

SISTEMA DE LOCALIZAÇÃO DE FONTES EMISSORAS DE SONS PARA O ROBÔ DE SERVIÇO

Thiago Spilborghs Bueno Meyer¹, Plínio Thomaz Aquino Junior²

^{1,2} Ciência da Computação, Centro Universitário FEI

thiago.spilborghs@gmail.com, plinio.aquino@fei.edu.br

Resumo: A utilização de robôs é cada vez mais discutida no mundo, com diversidade de cenários de aplicação, sendo a comunicação por voz o formato de interação mais representativo, na qual é desejado a localização do usuário. O objetivo deste trabalho foi realizar a comparação e avaliação da localização sonora com a variação da distância e angulação da fonte sonora com relação ao sistema.

1. Introdução

A interação por voz está aplicada em softwares, máquinas de autoatendimento, aplicativos em *smartphones* e robôs. Porém, o ciclo de interação com os humanos no robô possui maior complexidade, pois os robôs estão inseridos em ambientes e contextos de uso que motivam o deslocamento do robô. Por esse motivo, quando ocorre um ciclo de interação entre humanos e robô, o comando do humano pode ocorrer quando o robô não está virado de frente para aquele que manterá uma sequência de trocas de informações. O robô considerado nesta pesquisa, trata-se do Robô HERA (*Home Environment Robot Assistant*). Esse robô foi construído pela Equipe RoboFEI [1] e participa de competições de robótica que valorizam tarefas executadas com comando de voz para acionamento autônomo do robô.

2. Contextualização

Em muitos cenários o robô necessita de um sistema de localização e rastreamento sonoro para localizar um usuário [2]. Neste sistema 3 dimensões são necessárias, sendo essas: 2 ângulos e 1 distância. Um ângulo é correspondente à uma angulação horizontal, de 0 a ± 180 graus, o outro ângulo é correspondente à altura, de 0 a ± 90 graus na vertical, e a distância é a profundidade podendo variar de $0 \pm \infty$ centímetros. Três parâmetros físicos principais são utilizados pelo sistema auditivo para a localização sonora, sendo eles: tempo, intensidade e formato espectral [3]. Horizontalmente, a angulação é definida por fatores biauriculares envolvendo ambos os ouvidos ou receptores sonoros, como, por exemplo, o diferencial de intensidade e tempo interauricular. Já na vertical, a altura é definida por fatores monoauriculares verificados por apenas um ouvido ou receptor sonoro, como alterações na forma espectral das ondas sonoras. A distância também é definida por fatores monoauriculares [3]. O robô HERA tem sua arquitetura computacional baseada em um minicomputador utilizando o sistema operacional Ubuntu, em conjunto com o ROS (*Robot Operating System*) [1] diversos sensores que auxiliam o robô no exercício de funções autônomas e na interação com o usuário. O reconhecimento de localização é uma área de pesquisa muito importante que pode ser crucial em ambientes reais com situações de perigo, como por

exemplo em incêndio para a localização de pessoas através de robôs de resgate e para isso são utilizadas técnicas de reconhecimento de localização de múltiplas fontes sonoras [4]. Um método muito popular para a estimativa de localização com apenas uma dimensão é a diferença de tempo de chegada, em conjunto com o cálculo da correlação cruzada de vetores. Porém este método é muito sensível a reverberações e outros sons que podem ocorrer durante o recolhimento do áudio. Para que esses problemas não criem distúrbios no cálculo, o método de correlação cruzada com base na frequência dominante pode ser utilizado. A estimativa é trabalhada em um único espaço de tempo, onde é calculada a localização de uma fonte em uma janela de tempo única e, quando esses cálculos são gerados, é possível, através de métodos de filtragem, realizar o rastreamento da fonte sonora que está comunicando-se com o sistema. Entre os métodos mais utilizados estão os métodos: Filtro de Kalman, Filtro de Partículas. O filtro de Kalman é um filtro de Bayes que possui como foco realizar o cálculo através do som, levando em conta que este cálculo satisfaz uma distribuição normal, resultando em uma curva suave estimada para a trajetória do alvo. O filtro de Partículas baseia-se no método de Monte Carlo para o cálculo probabilístico da localização da fonte tendo como base pesos, ou partículas, ajustando a mesma conforme o tempo [4].

3. Localização Sonora e Rastreamento

Com o sistema a ser utilizado definido, foi realizado uma pesquisa quanto ao método de localização já presente em códigos base para testes das funcionalidades do sistema. Foi verificado que o sistema utilizava o método de diferença de tempo e intensidade, porém esse sistema apenas resultava no microfone mais próximo à possível localização da fonte sonora, tendo como erro estimado 26,47% devido à disposição dos microfones no sistema. Tal erro não cumpriria com um dos requisitos, que seria obter um erro de no máximo 10%. Como não seria viável utilizar esta solução foi então implementado outro método de localização e rastreamento de fontes sonoras, o método SRP-PHAT. O *Steered Response Power using the PHASE Transform* (SRP-PHAT, Potência de Resposta Guiado usando Transformação de Fase) é um algoritmo proposto por DiBiase [5], que combina a estratégia de localização do SRP em conjunto com os pesos para cálculo do PHAT [6]. Como visto em [7], uma solução para um melhor reconhecimento de localização das fontes sonoras é utilizar em conjunto com o SRP-PHAT, o *Hierarchical Search with Directivity model and Automatic calibration* (HSDA), e para o rastreamento da fonte sonora um método de filtro 3D de Kalman modificado (M3K).

4. Metodologia

No início do trabalho, as regras de ambas as competições foram revisadas, com enfoque na interação e localização de usuários que se comunicam com o robô indiretamente. O foco principal da localização de fontes sonoras nesse caso foi da angulação horizontal, visto que o desafio é o reconhecimento de qual usuário, localizado ao redor do robô, que está realizando a comunicação com o sistema, além da precisão da localização conforme a distância da fonte sonora. A partir do trabalho verificado em [2] foi realizada a implementação da nova funcionalidade ao robô, integrando o sistema ao conjunto de pacotes já usados, utilizando como referência, a base de módulos existentes no robô. O sistema para realizar a localização sonora foi a *Matrix Creator*. Verificando as regras e o estudo já realizado no artigo, foi definido alguns testes e cenários possíveis para observação e comparação da qualidade da localização sonora proposta.

5. Testes

Para os testes, utilizou-se um som contínuo com intensidade média de 63 dB, que representa uma conversa normal [8]. Dois ângulos foram escolhidos para a realização da verificação de precisão conforme a distância, os ângulos de 85° e 115° que representam, no sistema utilizado, um ponto entre microfones e outro na direção de um microfone, respectivamente, e para a distância foi utilizado distâncias de 30 centímetros a 190 centímetros, variando 20 centímetros entre as medições.

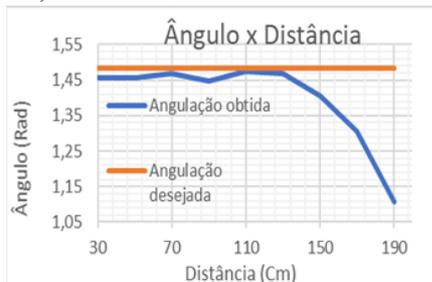


Gráfico 1 - Teste Localização utilizando um ângulo de 85 Graus. Fonte: Autor.

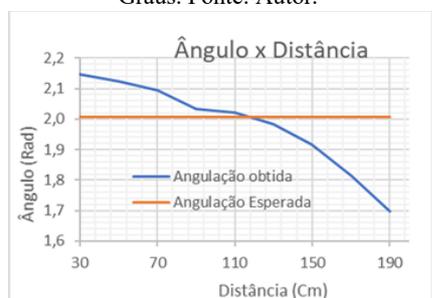


Gráfico 2 - Teste Localização utilizando um ângulo de 115 Graus. Fonte: Autor

No Gráfico 1 é representado o valor desejado para o reconhecimento e o valor real durante os testes para 85° para com o sistema. No Gráfico 2 é representado o valor desejado para o reconhecimento e o valor real durante os testes para 115° para com o sistema. Para uma realização de média e erro estimado no cálculo da localização da fonte sonora, foram realizados 30 testes para cada posição definida anteriormente. Com os dados sendo recolhidos, dois gráficos para verificar a

aproximação dos dados colhidos com o valor esperado foram gerados.

6. Resultados

A partir dos dados colhidos nos testes foi verificado, para 85 graus, um erro estimado de 5,7% em sua totalidade, e quando verificado o resultado dos dados entre as distâncias de 70 e 130 centímetros, que correspondem à distância exemplificada nas exigências do projeto, obtemos um erro estimado de 1,3%. Para o ângulo de 115 graus, foi obtido um erro estimado de 5,5% e realizando a mesma verificação feita para o teste em 85 graus, foi constatado um erro de 1,9%. Ambos os resultados se demonstram satisfatórios, cumprindo a exigência estabelecida durante o início do projeto.

7. Conclusões

É possível concluir que conforme a distância entre o sistema e a fonte sonora aumenta, o reconhecimento de sua localização possui uma tendência a perder sua precisão, porém, em sua maioria o resultado foi satisfatório, obtendo um valor muito aproximado do valor real esperado, possuindo um erro estimado de 5,53% para o ângulo de 115° e 1,89% para ângulo de 85°. Independente da fonte sonora estar localizada na direção de algum microfone do sistema ou estar entre eles, a localização pode ser realizada e o robô terá a possibilidade de interagir diretamente com o usuário. É desejado para trabalhos futuros, realizar a localização da altura do usuário, além de fazer uso de outros sensores disponíveis no sistema escolhido como base do projeto.

8. Referências

- [1] P. T. A. Junior et al., "HERA: Home Environment Robot Assistant," in II BRAHUR and III Brazilian Workshop on Service Robotics, 2019.
- [2] T. S. B. Meyer and P. T. A. Junior, "Sound Source Localization and Tracking for the @Home Service Robot," in II BRAHUR and III Brazilian Workshop on Service Robotics, 2019, pp. 59–64.
- [3] M. Risoud et al., "Sound source localization," European Annals of Otorhinolaryngology, Head and Neck Diseases, v135, n4. Elsevier Masson SAS, p 259–264, 2018.
- [4] C. Rascon and I. Meza, "Localization of sound sources in robotics: A review" Rob. Aut. Syst., v96, p184–210, 2017.
- [5] DIBIASE and J., "A High-Accurate, Low-Latency Technique for Talker Localization in Reverberation Environments Using microphone Array," PhD thesis, Brown Univ., 2000.
- [6] L. A. Seewald, L. Gonzaga, M. R. Veronez, V. P. Minotto, and C. R. Jung, "Combining SRP-PHAT and two Kinects for 3D Sound Source Localization," Expert Syst. Appl., vol. 41, no. 16, pp. 7106–7113, Nov. 2014.
- [7] F. Grondin and F. Michaud, "Lightweight and optimized sound source localization and tracking methods for open and closed microphone array configurations," Rob. Auton. Syst., vol. 113, pp. 63–80, Mar. 2019.
- [8] Y. University, "Decibel Level Comparison Chart." 2019.

Agradecimentos

À instituição Centro Universitário FEI pelo empréstimo de equipamentos, a equipe de robótica RoboFEI@Home e ao orientador pelo suporte com o projeto.

¹ Aluno de IC do Centro Universitário FEI. Projeto com vigência de 06/18 a 12/18.