



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE CIÊNCIAS, TECNOLOGIA E SAÚDE - CAMPUS ARARANGUÁ
CURSO DE GRADUAÇÃO EM ENGENHARIA DE COMPUTAÇÃO

Augusto Scarduelli Prudencio

**Protótipo Vestível de Baixo Custo para Auxílio à Detecção de Obstáculos
para Pessoas com Deficiência Visual**

Araranguá

2024

Augusto Scarduelli Prudencio

**Protótipo Vestível de Baixo Custo para Auxílio à Detecção de Obstáculos
para Pessoas com Deficiência Visual**

Trabalho de Conclusão de Curso do Curso de Graduação em Engenharia de Computação do Centro de Ciências, Tecnologia e Saúde - Campus Araranguá da Universidade Federal de Santa Catarina para a obtenção do título de Bacharel em Engenharia de Computação.

Orientador: Prof. Antonio Carlos Sobieranski, Dr.

Araranguá

2024

Ficha catalográfica gerada por meio de sistema automatizado gerenciado pela BU/UFSC.
Dados inseridos pelo próprio autor.

Prudencio, Augusto Scarduelli

Protótipo Vestível de Baixo Custo para Auxílio à Detecção de
Obstáculos para Pessoas com Deficiência Visual / Augusto
Scarduelli Prudencio ; orientador, Antonio Carlos
Sobieranski, 2024.

26 p.

Trabalho de Conclusão de Curso (graduação) -
Universidade Federal de Santa Catarina, Campus Araranguá,
Graduação em Engenharia de Computação, Araranguá, 2024.

Inclui referências.

1. Engenharia de Computação. 2. Visão Computacional. 3.
Tecnologia Assistiva. 4. Baixo Custo. 5. Deficiente
Visual. I. Sobieranski, Antonio Carlos. II. Universidade
Federal de Santa Catarina. Graduação em Engenharia de
Computação. III. Título.

Augusto Scarduelli Prudencio

**Protótipo Vestível de Baixo Custo para Auxílio à Detecção de Obstáculos
para Pessoas com Deficiência Visual**

Este Trabalho de Conclusão de Curso foi julgado adequado para obtenção do Título de “Bacharel em Engenharia de Computação” e aprovado em sua forma final pelo Curso de Graduação em Engenharia de Computação.

Araranguá, 26 de junho de 2024.



Documento assinado digitalmente

Jim Lau

Data: 08/07/2024 10:19:26-0300

CPF: ***.464.702-**

Verifique as assinaturas em <https://v.ufsc.br>

Prof. Jim Lau, Dr.
Coordenador do Curso

Banca Examinadora:



Documento assinado digitalmente

Antonio Carlos Sobieranski

Data: 08/07/2024 10:11:21-0300

CPF: ***.305.809-**

Verifique as assinaturas em <https://v.ufsc.br>

Prof. Antonio Carlos Sobieranski, Dr.
Orientador



Documento assinado digitalmente

Analucia Schiaffino Morales

Data: 08/07/2024 10:16:59-0300

CPF: ***.256.420-**

Verifique as assinaturas em <https://v.ufsc.br>

Prof.^ª Analucia Schiaffino Morales, Dr.^ª.
Avaliadora
Universidade Federal de Santa Catarina



Documento assinado digitalmente

Jim Lau

Data: 08/07/2024 10:40:26-0300

CPF: ***.464.702-**

Verifique as assinaturas em <https://v.ufsc.br>

Prof. Jim Lau, Dr.
Avaliador
Universidade Federal de Santa Catarina

Protótipo Vestível de Baixo Custo para Auxílio à Detecção de Obstáculos para Pessoas com Deficiência Visual

Augusto Scarduelli Prudencio

26 de junho de 2024

Resumo

Com os recentes avanços em novas tecnologias de arquiteturas de hardware e software de visão computacional, diversas abordagens assistivas foram desenvolvidas para auxiliar, manter ou melhorar as capacidades funcionais dos indivíduos portadores de deficiências. Especificamente, para pessoas portadoras de deficiências visuais, múltiplos problemas surgem da necessidade de detectar obstáculos ao longo do caminho, principalmente os que ficam acima do chão, que não são facilmente identificados por meio de bengalas. O objetivo é apresentar uma solução baseada em tecnologias computacionais de baixo custo para auxiliar pessoas com deficiência visual na navegação de ambientes. Para tal, o objetivo deste trabalho é o desenvolvimento de um protótipo aparato vestível de baixo custo associado com sistema embarcado de visão computacional. O protótipo é um dispositivo equipado com um Raspberry Pi para o processamento de imagem dos quadros adquiridos por uma câmera de profundidade, que é acoplada ao peito do usuário. O *feedback* ao usuário é fornecido através de motores de vibração embutidos em uma pulseira colocada no pulso da pessoa, permitindo a diferenciação de obstáculos de acordo com a proximidade. Este protótipo demonstra que as soluções nesta área podem ser mais viáveis, leves, práticas e de menor custo do que os apresentados na literatura. O estudo realizado previamente considerou um levantamento bibliográfico onde foram selecionados os trabalhos similares à abordagem aqui proposta. Os resultados preliminares obtidos demonstraram a viabilidade do presente trabalho, colaborando para um melhor processo de mobilidade com configurações de baixo custo para auxiliar na detecção de obstáculos.

Palavras-chaves: Visão Computacional, Tecnologia Assistiva, Baixo Custo, Raspberry Pi, Deficiente Visual, Resposta Tátil, Vestível.

Low Cost Assistive Technology System for Obstacle Detection Assistance for Visually Impaired People

Augusto Scarduelli Prudencio

26 de junho de 2024

Abstract

With recent advances in new technologies for hardware and software computer vision architectures, various assistive approaches have been developed to aid, maintain, or improve the functional abilities of individuals with disabilities. Specifically, for visually impaired people, multiple problems arise from the need to detect obstacles along their path, especially those above the ground, which are not easily identified by canes. The objective is to present a solution based on low-cost computational technologies to assist visually impaired people in navigating environments. A low-cost wearable device associated with an embedded computer vision system was developed. The prototype is a device equipped with a Raspberry Pi for image processing of frames acquired by a depth camera, which is attached to the user's chest. Feedback to the user is provided through vibration motors embedded in a bracelet worn on the wrist, allowing differentiation of obstacles based on proximity. This prototype demonstrates that solutions in this area can be more viable, lightweight, practical and cost-effective than those presented in the literature. The prior study considered a literature review based on the PRISMA protocol, and the most similar works were selected to support this study. The preliminary results obtained demonstrated the feasibility of the proposed approach, contributing to improved mobility processes with low-cost configurations to assist in obstacle detection.

Key-words: Computer Vision, Assistive Technology, Low Cost, Raspberry Pi, Visually Impaired Person, Tactile Feedback, Wearable

1 Introdução

De acordo com a Organização Mundial de Saúde (OMS) (LISTE, 2023), estima-se que pelo menos 2,2 bilhões de pessoas apresentem algum tipo de deficiência visual, seja para visão de perto ou de longe. Esta condição acarreta um custo anual global de aproximadamente US\$411 bilhões de dólares, além de reduzir muito a qualidade de vida das pessoas portadoras de deficiência, impondo limites de mobilidade básicos para estas pessoas. No Brasil, segundo a Pesquisa Nacional de Saúde (PNS) (ESTATÍSTICA, 2019) de 2019 conduzida pelo IBGE, pessoas com deficiência visual representam 3,4% do total da população, o que totaliza um custo que gira em torno de 7 milhões de pessoas, já as que necessitam de óculos ou outro tipo de equipamento para visão são desta forma para 40,4% da população. Na PNS de 2010 (IBGE, 2011), foi registrado que aproximadamente 528 mil pessoas são incapazes de enxergar, apresentando e caracterizando deficiência visual total.

O dispositivo assistivo mais comum entre pessoas com deficiência visual significativa é a bengala, oferecendo um grande auxílio ao qual estão habituadas. Segundo (ISAKSSON-DAUN; JANSSON; NILSSON, 2020), embora soluções inovadoras tenham sido desenvolvidas desde a década de 1950, nenhuma delas alcançou uma taxa de aceitação superior à da bengala como principal auxiliar na locomoção. O custo de uma bengala convencional é aproximadamente R\$150, enquanto o modelo avançado da Bengala Inteligente WeWalk (WEWALK, 2021) pode atingir até R\$5,4 mil, sendo desta forma inacessível para a grande maioria dos portadores de deficiências.

Nos últimos anos, as soluções acadêmicas e industriais para a área da tecnologia assistiva têm empregado processamento de imagem por meio de inteligência artificial (IA), demandando considerável capacidade computacional, como destacado em (ISAKSSON-DAUN; JANSSON; NILSSON, 2020) e (MOHAMED; SIRLANTZIS; HOWELLS, 2021). Isso resulta em custos elevados para suas aplicações, devido às especificações necessárias. A visão computacional, tem se mostrado promissora no desenvolvimento de tecnologias assistivas inovadoras. Permitindo a interpretação e análise de informações visuais do ambiente, possibilitando a criação de dispositivos que podem identificar e sinalizar obstáculos, reconhecer rostos e objetos, e até mesmo ler textos em voz alta. Deste modo, as tecnologias supramencionadas possibilitam melhora e auxílio na qualidade de vida de pessoas com algum tipo de deficiência visual.

Além disso, a crescente disponibilidade de dispositivos móveis com câmeras de alta capacidade de resolução espectral e espacial, assim como o poder de processamento significativo abriram novas possibilidades para o desenvolvimento de soluções assistivas baseadas em visão computacional. Conforme discutido por (LIU et al., 2021), esses dispositivos podem ser equipados com aplicativos que utilizam redes neurais profundas para fornecer assistência em tempo real, como a detecção de obstáculos e leitura de texto, diretamente de dispositivos embarcados no próprio usuário ou no smartphone do mesmo.

Neste trabalho, com o intuito de auxiliar a locomoção de indivíduos com deficiência visual severa, desenvolver-se-á uma tecnologia assistiva de baixo custo e baixa complexidade para ajudar esse grupo na detecção de obstáculos situados a aproximadamente 1,5 metros do usuário e acima do chão. Vale ressaltar que esta abordagem visa não substituir o uso da bengala, mas sim se destina a complementá-lo. O objetivo do estudo é integrar a visão computacional em dispositivos acessíveis, combinando eficiência tecnológica com viabilidade econômica para ampliar o acesso e a aceitação entre os usuários.

Na Seção 2 são apresentados 12 trabalhos relevantes desenvolvidos na mesma área

de pesquisa. Na Seção 3 se discorre sobre toda a fundamentação teórica utilizada neste trabalho. Em seguida na Seção 4 se descreve o processo de desenvolvimento, os dispositivos utilizados e a implementação do algoritmo proposto. Na Seção 5 é realizado uma comparação entre os resultados de distância obtidos pela câmera e um *LIDAR* de celular. Por último, na Seção 6 são apresentados as conclusões provenientes do estudo e sendo discutidas algumas ideias de trabalhos futuros.

2 Trabalhos Relacionados

2.1 Metodologia

Para a investigação na literatura se utilizou o protocolo Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) para selecionar os trabalhos a serem analisados (PAGE et al., 2021). Tendo em vista que os principais assuntos são "tecnologia assistiva" e "visão computacional" e utilizando a estratégia PIE (*Population, Intervention, and Evaluation*) como definida a seguir:

- *Population*: todas as pessoas que necessitam de tecnologia assistiva;
- *Intervention*: Tecnologia assistiva que implementa visão computacional;
- *Evaluation*: Tecnologias e aplicações.

Para adquirir os artigos a string de busca "*computer vision AND assistive technology*" foi utilizada em dois bancos de dados (PUBMED and IEEE Xplore). O processo de seleção e triagem começou então com 1136 artigos dos quais 172 foram capturados da PUBMED e os outros 964 da IEEE Xplore. Nesta etapa foi utilizado a ferramenta *Rayyan* (OUZZANI et al., 2016), e com o auxílio de dois revisores "cegos", foi possível triar 108 destes artigos. Dos quais, após a leitura, mantiveram-se 78 sendo classificados em 5 categorias Deficiência Visual, Deficiência Auditiva, Deficiência na Mobilidade Física, Idosos e Outros.

Dos artigos selecionados, o grupo de interesse para este estudo são os categorizados como deficiência visual. Esta por sua vez fora dividida em 10 subcategorias: *translation, mobility, object recognition, face recognition, character detection, color blind, image caption generation, object location, prosthetic vision and task identification*. Por fim os artigos selecionados foram das subcategorias *translation, mobility*, respectivamente Codificação/-Transcrição e Mobilidade, totalizando 12 e são abordados na Seção 2.2.

2.2 Artigos Correlatos Selecionados

Nos últimos anos, houve um considerável desenvolvimento de artigos na interseção entre tecnologia assistiva e visão computacional. Deste modo é possível ter uma visão profunda e robusta do estado da arte atual deste assunto. Os artigos selecionados para este trabalho foram agrupados em duas categorias principais: Codificação/Transcrição e Mobilidade.

No primeiro grupo, destacam-se aqueles que exploram a codificação ou transcrição do ambiente para outros sentidos, permitindo uma interpretação mais acessível para usuários com deficiência visual. Em (RIBANI; LIMA; MARENGONI, 2019), propõe-se a codificação de uma câmera RGB em um dispositivo vestível com formato de camisa,

equipado com motores de vibração. Já em (KIM et al., 2019), a transcrição para um *display* em braille é realizada para representar o contorno dos objetos na cena, facilitando sua identificação. Em (KIM et al., 2016), é desenvolvido um dispositivo que auxilia pessoas com deficiência visual a mirar um alvo em um *display*, utilizando *feedback* vibrotátil. Além disso, são testadas três técnicas (aleatória, cruciforme e radial) para determinar a mais eficiente na localização do alvo. Em (MASCETTI et al., 2016), propõe-se uma técnica de reconhecimento de imagem que permite captar de maneira clara os sinais de trânsito de forma confiável, robusta e eficiente. Por fim, em (ISAKSSON-DAUN; JANSSON; NILSSON, 2020), são realizados cálculos para traduzir os *pixels* RGB-D do ambiente em frequências sonoras, fornecendo *feedback* por meio de um fone de condução óssea.

No grupo de iniciativas relacionadas à mobilidade, agrupam-se os trabalhos que oferecem assistência direta ou indireta na locomoção das pessoas. No estudo (MARTINEZ et al., 2020), aborda-se a manutenção do distanciamento social durante a pandemia de COVID-19, empregando algoritmos de segmentação semântica para localizar pessoas por meio de uma câmera RGB. Além disso, utilizando uma câmera de distância, identificam-se se estão a 1,5 metros de distância do indivíduo. No artigo (ELGENDY; SIK-LANYI; KELEMEN, 2021), utiliza-se o Tiny-YOLOv3 e marcadores para auxiliar esse grupo a se locomover em ambientes internos. Em (YOUNIS et al., 2019b), desenvolve-se um sistema para identificar perigos com base no contexto, abrangendo ambientes tanto internos quanto externos.

Uma comparação entre tecnologias que utilizam visão computacional e aquelas que empregam BLE (*Bluetooth Low Energy*) para auxílio na locomoção em ambientes internos é explorada em (KUNHOTH et al., 2019). No estudo de (YOUNIS et al., 2019a), os autores propõem um sistema utilizando óculos inteligentes e realidade aumentada para detectar objetos em movimento que podem representar perigos nas áreas de perda de visão periférica.

No trabalho (TAPU; MOCANU; ZAHARIA, 2017), é apresentado o sistema DEEP-SEE, que utiliza algoritmos de visão computacional e redes neurais convolucionais profundas (CNNs) para detecção, rastreamento e reconhecimento de objetos em tempo real. Por fim, em (MOHAMED; SIRLANTZIS; HOWELLS, 2021), oferece-se suporte à locomoção de pessoas com deficiência visual e cadeirantes, utilizando deep learning em ambientes tanto internos quanto externos.

O presente trabalho, em contraste aos acima mencionados, visa prototipar uma solução leve, minimalista, de baixo custo e de baixa complexidade para complementar o uso da bengala na detecção de obstáculos acima do chão para pessoas com deficiência visual. Diferentemente das soluções dos artigos mencionados acima esta se baseia no núcleo de processamento sendo um Raspberry Pi cujo tamanho e peso são muito menores e mais facilmente acopláveis que os notebooks postos em mochilas requeridos para soluções com IA. E para o *feedback* optou-se pelos motores de vibração para não obstruir a audição, essa sendo, para pessoas com baixa visão, o sentido que mais utilizam.

3 Fundamentação Teórica

3.1 Fundamentos da Visão Computacional e Processamento de Imagem

A análise de imagens digitais é uma área fundamental dentro do campo da visão computacional e do processamento de imagens, que se dedica ao estudo e desenvolvimento

de técnicas para a captura, processamento, análise e interpretação de imagens digitais. Segundo o livro (PEDRINI, 2007), uma imagem digital pode ser definida como uma representação bidimensional de uma cena ou objeto, codificada em uma matriz de *pixels*. Cada *pixel* contém informações sobre a intensidade luminosa e, em alguns casos, sobre a cor da imagem, permitindo a construção de representações visuais detalhadas e precisas de cenários do mundo real.

Para o computador reconhecer uma imagem RGB-D ele deve salvar as informações de seus *pixels* em uma matriz tridimensional e seu tamanho depende de sua resolução e da quantidade de canais. Uma resolução Full HD de 1920x1080 possuiria um total de 2.073.600 *pixels*. Sendo que cada *pixel* pode conter até 4 canais de dados: *red*, *green*, *blue* e *depth* e cada um deles registra um inteiro de 1 *byte*. Assim uma imagem RGB-D é aquela que salva todos os quatro canais de informação de cada elemento, como demonstrado na Figura 1.

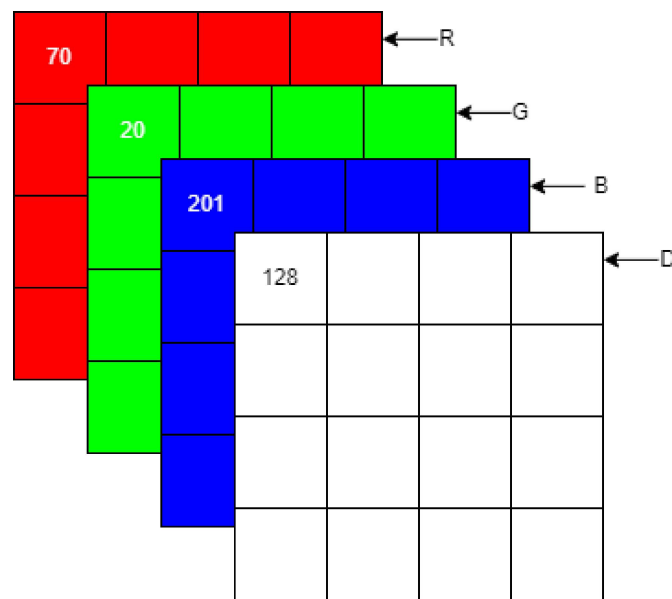


Figura 1 – Demonstração dos quatro canais RGB-D de uma imagem com resolução de 4x4.
Fonte: Autor

O processamento de imagens digitais envolve uma série de etapas e técnicas destinadas a melhorar a qualidade da imagem ou a extrair informações significativas dela. Conforme discutido em (PEDRINI, 2007), os principais objetivos do processamento de imagens incluem a melhoria da visualização para o olho humano, a redução de ruídos e a correção de distorções, bem como a segmentação e a classificação de objetos dentro da imagem. Essas técnicas são essenciais para aplicações em diversas áreas, como medicina, onde imagens de tomografia e ressonância magnética são analisadas para diagnósticos; em segurança, para reconhecimento facial e de placas de veículos; e em robótica, para navegação e reconhecimento de ambientes.

A visão computacional visa replicar a capacidade humana de interpretar e entender o mundo visual a partir de imagens digitais ou vídeos. Suas técnicas podem envolver o uso de processamento de imagem e de inteligência artificial. As aplicações de visão computacional são vastas e incluem desde a automação industrial até a análise de imagens médicas. Segundo (SZELISKI, 2010), os principais desafios da visão computacional envolvem a detecção, reconhecimento e rastreamento de objetos em diferentes condições ambientais.

Os avanços recentes em redes neurais convolucionais (CNNs) têm impulsionado significativamente o campo da visão computacional. Em um estudo, (KRIZHEVSKY; SUTSKEVER; HINTON, 2012) demonstraram que as CNNs podem superar métodos tradicionais em tarefas de reconhecimento de imagem, graças à sua capacidade de aprender características hierárquicas diretamente a partir dos dados. Essa abordagem tem sido fundamental para o desenvolvimento de sistemas de visão computacional mais precisos e eficientes.

Contudo estes trabalhos que utilizam inteligência artificial também necessitam de uma grande capacidade de processamento e em estudos como o (ISAKSSON-DAUN; JANSSON; NILSSON, 2020) e (TAPU; MOCANU; ZAHARIA, 2017) eles acoplam mochilas com notebooks nos pacientes podendo gerar desconforto e altos preços para estas soluções, como mostrado nas Figuras 2 e 3.

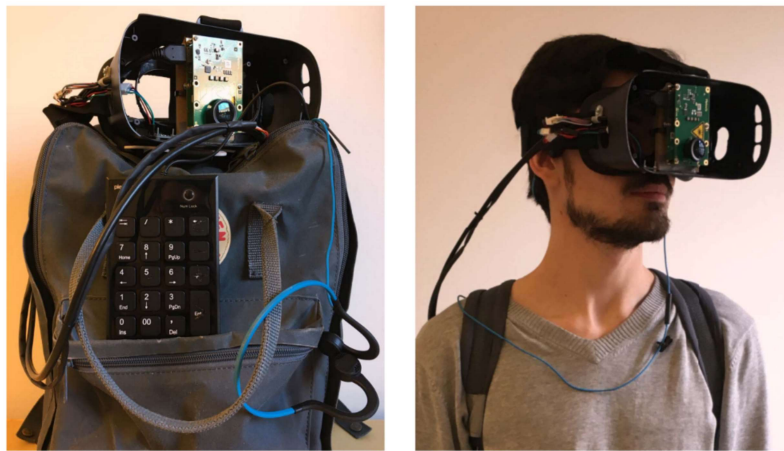


Figura 2 – Design físico do AUDOMNI. Fonte: (ISAKSSON-DAUN; JANSSON; NILSSON, 2020)



Figura 3 – Componentes físicos do DEEP-SEE. Fonte: (TAPU; MOCANU; ZAHARIA, 2017)

3.2 Sensores e Câmeras de Profundidade

O ser humano possui 5 sentidos: visão, audição, paladar, olfato e tato. Eles são de suma importância para a nossa vida e sobrevivência. A visão em específico fornece todas as informações de profundidade e capacidade de enxergar o mundo ao nosso redor.

Na área da computação transformar informações sobre o mundo para dados que podem ser analisados e processados é um grande avanço que pode fornecer soluções para problemas em acessibilidade e na indústria. Sensores e câmeras de profundidade visam transcrever o ambiente e mimetizar o sentido da visão para o computador.

Sensores baseados em ultrassom podem ser simples e medir apenas uma distância perante objetos e obstáculos à frente. HC-SR04 (MORGAN, 2014) é um exemplo desse equipamento e seu funcionamento, como pode ser visto na Figura 4, gira em torno do envio de uma onda sonora a qual reverbera na parede mais próxima e a distância é medida na diferença de tempo entre o envio e o recebimento de onda.

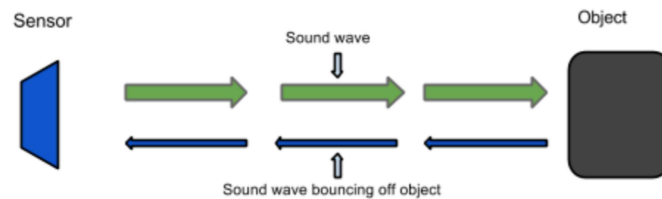


Figura 4 – Funcionamento de um sensor de distancia que utiliza ultrassom. Fonte: (MORGAN, 2014)

Os sensores de profundidade são componentes fundamentais no desenvolvimento de câmeras que possuem todos os quatro canais que podem compor um *pixel* em imagens de RGB-D (*red, green, blue e depth*). Entre as tecnologias mais usadas por este grupo de equipamento se pode destacar os que utilizam ToF (*Time of Flight*) a qual é utilizada também em *LIDAR* (*Light Imaging Detection And Ranging*).

De acordo com (HORAUD et al., 2016) *LIDAR* é uma tecnologia de sensoriamento remoto que estima a distância (ou profundidade) ao iluminar um objeto com um feixe de laser colimado, seguido pela detecção da luz refletida usando um fotodetector. Este princípio de medição remota também é conhecido como tempo de voo (ToF). Como os *LIDARs* utilizam um feixe de laser fino, eles podem estimar a distância com alta resolução. Os *LIDARs* podem usar luz ultravioleta, visível ou infravermelha. Eles podem direcionar uma ampla gama de materiais, incluindo objetos metálicos e não metálicos, rochas, vegetação, chuva, nuvens, e assim por diante – exceto materiais altamente especulares.

Um exemplo de equipamento que usa a tecnologia de ToF e *LIDAR* seria o Kinect. Popularizado em 2010, foi lançado ao Xbox 360 como um formato de interface que permitia o usuário controlar jogos, como o *Kinect Adventures*, através de movimentos físicos. O equipamento retornava uma nuvem de pontos de profundidade, demonstrada na Figura 5, que permitia ao console interpretar os movimentos e ações do indivíduo.

Todavia estes equipamentos podem ser custosos e não usuais para aplicações de tecnologia assistiva do tipo vestível, o caso do Kinect para Xbox One pode ser encontrado por R\$350,00 e medindo 30cm de comprimento. Outro caso é a Intel Realsense Depth Câmera D415 que pode ser encontrada por R\$2.000,00 e possui dimensões de 99mm x 20mm x 23mm.

3.3 Tecnologia Assistiva associada à Mecanismos de Visão

Tecnologia assistiva refere-se a produtos, equipamentos, dispositivos ou sistemas que aumentam, mantêm ou melhoram as capacidades funcionais de indivíduos com deficiência. De acordo com Cook e Polgar (COOK; POLGAR, 2015), essas tecnologias são

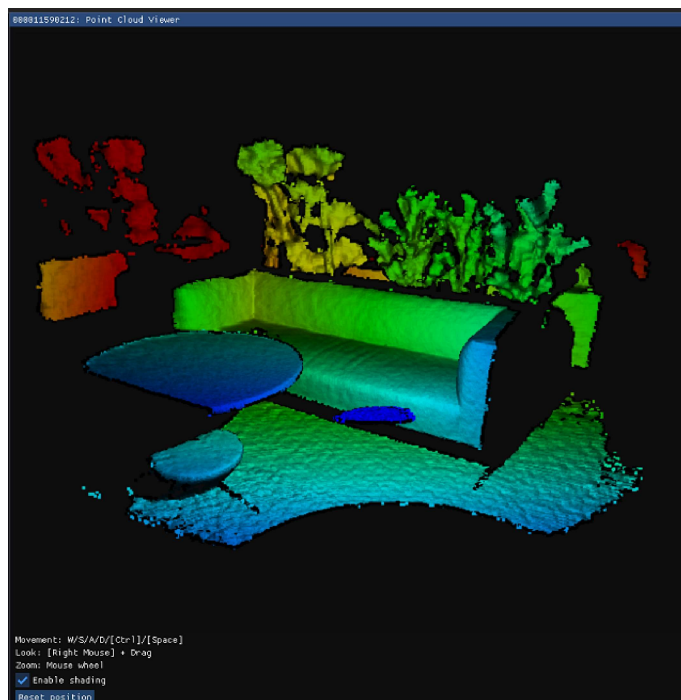


Figura 5 – Visualização da nuvem de pontos de profundidade do Kinect. Fonte: (PHILMEIA, 2024)

essenciais para promover a independência e a qualidade de vida das pessoas com deficiência, os permitindo, assim, participar de atividades diárias que de outra forma seriam inacessíveis.

De acordo com a Organização Mundial da Saúde (OMS), a tecnologia assistiva é de suma importância para pessoas portadoras de deficiência, pois melhora significativamente a sua mobilidade, comunicação e habilidades diárias (LINDSTRÖM, 2024). A OMS enfatiza que o acesso a essas tecnologias é um direito humano essencial e que a falta de disponibilidade e custo acessível são barreiras críticas que precisam ser superadas para garantir a inclusão e a equidade (LINDSTRÖM, 2024).

Dentre as tecnologias assistivas mais comuns se destacam a bengala e a cadeira de rodas. A primeira sendo muito útil no auxílio à detecção de obstáculos por pessoas com deficiência visual, ajudando na sua locomoção proporcionando segurança e independência. A segunda, por sua vez, contribuindo na mobilidade de pessoas portadoras de algum tipo de deficiência motora. Essas que possuem modelos simples até versões motorizadas.

Além disso, os cães-guia são um recurso valioso para indivíduos cegos ou com visão reduzida. Esses animais são treinados para auxiliar na navegação, evitando obstáculos e ajudando seus donos a se deslocarem de maneira segura e independente.

Um exemplo de aplicação de tecnologia assistiva baseada em visão computacional é o dispositivo OrCam MyEye demonstrado na Figura 6, que utiliza uma câmera acoplada em óculos para interpretar o ambiente e fornecer feedback auditivo ao usuário (ALIPRANTIS, 2020). Este dispositivo pode ler textos, reconhecer rostos e identificar objetos, facilitando a navegação e interação social de pessoas com deficiência visual.

Outro avanço significativo é a bengala eletrônica WeWalk, que integra sensores de ultrassom e uma unidade de processamento com visão computacional para detectar obstá-



Figura 6 – Dispositivo OrCam MyEye. Fonte: (ALIPRANTIS, 2020)

culos acima do nível do solo, como galhos de árvores e placas de sinalização (WEWALK, 2021). Este dispositivo é um exemplo de como a combinação de tecnologias pode criar soluções mais abrangentes e eficazes para a locomoção de indivíduos com deficiência visual.

Entretanto essas duas soluções são encontradas à preços altos podendo não ser acessíveis à todas as pessoas deste grupo de necessidades. O OrCam MyEye pode ser encontrado a partir de R\$ 14.900,00 enquanto o WeWalk pode ser encontrado por R\$ 5.400,00.

4 Desenvolvimento da Abordagem Proposta

Para desenvolver o trabalho e tendo em vista as dificuldades de mobilidade enfrentados por indivíduos com deficiência visual parcial ou total nas suas atividades diárias de locomoção, o presente trabalho propõe um sistema destinado a auxiliar a detecção de obstáculos. Os princípios deste sistema residem em tecnologias recentes, que combina uma câmera de distância para capturar informações essenciais do ambiente e um processamento realizado por meio de um Raspberry Pi 3B, demonstrado na Figura 7.

A câmera de distância, identificada com (i) na Figura 8, constitui o núcleo sensorial e de entrada do sistema, capturando dados precisos sobre a disposição dos objetos no ambiente. Por meio da análise dos *pixels* da imagem, foram extraídas informações cruciais acerca das distâncias entre o usuário e os obstáculos ao redor. Essa valiosa informação é, então, encaminhada ao Raspberry Pi 3B, identificado com (ii) na Figura 8, desempenhando um papel essencial no processamento e na tomada de decisões em tempo real.

Após a análise dos dados, o sistema proporciona um retorno imediato ao usuário através de três motores de vibração estrategicamente posicionados. Esses motores têm a função de comunicar a direção do obstáculo detectado: à esquerda, à frente ou à direita. Essa resposta tátil, identificada com (iii) na Figura 8, é fundamental para fornecer ao usuário informações vitais sobre o ambiente ao seu redor, possibilitando uma navegação mais segura e autônoma.



Figura 7 – Raspberry Pi modelo 3B. Fonte: Raspiberry Pi

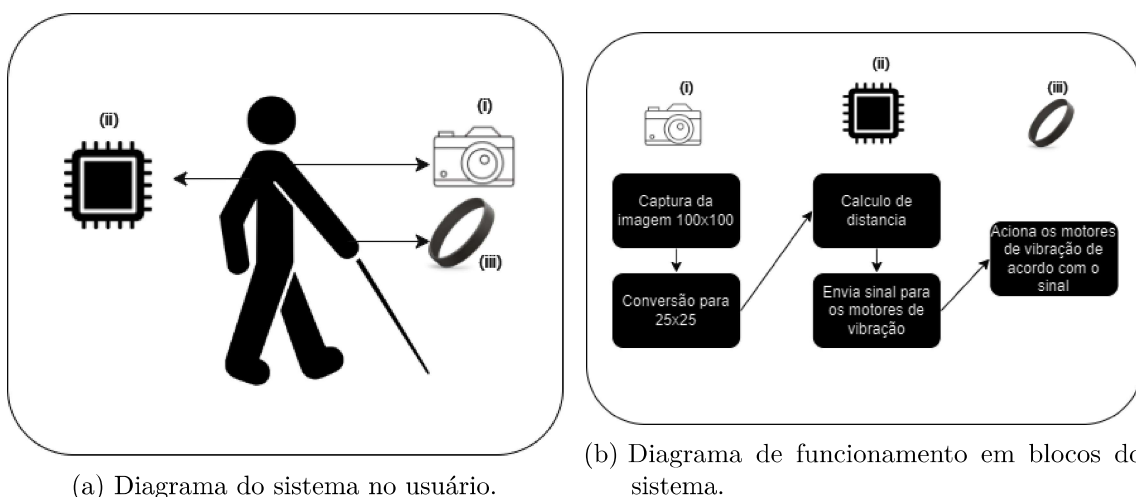


Figura 8 – (i) câmera acoplada no peito de entrada do sistema, (ii) núcleo do processamento do sistema e em (iii) retorno tátil por motores de vibração embutidos na pulseira. Fonte: Autor

O núcleo desse sistema reside na sinergia entre a tecnologia de detecção precisa de distâncias e o processamento rápido e eficaz do Raspberry Pi 3B, combinados para oferecer um auxílio confiável na detecção de obstáculos. A abordagem tátil, por meio dos motores de vibração, constitui uma interface intuitiva e eficaz para comunicar as informações necessárias ao usuário, permitindo uma maior independência e confiança em suas operações diárias de navegação.

4.1 Câmera de Profundidade

A câmera empregada desempenha a leitura das distâncias no ambiente e é denominada MaixSense-A010, produzida pela SiPEED. E suas especificações podem ser encontradas em (SIPEED, 2023b). Seu funcionamento envolve a emissão de uma onda no espectro eletromagnético e a detecção da onda refletida, com o cálculo da distância realizado por meio do tempo de “voo” (*time of flight*). Essa tecnologia é conhecida como *LIDAR* (*Light Imaging Detection and Ranging*), tal qual explicado na Seção 3.2.

O módulo apresenta dois botões laterais para a alteração manual de algumas



Figura 9 – Câmera MaixSense-A010. Fonte: SiPEED

configurações, como a taxa de quadros e a resolução. Além disso, conta com uma tela LCD que se revelou extremamente útil para verificar a correta reconstrução da imagem. As opções de conexão disponíveis são UART ou USB-C na parte inferior. A Figura 9 exhibe uma representação visual do referido módulo.

A resolução adotada foi de 25×25 , na qual regiões de 16 *pixels* da matriz 100×100 original são agrupadas para alcançar uma resolução inferior. Em outras palavras, a matriz maior é dividida por quatro, de modo que cada pixel da matriz menor corresponde a uma região 4×4 da matriz 100×100 . A taxa de quadros selecionada foi de 10 quadros por segundo, resultando em 10 medidas do ambiente por segundo. Essas configurações foram deliberadamente escolhidas em função de considerações de desempenho, dada a restrição de memória RAM de 1GB no Raspberry Pi. A resolução mais baixa se revelou suficiente para obter uma imagem clara e sem ruídos dos obstáculos no ambiente, enquanto a taxa de quadros foi otimizada para fornecer informações adequadas para uma velocidade de caminhada.

4.2 Resposta Tátil

Para a tradução dos dados captados pela câmera em uma resposta tátil, foram empregados pequenos módulos de motores de vibração. Esses módulos foram estrategicamente posicionados em uma pulseira improvisada para facilitar a identificação da região do pulso correspondente à vibração. Cada módulo já é equipado com um circuito que controla a alimentação e a geração de uma frequência de vibração.

A configuração de cada módulo inclui três pinos de conexão, sendo que um par é destinado à alimentação e ao aterramento, enquanto o terceiro pino é utilizado para fornecer a frequência de vibração. A Figura 10 ilustra o módulo do motor de vibração empregado nesse contexto.



Figura 10 – Módulo motor de vibração. Fonte: Fornecedor

Optou-se por utilizar uma única frequência para a resposta tátil, alcançada simplesmente inserindo um sinal de nível alto no pino de entrada. Em testes, essa abordagem demonstrou ser uma escolha eficiente, uma vez que não interferiu na identificação das origens das vibrações quando mais de um motor estava em funcionamento.

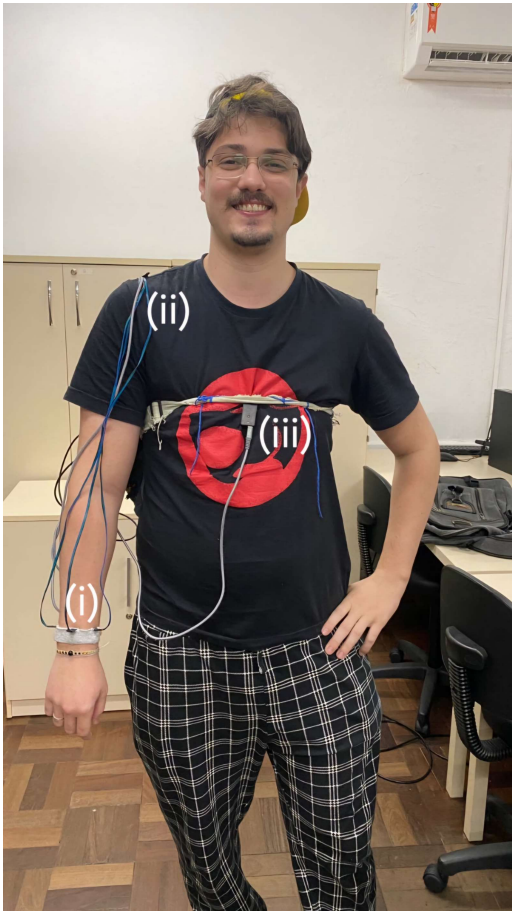
4.3 Implementação

Inicialmente, as configurações iniciais da câmera são enviadas via protocolo USB. Optou-se por uma resolução de 25×25 , para redução de ruído e uma taxa de 10 quadros por segundo, deste modo exigindo um menor processamento e atingindo uma resposta mais próxima em tempo real. A partir do quadro enviado pela câmera para o Raspberry Pi, são realizados os cálculos de distância na matriz quadrada de ordem 25.

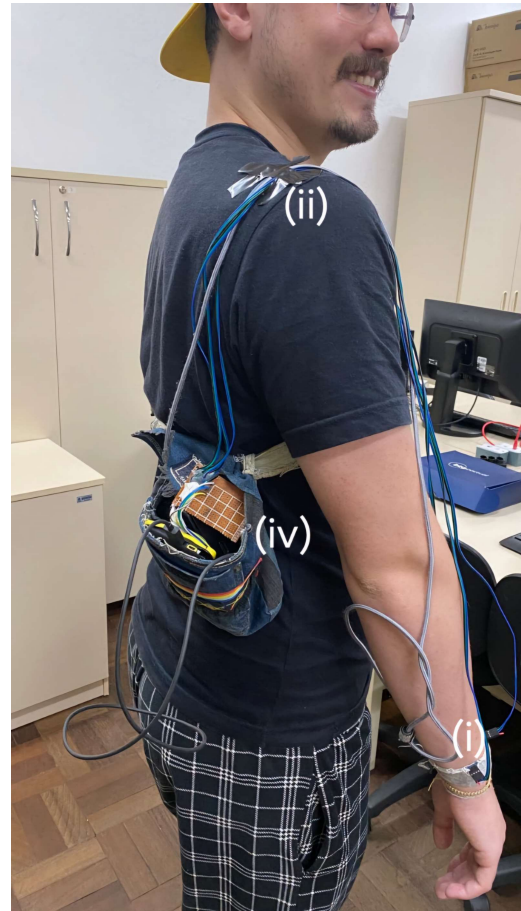
Os motores de vibração são ativados por meio de *threads*. Após a obtenção da matriz de distâncias, a *thread* é executada somente se houver detecção em uma distância de até 1,5 metros. Em seguida, verifica-se em quais regiões da matriz há algum elemento nessa faixa de distância: esquerda (da posição 0 a 7), centro (da posição 8 a 15) ou direita (da posição 16 a 25). Finalmente, o motor correspondente à região identificada na matriz é acionado.

O código foi desenvolvido em Python (FOUNDATION, 2024) e o funcionamento detalhado do algoritmo pode ser observado na Figura 12. Este diagrama descreve todos os passos desde a captura da imagem até o acionamento dos motores. Na Figura 11, apresenta-se o protótipo vestido por uma pessoa de 181cm de altura com uma visão fronta (a) e uma visão lateral (b).

Ainda na Figura 11, pode-se verificar no ponto identificado com (i) a pulseira com os motores de vibração, em (ii) os cabos que ligam toda a estrutura até o Raspberry Pi, em (iii), no peito, a câmera acoplada na alça da pochete e o cabo USB que conecta ao Raspberry Pi e em (iv) a pochete com o Raspberry Pi, uma *Power Bank* e uma placa de



(a) Visão frontal



(b) Traseira do equipamento montado

Figura 11 – (i) a pulseira com os motores de vibração, (ii) os cabos que ligam toda a estrutura até o Raspberry Pi, (iii), no peito, a câmera acoplada na alça da pochete e o cabo USB que conecta ao Raspberry Pi e em (iv) a pochete com o Raspberry Pi, uma *Power Bank* e uma placa de circuitos para ligar as entradas e saídas do vestível. Fonte: Autores

circuitos para ligar as entradas e saídas do vestível.

4.4 Custos envolvidos

No desenvolvimento deste trabalho são considerados os valores gastos dos recursos de hardware e software previamente comentados. Não podendo ignorar que em um ambiente real deveria ser considerado outros aspectos como as mãos de obra dos engenheiros e trabalhadores que envolvem o dispositivo, bem como os custos de tecidos e cabeamentos utilizados nas conexões do vestível. Assim a tabela 1 consta os valores da câmera, do Raspberry Pi, motores de vibração e da linguagem utilizada e o total deles.

5 Resultados Experimentais

Para a obtenção dos resultados, inicialmente é testado o desempenho do sistema com as configurações originais: uma matriz de 100×100 para a imagem proveniente da câmera e 15 *frames* por segundo. No entanto, nesse cenário, observa-se um atraso de até

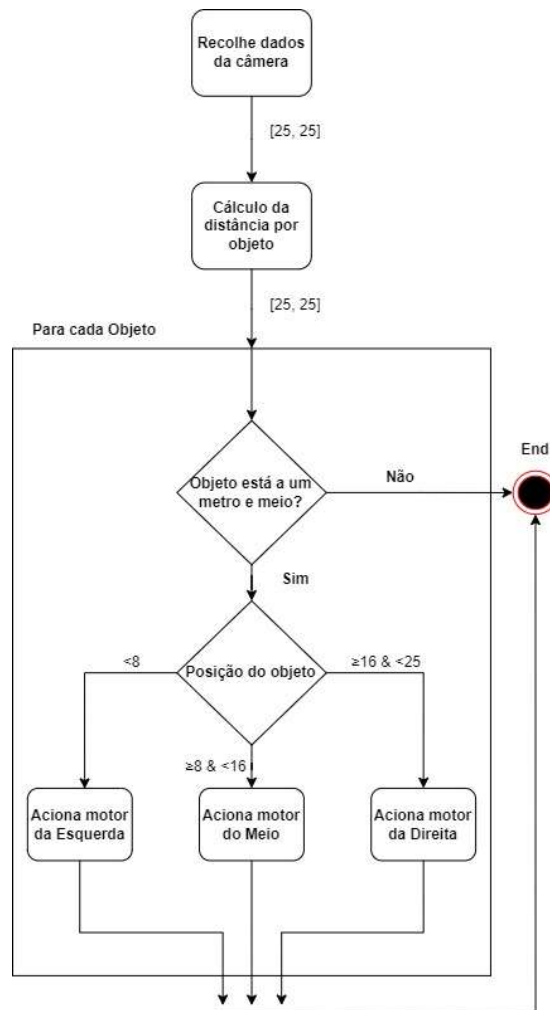


Figura 12 – Diagrama do algoritmo implementado. Fonte: Autor

Equipamento	Quantidade	Preço
Câmera MaixSense-A010	1	R\$210.00
Raspberry Pi 3B	1	R\$380.00
Motores de vibração	3	R\$0.43
Python	1	R\$0.00
Total		R\$591.29

Tabela 1 – Custo dos equipamentos usados no protótipo apresentado.

4 segundos para que o Raspberry Pi 3B pudesse processar cada *frame*, além de receber considerável ruído, o que dificultava a utilização do algoritmo implementado. Assim, tornou-se inviável, considerando a necessidade de uma resposta em tempo real de qualidade para auxiliar pessoas com deficiência.

Conseqüentemente, foi reduzida a taxa de quadros para 10 por segundo e a matriz de imagem para 25×25 . Essa adaptação resultou em uma resposta rápida o suficiente para ser utilizada em tempo real no cotidiano da pessoa, superando as limitações identificadas anteriormente.

O algoritmo e o sistema que acionam os motores de vibração são considerados triviais por não apresentarem nenhuma complexidade e serem sucintos, como apresentado na Seção 4.3. Portanto, a avaliação se concentra na acurácia da câmera e na qualidade dos *pixels* de distância.

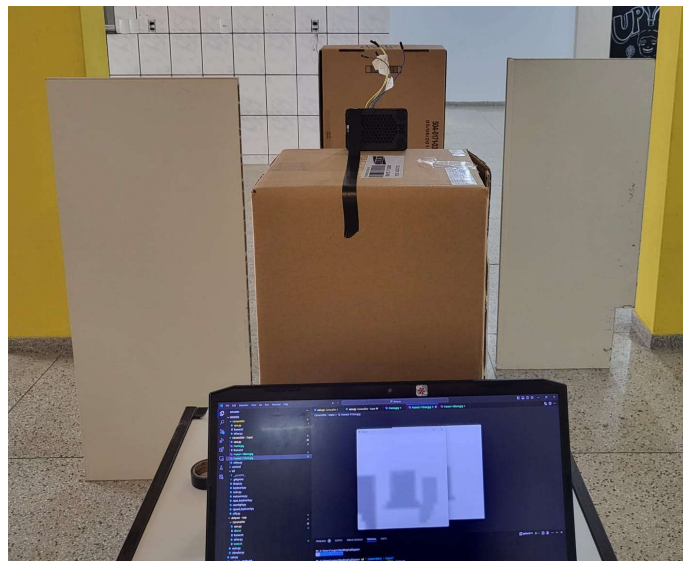


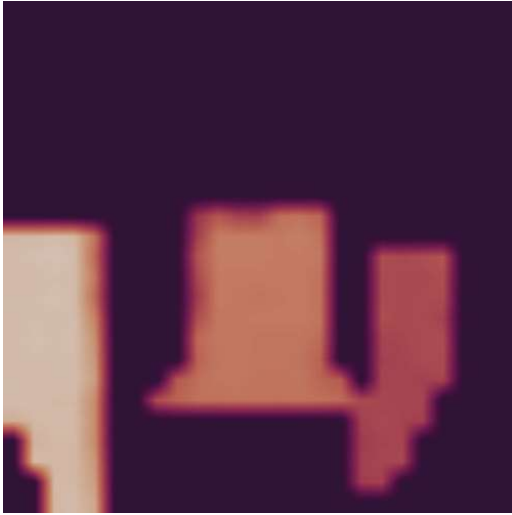
Figura 13 – Configuração para aquisição de resultados da câmera. Ela foi posicionada na altura do peito de uma pessoa de 1 metro e 81 centímetros de altura. Fonte: Autor

Para essa avaliação, conforme ilustrado na Figura 13, foi montado uma estrutura com obstáculos simulados e a câmera foi posicionada em um suporte à altura do peito de uma pessoa com 181 centímetros.

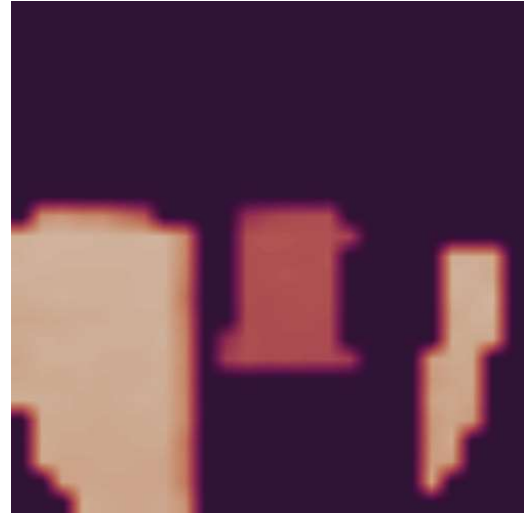
As imagens da Figura 14 foram adquiridas, e as distâncias (D) dos *pixels* centrais de cada objeto foram calculadas de acordo com a Equação 1, fornecida pelo fabricante da câmera por meio do *website* (SIPEED, 2023a). Para avaliar os resultados, procedeu-se à divisão da imagem em três partes iguais da esquerda para a direita, considerando a posição do objeto conforme ilustrado na Figura 12. Em seguida, os principais objetos em cada uma das partes foram isolados, e a distância de seus centros foi medida utilizando a câmera.

$$D = \left(\frac{p}{5.1} \right)^2 \quad (1)$$

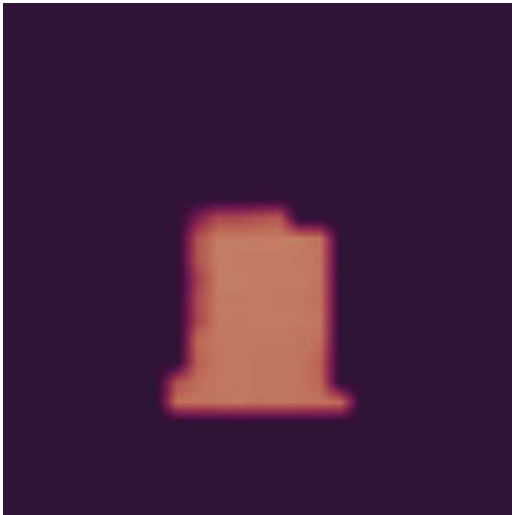
Além disso, empregou-se o *LIDAR* de um dispositivo celular para mensurar a distância real dos objetos na cena, com o propósito de realizar uma comparação com



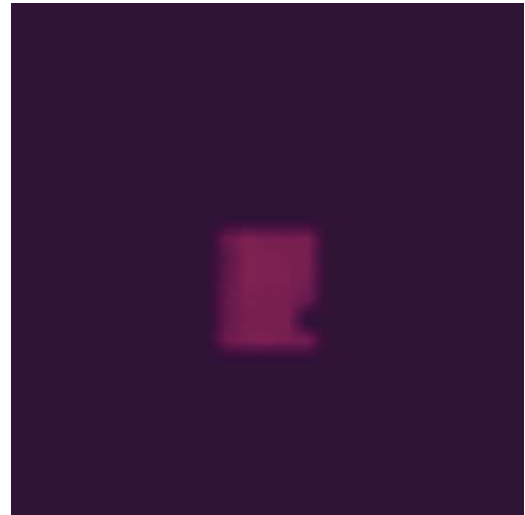
(a) Distancia dos objetos à 96cm, 115cm e 157cm respectivamente



(b) Distancia dos objetos à 100cm, 142cm e 100cm respectivamente



(c) Distancia dos objeto à 115cm



(d) Distancia dos objeto à 180cm

Figura 14 – Imagens capturadas pela câmera de distancia. Coloridas após a aquisição.
Fonte: Autor

os resultados obtidos pela câmera, como mostrado na Tabela 2. Com o intuito de facilitar a identificação das distâncias entre os objetos, aplicou-se um mapeamento de cores utilizando a função *applyColorMap* do OpenCV (ITSEEZ, 2015) com o esquema *COLOR_MAP_TWILIGHT_SHIFTED*. Desse modo, é possível observar que a imagem da Figura 14a acionaria os motores da esquerda e do centro; o da Figura 14b acionaria todos os motores; a imagem da Figura 14c acionaria apenas o motor do centro; e, na Figura 14d, nenhum motor seria acionado.

6 Conclusão e Discussão

Neste trabalho foi desenvolvido um protótipo como uma solução de baixo custo para a detecção de obstáculos a um metro e meio de distância, utilizando uma câmera acoplada ao peito de uma pessoa com deficiência visual. O protótipo, então, possui um

Figura	Distância da câmera dos objetos na parte da esquerda, centro e direita da imagem	Distância medida pelo <i>LIDAR</i>
Figura 14a	113.7cm, 117.7cm, 158.4cm	96cm, 115cm, 157cm
Figura 14b	91.2cm, 144.7cm, 92.4cm	100cm, 142, 100cm
Figura 14c	119.1cm	115cm
Figura 14d	186.1cm	180cm

Tabela 2 – Comparação das distancias medidas por *LIDAR* e pelas medidas pela câmera de distância

custo muito menor do que outras soluções encontradas nos artigos correlatos selecionados.

Em trabalhos relacionados foi aplicado o método prisma para recolher os artigos de interesse e depois da triagem separado os que são mais correlatos com este trabalho. Nos resultados apresentados, devido a baixa complexidade do algoritmo, foi verificado o funcionamento e qualidade da câmera e no desenvolvimento foi explicado o funcionamento do algoritmo de processamento de imagem em tempo real, e ativação da resposta tátil do sistema embarcado.

Esta solução destaca-se pela praticidade, exigindo menos equipamento e poder computacional, além de alcançar uma resposta em tempo real. Dessa forma, é possível mantê-la de maneira minimalista, contribuindo para a mobilidade ao dispensar a necessidade de um computador para processamento de imagens pesadas com inteligencia artificial.

6.1 Trabalhos Futuros

Para trabalhos futuros tem-se em vista a realização de testes com grupos de pessoas com alguma deficiência visual, em ambiente controlado, para coleta e comparação dos dados obtidos. Dessa forma, será possível desenvolver funcionalidades que melhor se adéquem às necessidades dos usuários.

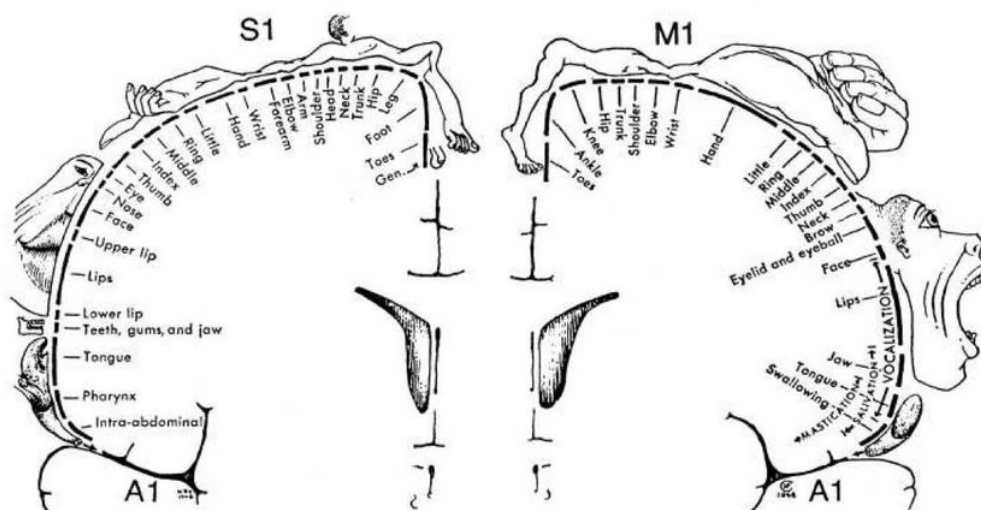


Figura 15 – Ilustração original do homúnculo sensorial de 1937. Fonte: Wilder Graves Penfield

Com relação ao retorno tátil, poderia ser aplicada uma variação na vibração de acordo com a distância do objeto. Sendo possível também deslocar os motores para o peito já que, como pode ser visto na Figura 15, o mapeamento do cérebro da região do peito é de igual tamanho, se não, maior que o mapeamento do pulso sendo assim tendo uma maior facilidade de distinguir qual motor estaria sendo acionado. Deste modo reduzindo a gestão e quantidade de fios e cabos no paciente.

Para diminuir o custo, o tamanho e o peso do protótipo seria necessário fazer testes aplicando o algoritmo desenvolvido em aparelhos mais baratos, leves e menores como um ESP32.

E caso opte-se por melhorias no algoritmo que necessitem de processamento mais pesado uma das soluções a serem implementadas seria o uso do smartfone conectado por (*bluetooth low energy*) BLE ao núcleo de processamento do equipamento seja este um ESP32 ou o Raspberry Pi.

Referências

- ALIPRANTIS, J. The orcam myeye device for aiding visually impaired patients. *Journal of Mobile Technology in Medicine*, v. 9, n. 1, p. 35–38, 2020. Citado 2 vezes nas páginas 12 e 13.
- COOK, A. M.; POLGAR, J. M. *Assistive Technologies: Principles and Practice*. [S.l.]: Mosby, 2015. Citado na página 11.
- ELGENDY, M.; SIK-LANYI, C.; KELEMEN, A. A novel marker detection system for people with visual impairment using the improved tiny-yolov3 model. *Computer Methods and Programs in Biomedicine*, v. 205, p. 106112, 06 2021. Citado na página 8.
- ESTATÍSTICA, I. Instituto Brasileiro de Geografia e. *Painel de Indicadores de Saúde – Pesquisa Nacional de Saúde*. 2019. Acessado dia 07/12/2023. Disponível em: <<https://www.pns.icict.fiocruz.br/painel-de-indicadores-mobile-desktop/>>. Citado na página 6.
- FOUNDATION, P. S. *Python 3.12.3 documentation*. 2024. Acessado dia 25/05/2024. Disponível em: <<https://docs.python.org/3.12/>>. Citado na página 16.
- HORAUD, R. et al. An overview of depth cameras and range scanners based on time-of-flight technologies. *Machine Vision and Applications*, Springer Science and Business Media LLC, v. 27, n. 7, p. 1005–1020, jun. 2016. ISSN 1432-1769. Disponível em: <<http://dx.doi.org/10.1007/s00138-016-0784-4>>. Citado na página 11.
- IBGE, I. B. de Geografia e E. *Sinopse do Censo Demográfico 2010*. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística - IBGE, 2011. Disponível em: <<http://www.ibge.gov.br/home/estatistica/populacao/censo2010/sinopse.pdf>>. Citado na página 6.
- ISAKSSON-DAUN, J.; JANSSON, T.; NILSSON, J. Audomni: Super-scale sensory supplementation to increase the mobility of blind and low-vision individuals – a pilot study. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, v. 28, p. 1187–1197, 05 2020. Citado 3 vezes nas páginas 6, 8 e 10.
- ITSEEZ. *Open Source Computer Vision Library*. 2015. Acessado dia 25/05/2024. Disponível em: <<https://github.com/itseez/opencv>>. Citado na página 20.
- KIM, K. et al. Assisting people with visual impairments in aiming at a target on a large wall-mounted display. *International Journal of Human-Computer Studies*, v. 86, p. 109–120, 2016. ISSN 1071-5819. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1071581915001603>>. Citado na página 8.
- KIM, S. et al. Towards tangible vision for the visually impaired through 2d multiarray braille display. *Sensors*, v. 19, n. 23, 2019. ISSN 1424-8220. Disponível em: <<https://www.mdpi.com/1424-8220/19/23/5319>>. Citado na página 8.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: *Advances in Neural Information Processing Systems*. [S.l.: s.n.], 2012. v. 25, p. 1097–1105. Citado na página 10.

KUNHOTH, J. et al. Comparative analysis of computer-vision and ble technology based indoor navigation systems for people with visual impairments. *International journal of health geographics*, v. 18, p. 29, 12 2019. Citado na página 8.

LINDSTRÖM, A. *Assistive technology*. 2024. Acessado dia 28/05/2024. Disponível em: <<https://www.who.int/news-room/fact-sheets/detail/assistive-technology>>. Citado na página 12.

LISTE, S. *Blindness and vision impairment*. 2023. Acessado dia 07/12/2023. Disponível em: <<https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment>>. Citado na página 6.

LIU, W. et al. Deep learning for generic object detection: A survey. *International Journal of Computer Vision*, Springer, v. 128, n. 2, p. 261–318, 2021. Citado na página 6.

MARTINEZ, M. et al. Helping the blind to get through covid-19: Social distancing assistant using real-time semantic segmentation on rgb-d video. *Sensors*, v. 20, n. 18, 2020. ISSN 1424-8220. Disponível em: <<https://www.mdpi.com/1424-8220/20/18/5202>>. Citado na página 8.

MASCETTI, S. et al. Robust traffic lights detection on mobile devices for pedestrians with visual impairment. *Computer Vision and Image Understanding*, v. 148, p. 123–135, 2016. ISSN 1077-3142. Special issue on Assistive Computer Vision and Robotics - "Assistive Solutions for Mobility, Communication and HMI". Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1077314215002611>>. Citado na página 8.

MOHAMED, E.; SIRLANTZIS, K.; HOWELLS, G. Indoor/outdoor semantic segmentation using deep learning for visually impaired wheelchair users. *IEEE Access*, v. 9, p. 147914–147932, 2021. Citado 2 vezes nas páginas 6 e 8.

MORGAN, E. J. *HC-SR04 Ultrasonic Sensor*. 2014. Acessado dia 28/05/2024. Disponível em: <<https://pdf1.alldatasheet.com/datasheet-pdf/view/1132204/ETC2/HCSR04.html>>. Citado na página 11.

OUZZANI, M. et al. Rayyan—a web and mobile app for systematic reviews. *Systematic reviews*, Springer, v. 5, p. 1–10, 2016. Citado na página 7.

PAGE, M. J. et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Systematic Reviews*, v. 10, n. 1, p. 89, mar. 2021. ISSN 2046-4053. Disponível em: <<https://doi.org/10.1186/s13643-021-01626-4>>. Citado na página 7.

PEDRINI, W. S. H. *Análise de Imagens Digitais: Princípios, Algoritmos e Aplicações*. São Paulo, Brasil: Thomson Learning, 2007. Citado na página 9.

PHILMEA, t. *Azure Kinect Viewer*. 2024. Acessado dia 04/07/2024. Disponível em: <<https://learn.microsoft.com/pt-br/azure/kinect-dk/azure-kinect-viewer>>. Citado na página 12.

RIBANI, R.; LIMA, R. F.; MARENGONI, M. Toward encoding vision-to-touch with convolutional neural networks. *IEEE Access*, v. 7, p. 161175–161183, 2019. Citado na página 7.

SIPEED. *MaixSense-A010 Development*. 2023. Acessado dia 09/03/2024. Disponível em: <https://wiki.sipeed.com/hardware/en/maixsense/maixsense-a010/at_command_en.html#Image-Packet-Description>. Citado na página 19.

SIPEED. *MaixSense series*. 2023. Acessado dia 25/05/2024. Disponível em: <<https://wiki.sipeed.com/hardware/en/maixsense/index.html#What-is-MaixSense%3F>>. Citado na página 14.

SZELISKI, R. *Computer Vision: Algorithms and Applications*. [S.l.]: Springer, 2010. Citado na página 9.

TAPU, R.; MOCANU, B.; ZAHARIA, T. Deep-see: Joint object detection, tracking and recognition with application to visually impaired navigational assistance. *Sensors*, v. 17, n. 11, 2017. ISSN 1424-8220. Disponível em: <<https://www.mdpi.com/1424-8220/17/11/2473>>. Citado 2 vezes nas páginas 8 e 10.

WEWALK. Wewalk smart cane: Enhancing the mobility of the visually impaired. *Assistive Technology Journal*, v. 33, n. 4, p. 217–225, 2021. Citado 2 vezes nas páginas 6 e 13.

YOUNIS, O. et al. A hazard detection and tracking system for people with peripheral vision loss using smart glasses and augmented reality. *International Journal of Advanced Computer Science and Applications*, The Science and Information Organization, v. 10, n. 2, 2019. Disponível em: <<http://dx.doi.org/10.14569/IJACSA.2019.0100201>>. Citado na página 8.

YOUNIS, O. et al. A smart context-aware hazard attention system to help people with peripheral vision loss. *Sensors*, v. 19, n. 7, 2019. ISSN 1424-8220. Disponível em: <<https://www.mdpi.com/1424-8220/19/7/1630>>. Citado na página 8.