

Estudo bibliográfico de estimativa da distância entre um humano e uma câmera omnidirecional

José Eduardo Santos Rabelo¹, Helio Azevedo²

¹Engenharia de Computação
Universidade Estadual de Campinas – Campinas, SP

²Divisão de Sistemas Ciberfísicos – DISCF
Centro de Tecnologia da Informação Renato Archer – Campinas, SP

j260551@dac.unicamp.br, hazevedocti@gmail.com

Abstract. *The distance between a human and a robotic agent represents one of the first steps in the development of machines that can interact socially and respect distance appropriately. A commonly used strategy to measure this distance is using an omnidirectional camera. We made a systematic search with the keyword-string: “human AND distance AND camera AND (fisheye OR 360 OR omnidirecional)” in the Web of Science database. In the search, 65 studies were returned of which only 16 were accepted in the end. The accepted studies used different approaches based on: Machine Learning or Deep learning (31.3%), Matematica (25%), Procedural (25%) and Statistical (18.8%).*

Resumo. *A distância entre um humano e um agente robótico representa um dos primeiros passos no desenvolvimento de sistemas que possam interagir socialmente. Uma estratégia utilizada para medir essa distância é através de câmera omnidirecional. Realizamos uma busca sistemática na base de dados Web of Science com a seguinte sequência de palavras-chave: “human AND distance AND camera AND (fisheye OR 360 OR omnidirecional)”. Essa busca identificou 65 estudos dos quais 16 foram aceitos para leitura completa. A análise das referências aceitas revela 4 abordagens distintas para a questão: Aprendizado de máquina ou aprendizagem profunda (31,3%), Matemática (25%), Procedural (25%) e Estatística (18,8%).*

1. Introdução

A capacidade humana de se locomover em meio a sociedade exige o uso de sentidos apurados como: visão, audição, tato, etc. Nos últimos anos, observamos a busca pelo desenvolvimento de robôs que sejam capazes de realizar a mesma navegação que o ser-humano utilizando basicamente o poder de visão por meio de câmeras. O uso de câmeras omnidirecionais com campo de visão amplo (360 graus) têm sido pesquisado com o objetivo de extrair o máximo de informações visuais do ambiente em que a máquina se locomove [TANAKA, 2020].

Um importante parâmetro na navegação de robôs móveis é a distância entre os objetos. Esse parâmetro aumenta de complexidade quando envolve o espaço entre robôs e humanos, pois ele não apenas existe para evitar as colisões, mas também para respeitar o espaço de cada indivíduo. Essa distância individual, definida como proxêmica por Edward T. Hall em sua obra “A Dimensão Oculta” [Hall, 1966], depende de fatores pessoais, sociais, geográficos, etc. e permite inferir como uma máquina deve se

comportar em um meio social de forma que gere a menor estranheza possível às pessoas ao redor.

Este artigo apresenta uma análise bibliográfica das abordagens utilizadas para realizar a estimativa de distância entre humano e robô exclusivamente com o uso de câmera omnidirecional. Com esse intuito, realizamos uma pesquisa na base de dados *Web Of Science* utilizando a string de pesquisa “*human AND distance AND camera AND (fisheye OR 360 OR omnidirectional)*” sem restrição temporal, ou seja, foram analisados todos artigos independente da data de publicação. Essa busca retornou 65 artigos, sendo que 16 foram aceitos para análise e leitura completa, 4 estão indisponíveis e 45 rejeitados devido ao escopo. O resultado final, possibilita atingir o objetivo de análise de características como: autores, jornais/revistas, datasets, número de participantes, abordagem utilizada, ano publicação e a disponibilidade do código fonte.

Ao observarmos a proporção de artigos que foram removidos devido ao escopo, pode-se inferir que a utilização apenas de câmeras omnidirecionais no desenvolvimento de sistemas capazes de estimar a distância ainda é um abordagem pouco utilizada. De fato, na pesquisa bibliográfica realizada, algumas das referências obtidas utilizam outros equipamentos para cálculo da distância. Esses trabalhos não foram considerados na pesquisa bibliográfica, considerando nosso objetivo de redução de custos de implementação restringindo o reconhecimento do ambiente somente ao uso de câmeras omnidirecionais.

Este artigo está organizado da seguinte forma: na seção 2 apresentamos a abordagem utilizada na pesquisa bibliográfica; o resultado da pesquisa bibliográfica é detalhada na seção 3; na sequência realizamos uma análise detalhada acerca dos métodos utilizados por 3 referências mais relevantes; finalmente o artigo é finalizado na seção 5 onde as conclusões são apresentadas.

2. Método

Durante o desenvolvimento da pesquisa bibliográfica, utilizamos uma sequência de passos [GOUGH, 2017], que envolve: 1) estabelecer o tema de pesquisa, 2) construir a lista de referências, 3) importar para uma base local e, por fim, 4) analisá-las e classificá-las. Para isso utilizamos os seguintes critérios de busca:

- Chave: “*human AND distance AND camera AND (fisheye OR 360 OR omnidirectional)*”;
- Base: *Web Of Science*, pesquisada em 20/10/2023
- Período: Toda a dimensão histórica da base foi considerada.

As referências obtidas pela busca foram classificadas considerando os grupos e subgrupos presentes na Tabela 1.

Tabela 1. Grupos e Subgrupos em análise: critérios utilizados na classificação das referências.

Grupo	Subgrupo	Observação
Base de dados	<ul style="list-style-type: none"> • <nome da base> • Anônimo • Indisponível • Ausente 	Se os resultados não foram disponibilizados classifica-se Indisponível Se o dataset foi disponibilizado mas não possui nome definido então é classificado como “Anônimo”.
Sujeitos	<ul style="list-style-type: none"> • <número de Sujeitos> • Ausente 	Número de participantes pesquisados(as), individual ou coletivamente, de caráter voluntário.
Idade	<ul style="list-style-type: none"> • <faixa de idade> • Ausente 	Apresenta a faixa etária como por exemplo 4-12, 13-17, 18-24.
Recurso	<ul style="list-style-type: none"> • Aprendizado máquina • Procedural • Estatístico • Matemática • Outro 	Define como o problema foi abordado considerando as diferentes estratégias computacionais
Código	<ul style="list-style-type: none"> • Disponível • Indisponível • Ausente 	Indica se algoritmo e/ou código foram disponibilizados.
Ano	<ul style="list-style-type: none"> • <ano de publicação> 	Ano de publicação da referência.

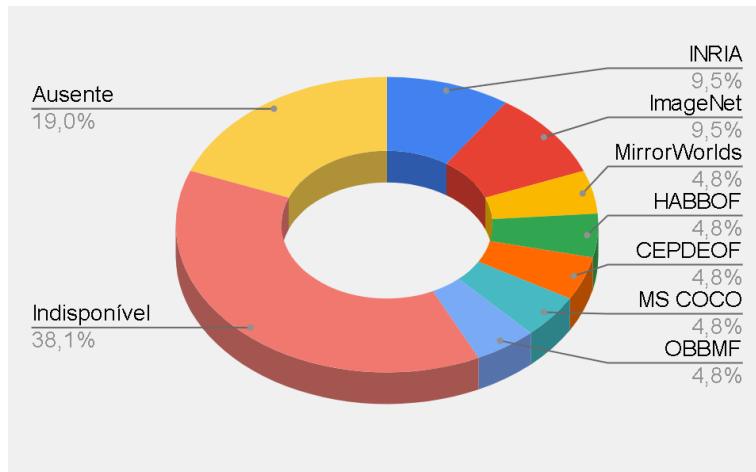
3. Análise da pesquisa bibliográfica

Obtidos os resultados da pesquisa, analisamos os artigos através do título e do *abstract* para selecionar os artigos mais relevantes. Inicialmente foram rejeitados 43 artigos por não estarem aderentes ao escopo da pesquisa. Após a leitura completa dos 22 artigos restantes, 2 foram removidos por escopo e 4 por estarem indisponíveis. Portanto, aceitamos 16 artigos como pertinentes para estudo e análise completa.

3.1. Grupo “Base de dados”

Durante o estudo de cada referência, analisamos os *datasets* (conjuntos de dados) utilizados. Podemos observar no Gráfico 1 que 38,1% dos artigos não disponibilizam os dataset. De fato, em geral, esses artigos utilizam *datasets* desenvolvidos pelos próprios autores durante a pesquisa. O subgrupo Ausente representa referências que não fazem uso de *uma base de dados*. Novamente, no Gráfico 1, do total 42,9% das referências que disponibilizam os *datasets* é visível a diversidade na escolha do conjunto de dados utilizados.

Gráfico 1. Porcentagem de uso de cada base de dados das publicações de referência.

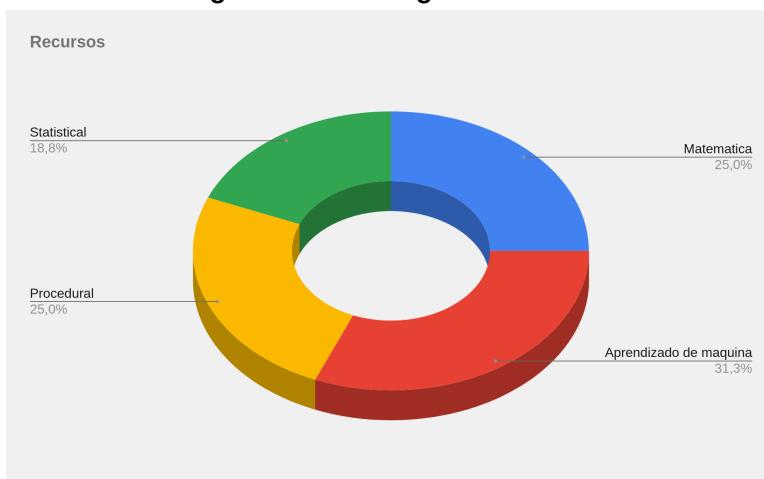


3.2. Grupo “Recurso”

Este grupo representa os tipos de abordagem utilizados no desenvolvimento do sistema. Analisando o Gráfico 2, observamos as 4 abordagens mais utilizadas nas referências e suas respectivas frequências. “Estatístico”, um dos subgrupos, representa artigos que se baseiam em métodos e princípios estatísticos no desenvolvimento geral do trabalho. O subgrupo definido como “Procedural” representa artigos que se baseiam em abordagens de procedimentos, onde algoritmos são desenvolvidos em etapas específicas para chegar ao objetivo. “Aprendizado de máquina” engloba todos artigos que utilizam aprendizado de máquina e, por fim, “Matemática” são os artigos que utilizam métodos puramente matemáticos para construir sua solução. Podemos notar que os métodos de aprendizado de máquina representam a maior parte das referências com 31,3% e isso indica um maior interesse no desenvolvimento desse tipo de abordagem.

A escolha dos critérios de classificação se baseia na observação dos métodos predominantes nos artigos e na necessidade de organizar as diferentes abordagens de maneira clara e coerente. Esta taxonomia permite uma visão estruturada das tendências e preferências metodológicas na área de estudo, facilitando a identificação de padrões e lacunas na literatura.

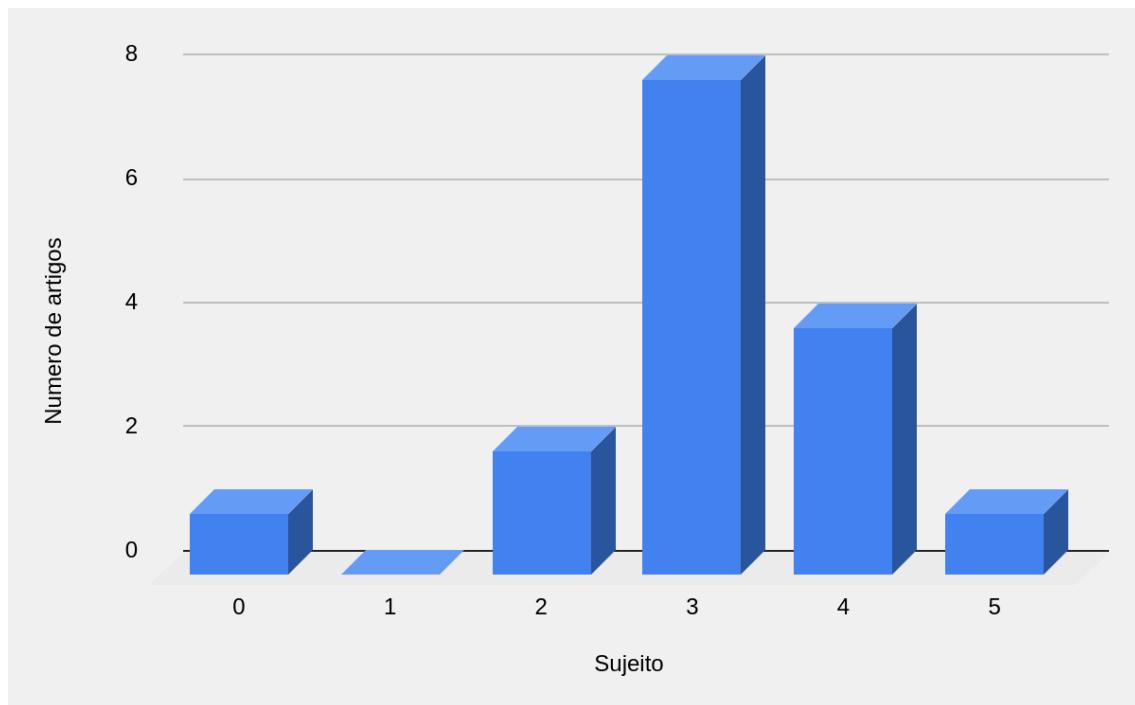
Gráfico 2. Porcentagem das abordagens utilizadas nas referências.



3.3. Grupo “Sujeito”

A análise da quantidade de pessoas que participam de um experimento é relevante, pois mostra como diferentes tipos de corpos e silhuetas são modelados. Naturalmente, quanto maior o número de sujeitos maior será o grau de generalidade considerando o aspecto de análise visual. A maior parte dos artigos realiza os experimentos com 3 indivíduos, o que indica uma pequena amplitude de análise.

Gráfico 3. Quantidade artigos para cada quantidade de participantes.



3.4. Grupo “Idade”

Observar a idade do indivíduo no experimento é importante para determinar o comportamento do método considerando diferentes faixas etárias. Apesar disso, nenhum dos artigos informa a faixa etária dos indivíduos analisados. Entretanto, consideramos que esse aspecto deva ser abordado em trabalhos futuros devido a sua relevância considerando as variações de características físicas presentes nas diversas faixas etárias.

3.5. Grupo “Código”

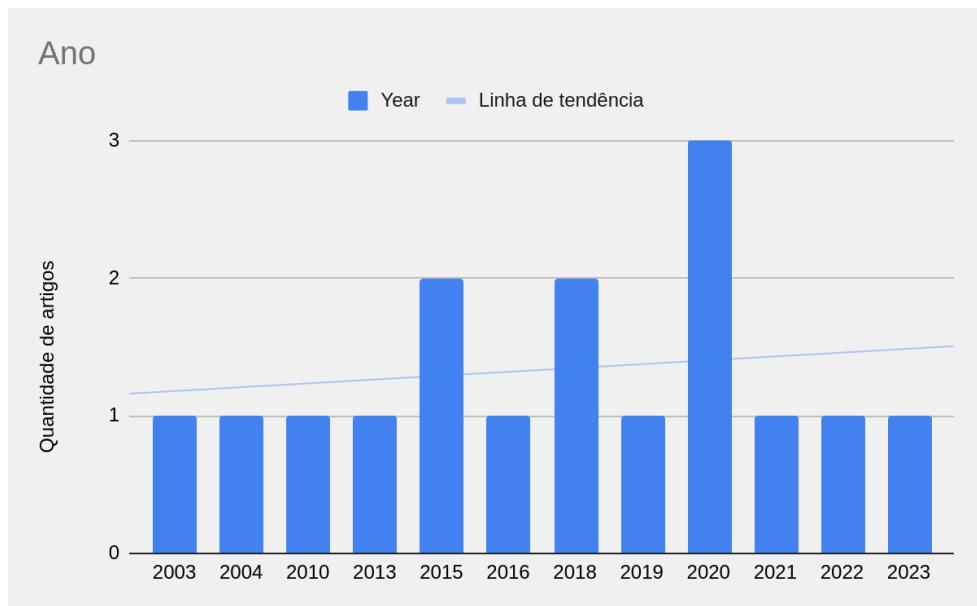
Ao analisarmos as referências, podemos notar que nenhum dos artigos disponibiliza o código fonte utilizado para o desenvolvimento da pesquisa. Isso impossibilita a análise e/ou replicação dos algoritmos e métodos utilizados.

3.6. Grupo “Ano”

A quantidade de publicações em um determinado tema indica a evolução da área de pesquisa com a evolução do tempo. Analisando o Gráfico 4, podemos concluir que a tendência das publicações apresenta um aumento entre os anos de 2015 a 2020. Por outro lado, devido à pequena quantidade de publicações, podemos concluir que se trata

de uma pesquisa com um elevado grau de complexidade ou então uma área com pouco interesse de pesquisadores.

Gráfico 4. Quantidade de artigos de referência publicados por ano. Linha de tendência das publicações de artigos.



4. Análise de referências

De todos os artigos referenciados, selecionamos 3 que consideramos relevantes a serem citados e analisados. O primeiro é intitulado “*Omnidirectional Vision for Mobile Robot Human Body Detection and Localization*” e foi publicado na conferência *IEEE International Conference on Systems, Man and Cybernetics* no ano de 2010. Neste trabalho Liu et al. apresenta um método que detecta, rastreia humanos e estima a distância do humano para a câmera omnidirecional. Esse artigo é importante, pois é um dos precursores no uso de aprendizado de máquina neste tipo de problema. Para determinar a distância estimada entre o humano e o robô, são utilizados procedimentos matemáticos e aspectos geométricos da construção dos objetos [LIU, 2010].

O segundo artigo é “*Outdoor Human Detection with Stereo Omnidirectional Cameras*” publicado no *Journal of Robotics and Mechatronics* em 2020. Nessa pesquisa, com os mesmos objetivos da anterior, são utilizados algoritmos que se baseiam em *Deep Learning* com o uso do algoritmo denominado YOLO. Neste caso, como são utilizadas duas câmeras a perspectiva de espaço se torna maior e, por conta disso, é possível aproveitar da geometria da construção feita (com duas câmeras alinhadas lado a lado numa distância fixa) para estimar a distância desejada [TANAKA, 2020].

O último trabalho é intitulado “*Human Tracking in Top-view Fisheye Images with Color Histograms via Deep Learning Detection*” publicado no *Journal of Imaging* no ano de 2021. Neste artigo, utiliza-se do método RAPiD (Rotation-Aware People Detection in Overhead Fisheye Images) onde o objetivo é determinar uma caixa de

contorno com centro, dimensões e o ângulo de visão. Baseado em algoritmos estatísticos, os humanos são rastreados utilizando histogramas de cores com base em espaços de cores RGB e HSV. No método de estimativa de distância são utilizados métodos procedurais e matemáticos [HAGGUI, 2021].

5. Desenvolvimento de sistema para cálculo da distância

Através dos resultados obtidos na pesquisa bibliográfica, iniciamos o processo de construção de um ambiente para aplicar os métodos estudados e analisados durante a pesquisa. Isso permitirá analisar experimentalmente métodos e abordagens que foram estudados durante a pesquisa.

Para isso, selecionamos a câmera Insta 360 ONE X2 (INSTA360 X2, 2021) capaz de captar imagens em ângulos de até 360 graus. Ela captura fotos numa resolução de 6080x3040 (2:1) para as imagens de 360 graus e 4320x1440 (3:1) para uma imagem panorâmica (*instapano*) de 200 graus de amplitude.

Com objetivo de usar a câmera para estimar a distância entre ela e um humano, está sendo desenvolvido um código para acessar o SDK do fabricante. A base deste código faz referência ao demo fornecido pela empresa fabricante no Git (INSTA360 GIT, 2022).

Observando a Figura 1, Apresentamos uma imagem com 200º de amplitude no ambiente de testes. A foto tem o propósito de ser o ponto de partida para a aplicação de abordagens já utilizadas nas referências.

Figura 1. Foto capturada no ambiente de teste utilizando a câmera Insta 360 ONE X2.



A partir desses insumos, iniciamos um conjunto de implementações dos algoritmos mais importantes revelados pela pesquisa bibliográfica, com objetivo de aprimorar técnicas e obter resultados mais efetivos dentro do contexto de pesquisa.

6. Conclusão

A análise realizada através dessa pesquisa bibliográfica permite compreender o grau de maturidade do processo de estimativa de distância, detecção e rastreamento de humanos com o uso de câmeras omnidirecionais. Ainda existem diversos desafios neste processo para torná-lo eficiente e suficiente para utilização em robôs móveis atuando em sociedade respeitando as normas de proxémica.

Vale mencionar, que nenhum dos projetos disponibilizou características dos objetos de pesquisa (sexo, idade, altura, etc), códigos fonte e para algumas referências nem mesmo o dataset. A omissão no detalhamento desses aspectos acrescenta uma dificuldade adicional para a evolução da área de pesquisa e deve ser revisada em divulgações futuras.

Por fim, o número reduzido de artigos que utilizam somente câmeras omnidirecionais em suas abordagens mostram a incipienteza do campo. Isso permite inferir que ainda existem diversas linhas de pesquisa a serem exploradas e utilizadas em projetos futuros.

Referências

- (ABABSA, 2019) ABABSA, F. et al. 3D Human Tracking with Catadioptric Omnidirectional Camera. ICMR'19: PROCEEDINGS OF THE 2019 ACM INTERNATIONAL CONFERENCE ON MULTIMEDIA RETRIEVAL, Univ Evry, Ibisc Lab, Evry, France, v. 19, p. 73-77, 5 jun. 2019.
- (ABABSA, 2020) ABABSA, F. et al. 3D Human Pose Estimation with a Catadioptric Sensor in Unconstrained Environments Using an Annealed Particle Filter. SENSORS, Univ Evry, IBISC Lab, F-91000 Evry, France, v. 20, p. 6985-6708, 7 dez. 2020.
- (DENIZ, 2004) DENIZ, O. et al. Useful computer vision techniques for human-robot interaction. IMAGE ANALYSIS AND RECOGNITION, PT 2, PROCEEDINGS, Univ Las Palmas Gran Canaria, Dept Informat & Sistemas, Las Palmas Gran Canaria 35017, Spain, v. 3212, p. 725-732, 2004.
- (FURUKAWA, 2003) FURUKAWA, M. et al. Human behavior interpretation system based on view and motion-based aspect models. IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, Osaka Univ, Grad Sch Engn, Dept Adapt Machine Syst, Suita, Osaka 5650871, Japan, v. 1-3, p. 4160-4165, 10 nov. 2003
- (GOUGH, 2017) Gough, D, Oliver, S., & Thomas, J. (2017). An introduction to systematic reviews (2nd ed.). Sage Publications Ltd.
- (HAGGUI, 2021) HAGGUI, O. et al. Human Tracking in Top-view Fisheye Images with Color Histograms via Deep Learning Detection. IEEE INTERNATIONAL CONFERENCE ON IMAGING SYSTEMS AND TECHNIQUES , Univ Montpellier, IMT Mines Ales, EuroMov Digital Hlth Mot, F-30100 Ales, France, p. 1-6, 27 dez. 2021.
- (HAGGUI, 2023) HAGGUI, O. et al. Centroid human tracking via oriented detection in overhead fisheye sequences. VISUAL COMPUTER, Univ Montpellier, EuroMov Digital Hlth Mot, IMT Mines Ales, Ales, France, p. 407-425, 17 fev. 2023.
- (INSTA360 X2, 2021) One-X2 – Insta 360 Brasil. Disponível em: <<https://insta360brasil.com.br/one-x2/>>. Acesso em: 20 mar. 2024.

- (INSTA360 GIT, 2022). Disponível em: <<https://github.com/Insta360Develop>>. Acesso em: 22 mar. 2024.
- (HALL, 1966) Hall, E. T. (1966). *The hidden dimension*. Anchor Books.
- (HARIYONO, 2015) HARIYONO, J. et al. Tracking Failure Detection Using Time Reverse Distance Error for Human Tracking. *CURRENT APPROACHES IN APPLIED ARTIFICIAL INTELLIGENCE*, Univ Ulsan, Grad Sch Elect Engn, Ulsan 680749, South Korea, v. 9101, p. 611-620, 30 abr. 2015.
- (LIU, 2010) LIU, H. et al. Omnidirectional Vision for Mobile Robot Human Body Detection and Localization. *IEEE INTERNATIONAL CONFERENCE ON SYSTEMS, MAN AND CYBERNETICS*, Peking Univ, Shenzhen Grad Sch, Key Lab Machine Percept & Intelligence, Beijing, Peoples R China, p. 2186-2191, 13 out. 2010.
- (NISHIMURA, 2020) NISHIMURA, H. et al. Multiple Human Tracking Using an Omnidirectional Camera with Local Rectification and World Coordinates Representation. *IEICE TRANSACTIONS ON INFORMATION AND SYSTEMS*, Nagoya Univ, Grad Sch Informat, Nagoya, Aichi 4648601, Japan, v. E103D, p. 1265-1275, 30 abr. 2020.
- (SEO, 2014) SEO, D. et al. Omnidirectional Stereo Vision based Vehicle Detection and Distance Measurement for Driver Assistance System. *39TH ANNUAL CONFERENCE OF THE IEEE INDUSTRIAL ELECTRONICS SOCIETY (IECON 2013)*, FuturIST Co Ltd, Ulsan, South Korea, v. 39, p. 5507-5511, 2 jan. 2014.
- (TALAOUBRID, 2022) TALAOUBRID, H. et al. Human Tracking in Top-View Fisheye Images: Analysis of Familiar Similarity Measures via HOG and against Various Color Spaces. *JOURNAL OF IMAGING*, Univ Nizwa, Coll Arts & Sci, Nizwa 616, Oman, v. 8, p. 8(4), 115, 16 abr. 2022.
- (TANAKA, 2020) TANAKA, S. et al. Outdoor Human Detection with Stereo Omnidirectional Cameras. *JOURNAL OF ROBOTICS AND MECHATRONICS*, Univ Nizwa, Coll Arts & Sci, Nizwa 616, Oman, v. 32, p. 1193-1199, 20 dez. 2020.
- (TANG, 2015) TANG, Y. Z. et al. Parametric distortion-adaptive neighborhood for omnidirectional camera. *APPLIED OPTICS*, City Univ Hong Kong, Dept Mech & Biomed Engn, Hong Kong, Hong Kong, Peoples R China, p. 6969-6978, 10 jun. 2015.
- (TANG, 2016) TANG, Y. Z. et al. Parameterized Distortion-Invariant Feature for Robust Tracking in Omnidirectional Vision. *IEEE TRANSACTIONS ON AUTOMATION SCIENCE AND ENGINEERING*, Natl Univ Singapore, Dept Biomed Engn & Adv Robot Ctr, Singapore 117575, Singapore, v. 13, p. 743-756, 2 abr. 2016.
- (YAMANO, 2018) YAMANO, F. et al. Improving the accuracy of a fisheye stereo camera with a disparity offset map. *PROCEEDINGS 2018 12TH FRANCE-JAPAN AND 10TH EUROPE-ASIA CONGRESS ON MECHATRONICS*, Clarion, Saitama, Japan, p. 94-97, 18 out. 2018.

(ZAENI, 2018) ZAENI, I. A. E. et al. An Implementation Of Multi-object Tracking Using Omnidirectional Camera For Trash Picking Robot. ELECTRICAL POWER, ELECTRONICS, COMMUNICATIONS, CONTROLS, AND INFORMATICS SEMINAR (EECCIS), Waseda Univ, Dept Informat Prod & Syst, Kitakyushu, Fukuoka, Japan, p. 154-158, 10 jun. 2018.