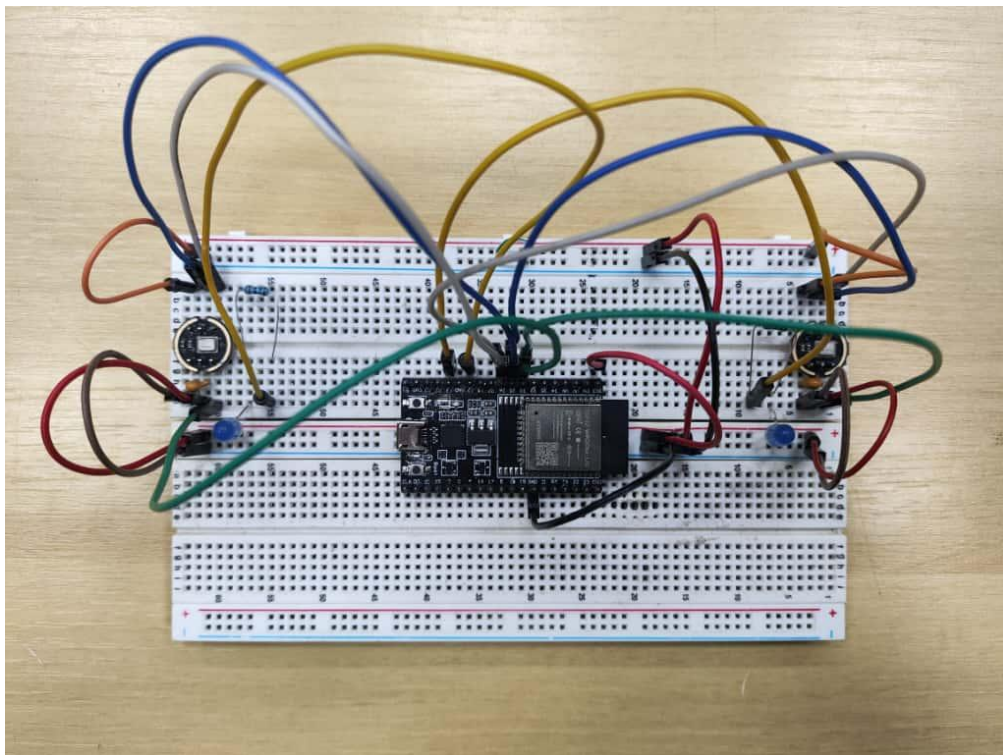


Figura 4. Foto do protótipo montado em uma *protoboard*.

O ESP32 mostrou-se capaz de realizar o processamento em tempo real, sem atrasos perceptíveis, o que confirma a adequação do hardware à aplicação. A introdução de filtros de software reduziu interferências de ruído, aumentando a robustez do sistema.

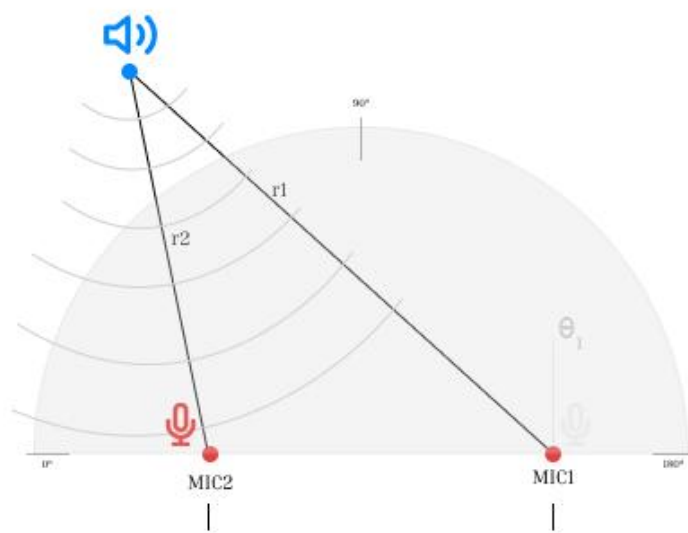


Figura 5. Ilustração do funcionamento do sistema com uma fonte sonora à esquerda do protótipo.

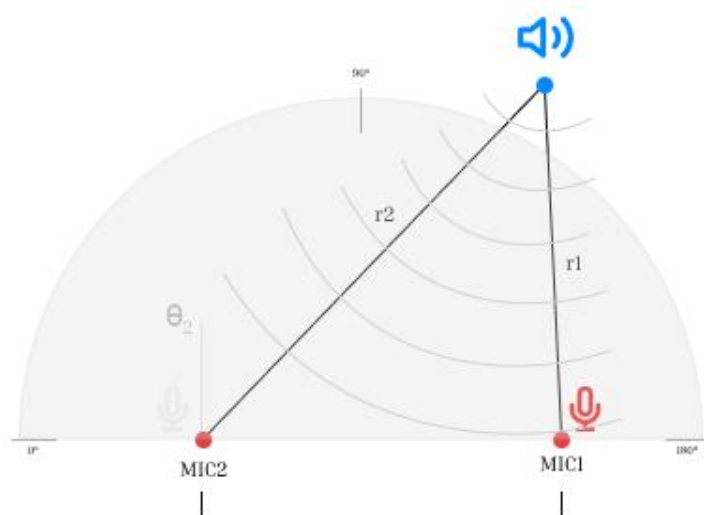


Figura 6. Ilustração do funcionamento do sistema com uma fonte sonora à direita do protótipo.

Durante a pesquisa, foram também exploradas possibilidades de expansão para *arrays* com mais de dois microfones. Testes preliminares com quatro canais indicaram maior precisão no que diz respeito à localização espacial, mas também revelaram limitações relacionadas à memória e à complexidade de sincronização das leituras no ESP32. O uso de algoritmos mais avançados,

como o SRP-PHAT^[4], mostrou-se promissor, mas exigirá maior otimização, a qual pode ser aprimorada em pesquisas futuras.

5. CONCLUSÕES

O trabalho atingiu o objetivo de desenvolver e validar um protótipo de baixo custo para localização de fontes sonoras. A implementação do GCC-PHAT^[2] no ESP32 demonstrou a viabilidade de executar algoritmos de processamento de sinais complexos em plataformas embarcadas acessíveis, obtendo resultados coerentes e funcionais.

Como perspectiva futura, destaca-se a expansão para *arrays* com quatro ou mais microfones, aliada à otimização do código e à aplicação de técnicas mais avançadas como o SRP-PHAT^[4]. Isso possibilitará não apenas maior precisão, mas também uma cobertura de 360°, eliminando a ambiguidade frente-trás do sistema atual.

A pesquisa realizada, além de validar a abordagem, abre caminho para aplicações em robótica, automação e tecnologias assistivas, consolidando a relevância do tema e incentivando a continuidade de estudos na área.

6. REFERÊNCIAS

- [1] Kaune, R., "Accuracy Studies for TDOA and TOA Localization", 15th International Conference on Information Fusion, (2012).
- [2] Chung, M.-A.; Chou, H.-C; Lin, C.-W., "Sound Localization Based on Acoustic Source Using Multiple Microphone Array in an Indoor Environment", Electronics, v.11, p. 890, (2022).
- [3] So, H. C., "Source Localization: Algorithms and Analysis", em: Zekavat, S. A.; Buehrer, R. M. (eds.), Handbook of Position Location: Theory, Practice, and Advances, Wiley, cap. 2, p. 25–66, (2011).
- [4] Grondin, F.; Michaud, F., "Lightweight and optimized sound source localization and tracking methods for open and closed microphone array configurations", Robotics and Autonomous Systems, v.113, p. 63–80, (2019).

7. ANEXOS

- Anexo A – Código Fonte:

Github do projeto: https://github.com/yurigabrielreiss/TDOA_ESP32_I2S.git

- Anexo B – Datasheet dos Componentes:

ESP32-WROOM-32D:

https://www.espressif.com/sites/default/files/documentation/esp32-wroom-32d_esp32-wroom-32u_datasheet_en.pdf

Microfone Digital I2S INMP441: <https://invensense.tdk.com/wp-content/uploads/2015/02/INMP441.pdf>

Orientador

Prof. Dr. Johnny Werner

Estudante

Yuri Gabriel dos Reis Souza