

CURSO DE CIÊNCIA DA COMPUTAÇÃO – TCC		
(x) PRÉ-PROJETO	() PROJETO	ANO/SEMESTRE:2021/1

APLICABILIDADE DO SENSOR LIDAR NA DETECÇÃO DE AMBIENTES E OBJETOS PARA ORIENTAÇÃO DE PESSOAS COM DEFICIÊNCIAS VISUAIS

Bruno Henrique de Borba

Dalton Solano dos Santos - Orientador

1 INTRODUÇÃO

De acordo com a Organização Mundial da Saúde, em 2021 pelo menos 2,2 bilhões de pessoas tem algum tipo de deficiência visual para perto ou para longe. De todas as 2,2 bilhões de pessoas citadas, pelo menos 1 bilhão de casos poderiam ter sido prevenidos, ou ainda não foram tratados. Grande parte dos casos de deficiência visual e cegueira atingem pessoas com idade acima dos 50 anos, mas podem acontecer em todas as idades (WORLD HEALTH ORGANIZATION, 2021).

A classificação internacional de doenças (IDC-11, 2020) separa a deficiência visual em para longe, sendo estas de níveis leve, moderada, severa e cegueira, e para perto. A deficiência visual pode ocorrer devido a alguma doença, como catarata, glaucoma, opacidade da córnea, retinopatia diabética, tracoma e até erro refrativo não resolvido (BOURNE, 2021). No geral, pessoas com deficiência visual podem sofrer mais com outros problemas quando comparados com outras pessoas, como com a solidão (BRUNES; HANSEN; HEIR, 2019) e depressão (RENAUD; BÉDARD, 2013).

Outro aspecto importante a ser destacado são as dificuldades na execução de tarefas diárias em ambientes fechados, sendo as principais: encontrar um quarto pelo número, encontrar um elevador, ler números em um banco, encontrar escadas e reconhecimento de objetos (PLIKYNAS, D. *et al.*, 2020). Para a solução de problemas como estes, existem sistemas que se propõem a auxiliar na detecção de objetos e ambientes, mas que contam com a necessidade de um servidor para processar o reconhecimento (JIANG; LIN; OU, 2016), de câmeras ou equipamentos acoplados no corpo do usuário (PHAM; LE; VUILLERME, 2015), ou que precisam ser vestíveis e utilizados em conjunto com outros equipamentos (BAI, J. *et al.*, 2019).

Com o advento dos dispositivos móveis e o aumento da sua capacidade de processamento constante (APPLE, 2021), novos hardwares são adicionados aos dispositivos. É o caso do sensor Light Detection And Ranging (LiDAR), que é capaz de coletar uma nuvem de pontos com coordenadas x, y e z de objetos ao redor do ambiente. Um dos grandes benefícios do sensor, é não ter influência da condição de luz, podendo trabalhar em dias e noites por exemplo (WU, 2018).

Diante dessas informações, esse trabalho propõe a implementação de um aplicativo móvel, que utilizará a tecnologia do sensor LiDAR para efetuar a detecção de objetos e ambientes internos, e auxiliar pessoas com algum tipo de deficiência visual a se locomoverem nestes locais.

1.1 OBJETIVOS

Este trabalho possui como objetivo disponibilizar um aplicativo móvel para auxiliar pessoas com deficiência visual a se orientarem em ambientes fechados utilizando o sensor LiDAR.

- 1

Os objetivos específicos são:

a)

utilizar o sensor LiDAR para efetuar o reconhecimento do ambiente e objetos;

b)

utilizar recursos do próprio dispositivo móvel para orientar o usuário no deslocamento;

c)

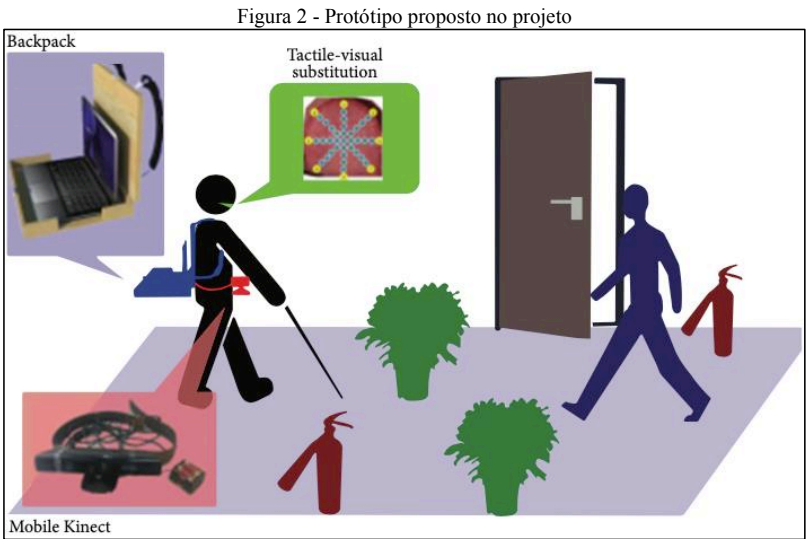
validar se os resultados apresentados são satisfatórios para conduzir com segurança em ambientes fechados, uma pessoa que seja portadora de alguma deficiência visual.

2 TRABALHOS CORRELATOS

Nesta seção serão apresentados alguns trabalhos que apresentam semelhanças com os principais objetivos do trabalho proposto. O primeiro trabalho é um sistema para reconhecimento em tempo real de objetos, com seus resultados convertidos em um som tridimensional (JIANG; LIN; QU, 2016). O segundo trata da detecção em tempo real de objetos e obstáculos utilizando o Microsoft Kinect (PHAM; LE; VUILLERME, 2015). Por último, um dispositivo vestível capaz de detectar obstáculos e objetos, e orientar o usuário em ambientes internos e externos (BAI, J. *et al.*, 2019).

1

alinhamento dos itens.



Fonte: Pham, Le e Vuillerme (2015).

- 1 O funcionamento do projeto de Pham, Le e Vuillerme (2015) foi separado em 5 etapas:
- a) aquisição de dados: utilizando o Kinect;
 - b) reconstrução: combinando cor e profundidade para criar uma nuvem de pontos;
 - c) filtragem: diminuir a amplitude da nuvem de pontos e ter uma resposta mais rápida do sistema;
 - d) segmentação: utilização do algoritmo RANdom Sample Consensus (RANSAC);
 - e) detecção de obstáculos: considerando ambientes onde os obstáculos sejam objetos no chão, escadas, portas ou paredes.

O projeto apresentou bons resultados com a detecção de objetos. Porém, não se mostrou tão confiável em ambientes com muita luz. Como o sistema se baseia na detecção do chão para então detectar os objetos, caso não seja possível mapear ou detectar o chão, todo o sistema irá falhar. Por último, a performance na detecção de escadas se mostrou baixa (Pham, Le e Vuillerme, 2015).

2.3 WEARABLE TRAVEL AID FOR ENVIRONMENT PERCEPTION AND NAVIGATION OF VISUALLY IMPAIRED PEOPLE

O trabalho de BAI, J. *et al.* (2019), propõe a utilização de um óculos com uma câmera RGB-D, uma Unidade de Medição Inercial (Inertial Measurement Units - IMU) e um smartphone. O dispositivo em questão utiliza a continuidade da altura do solo com quadros de imagem adjacentes para segmentar o solo com precisão e rapidez, para pesquisar a direção do movimento em relação ao chão.

Um sistema de reconhecimento de objetos baseados em uma Rede Neural Convolutacional (Convolutional Neural Network - CNN) executada no smartphone é utilizada para aumentar a habilidade de percepção do usuário e promover o sistema de navegação, com informações semânticas dos arredores, como localizações e orientações dos objetos. Como interação humana para o alerta do usuário, é utilizado um som de bipe para obstáculos, reconhecimento de voz para entender comandos do usuário, e síntese de fala para dar informações semânticas dos objetos nos arredores do ambiente (BAI, J. *et al.* 2019). Mais detalhes do protótipo podem ser vistos na Figura 3.

1 Alinhamento dos itens.

3 PROPOSTA DA APLICAÇÃO

Nesta seção é descrita a proposta desta aplicação, definindo suas justificativas científicas e sociais, requisitos funcionais e não funcionais, bem como a metodologia utilizada.

3.1 JUSTIFICATIVA

No Quadro 1 são apresentados comparativos entre os trabalhos correlatos:

Quadro 1 - Comparativo dos trabalhos correlatos

Trabalhos Correlatos Características	Jiang, Lin e Qu (2016)	Pham, Le e Vuillerme (2015)	BAI, J. <i>et al.</i> , (2019)
Depende de um computador para o processamento	Sim	Sim	Não
Utiliza apenas um equipamento para a detecção	Não	Não	Não
Funciona em ambientes de baixa iluminação	Não	Sim	Sim
Possui equipamento capaz de reconhecer profundidade	Não	Sim	Sim
Faz o reconhecimento de objetos	Sim	Não	Sim
Dispositivo para captura de imagens	GoPro	Microsoft Kinect	Câmera RGB-D

Fonte: elaborado pelo autor.

No Quadro 1, é possível observar vários aspectos em comum dos trabalhos correlatos, e que consequentemente justificam o desenvolvimento deste, a começar pela questão do processamento. Apenas o trabalho de BAI, J. *et al.*, (2019) é capaz de efetuar o processamento e reconhecimento dos objetos utilizando um aplicativo instalado em um smartphone.

Também é possível observar que se faz necessário o uso de mais equipamentos para a detecção do ambiente e/ou dos objetos também em todos os trabalhos. O trabalho de Jiang, Lin e Qu (2016) por exemplo, utiliza-se de uma câmera GoPro e de um dispositivo que fará a transmissão das imagens para um computador, e por consequência, o usuário deve carregar consigo de alguma forma ambos os equipamentos.

O trabalho de Jiang, Lin e Qu (2016) não funciona em ambientes que possuem baixa luminosidade por utilizar uma câmera simples GoPro. Consequentemente, este também não apresenta um equipamento capaz de reconhecer a profundidade ou a distância dos ambientes, sendo necessário realizar uma estimativa via software, o que pode ocasionar numa certa imprecisão.

Por fim, apenas o trabalho de Pham, Le e Vuillerme (2015) não faz o reconhecimento de objetos, podendo ser um impeditivo caso alcançar determinado objeto seja o objetivo do usuário ao invés de tratá-lo como obstáculo. O trabalho de BAI, J. *et al.*, (2019) por exemplo, cita uma cadeira como um obstáculo que pode ser removido do caminho para permitir a passagem, ao invés de direcionar o usuário por outro percurso para continuar a locomoção.

Perante as informações apresentadas, este trabalho apresenta relevância pela sua proposta mais voltada ao segmento de dispositivos móveis. A aplicação exigiria apenas o próprio smartphone do usuário com o sensor LiDAR para executar todas as ações, sem a necessidade de ter mais algum dispositivo de certa forma “invasivo” ao corpo, nem o uso de um servidor dedicado para executar todos os processamentos necessários na detecção de ambientes e objetos, trazendo de certa forma uma maior acessibilidade a este tipo de recurso. Além do mais, por consequência da utilização do sensor LiDAR, este também funcionaria em ambientes com baixa iluminação, podendo trazer desta forma uma maior flexibilidade nos cenários de uso da aplicação.

3.2 REQUISITOS PRINCIPAIS DO PROBLEMA A SER TRABALHADO

O aplicativo a ser desenvolvido deverá:

- a) realizar a detecção de objetos (RF – Requisito Funcional);
- b) realizar a detecção de ambientes (RF);

Ponto final.

- c) orientar o usuário para evitar obstáculos no deslocamento em ambientes fechados (RF);
- d) ser desenvolvido utilizando a linguagem Swift (RNF – Requisito Não Funcional);
- e) utilizar o sensor LiDAR para o processo de detecção (RNF).

3.3 METODOLOGIA

O trabalho será desenvolvido observando as seguintes etapas:

- a) Levantamento bibliográfico: levantar fontes bibliográficas relacionadas a processamento de imagens, cálculos de distância e profundidade, bem como trabalhos correlatos;
- b) Levantamento de requisitos: com base na etapa citada anteriormente, levantar os requisitos necessário para a aplicação;
- c) Definir técnicas para detecção de ambientes e objetos: avaliar algoritmos disponibilizados pela linguagem de programação Swift para a detecção de ambientes e objetos, bem como a necessidade da construção do algoritmo e técnicas por conta própria;
- d) Definir formas de alerta de obstáculos para o usuário: caso exista a detecção de um objeto, o usuário deve ser alertado de alguma forma para qual direção seguir. Nesta etapa, serão avaliadas possibilidades de alertar o usuário através da aplicação;
- e) Implementação: implementar a aplicação para a plataforma iOS utilizando a linguagem de programação Swift e o ambiente Xcode;
- f) Testes: validar os resultados obtidos juntamente com possíveis usuários da aplicação.

As etapas serão realizadas nos períodos relacionados no Quadro 2.

Quadro 2 - Cronograma											
etapas / quinzenas	2021										
	ago.		set.		out.		nov.		dez.		
	1	2	1	2	1	2	1	2	1	2	
levantamento bibliográfico											
levantamento de requisitos											
definir técnicas para detecção de ambientes e objetos											
definir formas de alerta de obstáculos para o usuário											
implementação											
testes											

Fonte: elaborado pelo autor.

4 REVISÃO BIBLIOGRÁFICA

Neste capítulo, é descrito brevemente os seguintes assuntos que fundamentarão o estudo a ser realizado: detecção de objetos, sensor LiDAR e o ARKit.

De acordo com ZOU, Z. *et al.*, (2019), a detecção de objetos é uma tarefa da computação que lida com a detecção de instâncias de objetos visuais de certas classes, como humanos, animais ou carros, em imagens digitais, tendo como objetivo desenvolver modelos computacionais e técnicas que provenham a informação necessária para detectar o objeto necessário. De uma forma geral, é possível dizer que a detecção de objetos pode ser dividida em técnicas de aprendizagem de máquina utilizando visão computacional, e aprendizagem de profunda utilizando redes neurais convolucionais (FRITZ AI, 2021).

A visão computacional imita os processos naturais da visão do ser humano, recuperando as informações visuais, as manipulando e interpretando. Graças ao uso de aprendizagem profunda no reconhecimento de imagens e classificação de objetos, computadores podem gerar e aprender recursos, características e propriedades distintas para cada objeto. Baseado em diversos recursos, as máquinas preveem o que está na imagem e indicam o nível de probabilidade (INDATA LABS, 2021).

Redes neurais convolucionais são compostas de neurônios que se auto otimizam por meio do aprendizado, sendo primariamente utilizadas no reconhecimento de padrões em imagens. Desta forma, é possível codificar

1

aprendizagem profunda

recursos específicos de imagem na arquitetura, tornando a rede mais adequada para tarefas focadas em imagem, e reduzindo os parâmetros necessários para configurar o modelo (O'SHEA, Keiron; NASH, Ryan, 2015).

O sensor LiDAR é capaz de detectar uma nuvem de pontos com as coordenadas x, y e z de objetos ao seu redor (WU, 2018). Utilizado comumente em carros autônomos, o sensor fornece informações valiosas em condições que outros sensores não são capazes de fornecer, sendo seu principal foco a informação de distância, medindo o tempo de “viagem” da luz emitida (MAKSYMOVA, Ievgeniia; STEGER, Christian; DRUML, Norbert, 2018).

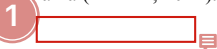
Inicialmente o LiDAR foi desenvolvido com o objetivo de mapear áreas aeronáuticas, como florestas, mantos de gelo, oceanos e atmosfera, e áreas aeroespaciais, como a superfície da lua na missão Apollo 15 (BLICKFELD, 2021). Somente após a disponibilidade do Sistema de Posicionamento Global (Global Positioning System – GPS) e das Unidades de Medição Inerciais (Inertial Measurement Units – IMUs) no final dos anos 1980, foi possível obter uma maior precisão dos dados do LiDAR (BRITANNICA, 2021).

Atualmente o LiDAR é aplicado nas mais diversas áreas, como por exemplo na agricultura para análise da qualidade do solo, veículos autônomos para percepção do mundo ao redor, arqueologia para definir variação topográfica, e na legislação para conduzir análises forenses (EMERLINE, 2021). Áreas como infraestrutura, cidades inteligentes, internet das coisas, transportes, logísticas e aplicações industriais certamente irão, ou já estão sendo altamente beneficiadas com as tecnologias do LiDAR (BLICKFELD, 2021).

ARKit é um framework desenvolvido pela Apple, focado em realidade aumentada. O ARKit combina rastreamento de movimento do dispositivo, captura de cena de câmera, processamento avançado de cena e conveniências de exibição, facilitando a construção de uma experiência de realidade aumentada ao utilizar a câmera frontal ou traseira de um dispositivo iOS (APPLE, 2021).

A versão 4 do ARKit introduziu uma nova Depth API com foco no sensor LiDAR, sendo dessa forma possível acessar informações detalhadas de profundidade geradas pelo sensor. A geometria de cena do ARKit permite criar um mapa topológico de algum espaço com rótulos, identificando pisos, paredes, tetos, janelas, portas e assentos, o que também contribui para a coleta de mais informações necessárias em trabalhos de realidade aumentada (APPLE, 2021).

Para visualizar e interagir com uma cena reconstruída, o ARKit usa o sensor LiDAR para criar um modelo poligonal do ambiente físico. Dessa forma, o sensor obtém as informações de profundidade de uma grande área na frente do usuário, e o ARKit a converte em uma série de vértices que se conectam para formar uma espécie de malha (APPLE, 2021).



REFERÊNCIAS

APPLE. **ARKit**. Disponível em: <https://developer.apple.com/documentation/arkit>. Acesso em: 11 abr. 2021.

APPLE. **Visualizing and Interacting with a Reconstructed Scene**. Disponível em: https://developer.apple.com/documentation/arkit/content_anchors/visualizing_and_interacting_with_a_reconstructed_scene. Acesso em: 11 abr. 2021.

APPLE. **iPhone 12 Pro e iPhone 12 Pro Max**. Disponível em: <https://www.apple.com/br/iphone-12-pro/>. Acesso em: 21 mar. 2021.

APPLE. **Introducing ARKit 4**. Disponível em: <https://developer.apple.com/augmented-reality/arkit/>. Acesso em: 11 abr. 2021.

BAI, J. *et al.* Wearable Travel Aid for Environment Perception and Navigation of Visually Impaired People. **Electronics**, MDPI, v. 8, n. 6, p. 697, jun./2019. Disponível em: <https://doi.org/10.3390/electronics8060697>. Acesso em: 13 mar. 2021.

BLICKFELD. **THE BEGINNINGS OF LIDAR – A TIME TRAVEL BACK IN HISTORY**. Disponível em: <https://www.blickfeld.com/blog/the-beginnings-of-lidar/>. Acesso em: 11 abr. 2021.

BOURNE, Professor Rupert. Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: the Right to Sight: an analysis for the Global Burden of Disease Study. **Lancet Global Health**, Anglia Ruskin University, v. 9, n. 2, p. 144-160, fev./2021. Disponível em: [https://www.thelancet.com/journals/langlo/article/PIIS2214-109X\(20\)30489-7/fulltext](https://www.thelancet.com/journals/langlo/article/PIIS2214-109X(20)30489-7/fulltext). Acesso em: 21 mar. 2021.

BRITANNICA. **Lidar**. Disponível em: <https://www.britannica.com/technology/lidar>. Acesso em: 11 abr. 2021.

1

Ficaria perfeito ... se colocasse em algum lugar aqui uma imagem com pontos coletados pelo LiDAR de uma cena.