

EVOLUÇÃO DA ASTRA COMO PLATAFORMA MÓVEL PARA NAVEGAÇÃO VISUAL INDOOR

Lucas Montoya Correia¹, Marcelo Xavier Guterres²,

¹ Universidade Presbiteriana Mackenzie

² Instituto Tecnológico de Aeronáutica

Resumo

Este trabalho apresenta a evolução da ASTRA no ciclo PIBIC 2025–2026. O protótipo anterior utilizava uma Raspberry Pi Zero 2 W conectada ao hotspot do celular, diálogo sequencial e navegação qualitativa. Neste ciclo, um aplicativo Android tornou o telefone a plataforma de interação e captura do sistema. A solução reúne áudio em tempo real, fotografias sob demanda e navegação visual guiada por um mapa de conexões. Em 25 turnos, a mediana entre o fim da fala e o primeiro áudio foi de 821 ms. O mapeamento experimental reuniu 17 locais, 26 conexões e 150 referências; uma rota de sete pontos recuperou cinco perdas temporárias. O projeto eletrônico dos óculos avançou até o esquemático funcional da Rev A. Os ensaios demonstram integração técnica e fornecem uma base mensurável para a próxima fase, mas ainda não avaliam eficácia assistiva, pois foram conduzidos somente pelo desenvolvedor.

Palavras-chaves: tecnologia assistiva. navegação indoor. localização visual. interação por voz. aplicativo móvel.

1. Introdução

A navegação em interiores ainda desafia pessoas cegas ou com baixa visão, pois o GPS não oferece precisão adequada. Uma revisão de 58 estudos destaca custo, energia, acessibilidade e validação experimental como problemas em aberto [1]. O UNav demonstrou, com 20 participantes, que imagens, mapas e instruções faladas podem apoiar essa tarefa [2]. A ASTRA partiu de uma Raspberry Pi Zero 2 W conectada pelo hotspot do celular. Para uma pergunta simples, o sistema precisava reconhecer a fala, elaborar a resposta e só então sintetizá-la; perguntas visuais ainda acrescentavam uma etapa de análise. Neste ciclo, o objetivo foi reduzir a espera percebida, dar ao aplicativo papel central e registrar evidências suficientes para orientar testes futuros.

2. Metodologia

O aplicativo Android reúne microfone, câmera, GPS e sensores de movimento. A conversa utiliza o modelo gpt-realtime-2.1-mini por WebRTC [3], enquanto um detector local de voz reduz interrupções por ruídos curtos. A câmera só produz uma fotografia quando a pergunta ou a navegação precisa de informação visual. No mapeamento, o telefone registra fotografias em várias direções e sua orientação. Para localizar o usuário, o MegaLoc recupera as referências mais parecidas [4]; em seguida, um modelo visual compara poucas alternativas, como na Fig. 1. Somente após essa confirmação o mapa de conexões libera o próximo passo da rota. Assim, a orientação não depende de uma única imagem ou de um ângulo isolado. Na frente vestível, a primeira revisão dedicada, chamada Rev A, atingiu um esquemático funcional com 163 posições de montagem para câmera, áudio, sensores, alimentação e comunicação. O contorno mecânico da haste principal e modelos dos componentes críticos foram integrados no Autodesk Fusion [5]. O servidor permaneceu local durante o desenvolvimento, evitando custo prematuro com nuvem.



Fig. 1 – Localização em duas etapas: captura atual e três referências recuperadas são comparadas antes da orientação.

3. Resultados e Discussão

Nos ensaios de voz, 25 respostas concluídas em sete sessões apresentaram mediana de 821 ms entre o fim da fala e o primeiro áudio; em 90% delas, esse início ocorreu em até 1,407 s. Para a pessoa usuária, esse é o tempo até perceber que a ASTRA ouviu e começou a responder, não a duração total da fala da assistente. O pôster de 2025 registrava aproximadamente 5 s, em média, do início da fala ao início do TTS no pipeline anterior. As métricas não são equivalentes, e as solicitações visuais tinham etapas adicionais; por isso, não se calcula uma redução percentual. Ainda assim, a resposta passou a começar em menos de um segundo na mediana observada. Em 22 respostas com dados completos, o custo mediano estimado foi de US\$ 0,0043, menos de meio centavo de dólar por resposta, sem ferramentas, transcrição ou impostos.

O mapa reuniu 17 locais, 26 conexões e 150 referências. A Fig. 1 mostra a etapa decisória: a captura atual é confrontada com três referências recuperadas; somente a alternativa visualmente coerente confirma o local. Em vídeo independente, elevar a resolução de 368 para 1440 pixels levou a proporção de quadros com escore igual ou superior a 0,50 de 36,6% para 49,3%. Isso mede a busca inicial, não a acurácia da rota.

No percurso exploratório, seis pontos intermediários foram confirmados e cinco perdas temporárias foram recuperadas. Como houve um único operador e nenhum protocolo com participantes, trata-se de prova funcional, não de avaliação assistiva. Os testes com o público-alvo dependem de aprovação ética e de um dispositivo fabricado [2]. A fabricação eletrônica de cinco unidades, incluindo câmeras, foi estimada entre R\$ 4,1 mil e R\$ 7,5 mil, sem tributos e sem a estrutura mecânica. A falta de verba adiou fabricação e ensaio; layout e roteamento permanecem em desenvolvimento.

4. Conclusões

A ASTRA evoluiu de um protótipo de laboratório para uma plataforma móvel capaz de conversar, observar uma cena quando necessário e retomar uma rota após incerteza. O principal avanço foi reduzir a espera percebida sem transformar a navegação em uma decisão automática de uma única camada. A mesma arquitetura também passou a ser explorada em uma interface web acessada por QR Code, para fornecer um ponto inicial em eventos e centros comerciais. Os ensaios mediram latência, custo e um percurso real, mas ainda não validam o uso assistivo. As próximas etapas são concluir o layout e o roteamento, fabricar o hardware vestível, obter aprovação ética e avaliar segurança, eficiência e carga cognitiva com participantes.

5. Agradecimentos

Ao CNPq, ao Instituto Tecnológico de Aeronáutica e aos colaboradores do projeto ASTRA.

Referências

- [1] CASANOVA, E.; GUFFANTI, D.; HIDALGO, L. Technological Advancements in Human Navigation for the Visually Impaired: A Systematic Review. *Sensors*, v. 25, n. 7, art. 2213, 2025. DOI: 10.3390/s25072213.
- [2] YANG, A. et al. Evaluating the efficacy of UNav: A computer vision-based navigation aid for persons with blindness or low vision. *Assistive Technology*, v. 37, n. 6, p. 459–473, 2025. DOI: 10.1080/10400435.2024.2382113.
- [3] OPENAI. GPT-Realtime-2.1 mini. 2026. Disponível em: <https://developers.openai.com/api/docs/models/gptrealtime-2.1-mini>. Acesso em: 16 ago. 2026.
- [4] BERTON, G.; MASONE, C. MegaLoc: One Retrieval to Place Them All. In: 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Nashville, 2025. p. 2852–2858. DOI: 10.1109/CVPRW67362.2025.00269.
- [5] AUTODESK. Fusion Electronics tutorials: schematic and PCB design. Disponível em: <https://help.autodesk.com/cloudhelp/ENU/FusionECAD/files/ECD-TUTORIALS-CPT.htm>. Acesso em: 16 ago. 2026.