



ESCOLA SUPERIOR
DE TECNOLOGIA
E GESTÃO

Politécnico de Leiria
Escola Superior de Tecnologia e Gestão
Departamento de Engenharia Eletrotécnica
Mestrado em Eng.^a Eletrotécnica

SISTEMA DE LOCALIZAÇÃO INTERIOR PARA
OPERAÇÕES LOGÍSTICAS COM REALIDADE
AUMENTADA EM ARMAZÉNS

JAIME MIGUEL CONCEIÇÃO SILVA

Leiria, Setembro de 2023



ESCOLA SUPERIOR
DE TECNOLOGIA
E GESTÃO

Politécnico de Leiria
Escola Superior de Tecnologia e Gestão
Departamento de Engenharia Eletrotécnica
Mestrado em Eng.^a Eletrotécnica

SISTEMA DE LOCALIZAÇÃO INTERIOR PARA
OPERAÇÕES LOGÍSTICAS COM REALIDADE
AUMENTADA EM ARMAZÉNS

JAIME MIGUEL CONCEIÇÃO SILVA

Número: 2182458

Projeto realizado sob orientação do Professor Doutor Hugo Filipe Costelha de Castro (hugo.costelha@ipleiria.pt) e Professor Doutor Carlos Fernando Couceiro de Sousa Neves (carlos.neves@ipleiria.pt).

Leiria, Setembro de 2023

AGRADECIMENTOS

Gostaria de prestar os meus agradecimentos ao Politécnico de Leiria pela oportunidade de poder fazer parte do Mestrado de Eng. Eletrotécnica, bem como aos docentes que me acompanharam neste percurso, em especial aos meus orientadores, Doutor Hugo Filipe Costelha De Castro e ao Doutor Carlos Fernando Couceiro de Sousa Neves pela orientação que me transmitiram ao longo deste projeto. Agradeço ainda aos meus colegas de curso, aos meus amigos, aos meus familiares e a todos que, de uma forma ou outra, contribuíram para a realização deste projeto.

RESUMO

Ao longo dos últimos anos a realidade aumentada (RA) tem sido cada vez mais popularizada a nível mundial, podendo ser utilizada de diversas formas em vários tipos de dispositivos, como tablets, smartphones e óculos de RA.

Esta tecnologia permite sobrepor elementos virtuais à nossa visão da realidade. Estes sistemas já têm sido usados com sucesso em vários setores como, por exemplo, na saúde, no entretenimento, na educação e na indústria. Ao nível dos armazéns industriais novas tecnologias têm sido testadas para melhorar o desempenho das operações.

A melhoria da eficiência dos armazéns desperta interesse a nível académico e industrial, uma vez que vários estudos realizados indicam que este mercado pode atingir valores financeiramente elevados. A logística de um armazém industrial representa uma parte importante dos custos de um produto que tenha de ser armazenado.

Os sistemas com RA utilizados em armazéns pretendem ajudar os operadores a executar as suas tarefas com mais facilidade e rapidez, utilizando a sua localização, reconhecimento de identificadores e sincronização da informação com os demais sistemas da empresa.

Este documento contém uma componente desenvolvida no âmbito do projeto *ARware-Augmented Reality for intelligent WAREhouse management* e tem como base a utilização de sistemas de localização em dispositivos de RA nas operações de armazenamento industrial.

Este sistema de localização é baseado na deteção de marcadores, navegação inercial e odometria monocular. Esta solução permite que, em tempo real, seja determinada a posição e orientação de um dispositivo de RA.

No desenvolvimento deste sistema foram abordadas vários requisitos funcionais, arquiteturas, equipamento, algoritmo de localização interior (*tracking*), sensores, integração de dados e, por fim, testes da componente de Localização Fina (LF).

Tendo em conta os resultados obtidos, embora tenham sido realizados avanços significativos na área da RA, a usabilidade destes sistemas em aplicações de utilização contínua, com projeção 3D baseada no ambiente em torno do operador, ainda tem

que ser melhorada para uma efetiva comercialização e utilização destes sistemas, mas demonstra um potencial elevado perante as indústrias. Apesar desta conclusão, pode afirmar-se que o seu potencial para a indústria pode refletir-se na redução do tempo despendido na procura dos objetos, na redução de erro humano, bem como na facilidade de interligação com os sistemas de gestão.

ABSTRACT

Over the last few years, augmented reality (AR) has been increasingly popularized worldwide, and can be used in different ways on various types of devices, such as tablets, smartphones and AR glasses.

This technology allows us to superimpose virtual elements on our vision of the reality in real-time. These systems have already been used successfully in several sectors, such as healthcare, entertainment, education and industry. At the level of industrial warehouses, new technologies have been tested to improve the performance of logistic operations.

Improving the efficiency of warehouses is of interest at an academic and industrial level, as several studies indicate that this market can reach high financial values. The logistics of an industrial warehouse represents an important part of the costs of a product that has to be stored.

The systems with AR used in warehouses aim to help operators perform their tasks more easily and quickly, using their location, recognition of identifiers and synchronization of information with the othercompany systems.

This document describes a component developed within the *ARware-Augmented Reality for intelligent WAREhouse management* project, and is based on the use of location systems on AR devices for industrial storage operations.

This localization system is based on marker detection, inertial navigation and monocular odometry. This solution allows, in real-time, to determine the position and orientation of an AR device

The system developmentincluded the specification, evaluation and development of several architectures, functional requirements, technical restrictions, equipment, usability, hardware, software, location algorithm *indoor, tracking*, markers, sensors, data integration and, finally, testing the Fine Location (LF) component.

Taking into account the results obtained, although significant advances have been made in the area of AR, the usability of these systems in applications for continuous use, with 3D projection based on the environment around the operator, still has to be improved for an effective commercialization and use of these systems. Nevertheless, it demonstrated a high potential for industrial applications in logistics. Despite this

conclusion, it can be said that its potential for the industry can be reflected in the reduction of time spent looking for objects, in the reduction of human error, as well as in the ease of interconnection with the companies' management systems.

ÍNDICE

Agradecimentos	i
Resumo	iii
Abstract	v
Índice	vii
Lista de Figuras	ix
Lista de Tabelas	xiii
Lista de Abreviaturas	xv
1 Introdução	1
1.1 Motivação	2
1.2 Enquadramento do Projeto	5
1.3 Realidade Aumentada a nível industrial	6
2 Trabalho Relacionado e Tecnologias de Suporte	7
2.1 Início da Realidade Aumentada	7
2.2 Análise de mercado da Realidade Aumentada	7
2.3 Trabalho Relacionado	10
2.4 Tecnologias de Localização interior	14
2.5 Hardware de Suporte	17
2.6 Software de Suporte	20
3 Desenvolvimento de um Sistema de Localização para Realidade Aumentada	27
3.1 Requisitos funcionais	28
3.2 Arquitetura	29
3.3 Equipamento Utilizado	30
3.4 Sensores	34
3.5 Calibração do sistema	35
3.6 Tecnologias de localização em espaços interiores	36
3.6.1 Utilização de marcadores	36
3.6.2 Utilização de pontos-chave	38
3.7 Desenvolvimento algoritmo de localização	39
3.7.1 Recolha de dados dos sensores do Moverio BT-35E	42

3.8	Algoritmo de Estimação de Pose	43
4	Testes e Resultados	47
4.1	Primeira fase de testes	48
4.1.1	Intel® RealSense™ T265 com SDK 2.0	48
4.1.2	Smartphone Huawei P20 com SDK ARCore®	49
4.1.3	Epson Moverio BT-35E com SDK MAXST®	50
4.1.4	Conclusões da primeira fase de testes	51
4.2	Segunda fase de testes	53
4.2.1	Testes com marcadores em ambiente simulação	55
4.2.2	Testes de marcadores em ambiente real	58
4.2.3	Conclusões da segunda fase de testes	62
4.3	Terceira fase de testes	62
4.3.1	Conclusões terceira fase de testes	63
4.4	Teste e demonstração do sistema em ambiente real	65
4.5	Limitações do sistema	72
5	Conclusões e Trabalho Futuro	73
	Bibliografia	75
	<i>Anexos</i>	
A	Análise Comparativa de Óculos de Realidade Aumentada	83
B	Resultados dos testes com marcadores em ambiente simulação (trajetórias)	87
C	Testes de marcadores em ambiente real (trajetórias)	93
	Declaração	99

LISTA DE FIGURAS

Figura 1	Receita global do mercado de aplicações de RA móvel ¹	8
Figura 2	Gráfico relatórios da Digi-Capital®	9
Figura 3	Microsoft® HoloLens.	17
Figura 4	ODG® R-7.	17
Figura 5	Vuzix® M300.	18
Figura 6	Epson® Moverio BT-300.	18
Figura 7	Resultado da estimação da localização de marcadores ArUco	25
Figura 8	Arquitetura geral do projeto ARWare. No "Dispositivo RA", "X,Y" representa os dados enviados originalmente do mó- dulo de localização aproximada, enquanto os dados "X, Y, Z, Q" representam a posição (X,Y e Z) e orientação (Q - quaternião) estimada pelo módulo de LF.	30
Figura 9	Arquitetura da componente de realidade aumentada e loca- lização fina.	31
Figura 10	Epson® Moverio BT-35E.	33
Figura 11	Epson® Moverio BT-300.	33
Figura 12	Logitech® Brio 4k.	33
Figura 13	Intel® RealSense™ T265.	33
Figura 14	Total Leica® TS02.	34
Figura 15	Marcador do tipo ArUco.	37
Figura 16	Marcador do tipo MAXST®.	38
Figura 17	Diagrama do algoritmo localização fina.	40
Figura 18	Diagrama do algoritmo BT-35E Sensors.	43
Figura 19	Diagrama detalhado do filtro de Kalman.	45
Figura 20	Ambiente testes com equipamentos a avaliar.	47
Figura 21	Circuito de testes Intel® RealSense™ T265 com SDK 2.0 e Huawei P20 com SDK ARCore®	48
Figura 22	Pose estimada usando a Intel® RealSense™ T265	49
Figura 23	Resultados de estimação da pose utilizando o <i>smartphone</i> Huawei P20 com SDK ARCore®.	50
Figura 24	Circuito de teste com os óculos Epson Moverio BT-35E com SDK MAXST®.	51

Figura 25	Teste de caminho linear (Epson Moverio BT-35E + MAXST® + Unity®).	52
Figura 26	Dados da pose para o caminho linear (Epson Moverio BT-35E + MAXST® + Unity®).	53
Figura 27	Visão geral da posição obtida com a Intel® T265, o Samsung® Galaxy A7 (ARCore®) e o Huawei® P20 (ARCore®) ao longo da realização de 3.5x um percurso circular no laboratório.	54
Figura 28	Ambiente de testes no software RobotStudio.	55
Figura 29	Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).	56
Figura 30	Gráfico de posição em 3D dos marcadores relativamente à câmara.	57
Figura 31	Gráfico do erro da posição e orientação estimadas, por eixo, de cada marcador relativamente à câmara.	57
Figura 32	Gráfico de erro de estimação de posição e orientação por marcador.	58
Figura 33	Ambiente real de testes de marcadores.	59
Figura 34	Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).	60
Figura 35	Gráfico de posição em 3D dos marcadores relativamente à câmara.	60
Figura 36	Gráfico de erro de posição e orientação, por eixo, de cada marcador relativamente à câmara.	61
Figura 37	Gráfico de erro de posição e orientação por marcador.	61
Figura 38	Ambiente real para testes com o ser humano equipado com os óculos de RA.	63
Figura 39	Vista de topo do ambiente real com o percurso a realizar pelo ser humano no laboratório (as cores dos círculos, acima da imagem, corresponde a legenda da setas na imagem).	64
Figura 40	Vista de topo do ambiente real com os marcadores espalhados no percurso a realizar pelo ser humano no laboratório.	64
Figura 41	Gráfico de posição em 3D (teste com o utilizador).	65
Figura 42	Vista da projeção das setas no ecrã dos óculos, com os marcadores ao longo dos corredores do laboratório e projeção gráfica algoritmo de otimização de rotas.	66

Figura 43	Vista com os marcadores ao longo dos corredores do armazém da empresa Matceramica.	67
Figura 44	Teste em ambiente real, armazém da Matceramica.	67
Figura 45	Mapa 2D (vista de topo) do ambiente real com o percurso a realizar pelo ser humano.	68
Figura 46	Vista com os marcadores ao longo das prateleiras do armazém da empresa Matceramica.	69
Figura 47	Visão geral da posição estimada ao longo do percurso realizado no armazém da Matceramica.	70
Figura 48	Visão eixo z e y da posição estimada com o algoritmo ao longo do percurso realizado no armazém da Matceramica.	70
Figura 49	Visão eixo z e x da posição estimada ao longo do percurso realizado no armazém da Matceramica.	71
Figura 50	Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).	87
Figura 51	Gráfico de posição em 3D dos marcadores relativamente à câmara.	88
Figura 52	Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).	88
Figura 53	Gráfico de posição em 3D dos marcadores relativamente à câmara.	89
Figura 54	Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).	89
Figura 55	Gráfico de posição em 3D dos marcadores relativamente à câmara.	90
Figura 56	Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).	90
Figura 57	Gráfico de posição em 3D dos marcadores relativamente à câmara.	91
Figura 58	Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).	91
Figura 59	Gráfico de posição em 3D dos marcadores relativamente à câmara.	92

Figura 60	Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).	93
Figura 61	Gráfico de posição em 3D dos marcadores relativamente à câmara.	94
Figura 62	Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).	94
Figura 63	Gráfico de posição em 3D dos marcadores relativamente à câmara.	95
Figura 64	Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).	95
Figura 65	Gráfico de posição em 3D dos marcadores relativamente à câmara.	96
Figura 66	Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).	96
Figura 67	Gráfico de posição em 3D dos marcadores relativamente à câmara.	97
Figura 68	Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).	97
Figura 69	Gráfico de posição em 3D dos marcadores relativamente à câmara.	98

LISTA DE TABELAS

Tabela 1	Comparação de óculos de realidade aumentada.	18
Tabela 2	Características de <i>Software development kit</i> (SDKs) para RA	20
Tabela 3	Características dos SDK com relação com operações logísticas	21
Tabela 4	SDKs e respectivas plataformas	22
Tabela 5	Análise comparativa de modelos de óculos de realidade aumentada (-"significa que não foi possível apurar).	83

LISTA DE TABELAS

LISTA DE ABREVIATURAS

ESB	<i>Enterprise Service Bus.</i>
EULA	<i>End-User License Agreement.</i>
FOV	<i>Field Of View.</i>
FPS	<i>Frames Per Second.</i>
GPS	<i>Global Positioning System.</i>
HMD	<i>Head Mounted Display.</i>
ICP	<i>Iterative Closest Point.</i>
IDE	<i>Integrated Development Environment.</i>
IV	<i>Infra-Vermelho.</i>
JSON	<i>JavaScript Object Notation.</i>
LF	Localização Fina.
LOF	<i>Local Outlier Factor.</i>
LRAFI	<i>Laboratório de Robótica Avançada e Fábricas Inteligentes.</i>
RA	Realidade Aumentada.
RFID	<i>Radio Frequency Identification.</i>
RSSI	<i>Received Signal Strength Indicator.</i>
RTLS	<i>Real Time Location System.</i>

Lista de Abreviaturas

SDK	<i>Software Development kit.</i>
SLAM	<i>Simultaneous Localization And Mapping.</i>
UDP	<i>User Datagram Protocol.</i>
UHD	<i>Ultra-High Definition.</i>
USB	<i>Universal Serial Bus.</i>
UWB	<i>Ultra-Wideband.</i>
VIO	<i>Visual Inertial Odometry.</i>
VPU	<i>Vision processing unit.</i>
WMS	<i>Warehouse Management Software.</i>

INTRODUÇÃO

Atualmente a indústria está a ser alvo de uma nova revolução, a chamada Indústria 4.0, resultado da integração e utilização de novas tecnologias, como é o caso da realidade aumentada (RA). Nesta designação existem várias tecnologias e métodos disruptivos, tais como a *Internet Of Things* (IOT), robótica colaborativa, sistemas de ciber-segurança, *big data* e a melhoria da eficiência dos processos, entre outros. Estas alterações levam à mudança da forma como o colaborador executa determinadas tarefas e, para que isto seja possível, torna-se necessária a interligação entre os colaboradores e os equipamentos.

Apesar de já existirem elementos de gestão que permitem a melhoria significativa da organização de um armazém, mesmo assim, em muitos casos as operações no armazenamento industrial são ainda pouco automatizadas, levando assim a um aumento do tempo gasto nas diversas operações (François, 1989). Com o surgimento da Indústria 4.0 passou a ser possível a utilização de novas tecnologias, como é o caso da RA, que tem vindo a ser explorada e desenvolvida nos últimos tempos, existindo uma maior adoção e um maior investimento, fruto dos avanços no hardware e no software (Weeks, 2021).

Através dos estudos desenvolvidos no âmbito deste projeto, percebeu-se que existem ainda vários aspetos a serem abordados e melhorados como, por exemplo, o seu desempenho em termos de localização. Este é o principal foco deste trabalho, a "localização fina"(LF) do equipamento de RA (doravante chamada de localização do utilizador), estimando a posição e orientação, tridimensional, do utilizador, com recurso à RA, informação fundamental, que serve de base para a projeção dos elementos de RA de forma coerente, confortável e contextualizada.

O trabalho realizado foi dividido em cinco etapas: ao longo da primeira etapa foram realizadas pesquisas de forma a entender a necessidade da implementação da tecnologia para operações de armazéns; na segunda etapa foi estudado o estado da arte relativamente aos projetos e tecnologias existentes relacionados com a localização interior; o desenvolvimento da aplicação decorreu ao longo da terceira etapa, onde foram abordados vários equipamentos, hardware, software e ainda foi construído um sistema de localização com base nos óculos inteligentes (dispositivos

eletrônicos que combinam tecnologias imersivas e interativas aos seus utilizadores, sendo também designados como óculos de RA neste documento); de seguida foram realizados vários testes com os desenvolvimentos da componente de localização; na última etapa foi feita a conclusão sobre o trabalho desenvolvido.

O restante documento está organizado da seguinte forma:

- No Capítulo 1 é apresentado a motivação do projeto, o seu enquadramento e, por fim, uma breve análise da RA a nível industrial;
- No Capítulo 2 encontra-se uma descrição dos estudos referentes ao início da RA, estudo de mercado da RA, trabalhos relacionados, tecnologias de localização interior, hardware e softwares de suporte;
- No Capítulo 3 são enunciados os desenvolvimentos, tais como, requisitos funcionais, arquiteturas, equipamentos, tecnologias de localização em espaços internos, sensores e, por fim, desenvolvimento de base e integração dos dados;
- No Capítulo 4 são apresentadas as restrições técnicas, os testes realizados e a demonstração final do sistema em ambiente real;
- No Capítulo 5 são apresentadas as conclusões e trabalho futuro.

1.1 MOTIVAÇÃO

As empresas que realizam operações de logística em armazéns industriais, têm-se deparado com novos desafios na gestão dos processos. O aumento do poder de compra e o aparecimento de métodos tecnológicos (compras *online*), tem levado a um aumento dos custos associados às operações de logística. Estes custos mantêm-se pelo facto das operações ainda serem realizadas recorrendo a métodos físicos, como por exemplo, a utilização de papel. Estas operações de armazenamento representam cerca de 20% dos custos em termos logísticos, sendo que a tarefa de encontrar um objeto representa cerca de 65% do custo total das operações em armazenamento (Carlos e Figueiredo, 2016).

Os atuais métodos de operação em armazém necessitam de melhorias, de forma a facilitarem as operações aos seus operadores. Nos últimos anos têm surgido novas iniciativas com vista a otimizar as operações em armazéns industriais simplificando as suas tarefas.

Ao longo dos anos a gestão de armazéns tem sofrido várias mudanças de forma a melhorar a eficiência das suas operações. Uma das maiores necessidades nas

empresas que realizam a gestão intralógica de armazéns começa imediatamente no planeamento e mapeamento, que se traduz numa maior necessidade de melhores ferramentas de mapeamento e organização dos seus espaços. O relatório elaborado pela Vaisala, apelidado de “GMP *Warehouse Mapping*” aponta que o mapeamento de um armazém industrial deverá ser dividido em nove fases cruciais (VAISALA, 2020):

1. A criação de um plano de validação;
2. Identificação de áreas de risco;
3. Desenvolvimento/integração de protocolos de informação;
4. Planeamento da distribuição de sensores;
5. Escolha das tecnologias apropriadas;
6. Realização da configuração de equipamentos de mapeamento;
7. Realização de testes e revisão de resultados dos mesmos;
8. Realização de modificações;
9. Documentação e agendamento de testes de eficiência de mapeamento.

Atualmente o mapeamento de armazéns é realizado sobretudo com recurso a software de modelação 2D ou 3D como, por exemplo, o Autocad Architecture¹ ou o Smartdraw², que não estão preparados para ter em conta todos os passos recomendados para um planeamento eficaz dos armazéns, limitando o utilizador à conceção do *design/layout* do armazém e não permitindo a localização de sensores, os protocolos de informação, etc. No entanto, as tecnologias têm vindo a evoluir ao longo das últimas décadas e as empresas que realizam a gestão logística e o *picking* (processo de separação dos pedidos) dentro de armazéns começaram a acompanhar de perto o estado da arte tecnológico, nomeadamente ao nível da eletrónica e informática, de modo a modernizar e a tornar os seus armazéns cada vez mais eficientes.

Ao nível das operações de *picking* nos armazéns, hoje em dia é ainda bastante comum que estas sejam realizadas recorrendo à utilização de instrumentos de trabalho como papel ou leitores de códigos de barras, tornando assim os processos lentos e sujeito a erro humano. No entanto, a existência do fator humano dentro da cadeia de operações não pode ser ignorada, sobretudo pela sua capacidade em detetar e resolver conflitos e ambiguidades que possam surgir durante a operação.

1 <http://www.autodesk.com/products/autocad-architecture/overview>

2 <https://www.smartdraw.com/floor-plan/warehouse-layout-design-software.htm>

Estes meios de organização são ainda bastante utilizados em vários armazéns à escala mundial. Apesar disso, têm vindo a surgir novos paradigmas como a Indústria 4.0 ou a Internet das Coisas (*Internet Of Things*), que têm vindo a revolucionar o modo como as tecnologias e as indústrias se regem, ao mesmo tempo que contribuem para a sua eficiência. Neste âmbito, uma das tecnologias que tem mostrado grande promessa em revolucionar as operações de várias áreas, nomeadamente dentro dos armazéns comerciais, é a RA. A RA ganhou algum destaque no âmbito da indústria dos jogos, sendo o jogo Pokémon Go um dos primeiros exemplos de uma aplicação que apostou fortemente nesta tecnologia, que atualmente já se encontra em diversos setores industriais, podendo assim agregar valor aos armazéns de logística (Paavilainen et al., 2017).

Segundo um estudo realizado pela DHL® Trend Research (Cirulis e Ginters, 2013), esta tecnologia poderá ser aplicada em quatro categorias dentro da logística: as operações de armazém, a otimização de transportes, as entregas de última hora e serviços avançados de valor acrescentado. De facto, a DHL®, (Cirulis e Ginters, 2013), um dos maiores operadores inseridos na área da logística, idealizou um sistema de *picking*, em parceria com a empresa Ubimax, que permite identificar itens do armazém recorrendo a óculos inteligentes. Esta solução ficou apelidada de xPick³ e, segundo a DHL®, melhorou a sua eficiência da operação em cerca de 25%. Este sistema recorre ao uso de óculos inteligentes e a um software que apresenta graficamente qual o item a ser recolhido, a sua localização, quantidades e o próximo item em fila de espera para recolha. Ademais, o software tem em conta uma representação gráfica do *trolley* de recolha que permite indicar ao utilizador o local onde o item recolhido deve ser colocado para manter a organização.

Um outro exemplo é o sistema de *Visual Manual Picking* desenvolvido pela empresa Knapp®. Este consiste em realizar operações de *pick and place* de itens numa quantidade específica (More e More, 2020). Também a Picavi® criou um sistema que fornece informações em tempo real, guiando o utilizador para os elementos a buscar, constantemente conectado ao WMS ou ERP. A Generix® desenvolveu também um sistema utilizando óculos inteligentes, orientando o utilizador de forma a localizar itens num armazém industrial, também esta conectada ao WMS (M. K. Williams, 2019).

As soluções desenvolvidas permitem a utilização de óculos inteligentes para escolhas rápidas de listas de equipamentos dispostos num armazém. No entanto estes desenvolvimentos têm apresentado diversas dificuldades, nomeadamente na

3 <http://www.ubimax.de/index.php/en/products#solutions>

leitura dos códigos de barras e na projeção da imagem 3D em tempo real (Skuirowski et al., 2022).

1.2 ENQUADRAMENTO DO PROJETO

O trabalho aqui descrito insere-se no projeto ARWare, do qual fiz parte. Este projeto tem como base o desenvolvimento de uma plataforma para apoio à gestão das operações de logística em armazéns, tendo por recorrente a sistemas IoT, RA, processamento de imagem e algoritmos inteligentes. A maioria do desenvolvimento foi realizado no Laboratório de Robótica Avançada e Fábrica (LRAFI) do Departamento de Engenharia Eletrotécnica da Escola Superior de Tecnologia e Gestão do Politécnico de Leiria, Portugal. Parte dos testes finais foram realizados nos armazéns da empresa Matcerâmica, na zona de Fátima, Portugal.

O projeto *ARWARE* está direcionado para a área industrial, tendo como objetivo desenvolver uma plataforma para apoio à organização, gestão e melhoria da eficiência das operações de logística e *picking* em armazéns de média ou grande dimensão de empresas fabricantes, distribuidoras e/ou operadores logísticos. Pretende-se que a combinação de todas as componentes do projeto *ARWARE* permita criar informação contextualizada com a ação que o operador esteja a desempenhar naquele exato momento, em tempo real, mediada pelo uso de equipamentos compatíveis com tecnologia de RA.

Uma das previsões para o futuro da distribuição é exatamente essa: a *Internet das Coisas*. Agora, já é possível vestir a tecnologia, isto é, o projeto tem como objetivo desenvolver ferramentas capazes de otimizar as operações nos armazéns. Ferramentas idênticas já se apresentam disponíveis no mercado e são capazes de guiar os operadores dentro de um armazém, permitindo que o operador seja orientado em tempo real, não existindo ainda um produto comercial totalmente apto para suprimir as necessidades das operações logísticas em armazéns, especificamente.

A grande marca da investigação neste tipo de tecnologias é a posição e orientação dos óculos de RA que o operador utiliza. De forma a conseguir realizar os percursos, pré definidos dentro do armazém de uma forma mais eficiente, é necessário que a localização seja fina.

1.3 REALIDADE AUMENTADA A NÍVEL INDUSTRIAL

A RA pode ser definida como uma tecnologia de sobreposição de camadas de informação digital sobre o ambiente real, podendo o utilizador recorrer a esta tecnologia por meio de smartphones, tablets, óculos, lentes e outros dispositivos.

Embora a RA seja atualmente utilizada em inúmeras aplicações, há um forte investimento para a melhoria contínua dessas aplicações, bem como para alargar a utilização da RA para outras aplicações, com outros âmbitos como, por exemplo, redes sociais, saúde e bem-estar, jogos, educação e indústria (Jha et al., 2019).

Graças à utilização destas novas tecnologias, muitos investigadores estão a apresentar aplicações bastante interessantes, que ajudam no dia a dia dos utilizadores. Por exemplo, o IKEA® apresenta uma aplicação de decoração⁴ que auxilia no planeamento de um espaço com a inserção de móveis, cuja utilização permite aos compradores testar como um determinado produto se encaixará no seu espaço, tanto em termos de adequação ao espaço físico quanto em estilo (Ozturkcan, 2021).

Também o CEO do Facebook®⁵, Mark Zuckerberg, é um exemplo disso, dizendo que a realidade virtual (RV) e RA são a próxima plataforma computacional. A empresa pretende investir o seu capital, com uma quantia de 3 biliões de dólares americanos na próxima década para incentivar o avanço dessas tecnologias.

No setor industrial, existem várias empresas com bastante reputação como a General Electric®, a FedEx®, a Porsche® e a Mitsubishi Electric®⁽⁶⁾ que já utilizaram RA em diversas áreas, como no fabrico, no controlo de qualidade e na manutenção (Cohen et al., 2018). Ainda a The New Yorker®, a Vespa®, a Volkswagen®, a Manor®, a Moosejaw® e a Lego®, já têm aplicações que permitem ao utilizador interagir com os produtos que estas marcas oferecem, utilizando RA, beneficiando de uma melhor experiência de utilizador, sendo estas, na sua maioria, estratégia de marketing (Tsai, 2019).

4 <https://www.architectmagazine.com/technology/ikea-launches-augmented-reality-application>

5 <https://www.bbc.com/news/technology-57942909>

6 <https://www.thevrara.com/blog2/2021/9/14/virtual-reality-is-revolutionizing-enterprise-and-industrial-training-see-success-stories-from-walmart-verizon-porsche-bp-henkel-fedex>

TRABALHO RELACIONADO E TECNOLOGIAS DE SUPORTE

O crescimento e os avanços da indústria introduziu sistemas de RA no seu dia-a-dia, com a possibilidade de detecção de marcadores que, além de ajudar na projeção/sobreposição de imagens na vista do utilizador, ajuda também a geolocalização do equipamento em causa. Assim, foram estudados diversos equipamentos, softwares e algoritmos de localização interior para utilização com RA.

2.1 INÍCIO DA REALIDADE AUMENTADA

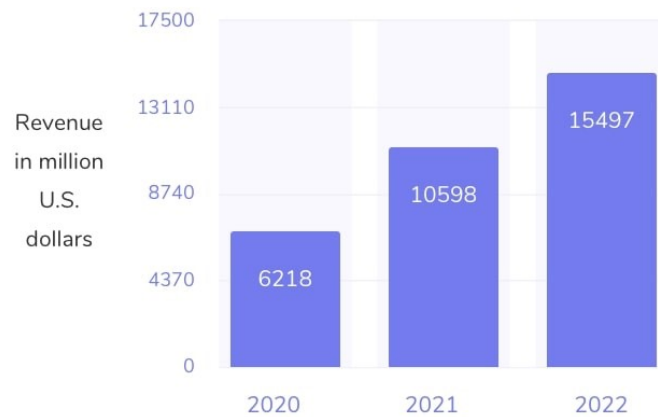
O início do estudo da RA ganha forma quando Ivan Sutherland publicou um artigo em 1968, *A Head-Mounted Three Dimensional Display*. (Sutherland, 1968), sendo que o termo RA foi utilizado pela primeira vez por Tom Caudell na empresa Boeing®, devido ao desenvolvimento de um sistema de fabrico de máquinas de voo, por volta dos anos 90 (Janin et al., 1993). Ainda nessa altura, devido ao projecto ARVIKA, a RA começou a ganhar proporção. Daí em diante houve um crescimento considerável da área, e empresas como a Airbus®, a EADS®, a BMW® e a Ford®, desenvolveram e promoveram as suas aplicações em RA (Cohen et al., 2018). O lançamento destas aplicações fez com que vários investigadores ganhassem interesse nesta área, desenvolvendo vários projetos. Um dos principais acontecimentos, que terá despertado a indústria e potenciado um maior crescimento da área, é apontado em 2013, quando foram lançados os óculos inteligentes da Google®.

2.2 ANÁLISE DE MERCADO DA REALIDADE AUMENTADA

A RA tem beneficiado muito com os avanços tecnológicos, sendo hoje em dia bastante mais fiável, mais eficiente do ponto de vista energético e de capacidade computacional, além de ter um custo mais baixo, comparativamente às soluções inicialmente disponíveis (Reljić et al., 2021).

Devido a estes avanços, tem sido possível observar uma maior tendência para a utilização de aplicações que recorrem a RA. Este facto deve-se à grande utilização das *app stores* neste tipo de aplicações. Com isto, as empresas de vários ramos têm aumentado o seu interesse pela RA. A empresa Statista®¹ especializada em dados de mercado e consumidores, apresenta a evolução da receita em aplicações móveis de RA dos últimos anos. Como se pode observar na Figura 1, os dados recolhidos mostram que, no ano 2022, o mercado atingiu 15,497 milhões de dólares Americanos.

Global revenue of the consumer mobile AR app market (2020 to 2022)



Source: Statista 2020

Visualized by Riseapps

Figura 1: Receita global do mercado de aplicações de RA móvel¹.

As indústrias fortemente impactadas pelas iniciativas de transformação digital e da IoT estão a liderar a adoção da RA. Uma das principais razões é a capacidade da RA servir como uma instrução avançada e ferramenta de orientação. Um exemplo disso é a aplicação do Google® Maps AR ², que projeta indicações para a navegação das trajetórias pedidas, de forma a auxiliar o utilizador nas suas caminhadas ³.

Relatórios de analistas da indústria têm previsto um crescimento para este mercado, principalmente por causa da sua capacidade de transformar fundamentalmente a maneira como as pessoas interagem com os mundos físico e digital. Diversas indústrias que envolvem processos complexos e de alta intensidade de capital, como produtos industriais, automóveis, eletrónicos, alta tecnologia, aeroespacial e defesa, estão a apostar na RA.

1 <https://www.statista.com/statistics/608990/mobile-ar-applications-installed-base-worldwide-by-type/>

2 <https://arvr.google.com/>

3 <https://www.pocket-lint.com/what-is-google-maps-ar-navigation-live-view/>

As projeções dos analistas para o crescimento desse mercado variam, no entanto. Por exemplo, relatórios da Digi-Capital®⁴ (Figura2), estimaram que o mercado da RA móvel possa atingir 3.5 mil milhões de dólares americanos.

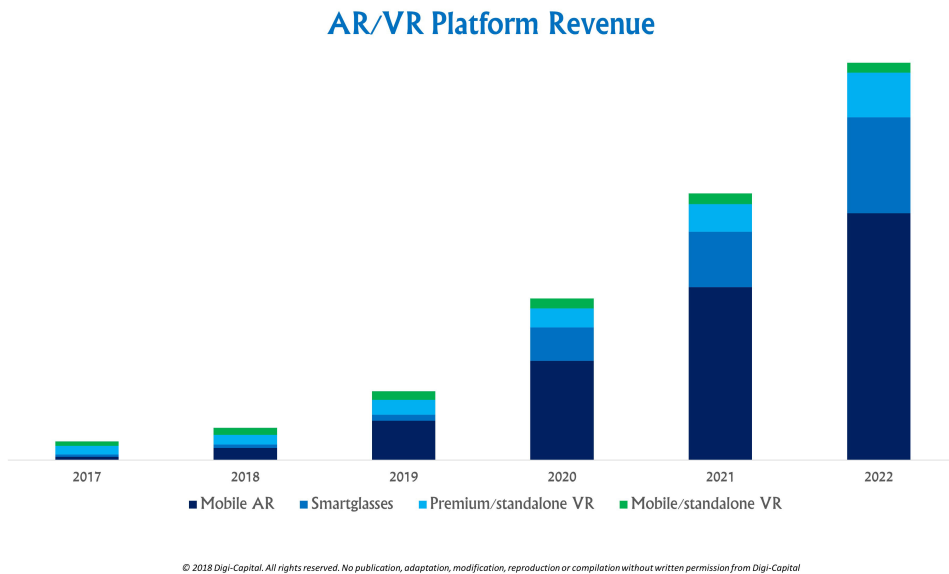


Figura 2: Gráfico relatórios da Digi-Capital®.⁵

De acordo com dados da PTC, 21% dos fabricantes do setor industrial, 11% do setor automóvel, 8% do setor aeroespacial e de defesa lideram a adoção da RA, tendo um grande número de fornecedores de serviços de software emergido para suportar essa tecnologia. Já existem exemplos vindos dos diferentes tipos de indústria, demonstrando a aplicabilidade horizontal desta tecnologia (Campbell et al., 2017).

Em setores que incluem processos, configurações e procedimentos, onde pequenos erros ou desvios podem incorrer em custos ou causar sérios danos, esta capacidade pode fazer uma diferença significativa. Por exemplo, a Ford®, que contém um conjunto de modelos de carros, com os mais diversos estilos, faz com que a produção tenha de mudar rapidamente e a RA veio ajudar na montagem dos revestimentos dos tetos dos carros. Este desenvolvimento auxilia os operadores a colocar no local correto diversos componentes de cada modelo como, por exemplo, os fios elétricos, diminuindo assim o erro humano, aumentando a produtividade e a eficiência do processo⁶.

⁴ <https://techcrunch.com/2018/01/25/ubiquitous-ar-to-dominate-focused-vr-by-2022/>

⁵ <https://techcrunch.com/2018/01/25/ubiquitous-ar-to-dominate-focused-vr-by-2022/>

⁶ <https://www.lightguidesys.com/resource-center/blog/6-uses-of-augmented-reality-for-manufacturing-in-every-industry/>

Outros estudos de mercado revelam ainda que a utilização da RA nas empresas aumenta significativamente os seus lucros. Um dos estudos onde é demonstrada esta afirmação é fornecido pela Juniper Research⁷, tendo este por conclusões finais que os lucros de cada empresa que usa RA aumentaram até cerca de 2,4 biliões de euros em 2019.

Um outro exemplo relevante é o da empresa DHL®, que realizou um sucesso com um conjunto de experiências no ano de 2018, usando RA para melhorar a eficiência da operação, obtendo melhorias em cerca de 25%, relativamente à não utilização deste tipo de tecnologia ⁸.

2.3 TRABALHO RELACIONADO

A RA atingiu níveis altos de popularidade nos últimos anos, sendo que um dos primeiros projetos mais famosos foi lançado em 1992 pela Boeing® (Caudell e Mizell, 2003). A visualização dos fios no casco da aeronave e a exibição da distribuição dos cabos facilitou as atividades de montagem e reduziu a probabilidade de erros.

O Google® Glass foi testado num conjunto de grandes empresas como, por exemplo, a General Electric® (GE) ⁹ e a Boeing® ¹⁰. Outros utilizadores, como a OneBonsai Inc.® e a BANC Engineering Inc.®, já mediram uma redução de 30% no tempo de treino dos funcionários através da utilização dos óculos (François, 1989).

O dispositivo Wearable Real foi projetado para ser utilizado em plataformas de refinaria de petróleo e co-desenvolvido pela Baker Hughes® (Jacobs, 2018). O dispositivo de RA Real é um equipamento mais robusto em comparação aos dispositivos como o Google® Glass, mas também é um sistema muito volumoso, projetado para projetos de indústria pesada.

Também a ODG® realizou testes com a FedEx® Express num novo produto chamado SAVED, produto este que consiste essencialmente numa máscara de oxigénio equipada com um mostrador *heads-up* usando a tecnologia de óculos

⁷ <https://www.juniperresearch.com/press/enterprise-ar-app-revenues-reach-2-4bn-by-2019>

⁸ <https://www.dhl.com/global-en/delivered/digitalization/dhl-successfully-tests-augmented-reality-application-in-warehouse.html>

⁹ <https://www.ge.com/news/reports/looking-smart-augmented-reality-seeing-real-results-industry-today>

¹⁰ <https://www.theverge.com/2016/7/14/12189574/boeing-google-glass-ar-building-airplane-parts>

ODG®. O objetivo é ajudar os pilotos a pilotar aviões no caso de a cabine se encher de fumo¹¹.

Outro exemplo do processo de treino de funcionários é a colaboração da Jaguar® Land Rover com a Bosh® e a RE‘FLEKT®. A aplicação permite treinar os funcionários sem desmontar os veículos, observando tudo “atrás do painel” através da RA¹².

Em 2015 a Airbus® começou a usar a ferramenta *AR - Smart Augmented Reality Tool* (SART) para o controlo da qualidade na linha de montagem. Desde então, o SART foi instalado em cerca de cem *tablets* e mais de mil funcionários utilizam-no diariamente (Kohn e Harborth, 2018).

A Mitsubishi Electric® criou a tecnologia RA de suporte de manutenção, que confirma a ordem de revisão na exibição de RA (Study, 2021). A ThyssenKrupp, seguindo a mesma filosofia, mudou a aparência clássica da manutenção, usando o Hololens da Microsoft® para criar um serviço de fornecimento preditivo para elevadores - MAX. Esta aplicação ajuda os funcionários do serviço a visualizar e identificar problemas com elevadores e dá acesso sem intervenção manual a informações técnicas (Stoltz et al., 2017).

Outro exemplo de aplicação, ainda em desenvolvimento, é apresentado pela Inglobe Technologies®, em conjunto com a empresa italiana de logística Metra®¹³. Neste caso, é referida a integração do fluxo de trabalho típico de *Warehouse Management System* (WMS), nomeadamente os processos de *picking*, *sorting* e *cargo loading* com as plataformas IoT com que a empresa já trabalha. Este exemplo é bastante idêntico ao projeto apresentado neste relatório, sendo que o mesmo ainda se encontra em desenvolvimento e com pouca informação relativamente à componente de localização do utilizador.

O potencial da RA aplicado ao setor logístico foi também reconhecido por empresas de relevo mundial, como a SAP®. Esta desenvolveu uma solução chamada de *ARWarehouse Picker*¹⁴ que permite realizar essencialmente as mesmas atividades que a solução criada pela DHL®, permitindo utilizar funções de reconhecimento de voz e comandos de áudio para guiar os utilizadores passo a passo nas tarefas de *picking* nos armazéns. Esta solução é também bastante interessante no ponto de

11 <https://www.businesswire.com/news/home/20180412005288/en/ODG-Smartglasses-Become-Perfect-Wingman-in-Emergency-Landings>

12 <https://www.re-flekt.com/portfolio-item/augmented-reality-training>

13 <https://www.inglobetechnologies.com/towards-the-smart-warehouse-logistics-4-0-with-augmented-reality/>

14 <https://pdf.directindustry.com/pdf/sap/read-sap-ar-warehouse-picker-product-brochure/9333-579204.html>

vista deste projeto, melhorando a eficiência nas operações em armazéns, mas não se foca na otimização de rota e por consequência na localização do operador.

Ao nível acadêmico surgem também algumas atividades de investigação com foco na RA aplicada à logística/intralógica. Por exemplo, Sarupuri Bhuvaneshwari (Sarupuri, 2016) propôs o desenvolvimento de um sistema que utiliza equipamentos de RA para ajudar condutores de empilhadores de armazém a manobrar e a acomodar cargas nos mesmos. Neste sentido, o sistema permitia reconhecer paletes em função da sua dimensão e da distância a que estas se encontravam do empilhador, para fornecer sugestões ao nível de como realizar a carga das paletes no empilhador. Por outro lado, Xu e Zhou (Xu e Zhou, 2010) propuseram um modelo de simulação baseado no sistema eM-Plant, em que otimizam o processo de *picking* usando algoritmos genéticos demonstrando ser possível reduzir o tempo e a distância da rota.

Jin et al. (Liu et al., 2010) apresentam uma solução híbrida para o problema de otimização de rotas, começando por gerar uma solução utilizando *Particle Swarm Optimization*, que depois é otimizada com recurso a algoritmos genéticos, demonstrando a viabilidade da solução num exemplo de um armazém navegado por um empilhador. Por sua vez, Alexander Diete et al (Diete et al., 2016) idealizaram também um sistema constituído por *smart glasses* e *smart wristbands*, munidos de diferentes tipos de sensores (ultrassom, pressão e inércia) para monitorizar as ações de recolha e depósito de encomendas nas atividades de *picking* em armazéns. Estes propuseram também o uso de algoritmos de aprendizagem automática para aprender a identificar o tipo de ação que o operador do armazém se encontra a realizar (ex.: se está a pegar num objeto ou a pousá-lo), permitindo também perceber quanto é gasto pelo operador a circular pelo armazém e a sua aceleração média. Para isto foi utilizada a recolha e análise de vídeo, processada através de *smart glasses*. Este vídeo era agregado em *datasets* e categorizado através dos ditos algoritmos de aprendizagem automática numa aplicação num *smartphone*.

De facto, os conceitos/algoritmos como aprendizagem automática irão também assumir um papel muito relevante nas aplicações da Indústria 4.0 e revelam já grande potencial em aplicações IoT. A aprendizagem automática é um campo da inteligência artificial que se foca na aprendizagem de máquinas de modo automático para que estas realizem melhor qualquer tipo de tarefa, baseando-se nas experiências do passado. Segundo a Deloitte, no estudo *The warehouse of tomorrow, today* (Deloitte, 2015), esta tecnologia irá no futuro ser utilizada para agregar toda a informação passada e presente na cadeia de operações dos armazéns, de modo a processá-la para prever e sugerir tomadas de ação, assim como para alertar sobre

ações a evitar, (tudo isto em tempo real) em prol da eficiência da otimização, organização, despacho e todas as outras operações em armazéns.

As tecnologias de aprendizagem automática aliadas à investigação e desenvolvimento têm permitido ultrapassar algumas barreiras nas áreas da visão computacional e do processamento de linguagem natural que, até há relativamente pouco tempo, não apresentavam solução.

A IoT tem vindo a potenciar a localização interior através do recurso a equipamentos como *wearables*. Ao nível da investigação, Zhice Yang e Zeyu Wang (Yang et al., 2015) propuseram recentemente um inovador sistema de localização interior através de luz visível que pode ser utilizado em qualquer dispositivo que permita refletir ou captar luz visível como, por exemplo, *smart glasses* com RA. Este desenvolvimento apresenta uma solução de localização compatível com a utilização de óculos inteligentes, no entanto este sistema de localização sobrecarrega os dispositivos receptores com o tratamento da oscilação da luz, esta transmite pulsos numa alta taxa (acima de 1KHz, para não perturbar os humanos) e com isto a taxa de pulsos excede a capacidade de amostragem da maioria das câmaras (30/60 fps), sendo que é necessário um sensor ótico de maior capacidade de atualização.

Os algoritmos de localização interior têm tido um papel relevante no sector de logística, em particular na localização dos operadores de armazéns industriais. Em 2015, Sivakumar et al (Sivakumar e Venkatesan, 2015) desenvolveram um novo método de minimizar os erros de localização por via de redes *wireless*, através do uso de métodos de localização móveis conhecidos como *Mobile Anchor Positioning* (MAP). No entanto, os seus estudos revelaram que esta aproximação por si só seria insuficiente e, portanto, foram adotadas aproximações meta-heurísticas como o *Bat Optimization Algorithm*, *Modified Cuckoo Search Algorithm* e o *Firefly Optimization Algorithm*. Estes revelaram otimizações na localização obtida, nomeadamente ao nível da precisão dos mesmos e na velocidade de convergência dos algoritmos (Pradhan e Bhende, 2019). Neste caso, a localização do utilizador irá apresentar ainda um erro superior a 0.5m, acima do valor pretendido para este projeto.

Ao longo dos últimos anos têm sido estudados vários métodos que permitam posicionar, por exemplo, *smartphones*, com taxas de erro de localização bastante pequenas. Estes avanços têm sido conseguidos, em boa parte, devido à utilização de algoritmos heurísticos.

2.4 TECNOLOGIAS DE LOCALIZAÇÃO INTERIOR

A localização do utilizador dos óculos de RA e a orientação do aparelho (no limite, a do olhar) no mundo, são muito importantes para a projeção de RA, para que a sobreposição da imagem com a realidade (registro), através dos óculos, seja precisa e confortável.

Esta localização (posição e orientação, daqui em diante denominada de pose), cuja determinação irá aqui ser referida como localização, pode ser calculada através de diversos métodos, tais como a localização relativa a marcadores (daqui em diante será utilizado a denominação de marcador no texto e nas imagens de *tag*, tendo ambas as denominações o mesmo significado), a localização sem marcadores (*markerless tracking*), *dead-reckoning*, *Simultaneous Localization and Mapping* (SLAM) entre outros. Todos os métodos supracitados já foram implementados nas mais diversas áreas da RA (Sandra et al., 2021). A maioria dos programas que suportam este tipo de métodos recorrem à informação disponibilizada pelos equipamentos a partir dos sensores embebidos (como câmaras, giroscópios, acelerómetros e magnetómetros).

Hoje em dia, existem várias tecnologias que podem ser usadas para obter a localização de um objeto, veículo ou pessoa num espaço interior (Mautz, 2012). Contudo, cada tecnologia tem as suas próprias características, o que se traduz na existência de inúmeras soluções com diferentes graus de precisão, com diferentes níveis de exigência em termos de infraestrutura no local, com funcionamento nem sempre uniforme em diferentes ambientes (influenciado pela luz, ocupação do espectro eletromagnético, etc.) e cuja aplicabilidade em grandes áreas nem sempre é possível. A maturidade de algumas das soluções existentes é ainda reduzida face a outras áreas tecnológicas, não existindo uma tecnologia dominante ou uma solução comercial abrangente, universal, para todos os casos.

O localização em ambiente interior é importante para a gestão de recursos, incluindo o *tracking* (seguimento) e a monitorização desses recursos dentro da organização. As tecnologias de localização em espaços interiores podem ser agrupadas em duas grandes categorias: baseadas em tecnologias com infraestrutura e em soluções tecnológicas sem infraestruturas.

Em ambientes interiores, em que não se pode recorrer ao GPS, as tecnologias baseadas na infraestrutura, incluem soluções usando radiofrequência, quer seja *Radio Frequency Identifier* (RFID), sensores *Radio Frequency RF*, *Bluetooth*, *ultra-wideband* (UWB) e Wi-Fi. Ao passo que as soluções sem infraestrutura incluem soluções baseadas em sinais magnéticos e sensores de movimento (unidade inercial).

As soluções sem infraestrutura são, normalmente, preferíveis devido ao seu custo ser inferior e à facilidade de utilização face às soluções com infraestrutura que exigem a pré-instalação e configuração de hardware específico dispendioso (Bolte e Goll, 2020).

Uma das tecnologias mais populares, nos dias de hoje, para a implementação de soluções de localização em espaços interiores, baseia-se na utilização das redes Wi-Fi, tirando partido do facto de estas redes estarem presentes na maioria dos locais. A técnica mais usada é designada de Wi-Fi *fingerprinting* e baseia-se na calibração, utilizando os valores do *Received Signal Strength Indicator* (RSSI). Esta técnica implica que, num primeiro momento, se faça um mapa do sinal de rádio no interior do edifício, procedendo a leituras do ambiente rádio nas várias divisões existentes. Assim, esse mapa consiste no conjunto das assinaturas rádio associadas a um determinado local. Este mapa rádio mostra a lista dos pontos de acesso captáveis em cada local e o respetivo nível de sinal rádio RSSI (Chen et al., 2019). Num segundo momento o mapa rádio, anteriormente construído, é usado para estimar a posição com base na similaridade entre uma assinatura rádio captada num instante e as assinaturas existentes no referido mapa.

Nos edifícios de maior dimensão a construção do mapa rádio é uma tarefa morosa e complexa. Para que o mapa seja detalhado e verdadeiramente útil para o processo de localização, conduzindo a soluções de localização com maior precisão, é necessário que distância física entre os pontos de recolha das amostras seja relativamente reduzida. Existem alguns autores que têm estudado a utilização de modelos matemáticos genéricos para evitar a construção do mapa de rádio (Mama, 2008). Existem ainda soluções que, para fazer face à imensa área de alguns edifícios e à necessidade de atualizar o mapa rádio com alguma frequência, usam soluções baseadas em *crowdsourcing*, em que os vários utilizadores ajudam, de forma colaborativa, a construir e a manter atualizado o mapa rádio.

De referir também o facto de existirem soluções híbridas que conjugam a utilização de várias tecnologias, com vista a aumentar a área de funcionamento e, acima de tudo, com vista a aumentar a precisão. A fusão de sensorial é, hoje em dia, muito usada para melhorar a precisão e acuidade dos sistemas de localização. Esta abordagem baseia-se no uso em simultâneo (fusão) de várias tecnologias que, conjugadas, permitam uma melhor estimativa da localização. Na *Track 3 da IPIN (International Conference on Indoor Positioning and Indoor Navigation) 2017 Indoor Localization Competition*, uma competição de localização em espaços interiores (Torres-Sospedra et al., 2018), organizada no âmbito da conferência internacional sobre este assunto, existem vários participantes que usam informação recolhida dos vários sensores

existentes num telemóvel, incluindo Wi-Fi, acelerómetro, giroscópio, magnetómetro, pressão atmosférica, temperatura, humidade, som e luz. Na edição de 2017, o melhor resultado alcançado pelas equipas foi de 3,48 m, de erro máximo, num modelo em que os participantes não tinham forma de obter os seus próprios dados e assim calibrar ou configurar de forma ótima os seus algoritmos para aquele espaço físico. No ano de 2018 (Renaudin et al., 2018), os melhores resultados foram de 0,9 a 2,1 metros de erro máximo, respetivamente para o 1º e 2º lugares.

Os sistemas ultrassónicos e de infra-vermelhos (IV) (Sauer et al., 2001), apesar de conseguirem bastante precisão, não são os mais indicados para esta aplicação, uma vez que a maioria dos óculos à venda no mercado não têm este tipo de sensores. Este tipo de sistemas são mais utilizados para localização dentro de espaços mais limitados, sendo menos aplicáveis em espaços com áreas maiores, como os armazéns industriais (Tejesh e Neeraja, 2018).

Outro sistema, este com grande precisão, é o RealTrac¹⁵, capaz de detetar e calcular os movimentos de pessoas e de veículos em espaços que podem atingir dimensões consideráveis (e.g portos marítimos), tendo ainda uma precisão em torno de 3 metros. Este sistema é utilizado para aplicações como, por exemplo, prevenção de colisões entre veículos e análise de necessidades de realocação de pessoal dentro de armazéns logísticos, interagindo através de sistemas Wi-Fi. Outras empresas trabalham de forma idêntica, mas utilizando BLE, tais como a Sonitor® (Comuniello et al., 2019).

A determinação da localização do utilizador foi também desenvolvida em outras aplicações como, por exemplo, o Posicionamento Visual Urbano na aplicação AR City¹⁶. Este sistema constrói um mapa do local e direciona o utilizador de acordo com a direção pretendida, projetando o caminho a seguir. Este sistema permite também a navegação dentro de edifícios, sendo o seu principal objetivo a aplicação ao comércio, ao turismo e ao entretenimento.

Algumas das soluções existentes baseiam-se na colocação de marcadores no ambiente (Sauer et al., 2001), recorrendo à imagem das câmaras e algoritmos de processamento dedicados à sua localização. A empresa Eliko®¹⁷ criou um sistema baseada em *Real-time locating system* (RTLS), que consiste em colocar pontos de referência fixos (também denominados de âncoras) em paredes e/ou estantes de um armazém e marcadores junto do objeto que se quer localizar. O sistema mede

15 <https://real-trac.com/en/solutions/logistics/>

16 <https://www.blippar.com/blog/2018/08/09/launching-indoor-visual-positioning-for-enhanced-location-based-augmented-reality>

17 <https://eliko.tech/uwb-rtls-ultra-wideband-real-time-location-system>

o tempo que o sinal via rádio leva a chegar dos marcadores às âncoras. Com esse tempo estima-se a distância a que o objeto que se pretende localizar se encontra de cada marcador e, recorrendo à distância a vários marcadores e ao conhecimento prévio da sua localização no mundo, estima-se as coordenadas do objeto.

2.5 HARDWARE DE SUPORTE

Atualmente existem aplicações de RA em equipamentos como *smartphones*, *tablets* e óculos de RA, ainda que, na sua maioria, estes estejam direcionadas para lazer e não para a indústria. Apesar de alguns dos primeiros equipamentos lançados, como, por exemplo, os Google® Glass, não terem tido sucesso e terem sido descontinuados, nesta altura existem já no mercado diversos equipamentos de RA, tal como os óculos de RA, que podem ser usados tanto para entretenimento como para a indústria.

No caso do projeto apresentado, foi realizada uma pesquisa sobre equipamentos direcionados à RA em armazéns, sendo os resultados apresentados de forma simplificada e resumida na Tabela 1. Na Tabela 5, no Anexo A, é apresentada a Tabela completa onde são apresentados todos os equipamentos analisados no âmbito desta pesquisa.

Desta análise destacam-se quatro tipos de óculos de RA das marcas ODG® (Figura 4), Vuzix® (Figura 5), Epson® (Figura 6) e Microsoft® (Figura 3). Estes enquadram-se na utilização de operações em armazéns, tendo em conta que contêm um conjunto alargado de sensores, tornando assim mais relevantes alguns modelos de óculos de RA destas diversas marcas (mais abaixo será feita uma análise mais detalhada destes equipamentos).



Figura 3: Microsoft® Hololens.



Figura 4: ODG® R-7.

A exigência deste conjunto de sensores nasce da necessidade de estimar a localização e orientação do utilizador para uma correta sobreposição da imagem com a realidade, do ponto de vista do utilizador, nas lentes dos óculos. O campo de visão, a existência de *trackpad* ou a capacidade de *tracking*, são também características a ter em conta na escolha de um equipamento deste tipo.

Modelo	Duração da bateria	Marcaador para navegação interior	Necessidade de dispositivo adicional	Peso	Campo de visão	Preço	SO
Epson Moverio BT-300	6 h com bateria externa	Sim	Sim	69g	23º	850€	Android
Google Glass	1 h	Sim	Sim	49g	14º	1500€	Android
Meta 2	Alimentado por cabo	Não	Sim	500g	90º	1700€	Windows
Microsoft HoloLens	4 h	Não	Não	580g	35º	5000€	Windows
Smart-Eyeglass SED-E1	2 h	Sim	Sim	77g	20º	800€	Android
Vuzix M300	2 h	Sim	Sim	110g	15º	1700€	Android

Tabela 1: Comparação de óculos de realidade aumentada.

Os sistemas de RA também necessitam obter informação com uma taxa de atualização elevada e da ordem de grandeza da persistência do olhar humano, na prática, na região entre as 20 e as 30 imagens por segundo (Walko, 2018). Além disso, para que a interação não se torne desconfortável para o utilizador, o atraso inerente ao processo tem também que ser muito baixo, no máximo na ordem de grandeza das poucas dezenas de milissegundos (Azuma, 1997).

Finalmente, o peso e a autonomia destes equipamentos são características a ter em atenção, sobretudo nos casos em que o utilizador tem necessidade de utilizar o equipamento durante várias horas.



Figura 5: Vuzix® M300.



Figura 6: Epson® Moverio BT-300.

Os Google® Glass foram os primeiros a ser analisados, sendo que, de acordo com os autores do artigo (Tully et al., 2015), estes apresentam uma solução desconfortável e difícil para ler, não tendo tido sucesso enquanto produto orientado para o consumidor e, desde então, está a ser levado para o mercado corporativo.

Os Epson® Moverio BT-300 (Figura 11) contêm um controlador (com a versão 5.1 do Android) para o utilizador interagir com os óculos, possui uma câmara monocular de 5 megapixels e duas telas OLED com resolução 1280p (visão binocular), sendo ainda um dos menos pesados desta análise. A capacidade de bateria e o preço dos óculos levam estes a ser uma das principais opções.

O modelo Hololens da Microsoft®, pelas suas características e capacidade, tem potencial para tornar mais fácil atingir os objetivos do projeto. No entanto, o custo por unidade leva a que uma solução baseada neste equipamento não possa ser facilmente escalada para uma utilização alargada pelos funcionários num típico armazém de logística. O modelo R-7 da ODG® apresentou-se como uma boa aposta, com um pouco menos potencialidade que os anteriores mas, em termos de qualidade/preço, encontravam-se bastante bem situados, podendo ser um dos possíveis equipamentos a comprar de acordo com a verba disponibilizada.

Por sua vez, o modelo M300 da Vuzix®, em termos de qualidade/preço, também se apresentavam muito bem relacionado, relativamente aos anteriores já apresentados, tendo um preço mais acessível. Esse preço permitiria a aquisição de duas unidades para testes mas, no que concerne aos equipamentos e tecnologias descritas anteriormente, este encontra-se um pouco abaixo dos restantes apresentados no que diz respeito à duração da sua bateria e ao seu campo de visão.

Durante a análise dos vários equipamentos disponíveis no mercado foram contactados os respetivos distribuidores e fornecedores. No caso da ODG®, dado que não tinham representante em Portugal, estes modelos foram descartados. Relativamente aos equipamentos da Vuzix®, estes apenas disponibilizavam equipamentos monoculares. Dado que um dos objetivos do projeto era a projeção 3D, tal implicava a utilização de uma solução binocular. Assim, com base na avaliação realizada, o foco foi para a Microsoft® e para a Epson®.

Sendo a nossa escolha pelas duas marcas referidas, foram reunidos esforços para realizar testes com esses equipamentos, tenho existido a interação com os óculos da Microsoft® através de uma instituição de ensino e os óculos da Epson® foram disponibilizados a título de empréstimo por um tempo limitado, pela própria empresa. Quando surgiu essa possibilidade, o principal objetivo foi a recolha de dados reais com cada um deles. Com a análise das informações recolhidas uma das diferenças

verificada foi na interação com o utilizador. Os óculos da Epson® têm um *trackpad* e os Microsoft® têm a possibilidade de interação através do reconhecimento de gestos das nossas mãos. Além disso, os óculos da Epson®, apesar de terem uma duração de bateria maior, são mais leves que os Microsoft®. Com isto, é importante referir que após esta análise foram utilizados os óculos da marca Epson®, sendo que não foi utilizado o modelo BT-300 da Epson® mas sim os BT-35E, sendo que estes contêm características idênticas, embora este último não tenha controlador, tendo assim de ser interligado a um equipamento externo.

2.6 SOFTWARE DE SUPORTE

A nível de software foi realizado um estudo comparativo cujo resultado se resume na Tabela 2, Tabela 3 e Tabela 4.

SDKs	Marcadores			Seguimento na nuvem	Tratamento de oclusões	SLAM	Extended Tracking	Detecção do plano do chão
	Fiduciário	Imagem	3D					
Vuforia	✓	✓	✓**	✗	✗	✓	✓	✓
Wikitude	✓	✓	✓**	✓**	✗	✓**	✓**	✓**
Kudan	✓	✓	✓	✗	✗	✓	✓	-
Maxst	✓	✓	✓	✗	-	✓	✓	✓
ARtoolkit 5	✓	✓		✗	✗	✗	-	✗
EasyAR	✓	-	✓**	✗	✗	✓**	0	-
ARkit 2	✓	✓	✓	✗	✗	✓	-	-
ARcore	✓	-	✗	✗	-	✓	-	-

Tabela 2: Características de *Software development kit* (SDKs) para RA, onde “✓” significa que inclui a funcionalidade, “✓**” significa que a funcionalidade é paga, “0” significa que não existe documentação disponível de raiz com o SDK, “✗” significa que não inclui a funcionalidade e “-” significa que não foi possível apurar.

SDK	Marcadores codificados	Seguimento na nuvem
Vuforia	✓	✓**
Wikitude	0	✓**
Kudan	0	✗
Maxst	✓	✓**
ARtoolkit 5	-	✗
EasyAR	✓	✓**
ARkit 2	0	✗
ARcore	✗	✗

Tabela 3: Características dos SDK com relação com operações logísticas, onde “✓” significa que inclui a funcionalidade, “✓**” significa que a funcionalidade é paga, “0” significa que não existe documentação disponível de raiz com o SDK, “✗” significa que não inclui a funcionalidade e “-” significa que não foi possível apurar.

SDKs	Última Versão	Plataforma alvo			Plataforma de Desenvolvimento			Preço para uso comercial	Licença
		Android	iOS	Windows	web	Game Engines	Outros		
Vuforia	8.0.x	6.0~	✗(9+ mobile)	UWP (10) x64 CPU	✗	Unity 2017.2+	✗	Pro >442.5€	EULA
Wikitude	8.2.x	4.4+ NDK dev supported	9.0+	UWP (10)	JS	2018.3(Beta) 2018.2(v8.1.0)	Cordova Xmarim Titanium	4490€ (Cloud) 2490€ por versão do SDK ou anual	EULA
Kudan	1.5.x	4.0.2+	8.0+	-	✗	Unity 5.2+	✗	1152€/ano	EULA
Maxst	4.0.x	4.3+	9.0+	-	✗	Unity 5.6.2 2017+	✗	531.60/ano 442.9€	EULA
ARtoolkit 5	5.x	4.0.3+	7.0+	Windows 8.1 x86 ou x86-64 CPU	✗	Unity (Unreal 4.19-4.21, unreal AR - 249€)	✗	0	LGPLv3
EasyAR	2.3.0	4.0+ NDK dev supported	7.0+	7+	✗	Unity 4.6+	✗	442.5€ / licença	EULA
ARkit 2	2	✗	12	✗	✗	Unity 2017.4+ unreal 4.20	✗	-	-
ARcore	1.6.x	~7.0+ NDK dev supported	11.0+	UWP (10)	✗	Unity 2017.4.15f1+ unreal 4.20	✗	-	Apache License 2.0 (API subject to GATS)

Tabela 4: SDKs e respectivas plataformas, onde “0” significa que não existe documentação disponível de raiz com o SDK, “✗” significa que não inclui a funcionalidade, “-” significa que não foi possível apurar e *End-user license agreement* (EULA) significa licença específica.

De entre as ferramentas mais utilizadas, destacam-se ferramentas como o ARCore®, Wikitude®, Vuforia e Unity®. Estas já têm inúmeros projetos e aplicações de exemplo, ainda que pouco direcionadas para a indústria. De todo o modo, há exemplos da utilização de diversos algoritmos de localização, reconhecimento e projeção.

Foram testados o ARCore® e o MAXST®. Ambos têm as características necessárias, tais como a capacidade de localizar pessoas ou objetos em movimento, a capacidade de detectar o tamanho e a localização de diferentes tipos de superfícies num espaço físico, a capacidade de lidar com variações de luz e adaptar o tempo de exposição e/ou o ganho à quantidade de luz, interação com o operador, possibilidade de utilizar pontos e/ou âncoras para obter a posição e orientação de um objeto ao longo do tempo e SLAM.

Relativamente ao ARCore®, foi utilizada a versão 1.10.0, descarregada diretamente da Play Store da Google®. Esta é compatível com diversos dispositivos Android, sendo expectável que o número de dispositivos compatíveis aumente com novas gerações de *smartphones*¹⁸. A ARCore® é a sucessora do projeto Tango, que foi cancelado pela Google® no final do ano 2017 (Riedlinger e Oppermann, 2018), sendo este último mais limitado ao nível dos dispositivos compatíveis que a versão inicial da ARCore®, pois exigia um hardware específico que estava presente num número reduzido de modelos disponíveis no mercado. O ARCore® foi sido considerada para os testes laboratoriais pela facilidade e performance anunciadas em termos de utilização com um *smartphones*.

O ARCore® foi projetado para funcionar numa ampla variedade de smartphones equipados com Android 7.0 (Nougat) e versões posteriores, fazendo com que nem todos os modelos Android mais antigos o suportem. A deteção de superfícies planas de referência (como o plano do chão, paredes, ou imobiliário) permite que o ARCore® coloque diferentes objetos digitais em diferentes escalas e posições, de modo a criar experiências RA mais realistas. Este assume que os planos estão imóveis, portanto é desaconselhável tentar ancorar um objeto digital num objeto do mundo real que esteja em movimento. Em geral, é melhor não colocar um objeto até que a sala tenha sido suficientemente mapeada e as superfícies estáticas tenham sido reconhecidas e designadas como pontos-chave (*keypoints*).

Tal como o ARCore®, o MAXST® SDK, permite criar aplicações de RA através do motor de jogo Unity®, software criado pela Unity® Technologies com a capacidade de criar elementos 2D e 3D. Neste âmbito o Unity® é utilizado com a finalidade de

18 <https://developers.google.com/ar/devices>

desenhar os objetos em simulação para posterior projeção no mundo real, visualizados através dos óculos de RA.

De forma a poder desenvolver funcionalidades adicionais, o MAXST® SDK é um *plugin* do Unity® que permite utilizar várias cenas, sendo possível abstrair funções fulcrais da aplicação como, por exemplo, desenvolver uma cena de RA e outra cena responsável pela localização do utilizador. Em termos de *tracking* esta permite desenvolver e integrar aplicações de reconhecimento de imagens/modelos, bem como algoritmos de deteção de marcadores. Não obstante a versão deste projeto estar a ser executada em Windows®, o MAXST® SDK (Unity®) é compatível também com a plataforma Android.

Devido às restrições técnicas encontradas durante os testes do SDK MAXST® ao nível da deteção de marcadores (descritas na Secção 4.1.3), optou-se por avaliar uma biblioteca de deteção genérica, independente deste fabricante, para a deteção de marcadores, nomeadamente a biblioteca ArUco (Duda e Frese, 2018). A utilização desta biblioteca teve por base a implementação incluída no OpenCV®¹⁹ (disponível também para computador e para smartphones Android), como se pode ver no exemplo de deteção ilustrado na Figura 7 (Bradski, 2000).

O OpenCV® (Open Source Computer Vision Library) é uma biblioteca de código aberto que contém software de visão computacional, com um conjunto de algoritmos de deteção, reconhecimento, identificação e classificação (Bradski, 2000). O módulo ArUco é um dos módulos incluídos no OpenCV®, e permite a deteção de marcadores, de forma a estimar a posição e orientação de uma câmara relativamente ao marcador detetado (ou vice-versa).

De forma a testar esta biblioteca, foram utilizados os marcadores ArUco, com o objectivo de estimar a pose de uma câmara, processo este baseado em encontrar correspondências entre pontos dos marcadores sintéticos (ou fiduciais) e a sua projeção na imagem 2D, como ilustrado na Figura 7.

Optou-se também pelo desenvolvimento maioritariamente em linguagem Python®, sobretudo em computador, de modo a reduzir o tempo de desenvolvimento, teste e depuração. No entanto, em desenvolvimentos futuros, por uma questão de desempenho, poder-se-á portar as aplicações para um hardware específico.

Ao nível de software, foi definido o seguinte ambiente de desenvolvimento:

- Windows® 10 Enterprise, versão 10.0.19041 Build 190;

¹⁹ <https://OpenCV.org/>

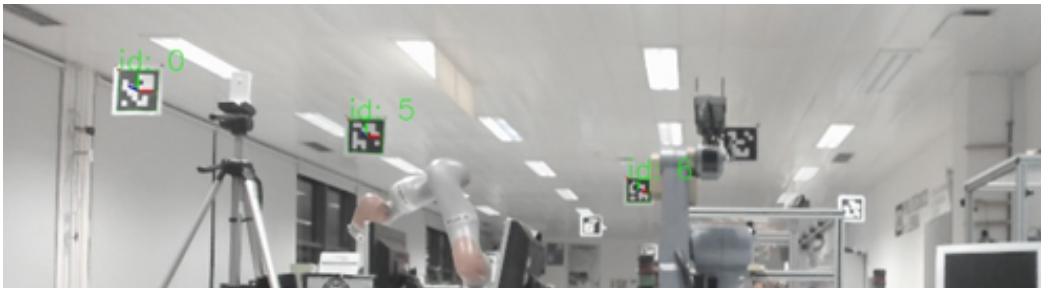


Figura 7: Resultado da estimação da localização de marcadores ArUco. Em cada marcador está colocado um referencial a indicar a posição e orientação estimadas, sendo que, em cada um dos referenciais exibidos, o eixo azul corresponde ao eixo X, o eixo verde ao eixo Y e o eixo azul ao eixo Z.

- Unity® (Versão 2019.4.9f1 LTS) com *Installs Windows Build Support (IL2CPP)* e *Android Build Support*, a selecionar e instalar com a ferramenta Unity® Hub;
- IDE Visual Studio® (2019) (pode ser selecionado nas opções de instalação do Unity®) ou IDE JetBrains Rider®;
- Python® 3.8.5 ou superior, adicionado ao PATH durante a instalação do Python®;
- Editor Visual Studio Code® com *plugin* para Python®.

DESENVOLVIMENTO DE UM SISTEMA DE LOCALIZAÇÃO PARA REALIDADE AUMENTADA

De uma forma geral a solução proposta pretende modernizar o modo como as operações logísticas em grandes armazéns de empresas fabricantes, distribuidoras e/ou operadoras logísticas são realizadas.

Neste sentido, do ponto de vista do projeto ARWARE, onde este desenvolvimento se insere, um dos principais avanços face ao atual estado da arte, será a conceção de uma plataforma que aborda diretamente os problemas relacionados com as operações de *picking* e a organização de armazém destes tipos de empresas. Desta forma não só se permite que os produtos tenham um *time-to-market* mais reduzido, como permite reduzir o número de itens perdidos no processo, garantindo uma melhor gestão dos armazéns e diminuindo o erro humano.

Um dos problemas a colmatar é a utilização de papel ou leitores de mão para os processos de *picking*, ainda muito utilizados para controlo logístico em armazéns e bastante propensos ao erro humano. A solução proposta, pelo projecto ArWare, passa por facilitar a tarefa de análise de itens por parte do colaborador do armazém, através do uso de óculos de RA, que permitem que a identificação de objetos seja feita automaticamente, através do simples olhar.

De igual modo, o projeto ARWARE pretende permitir a redução do tempo e o aumento da eficiência geral das operações em armazéns de logística. Para tal é necessário obter a localização do utilizador, através de métodos de localização interior nas operações de logística/*picking* em armazéns. Desta forma obtém-se não só a localização de equipamentos de mobilidade e operadores, mas também consegue-se criar pontos de rota que guiem os utilizadores do ponto A ao ponto B num armazém, para realizar tarefas laborais. No projeto ARWARE, a otimização de rotas, com recurso a heurísticas, meta-heurísticas ou Investigação Operacional, considerando condições e restrições como operadores (perfil do operador), corredores do armazém (largura e altura), veículos de armazém (empilhador, porta-paletes), encomendas (tipo de artigos, o seu volume e peso), horários (dos funcionários, da expedição da encomenda) e locais de descarga (tipo de cais, viaturas, foi desenvolvida por outra equipa no âmbito do projecto ArWare).

Ainda relacionado com a temática da localização, um dos parceiros do projecto ArWare, nomeadamente a empresa CODI, que se dedica ao desenvolvimento de produto, com soluções 3D como o Fabrico Aditivo, Maquetes, Digitalização 3D entre outras, realizou a implementação de um software de mapeamento 2D/3D de armazéns. Neste software é possível definir a localização de marcadores de RA nas respetivas prateleiras/*slots* de armazenamento de itens. Estes marcadores são depois relacionadas com as suas localizações reais e, com base nesta informação, é possível, através de algoritmos matemáticos, calcular a localização do utilizador. Ainda de forma a ajudar a obter a posição do utilizador, outro parceiro do projecto ARWare desenvolveu um sistema de localização aproximada para situações de inicialização, realocação *kidnapping* e verificação, com uma perspectiva de erro a rondar os 2m.

Por fim, o maior avanço do sistema desenvolvido no âmbito do projeto ARWARE resulta de, apesar de existirem soluções de RA para operações em armazéns, não existe qualquer outra que combine tecnologias de localização interior, IoT para interação entre equipamentos no âmbito da localização, mapeamento 2D/3D das plantas de armazém e RA. O foco do desenvolvimento aqui descrito, detalhado ao longo deste Capítulo, é o sistema de localização interior para o sistema de RA, denominado de LF.

3.1 REQUISITOS FUNCIONAIS

No módulo de LF, para assegurar que a informação seja corretamente apresentada ao utilizador, há que garantir um desempenho satisfatório em duas vertentes: a precisão dos resultados e a sua atualização atempada. Tal só é possível recorrendo a um conjunto de tecnologias e fontes de dados convenientemente integradas. Serão assim usadas várias fontes de dados sensoriais, diferentes equipamentos de processamento e algoritmos, que permitam integrar toda a informação de forma consistente e tendo em atenção a evolução no tempo e no espaço.

O módulo de LF tem os seguintes requisitos funcionais:

- Determinação da pose (posição e orientação) tridimensional do dispositivo de RA ou, por outras palavras, do operador, com uma resolução suficiente para garantir a coerência entre o mundo real e a informação projetada pelo módulo de RA, assegurando a operação em condições de conforto e contexto adequadas ao utilizador;

- Refletir esta informação para o módulo de RA a uma alta frequência, na região entre 20 e as 30 vezes por segundo, e baixa latência, no máximo 10 milissegundos (Azuma, 1997). Tal é necessário para que a interação não se torne desconfortável para o utilizador.
- Refletir esta informação, em baixa frequência, sensivelmente a, pelo menos, 10 vezes por segundo e baixa latência, por volta de 1 segundo, para os módulos com menor criticidade temporal, como por exemplo o otimizador de rota (módulo do projeto ARWare, desenvolvido por uma outra equipa, como ilustrado na Figura 8);
- Assegurar a portabilidade dos elementos de captura, processamento e interligação com os restantes sistemas, bem como a sua suficiência energética (pelo menos 4 horas de operação seguidas) para uma utilização típica em contexto de operações em armazéns.

3.2 ARQUITETURA

De forma a entender melhor os detalhes das várias componentes do projeto ARWare, nomeadamente toda a comunicação e interação que permite a sua integração, o diagrama da Figura 8 mostra os vários módulos do projeto ARWARE e a interligação entre estes. Como mostra a Figura 8, a aplicação de LF e a aplicação de RA são executadas simultaneamente no mesmo dispositivo de RA, podendo existir vários destes dispositivos, um por operador.

De acordo com a arquitetura geral do projeto, há dois tipos de recetores para a informação determinada pelo módulo de LF, com diferentes exigências em termos temporais. O mais exigente é a do módulo de RA, que necessita de informação com uma taxa de atualização elevada, na ordem de 20 a 30 vezes por segundo, de acordo com a informação descrita na Secção 3.1.

Tal exigência implica a utilização de ligações de dados diretas, o que é assegurado, pois as aplicações responsáveis por estes módulos são executadas no mesmo hardware e utilizam esquemas de partilha de memória. Os restantes módulos, não apresentando esta criticidade, têm acesso à informação de LF através do barramento de dados geral utilizando o *Enterprise Service Bus* (ESB) (Machado, 2021).

Será também através deste mesmo ESB que este módulo receberá a informação de configuração necessária, existente nas bases de dados do sistema, em particular os dados de localização absoluta de cada um dos marcadores, outras características

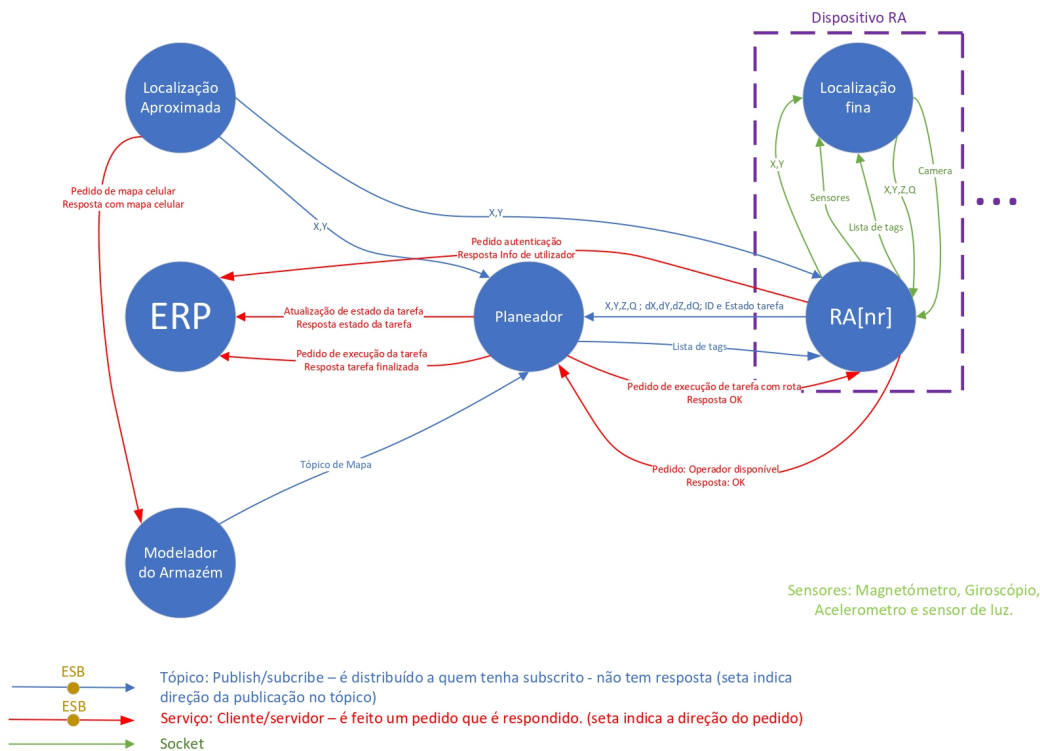


Figura 8: Arquitetura geral do projeto ARWare. No "Dispositivo RA", "X,Y" representa os dados enviados originalmente do módulo de localização aproximada, enquanto os dados "X, Y, Z, Q" representam a posição (X,Y e Z) e orientação (Q - quaternião) estimada pelo módulo de LF.

reconhecíveis no ambiente que sejam relevantes para a sua operação, assim como os dados antropométricos do utilizador (relevantes para o estimador).

A Figura 9 mostra um diagrama mais detalhado, focado somente na arquitetura aplicada às componentes de LF e RA, sendo que o “dispositivo RA” é o transmissor da informação destas duas componentes para os outros componentes do projeto ARWARE.

3.3 EQUIPAMENTO UTILIZADO

O projeto desenvolvido teve com base hardware já disponível no mercado de óculos RA, com as capacidades de localização, através da utilização de sensores (captura de imagens, acelerómetros, giroscópios e magnetómetros). Caso o equipamento não tenha estas capacidades, teremos de realizar as mesmas de outra forma, utilizando um sistema computacional externo. Os testes realizados numa fase inicial tiveram como base três telemóveis distintos - um Samsung Galaxy A7, um Nokia 6.1 e um

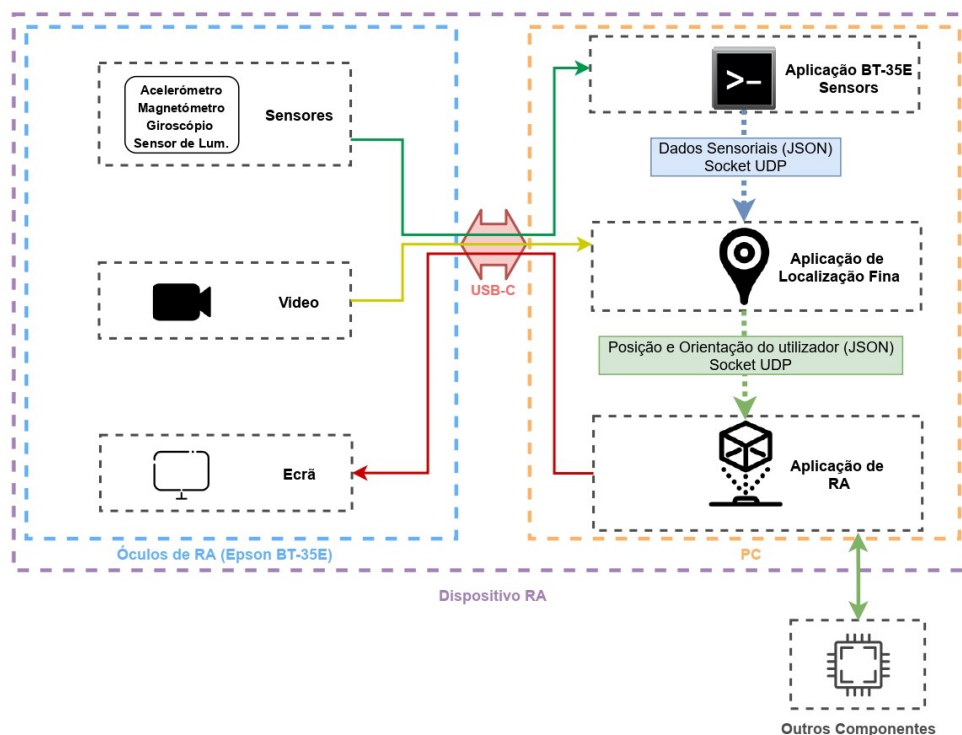


Figura 9: Arquitetura da componente de realidade aumentada e localização fina.

Huawei P20, de forma a analisar várias tecnologias de RA. Estes modelos apresentam uma boa qualidade de imagem, um bom campo de visão (por volta de 60 graus) mas, contrariamente aos óculos inteligentes, a utilização dos telemóveis demonstrou que, para a interação entre o utilizador e o equipamento, é necessário uma ou duas mãos para os agarrar.

A análise realizada com os óculos inteligentes demonstrou uma grande vantagem, pelo facto que o utilizador pode dispensar os elementos em papel, contribuindo para que este tenha as mãos livres, uma vez que os óculos exibem todas as instruções/informação necessária.

Dentro deste paradigma existiu a possibilidade de testar dois tipos de marcas. Nos testes com os óculos da Microsoft® a interação do utilizador com o sistema é realizada através da deteção dos movimentos das mãos. No caso dos óculos da marca Epson®, para que exista interação do utilizador com o equipamento tem de se utilizar um controlador. Os BT-300 têm um controlador que já vem integrado de origem com o equipamento, enquanto que os BT-35E são um modelo sem controlador que pode ser interligado a um equipamento externo, desde que o mesmo tenha a possibilidade de conexão *Universal Serial Bus* (USB) tipo-c 3.1 e/ou *High-Definition Multimedia Interface* (HDMI) com capacidade *DisplayPort Alternate mode*.

Em termos de campo de visão *field of view* (FOV), o dos óculos inteligentes é menor que no caso dos telemóveis considerados, sendo que os óculos da Microsoft® têm um FOV com mais 12º que os da Epson®, como mostra a Tabela 1.

Relativamente aos telemóveis utilizados, a escolha deveu-se ao facto das marcas e modelos selecionados apresentarem características bastante idênticas, suportarem a utilização o ARCore® da Google® (a calibração necessária para a utilização do ARCore® vem de fábrica e é disponibilizada pela própria Google®). Uma outra característica relevante relativamente aos telemóveis, é a necessidade de suportarem *DisplayPort Alternate mode* para poder suportar o acesso a todos os dados na interação com os óculos Moverio BT-35E da Epson®, no entanto, na altura da realização dos testes, apenas a Huawei® disponibilizava essa funcionalidade (a informação dos dispositivos que suportam essa tecnologia está disponível através da lista disponibilizada pelo site da *DisplayPort*¹).

A comunicação entre os dois dispositivos, telemóvel e os óculos BT-35E, tem de ser realizada através de um cabo USB 3.1 Type-C para *DisplayPort Alternate mode* com velocidade máxima teórica de 10Gb/s, com o comprimento suficiente que permita ao operador carregar o dispositivo móvel quando equipado com os óculos, e com capacidade para fornecer energia. Alternativamente, pode ser utilizada uma bateria externa ligada através da porta Micro-B USB dos BT-35E (por exemplo, um *power bank*).

Para o desenvolvimento e realização dos testes, apresentados no Capítulo 4, foi então selecionado o seguinte conjunto de equipamentos:

- **Óculos de realidade aumentada:**

- Epson® Moverio BT-35E (Figura 10), equipados com uma câmara monocular de 5Mpx, com um campo de visão de 23º e suporte de vários formatos de resolução (2592x1944/5Mp@15 fps, 1920x1080/FHD@30/15fps, 1280x720/HD@60/30/15 fps, 640x480/VGA@60 fps), bem como uma unidade inercial (giroscópio, acelerómetro e magnetómetro, todos de 3 eixos);
- Epson® Moverio BT-300 (Figura 11), equipados com uma câmara monocular de 5Mpx, com um campo de visão de 23º e de resolução (1280x720/RGB/30 fps), bem como uma unidade inercial (giroscópio, acelerómetro e magnetómetro, todos de 3 eixos);

- **Módulo de processamento:**

¹ <https://www.displayport.org/products-database/>



Figura 10: Epson® Moverio BT-35E.



Figura 11: Epson® Moverio BT-300.

- computador portátil Lenovo® Idealpad - 530 com o processador Intel® Core™ i7-8550U Quad Core, de 16 GB de memória RAM, uma placa gráfica da Nvidia™ GeForce MX150 | Intel® UHD Graphics 620;
- *smartphone* Huawei® P20, que inclui na parte traseira duas câmaras e uma unidade inercial (giroscópio, acelerómetro e magnetómetro, todos de 3 eixos);
- *smartphone* Samsung® Galaxy A7, que inclui na parte traseira três câmaras e uma unidade inercial (giroscópio, acelerómetro e magnetómetro, todos de 3 eixos);
- **Câmara de tracking:** Intel® RealSense™ T265 (Figura 13), que consiste num sensor equipado com duas câmaras (com lentes olho-de-peixe), com campo de visão horizontal de $163 \pm 5^\circ$, que admite vários formatos de resolução (848X800@30 fps, uma unidade inercial BMI055 (giroscópio e acelerómetro, ambos de 3 eixos), uma *Vision processing unit* (VPU) Movidius™ Myriad™ 2.0 com software embestado que implementa odometria inercial visual com localização e mapeamento simultâneos, incluindo assim o *tracking* contínuo da posição e orientação do sensor, relativamente à pose inicial;



Figura 12: Logitech® Brio 4k.



Figura 13: Intel® RealSense™ T265.

- **Câmara:** Logitech® (Figura 12) Brio 4k UHD (Ultra-High Definition), um equipamento com um sensor de imagem UHD, com campo de visão horizontal de 82.1°, que admite vários formatos de resolução (4097x2160/UHD@30 fps, 1920x1080/FHD@30/60 fps, 1280x720/HD@30/60/90 fps) e com um sensor de IV;
- **Equipamento topográfico:** Total Leica® TS02 (Figura 14), capaz de medir distâncias até 350 metros sem a necessidade de prismas, e até 3.500 metros com o uso de um único prisma, com uma precisão angular de 2 segundos;
- **Fonte portátil de energia:** Para assegurar a portabilidade do sistema, será integrado um sistema de baterias Powerbank ANKER, com capacidade de 20000 mAh, com 3 portas de saída de potência. Esta permite alimentar, em simultâneo, todos estes equipamentos, sem interferir com a capacidade de comunicação dos aparelhos.



Figura 14: Total Leica® TS02.

3.4 SENSORES

No que respeita às fontes sensoriais, em sistemas de RA são tipicamente conjugados dois tipos de fontes de informação para a estimação da pose dum dado sistema em ambiente interior: uma, ou mais, câmaras e uma unidade inercial.

Existe trabalho considerável para a localização de sistemas apenas baseados em informação recolhida a partir de uma unidade inercial, mas tipicamente este está limitado ao movimento de seres humanos (F. Li et al., 2012, Monteiro e Br, 2017), pois tira partido do padrão típico de movimentação de um ser humano (note-se que no projeto em questão o ser humano, ao deslocar-se, fará um percurso não linear).

Das unidades inerciais, é possível obter velocidades de rotação e acelerações triaxiais, o que permite uma estimação da pose atual a partir de uma pose inicial, apesar do seu erro acumulativo ser elevado. Todavia, mesmo em boas condições, não havendo referências externas, existe sempre um desvio considerável que se acumula ao longo do tempo, pelo que, para a aplicação em desenvolvimento neste projeto, a utilização apenas da informação da unidade inercial não é uma solução.

Já o uso de uma ou mais câmaras, com a eventual colocação de marcadores previamente referenciados num mapa conhecido, irá permitir uma localização sem acumulação ilimitada de desvios. Assim, a informação da unidade inercial, obtida com frequências elevadas (60 vezes por segundo) e de processamento mais simples, permite complementar a informação obtida a partir das imagens. Esta combinação designa-se *Visual-Inertial Odometry* (VIO) (P. Li et al., 2017) e (Fraga-Lamas et al., 2018), abrindo a possibilidade de obter um sistema final que é robusto e fornece informação de pose com a frequência adequada.

3.5 CALIBRAÇÃO DO SISTEMA

No caso das câmaras, e de forma a minimizar o erro de deteção dos marcadores e o erro na determinação da pose dos marcadores, é necessário que estas sejam calibradas previamente. Para esse fim, foi implementada uma aplicação² baseada na biblioteca OpenCV®, obtendo assim os parâmetros de calibração da câmara dos óculos de RA (parâmetros intrínsecos e extrínsecos) (Martins et al., 2020).

Resumidamente, como descrito em (Horaud e Dornaika, 1995), para realizar o processo de calibração é necessário ter uma imagem com um padrão planar conhecido, impressa, que possa ser vista de diferentes perspetivas (neste caso em particular foi utilizado um padrão em xadrez). Com estas imagens, o algoritmo procura o padrão e detetar os vértices entre cada dois quadrados pretos do padrão na imagem. Com base na posição determinada dos vértices na imagem, e no conhecimento da relação espacial real dos vértices no padrão impresso, o algoritmo determina os parâmetros intrínsecos e extrínsecos que garantem que o erro de reprojeção é minimizado, devolvendo esse erro, para cada padrão detetado, e os parâmetros determinados. Os parâmetros intrínsecos estão relacionados com a distorção da câmara, distância focal e dimensão do sensor ótico, enquanto os parâmetros extrínsecos descrevem a pose da câmara relativamente ao padrão detetado. Na documentação do OpenCV®³

² https://docs.opencv.org/4.x/dc/dbb/tutorial_py_calibration.html

³ https://docs.opencv.org/4.x/da/d13/tutorial_aruco_calibration.html

encontra-se uma descrição mais detalhada do processo de calibração. A calibração do sistema minimiza os erros relativamente aos dados obtidos a partir da análise da imagem, tendo um impacto direto na estimativa final da pose.

3.6 TECNOLOGIAS DE LOCALIZAÇÃO EM ESPAÇOS INTERIORES

Com base nos testes realizados (apresentados no Capítulo 4) foram avaliadas algumas tecnologias disponíveis, todas elas com o objetivo de determinar a localização do utilizador para esta ser aplicada na projeção da RA. O desenvolvimento e a escolha das tecnologias teve como referência a procura de uma solução com potencial de permitir um erro igual ou inferior a 0.5m.

Na realização destas avaliações foi realizado um *ground-truth* de comparação utilizando simultaneamente o telemóvel Samsung® Galaxy A7, um computador Lenovo® ideapad 530s com uma câmara da Logitech® BRIO 4K Ultra HD RightLight, um estação Total Leica TS02, os óculos inteligentes da Epson® Moverio BT-35E conectados ao telemóvel Huawei® P20, os óculos da Epson® Moverio BT-300 e uma câmara Intel® RealSense™ T265.

3.6.1 Utilização de marcadores

Das técnicas avaliadas em laboratório, a que mais se destacou foi a utilização de marcadores pois, apesar de todas as que foram testadas, apesar de funcionarem relativamente bem, a utilização de marcadores fornece a posição e orientação mais precisa, ou seja, com menos erro, erro este descrito no Capítulo 4.

Um marcador, no contexto de RA, refere-se ao que é mais geralmente conhecido como um marcador fiducial. Um marcador fiducial é qualquer objeto que possa ser colocado em cena e detetado para fornecer uma referência fixa relativamente à cena.

Os marcadores utilizados contêm padrões de códigos digitais. Estes códigos podem ser utilizados para localização, para a distinção de objetos, fornecimento de ponteiros para páginas web, ou texto em geral, bem como para fornecimento de conteúdos para serem projetados em aplicações de RA.

Os marcadores, tipicamente com uma área quadrada, geralmente consistem numa combinação de quadrados pretos e brancos, dentro de uma moldura branca de tamanho predefinido (Figura 15), que codificam a identidade do (ID) do marcador (Tørdal e Hovland, 2017).

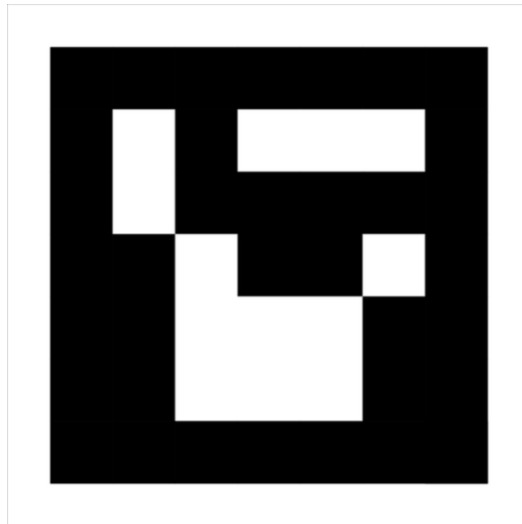


Figura 15: Marcador do tipo ArUco.

Nas aplicações de RA que recorrem a marcadores, a deteção destes passa, essencialmente, pelo processamento de imagem, para estimar a orientação e a posição do sistema de RA relativamente ao marcador, e relativamente à cena, caso o marcador seja fixo relativamente ao mundo. Esta é uma técnica que é também utilizada noutras aplicações que recorrem a marcadores como, por exemplo, aplicações de robótica (Tørdal e Hovland, 2017).

A utilização de marcadores, georreferenciados no espaço de trabalho, permite aumentar a taxa de sucesso na localização do equipamento de RA, comparativamente com outros métodos de visão computacional que não os utilizados. Existem vários estudos que analisam e comparam a qualidade da deteção de diferentes tipos de marcadores (Abawi et al., 2004, Fiala, 2005), como por exemplo, os marcadores ArUco ou MAXST® (Figura 16), com diferentes softwares de reconhecimento, nomeadamente o MAXST®⁴ e o ARToolkit⁵, a que se junta a biblioteca ArUco (Garrido-Jurado et al., 2014), através do SDK OpenCV®. Os requisitos principais na escolha do tipo de marcadores prendem-se com o número de marcadores diferentes que cada tipo de marcador permite, bem como a relação entre a sua dimensão e qualidade da sua deteção.

Foram realizados testes com vários tipos de marcadores, descritos na Secção 4.2, tendo se concluído que, para este projeto, os marcadores mais indicados seriam do tipo ArUco.

4 <https://maxst.com/>

5 <http://www.hitl.washington.edu/artoolkit/>

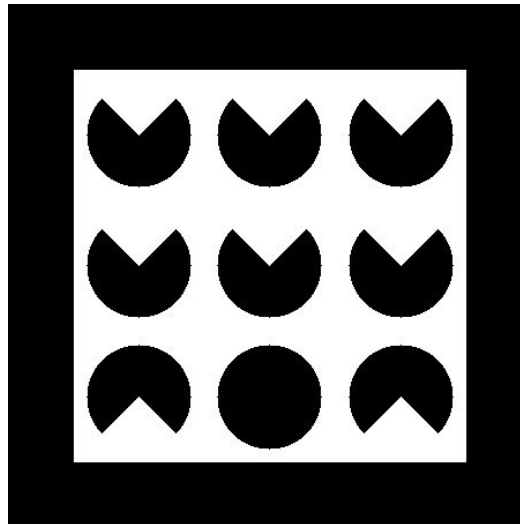


Figura 16: Marcador do tipo MAXST®.

3.6.2 Utilização de pontos-chave

Em ambientes sem marcadores, a solução passa por utilizar outras características de destaque, tais como segmentos, linhas retas, contornos, pontos sobre os contornos ou uma combinação destes (Ledwich e S. Williams, 2004). Com a captação sucessiva de imagens, é possível estimar a translação e rotação da câmara entre imagens consecutivas, tendo por base a calibração prévia do sistema de visão e a detecção de pontos característicos nessas imagens. Esta técnica denomina-se de odometria visual (DeTone et al., 2018).

Este tipo de sistema pode basear-se numa única câmara, denominado sistema monocular, à parte de um fator de escala (Mirabdollah e Mertsching, 2015). Esse fator de escala pode ser determinado com base na detecção de objetos conhecidos, ou pela combinação com outros sensores como, por exemplo, uma unidade inercial (Diete et al., 2016, Fraga-Lamas et al., 2018). A utilização de duas câmaras (visão estéreo), conjugada com conhecimento da pose relativa entre as duas câmaras, permite a odometria visual sem o problema do fator de escala, permitindo ainda determinar a profundidade da cena, embora à custa de um processamento mais elevado (Aqel et al., 2016).

Outras abordagens permitem calcular tanto o deslocamento da câmara e a estrutura da cena, usando técnicas baseadas em abordagens SLAM (Guerreiro et al., 2004).

O SLAM que, entre outras fontes de informação, recorre à odometria visual, armazena ao longo do tempo pontos-chave que lhe permitem construir um mapa do

espaço e se localizar relativamente a esse mapa. Esta técnica tem duas principais vantagens: a redução, ou mesmo eliminação, do erro cumulativo caso se utiliza a técnica de deteção de fecho de trajetória *loop closure* (deteção do retorno a um local previamente visitado); a determinação da pose relativamente a um referencial de base, e não apenas a determinação do deslocamento realizado, comparativamente à odometria visual, isto é, permite estimar uma localização absoluta e não apenas uma localização relativa. Podem também ser introduzidos marcadores específicos, em poses conhecidas, para assim obter uma referência absoluta (Fuentes-Pacheco et al., 2015).

Assim, o sistema concebido utiliza duas estratégias. Por um lado, recorre-se aos dados determinados pelo sensor Intel® RealSense™ T265, nomeadamente a pose relativamente à pose inicial. Por outro lado, recorre-se às imagens da câmara do dispositivo de RA (Epson® Moverio BT-35E) e a algoritmos de odometria visual implementados sobre a unidade de processamento (Huawei® EML-L29), complementados pela unidade inercial dos óculos de RA.

3.7 DESENVOLVIMENTO ALGORITMO DE LOCALIZAÇÃO

O objetivo do módulo de LF passa por estimar a pose dos óculos de RA que, na verdade, corresponde também à localização aproximada do utilizador que utiliza esses óculos. Dado que a pose estimada dos óculos será utilizada pela componente de RA, no projeto ARWARE, para a geração da informação a disponibilizar nos óculos de RA, é importante que o erro de estimação seja reduzido, e que a frequência de atualização seja elevada, caso contrário esta terá um impacto negativo na usabilidade do sistema, tornando a aplicação final desconfortável, ou mesmo não utilizável de todo.

Existem diversas possibilidades de implementação do sistema de LF, tendo por base a utilização de técnicas de visão computacional, a partir de uma das câmaras referidas anteriormente, complementada pela informação obtida pela unidade inercial, quer do smartphone, quer dos óculos de RA. Por outro lado, pode ainda utilizar-se diretamente a informação de pose fornecida pelo sensor Intel® T265.

Do ponto de vista do desenvolvimento do algoritmo de localização, cujo diagrama se apresenta na Figura 17, foi utilizado o algoritmo de deteção dos marcadores visuais ArUco, com a pose georreferenciado no espaço de trabalho, através da respetiva deteção dos mesmos nas imagens capturadas, como ilustrado na Figura 7.

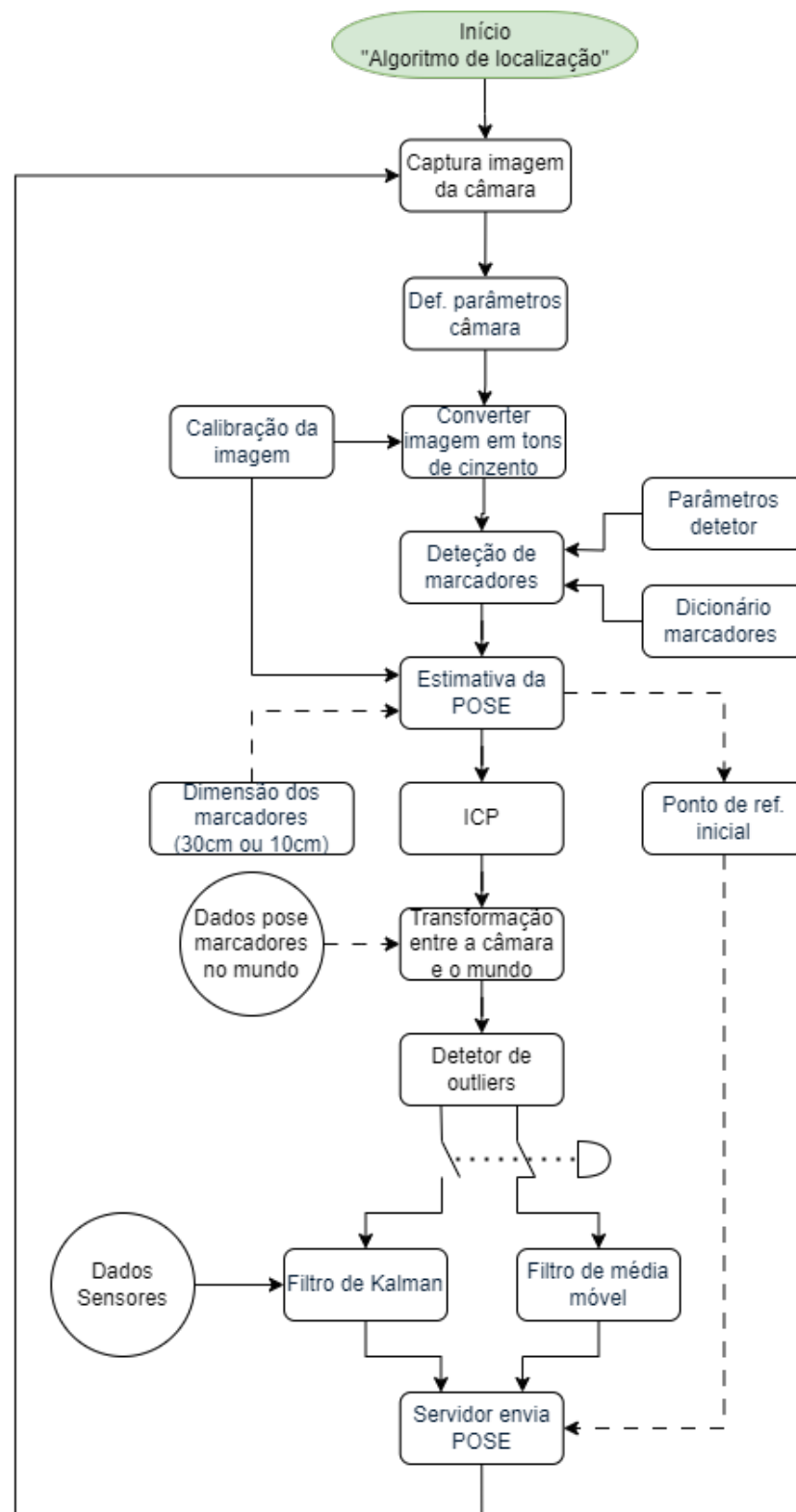


Figura 17: Diagrama do algoritmo localização fina.

De forma a obter a pose dos óculos de RA, a aplicação desenvolvida utiliza as imagens capturadas para detetar e reconhecer todos os marcadores visíveis. Dados os parâmetros resultantes da calibração da câmara e a dimensão real dos marcadores, é estimada a pose do utilizador (dispositivo de RA/câmara) relativamente a cada marcador (transformações geométricas entre o sistema de coordenadas do marcador e o sistema de coordenadas da câmara). Tendo a pose do utilizador relativa a um ou mais marcadores, e conhecendo a pose de cada um desses marcadores em coordenadas do mundo, pode estimar-se a pose 3D do utilizador em relação ao mundo.

Sabendo a pose inicial do sensor, é possível utilizar odometria visual para estimar a pose em cada instante, mesmo na ausência de marcadores. No entanto, esta estimativa resulta num acumular de erro, pelo que, com o tempo, a pose estimada do dispositivo diverge da pose real. A utilização de “âncoras” (pontos de referência absolutos conhecidos), permite que o sistema minimize esse erro. Além disso a utilização dessas referencias absolutas nos algoritmos de SLAM minimiza o desvio acumulado ao longo do tempo, no entanto interessa avaliar também a integração de referências globais (neste caso marcadores), pré-referenciadas a um mapa conhecido do espaço de trabalho, até porque estas referências podem servir de base para projecção de conteúdos no equipamento de RA. Estas referências permitem não só minimizar o erro na localização global, como podem servir de *ground-truth* (com o devido erro associado) para avaliar técnicas sem marcadores.

Através do registo das características relevantes detetadas nas imagens, é possível também criar um mapa do espaço percorrido, uma técnica denominada Visual SLAM, permitindo a localização nesse mapa sem recurso a marcadores . A Intel® RealSense™ T265 implementa esta técnica no hardware embebido, conjugado com os dados da unidade inercial. A utilização dos dados da unidade inercial, pela sua alta frequência de atualização, permite lidar melhor com movimentações mais bruscas do sensor, e permite um aumento da taxa de estimativa da pose, contribuindo assim para a redução do erro da estimativa no global aumentando a performance do sensor.

Ao nível do algoritmo *Iterative Closest Point* (ICP) (Exploration, 2022), este está implementado de forma a minimizar a diferença entre a estimativa da pose de cada marcador em relação à sua pose real.

Ao longo do desenvolvimento o algoritmo ICP foi melhorado com recurso ao desenvolvimento algoritmo kd-tree, tendo ainda sido feitas alterações na definição dos critérios de paragem, que incluem, na sua versão final, a convergência do erro,

além do número máximo de interações. No seu conjunto, as alterações realizadas tiveram um impacto relevante no desempenho da implementação do ICP, permitindo processar mais 3 FPS em média, sendo agora possível, em média, processar cerca de 21 FPS, sem impacto no erro obtido.

A detecção de *outliers* tem como objetivo detetar os valores de estimativa de pose de cada marcador que variam abruptamente dos restantes. Para isso foi implementado um detetor automático baseado no algoritmo *Local Outlier Factor* (LOF)⁶, comparando a estimativa da pose de cada marcador com a densidade das estimativas da pose dos restantes marcadores detetados (Breunig et al., 2000). A estimativa de pose dos marcadores que são considerados “incorretos” (*outliers*) são descartados, reduzindo desta forma o erro de estimativa da pose do utilizador. Ou seja, se houver um marcador cuja detecção não seja coerente com os restantes, essa detecção é descartada.

3.7.1 Recolha de dados dos sensores do Moverio BT-35E

Esta aplicação foi baseada na solução exemplo da Epson® para o dispositivo Moverio BT-35E⁷, a qual apresenta, continuamente, os valores dos vários sensores em gráficos numa janela em ambiente Windows®. De forma a reduzir a carga computacional da aplicação, foram removidos os métodos de desenho dos gráficos e o projeto convertido para uma aplicação de consola. Foram mantidos apenas a componente de aquisição dos dados dos sensores, juntando tudo num único objeto JSON (JavaScript Object Notation), enviado para a aplicação de LF através de um *socket User Datagram Protocol* UDP, como ilustrado na Figura 9.

A aplicação de recolha de dados dos sensores contém essencialmente três partes, sendo que na Figura 18 está apresentado um conjunto de Sensores, Observadores e Comunicação (por *socket*).

Relativamente à recolha de dados dos sensores, o utilizador tem a possibilidade de adicionar ou remover um *watcher* (elemento que permite observar a informação do sensor), ativando ou desativando a recolha de dados do sensor respetivo.

Ainda neste desenvolvimento, e para que haja partilha de dados entre as aplicações, foi decidido utilizar comunicação via Socket também para a partilha da informação que. No futuro, por uma questão de performance, poderá ser necessário integrar

6 <https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.LocalOutlierFactor.html>

7 https://tech.moverio.epson.com/en/bt-35e/sample_project_download.html

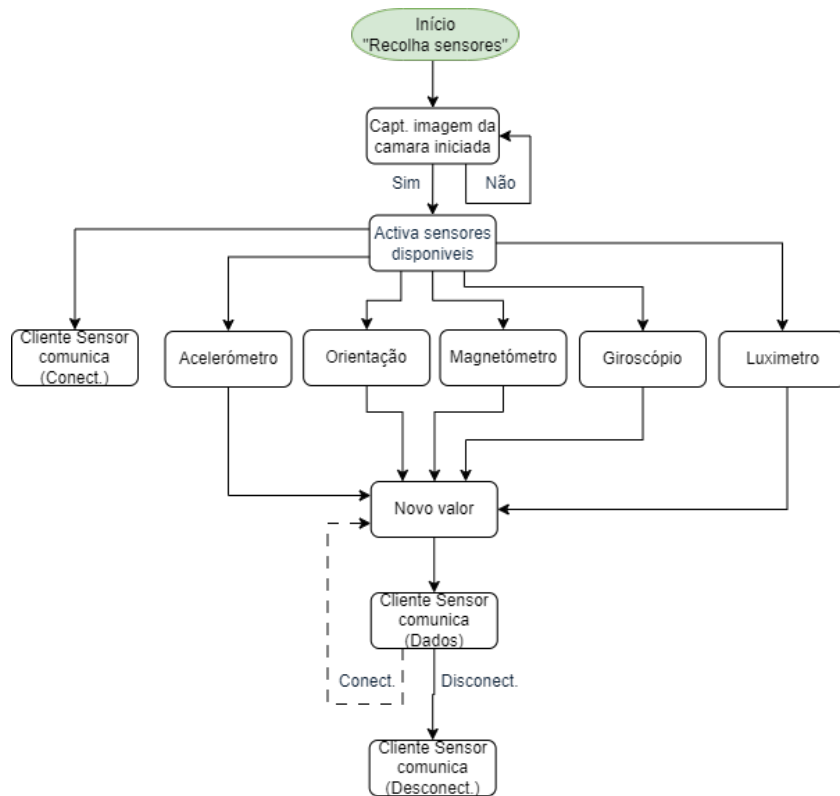


Figura 18: Diagrama do algoritmo BT-35E Sensors.

as aplicações numa única aplicação, permitindo uma melhoria do desempenho e redução da latência, e assim melhorando a qualidade de experiência por parte do utilizador.

3.8 ALGORITMO DE ESTIMAÇÃO DE POSE

Garantido o acesso aos dados dos diversos sensores como descrito anteriormente, torna-se necessário integrar toda a informação e estimar a pose com uma dada confiança associada. Para tal é necessário adaptar o algoritmo de VIO com a deteção dos marcadores para o caso em desenvolvimento.

Numa primeira fase foi implementado um algoritmo de localização interior com base na deteção de marcadores utilizando o SDK MAXST®. Ao longo dos testes realizados, este SDK continha bastantes oscilações na estimativa da pose e, por isso, passou-se a utilizar algoritmos baseados em OpenCV®, utilizando marcadores ArUco.

Numa segunda fase foi implementado um filtro de média móvel e, numa terceira fase, foi implementado um filtro de Kalman, sendo que, como mostra a Figura 17, a aplicação de LF permite utilizar uma opção ou outra.

O filtro de média móvel criado, calcula a média de um conjunto de estimativas da pose do utilizador, sendo que, sempre que o algoritmo obtém um novo valor de pose, este descarta o valor mais antigo e calcula uma nova pose do utilizador, baseado na média das restantes estimativas.

Além da integração das várias fontes sensoriais, a implementação baseado no filtro de Kalman (Santana et al., 2008) reduz também a oscilação e o erro de estimativa da pose do utilizador dos óculos de RA, comparativamente com a utilização do filtro de média móvel.

O filtro de Kalman consiste num estimador que, ao longo do tempo, com base nas observações realizadas e no seu estado, estima o próximo estado do sistema. Como se pode ver no diagrama apresentado na Figura 19, o detetor de *outliers*, após remoção destes, envia as poses do utilizador, geradas a partir de cada marcador detetado, assim como o valor dos sensores disponíveis nos óculos de RA, também usados como variáveis de entrada do filtro.

É importante que o filtro de Kalman seja inicializado com um valor de estado próximo do valor real. Dado que, do ponto de vista de utilização do equipamento pelo utilizador no armazém, é necessário para o arranque do algoritmo iniciar o mesmo num espaço específico definido para o efeito, nessa altura é feita a leitura de um marcador específico, a uma curta distância, aproveitando essa informação para inicializar o filtro de Kalman com a pose inicial do utilizador (futuramente esta inicialização poderá ser realizada no momento em que o utilizador fizer o seu *login*). Nesta operação determina-se também a altura aproximada do utilizador dos óculos de RA, pois esta informação é uma mais valia para o algoritmo de localização, sendo utilizada na filtragem da altura máxima da estimativa da pose do utilizador.

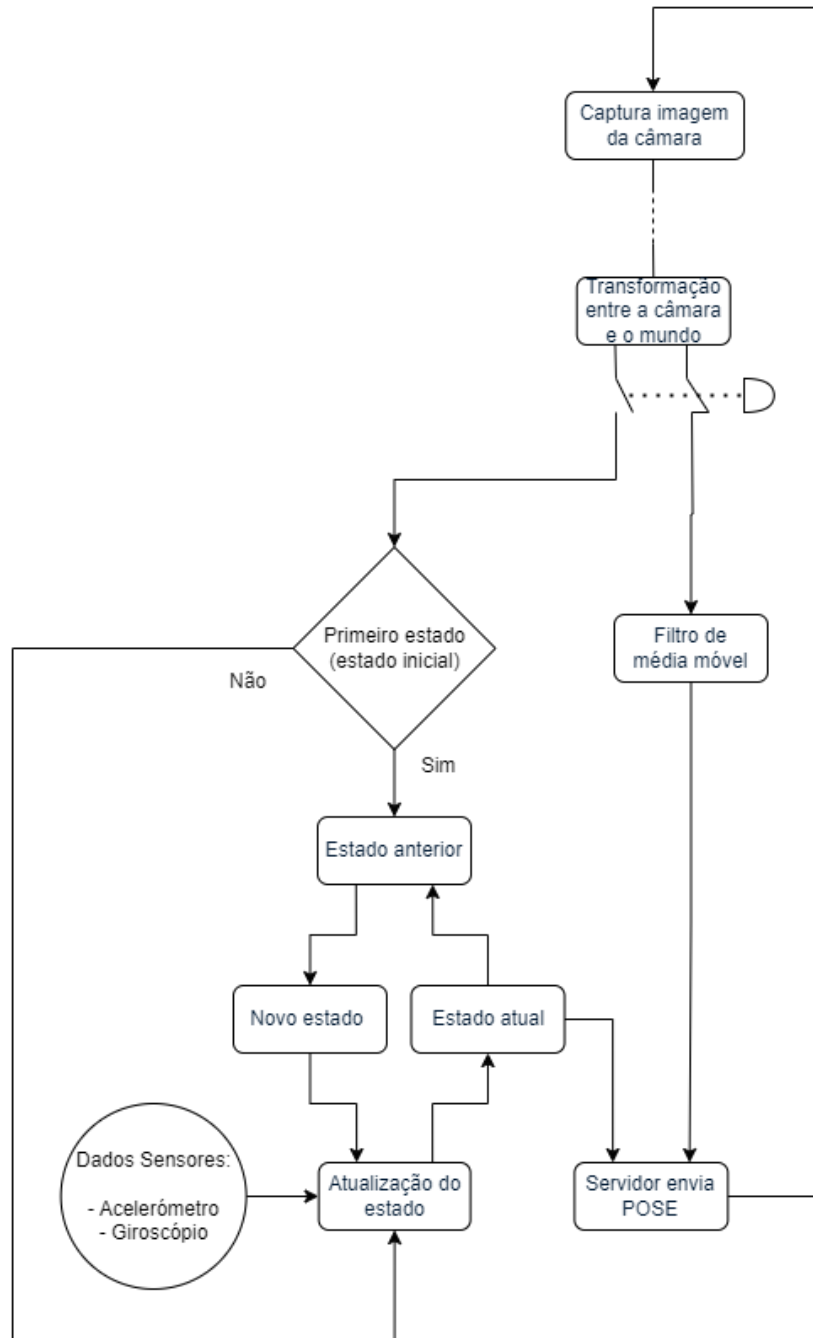


Figura 19: Diagrama detalhado do filtro de Kalman.

TESTES E RESULTADOS

Existiram três etapas relevantes na tarefa de testes da componente de localização interior. Numa fase inicial, foram efetuados testes em maior escala, ou seja, usando número maior de equipamentos e software (Figura 20), tais como os óculos de RA, *smartphone*, sensor Intel® T265 e a câmara Logitech® Brio. Estes testes foram realizados de forma individual, usando diretamente aplicações e os SDK fornecidos/suportados pelos seus fabricantes, com o objetivo de validar os mesmos perante os requisitos deste projeto e, como consequência, definindo assim desenvolvimentos futuros. Na segunda fase de testes, já com o equipamento e software definidos, direcionou-se o foco para a validação do algoritmo de localização, algoritmo este com base na deteção de marcadores, definindo nesta fase o tipo, tamanho, quantidade e distâncias entre marcadores. Na terceira e última fase, os testes foram orientados de forma a validar o algoritmo de localização interior na sua globalidade. Estes testes foram primeiro realizados em ambiente laboratorial, validando o cumprimento dos requisitos do projeto, e permitindo a realização de testes e demonstrações num ambiente real, ou seja, num armazém industrial.



Figura 20: Ambiente testes com equipamentos a avaliar.

4.1 PRIMEIRA FASE DE TESTES

Os testes realizados nesta fase, tinham como principal objetivo a análise dos equipamentos e software para utilização no projecto, sendo que foi testada a Intel® RealSense™ T265 com SDK 2.0, o *smartphone* Huawei P20 com o SDK ARCore®, e os óculos Epson Moverio BT-35E com SDK MAXST®. Estes foram testados em percursos pré-definidos em contexto laboratorial, procurando que fossem semelhantes a percursos num armazém industrial, mas neste caso específico. Desta forma foi possível obter um conjunto de dados que serviram como base para diversos algoritmos de determinação de pose, desenvolvidos numa segunda fase do projeto. .

4.1.1 Intel® RealSense™ T265 com SDK 2.0

Os testes realizados inicialmente com o sensor T265 centraram-se na avaliação dos dados fornecidos da estimação da pose, tendo como referência o caminho ilustrado a vermelho na Figura 21. Este caminho, criado ao longo do LRAFI, inclui a passagem por um corredor adjacente ao laboratório. Foram realizadas 3.5 voltas a pé no total deste teste, utilizando para isso os três equipamentos referidos anteriormente, sendo este percurso muito idêntico a um retângulo de 19m de comprimento e 6m de largura, sendo ele iniciado no *Start*, terminando em *End* após as 3.5 voltas.

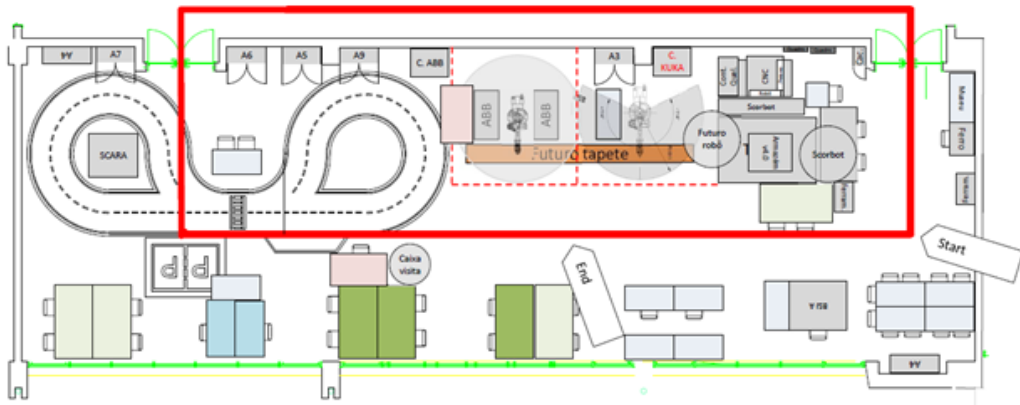


Figura 21: Circuito de testes Intel® RealSense™ T265 com SDK 2.0 e Huawei P20 com SDK ARCore®, onde “Start” e “End” representam o início e o final do percurso, respetivamente.

Os resultados obtidos neste teste, visíveis na Figura 22¹, apresentam um desempenho considerado bastante razoável relativamente à estimativa da pose da

¹ As linhas tracejadas apresentadas neste figura e nas seguinte semelhantes a esta representam o salto de um valor para o outro.

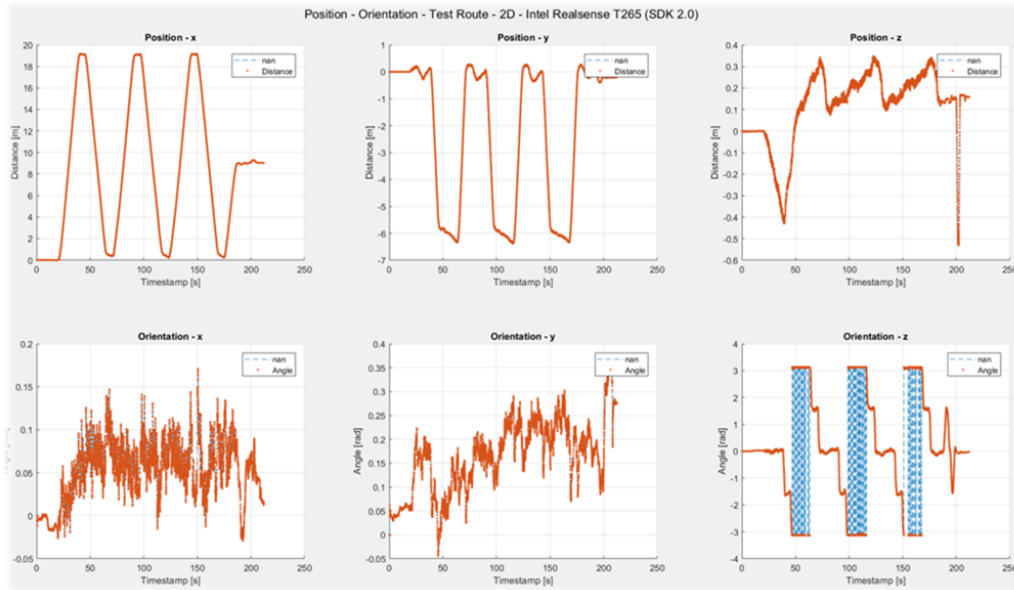


Figura 22: Pose estimada usando a Intel® RealSense™ T265 com SDK 2.0 ao longo da realização de 3.5 voltas no percurso identificado na Figura 21.

câmara, isto de acordo com os requisitos definidos. Pode observar-se que, ao longo do percurso, no eixo x a distância percorrida é por volta de 19m, enquanto no eixo y a distância percorrida é sensivelmente 6.5m, existindo desvios nestes dois eixos que, no pior caso, não superam o erro máximo admitido neste projeto (0.5m). Por fim, no eixo z, apesar de termos variações no erro, este é inferior aos restantes eixos. Em termos de orientação a mesma mantém-se sensivelmente constante no eixo x e y, apenas no eixo z existe uma variação de aproximadamente $\pi/2$ rads (90 graus), tal como esperado, a cada canto do retângulo da trajetória realizada. Nesta trajetória existem desvios temporários, mas relevantes, em situações específicas, sendo que isso acontece especialmente quando existem variações bruscas no movimento da câmara como, por exemplo, quando se realizam mudanças de direção. O algoritmo interno da Intel® RealSense™ SDK 2.0 procura corrigir esses desvios, como se pode ver na trajetória estimada (ver Figura 27). Estes resultados mostram que este sensor pode ser utilizado quer para complementar, quer para servir de referência na avaliação comparativa dos restantes algoritmos analisados.

4.1.2 Smartphone Huawei P20 com SDK ARCore®

O SDK da ARCore® integra a informação das duas câmaras traseiras (visão estéreo) e da unidade inercial do smartphone utilizado para estimar a pose do dispositivo.

De acordo com a literatura (Kim et al., 2022), as técnicas e algoritmos utilizados são idênticos aos da Intel® T265, nomeadamente utilizando VIO e SLAM.

Na Figura 23 podem ser vistos os dados obtidos nos testes realizados com o *smartphone*. Estes resultados são semelhantes aos obtidos com a Intel® RealSense™ T265 com SDK 2.0. Apenas estão a ser apresentados os testes realizados com o *smartphone* Huawei® P20, e não com os restantes *smartphones* como, por exemplo, o Samsung® Galaxy A7, devido ao facto deste originar erros mais elevados, como mostra a Figura 27.

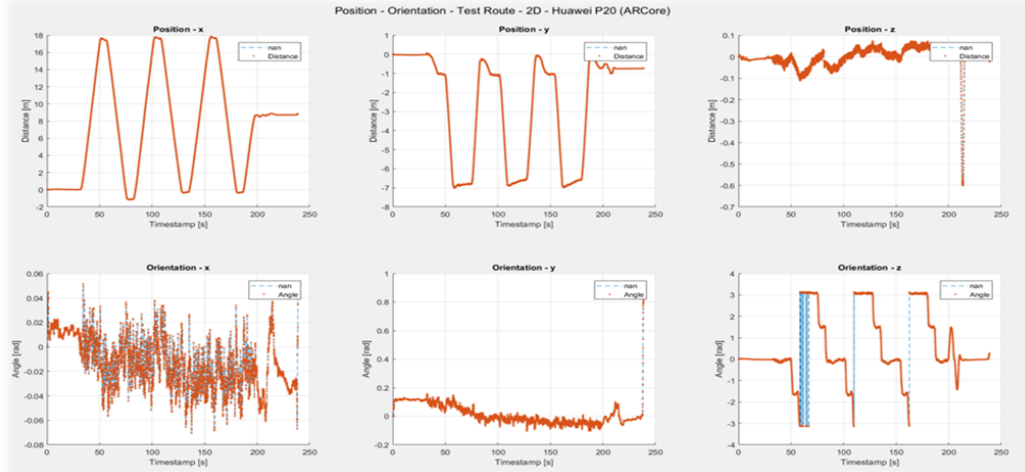


Figura 23: Resultados de estimação da pose utilizando o *smartphone* Huawei P20 com SDK ARCore®.

Os resultados ilustrados na Figura 23 mostram um desempenho inferior aos obtidos com a Intel® RealSense™ T265 com SDK 2.0, o que é expectável, dado que o FOV das câmaras utilizadas no Intel® T265 (163°) é superior ao das câmaras do *smartphone* (120°), o que permite obter um maior número de correspondências de pontos-chave entre capturas consecutivas, bem como lidar melhor com oclusões. Por exemplo, o eixo x e y apresenta um desempenho pior que o caso anterior, já no caso do eixo z, este apresenta melhores resultados comparativamente ao caso anterior. Em termos de orientação, os resultados são semelhantes aos resultados apresentados anteriormente.

4.1.3 Epson Moverio BT-35E com SDK MAXST®

O SDK MAXST® tem como base a deteção de marcadores, pelo que foram posicionados no ambiente marcadores como o ilustrado na Figura 16. Por este motivo, os testes foram realizados apenas ao longo de um troço linear (com 12m de com-

primento) da trajetória utilizadas nos testes anteriores, como ilustrado na Figura 24.

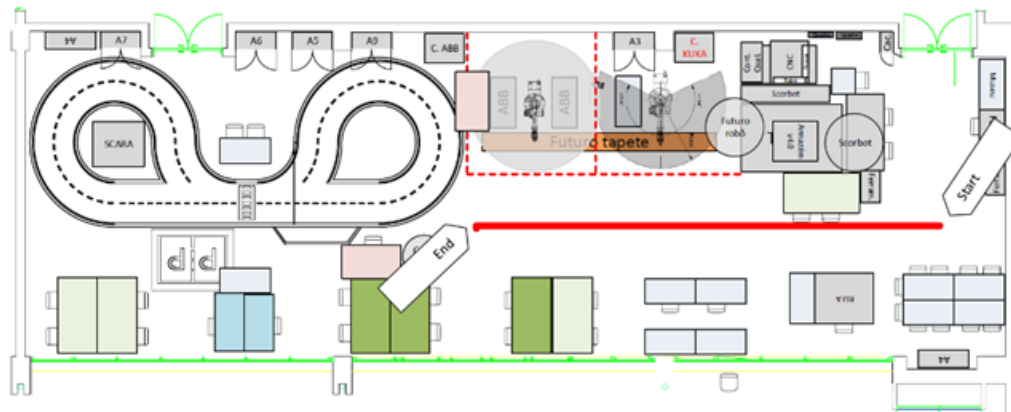


Figura 24: Circuito de teste com os óculos Epson Moverio BT-35E com SDK MAXST®.

Os resultados obtidos com base no algoritmo do SDK MAXST® foram obtidos utilizando a câmara dos óculos de RA (Epson Moverio BT-35E), com o processamento a decorrer num computador (Lenovo® Idealpad - 530) e utilizando o software Unity® (Versão 2019.4.9f1 LTS) para monitorização da experiência, estando ilustrados na Figura 25 e detalhados individualmente na Figura 26. De notar que estes dados contêm uma cor distinta para cada marcador, de forma a identificar qual o marcador que está a ser visualizado. Os resultados obtidos mostram uma distância percorrida estimada no eixo x de aproximadamente 11.5m, resultando assim num erro relativamente elevado, mas ainda dentro dos parâmetros pretendidos para este projeto, relativamente à distância total esperada de 12m. No caso do eixo y os erros são mais elevados, piora ainda no eixo z. Em termos de orientação, temos variações máximas de 0.1 rads (5°), valores estes aceitáveis. Além destes testes, foram realizados outros testes com o SDK MAXST®, mas que não são aqui apresentados por uma questão de espaço. Em todos esses testes foram sempre obtidos erros na ordem de grandeza dos aqui apresentados, razão pela qual se optou por não prosseguir com a utilização deste SDK.

4.1.4 Conclusões da primeira fase de testes

Foi possível testar com sucesso a obtenção das imagens das diversas câmaras (Intel® RealSense™ T265, Huawei P20 e os Epson Moverio BT-35E), diferentes dados inerciais de forma coerente nas diversas plataformas de programação, bibliotecas e sistemas de desenvolvimento, e para isso, utilizaram-se softwares (SDK da Intel®, SDK da ARCore® e SDK da MAXST®) através de diferentes sistemas operativos

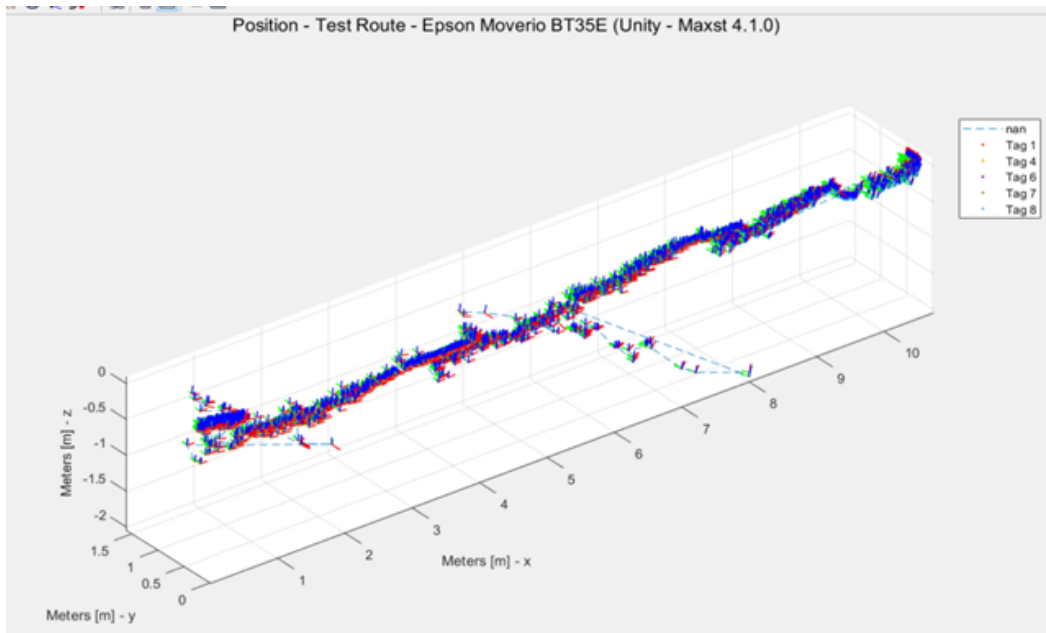


Figura 25: Teste de caminho linear (Epson Moverio BT-35E + MAXST® + Unity®).

(Windows®, Linux® e Android®) em diferentes linguagens (C#, C++, Python, Java), permitindo conhecer e modificar a parametrização interna dos diversos equipamentos.

Na Figura 27 é apresentado o resultado da utilização do SDK Intel® RealSense™ SDK 2.0 (aqui funcionando como referência) e do SDK ARCore®, este utilizando dois *smartphones* diferentes, um Samsung® Galaxy A7 e um Huawei® P20. De realçar que existe um ligeiro desvio entre os resultados apresentados na Figura 27 em especial no eixo X (distância percorrida) e no eixo Z (diferença de altura ente o teste da Intel® RealSense™ SDK 2.0 e os dois *smartphones*). Ainda nesta figura, próximo das coordenadas (8,-1) [m], podem ser vistos alguns "saltos", sendo estes relativos a um conjunto de movimentos mais específicos feitos pelo utilizador no final do teste, tais como um conjunto de agachamentos e um conjunto de movimentos de rotação em torno do eixo Z (rotação da cabeça).

Em resumo, nesta primeira fase foram realizados testes qualitativos de validação dos algoritmos para a determinação de pose, que permitiram avaliar diferentes soluções. Os testes realizados com a Intel® RealSense™ T265 com o SDK 2.0 são os que apresentam erros máximos dentro dos parâmetros admissíveis para este projeto. No caso do smartphone Huawei P20 com SDK ARCore®, o erro nos eixos x e y aumenta comparativamente com o anterior, sendo que o eixo z apresenta melhores resultados. Os Epson Moverio BT-35E com SDK MAXST® apresentam um erro no eixo x ainda dentro dos parâmetros pretendidos, mas os restantes eixos,

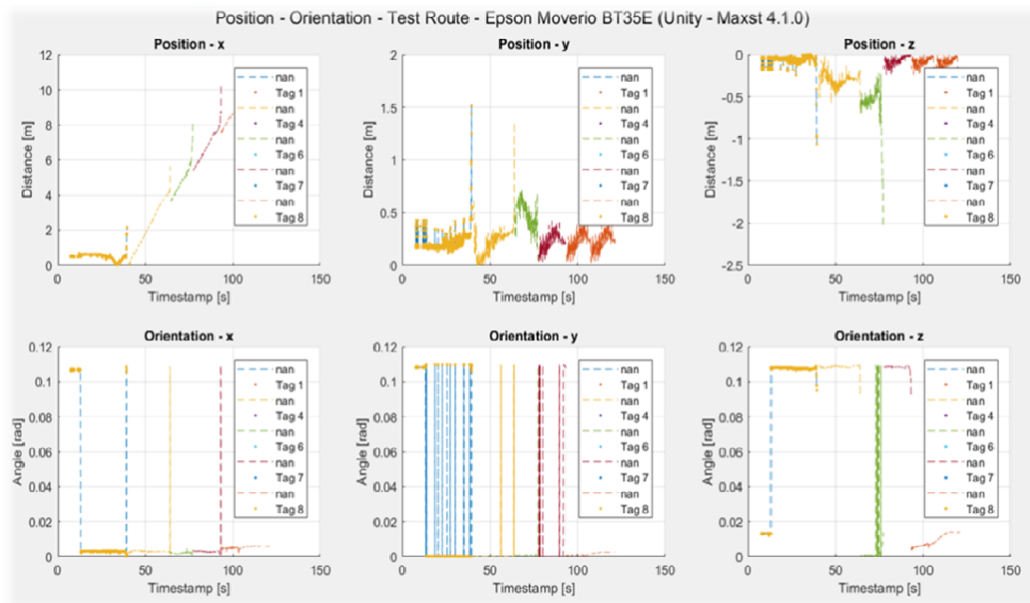


Figura 26: Dados da pose para o caminho linear (Epson Moverio BT-35E + MAXST® + Unity®).

y e z, contém erros bastante elevados, sendo esta a solução que apresentou piores resultados.

Na Seção seguinte descrevem-se os testes quantitativos recorrendo a sistemas de localização de maior precisão, incluindo robôs industriais, estações topográficas e um sistema de localização especificamente construído para o efeito.

4.2 SEGUNDA FASE DE TESTES

Os testes realizados nesta segunda fase foram direcionados a dois equipamentos específicos: os óculos de realidade aumentada (Epson® Moverio BT-35E) e uma webcam (Logitech® Webcam BRIO 4K Ultra HD). Estes equipamentos, vêm de acordo com a necessidade dos testes com sistemas de localização interior, tendo como principal foco a visão computacional, sendo que a escolha dos óculos teve como base a análise apresentada na Seção 2.6 e que, no caso webcam, este foi um equipamento adquirido para o projeto com o objetivo de ter um hardware dedicado apenas a aquisição de imagem. A utilização destes equipamentos foi realizada através de software desenvolvido utilizando o editor Visual Studio Code com Add-Ons para Python® (versão 3.8.5), com base na solução exemplo do OpenCV®².

² https://docs.opencv.org/4.x/d5/dae/tutorial_aruco_detection.html

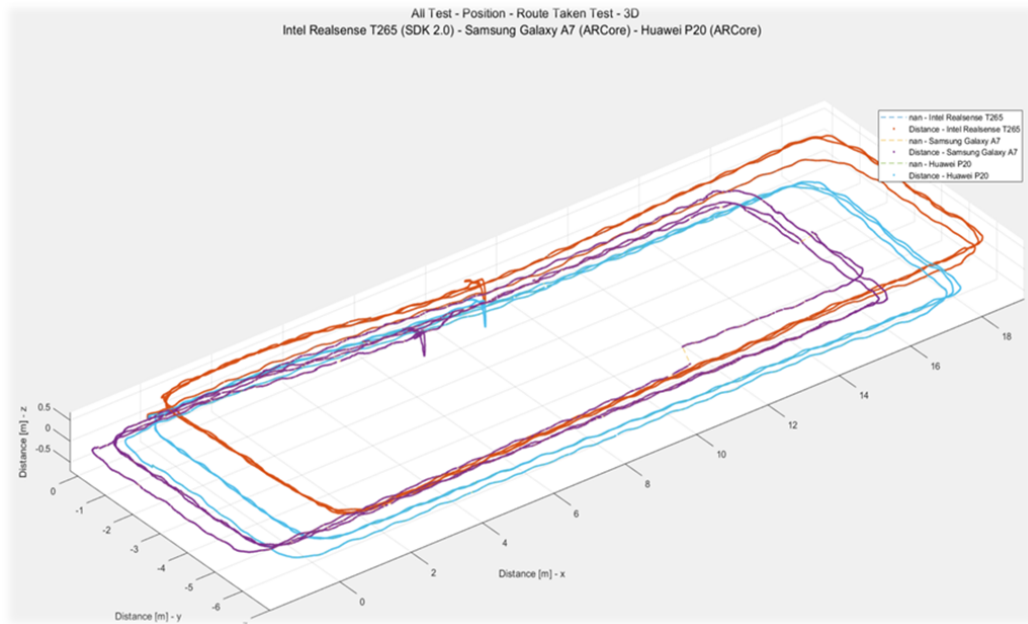


Figura 27: Visão geral da posição obtida com a Intel® T265, o Samsung® Galaxy A7 (ARCore®) e o Huawei® P20 (ARCore®) ao longo da realização de 3.5x um percurso circular no laboratório.

O estudo realizado inicialmente na tarefa de revisão do estado da arte e com base nos testes realizados na primeira fase deste projeto, conclui-se que os vários SDK de detecção de marcadores, nomeadamente o MAXST®, ARToolKit, ARCore®, ArUco e Vuforia®, estão de acordo com as necessidades deste projeto. O percurso que o utilizador terá de realizar, deverá ter o maior número de marcadores possível com curtas distâncias entre si. Por exemplo, o teste apresentado abaixo (Secção 4.2.1), onde foram colocados marcadores com as dimensões de 0.3m por 0.3m a uma distância de 1m entre eles, tanto longitudinalmente como transversalmente (eixo x e y), sendo que estes se encontram na mesma altura (eixo z), podemos observar na Figura 28 um exemplo do citado. Além disso, é necessário ter um mapeamento desses mesmos marcadores, de forma a saber a pose destes num referencial conhecido, permitindo assim a aplicação do algoritmo descrito na Secção 3.8.

Como se pode observar na primeira fase de testes foram utilizados os algoritmos e marcadores do SDK MAXST®, mas apesar das intervenções nesses algoritmos de forma a obter valores de erro mais baixos, não foi possível. Com isto depois de várias questões efetuadas ao fabricante deste mesmo SDK, o mesmo referiu que não seria possível alcançar erros inferiores a 0.5m. Com este entrave encontrado foi decidido direcionar os desenvolvimentos para os outros SDK. Tal como descrito acima, dados os resultados obtidos com o SDK MAXST, este não foi considerado na

segunda fase de testes. Assim, nesta segunda fase de testes, foram direcionados os novos trabalhos de desenvolvimento utilizando algoritmos e os marcadores ArUco.

4.2.1 Testes com marcadores em ambiente simulação

De forma a analisar o desempenho do algoritmo de detecção dos marcadores, por comparação dos valores estimados com os valores reais, optou-se por realizar testes recorrendo a um manipulador robótico industrial. Numa primeira fase os testes foram realizados em ambiente de simulação e posteriormente no robô real.

Para a realização dos testes em simulação foi utilizado o software RobotStudio 2020.1 para correr a aplicação baseada no manipulador robótico industrial, e o Python 3.8.5 (usando o Visual Studio Code), para a aplicação de detecção de marcadores, comunicando as duas entre si via *socket* TCP. No software da RobotStudio foi utilizado o robô IRB 2400, modelo disponível no laboratório, três marcadores ArUco (DICT_6X6_250), com os IDs 0, 1 e 2. Foi colocado, no elemento terminal do robô, uma câmara simulada com as características da câmara Logitech® Webcam BRIO 4K Ultra HD, resultando no ambiente de testes simulado visível na Figura 28.

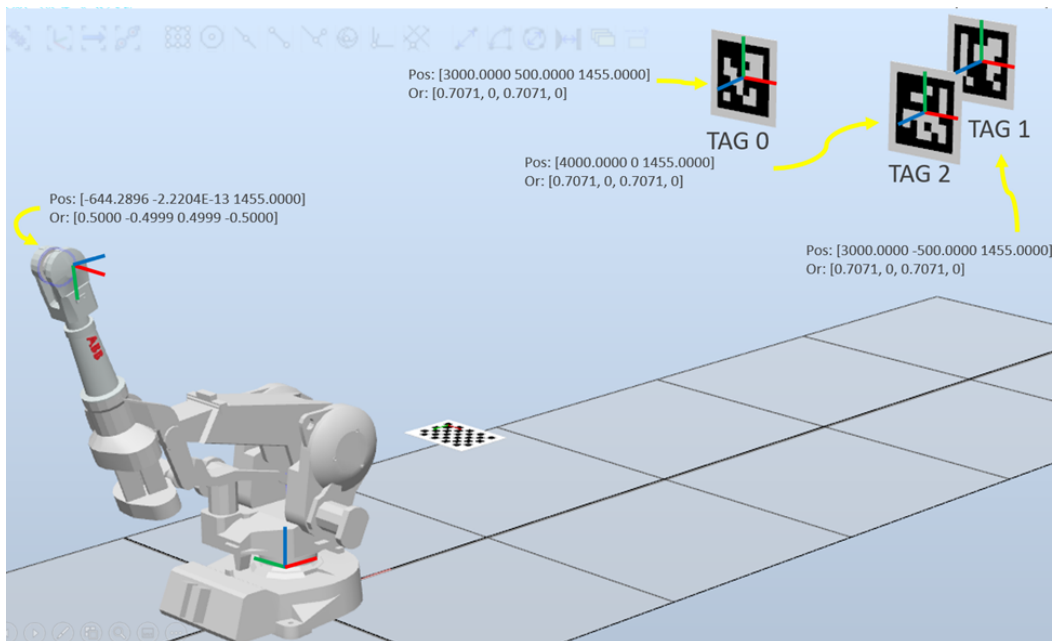


Figura 28: Ambiente de testes no software RobotStudio.

Foi criada uma trajetória para o robô com 6 tipos de movimentos, com translação em x, y e z, para os primeiros 3 tipos de movimentos, respetivamente, e rotação

em torno de x, y e z, respetivamente, para os 3 últimos movimentos. Em todos estes testes os 3 marcadores foram sempre mantidos no campo de visão da câmara, registando, para cada instante de tempo, a pose real da câmara, a pose estimada de cada marcador e a pose estimada da câmara. Os resultados obtidos para o movimento Rz estão ilustrados na Figura 29 e na Figura 30.

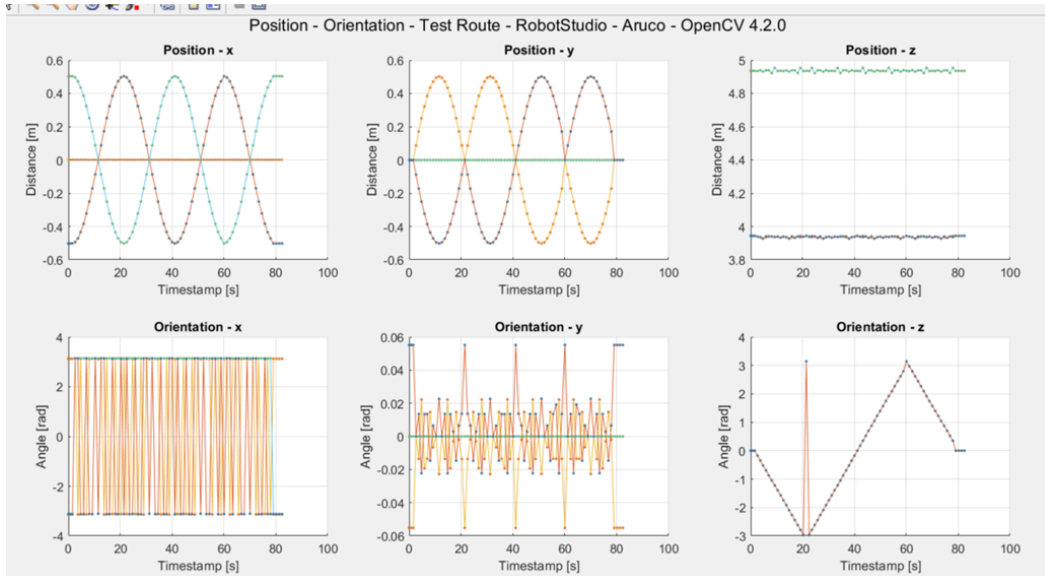


Figura 29: Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).

Na Figura 31 e Figura 32 estão visíveis os resultados por eixo da posição estimada de cada marcador relativamente à câmara, sendo que os resultados associados aos restantes movimentos estão disponíveis no Anexo B. De notar que a orientação em X e Z está sempre próxima de 180° e, por isso, altera frequentemente entre 180° e -180° .

Tendo em conta os resultados obtidos em simulação, conclui-se que o erro, em termos de posição na deteção dos marcadores, é mais elevado no eixo z do que nos restantes eixos, tendo sido obtido um valor máximo de erro de 0.02m para uma distância máxima entre a câmara e os marcadores de 4.8 m. Relativamente aos testes de rotação, o erro varia em função do eixo em torno do qual a rotação está a ser realizada, tendo-se obtido um valor máximo de erro de 0.1 rad (5.7°).

Na maioria dos testes, tanto em posição como em orientação, o marcador 0 apresenta um erro superior, comparado com os restantes. Para averiguar a origem deste erro, foram realizados testes adicionais com o referido marcador, um dos quais consistiu em alterar a localização do marcador 0 pelo marcador 2. Deste modo foi possível constatar que o erro existente se manteve idêntico para cada um dos marcadores, continuando a ser mais elevado para o marcador 0. Para além disso, foi

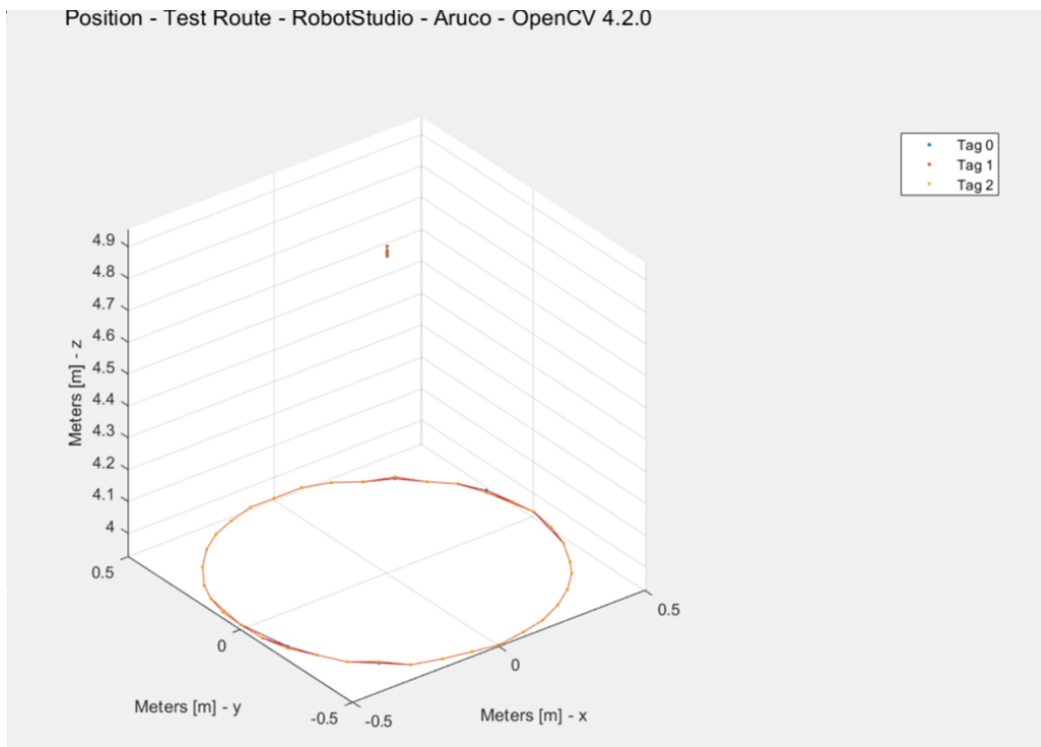


Figura 30: Gráfico de posição em 3D dos marcadores relativamente à câmara.

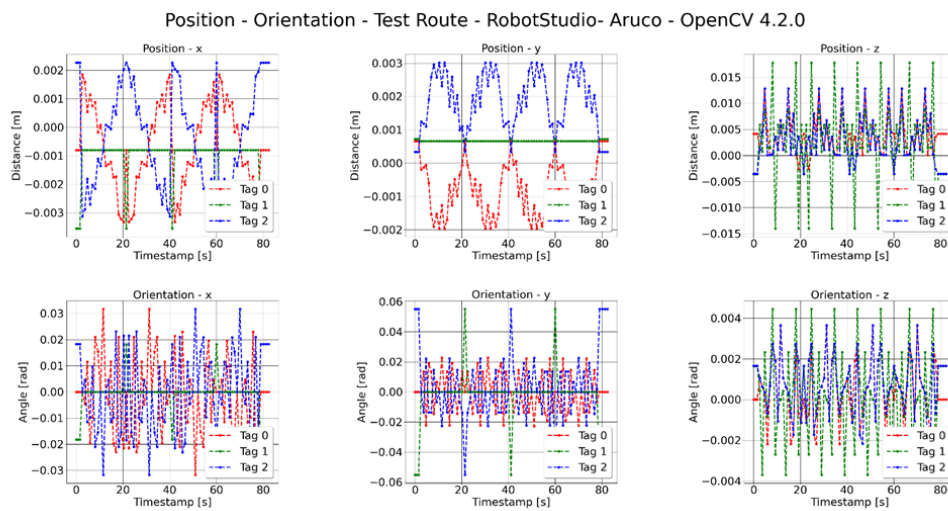


Figura 31: Gráfico do erro da posição e orientação estimadas, por eixo, de cada marcador relativamente à câmara.



Figura 32: Gráfico de erro de estimação de posição e orientação por marcador.

ainda feito um teste com o marcador 0, sem qualquer outro marcador, numa pose fixa da câmara relativamente ao marcador, sem qualquer movimento. De seguida colocou-se o marcador 2 nesse mesmo local, realizando o mesmo teste, podendo assim comparar o erro apresentado pelo mesmo e, mais uma vez, o erro associado à deteção do marcador 0 foi mais elevado. Com base nestes resultados, chegou-se à conclusão que a diferença obtida no erro de deteção seria proveniente da codificação do próprio marcador, tendo este um erro superior na estimativa da posição e da orientação, comparativamente com os restantes em estudo.

Numa análise geral, considerando os resultados de deteção obtidos no ambiente simulado, para as distâncias utilizadas, os erros de localização de cada um dos marcadores são bastante reduzidos, sendo que, como esperado, o erro de posição aumenta com o aumento da distância entre a câmara e o marcador, bem como quando existe rotação em torno de qualquer um dos eixos, caso em que o aumento da rotação entre a câmara e o marcador leva a um maior erro de orientação.

4.2.2 Testes de marcadores em ambiente real

Seguidamente, os testes de simulação foram replicados em ambiente laboratorial, utilizando o mesmo modelo de robô, o robô IRB 2400, equipado com a Webcam Logitech® BRIO 4K Ultra HD, e os três marcadores ArUco (DICT_6X6_250), com os mesmos IDs 0, 1 e 2 (ver Figura 33). Tal como no caso anterior, são apresentados apenas os resultados obtidos para o movimento Rz, os restantes movimentos estão disponíveis no Anexo C.

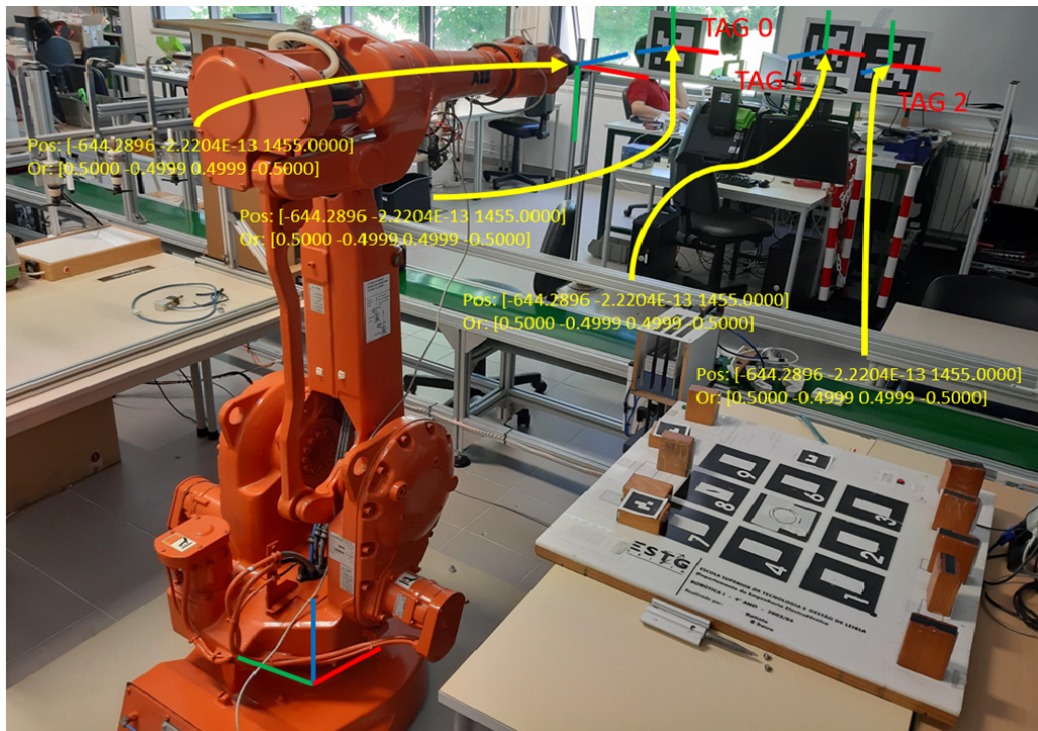


Figura 33: Ambiente real de testes de marcadores.

Os resultados obtidos estão ilustrados na Figura 34, Figura 35, Figura 36 e Figura 37. Para confirmação da localização de cada marcador foi utilizado o Teodolito (Total Leica TS02) como *ground-truth*, estando este numa pose conhecida e referência de acordo com o referencial do robô. Com isto, após retirar 3 pontos de cada marcador, foi possível determinar a pose de cada um dos marcadores.

Com base nos resultados obtidos com os testes em ambiente real pode concluir-se que, à semelhança dos testes em simulação, o erro em termos de posição na deteção de marcadores é mais elevado no eixo z, tendo obtido um erro máximo de 0.08m, para uma distância máxima entre câmara e marcadores de 4.8 m. Relativamente à orientação estimada na deteção de marcadores, obteve-se uma variação de erro ligeiramente mais elevado que nos testes em simulação, sendo que o valor máximo de erro obtido é aproximadamente 0.3 rad (17.2°).

À semelhança dos testes em ambiente simulado, foram realizados testes com o marcador 0, onde se chegou às mesmas conclusões, i.e., o erro da deteção da pose relativamente a este marcador é geralmente mais elevado.

Comparativamente aos testes em ambiente simulado, o erro é razoavelmente superior mas, tal como esperado, este varia sobretudo com o aumento da distância entre a câmara e o marcador, bem como com o aumento da orientação da

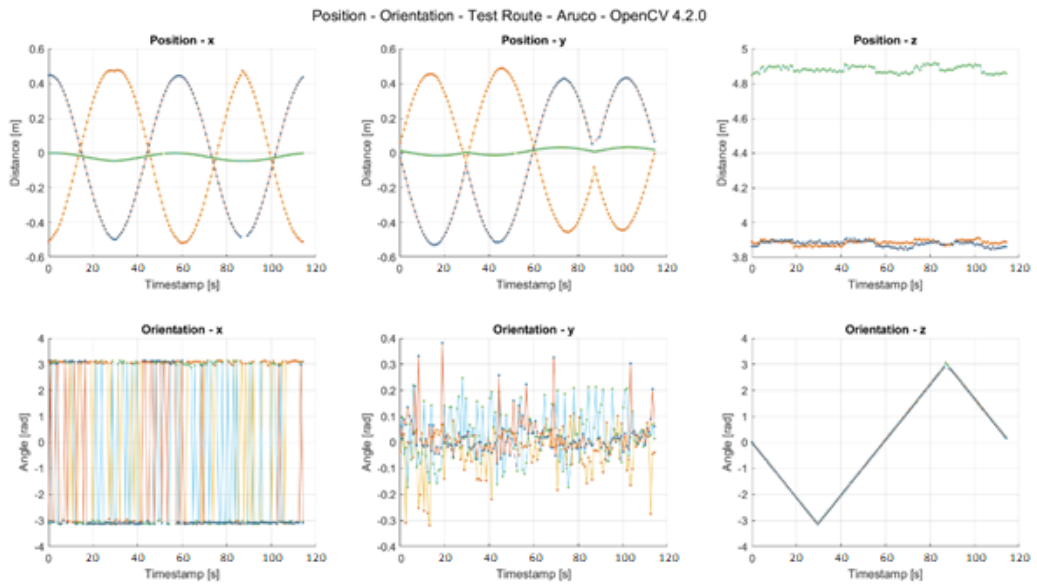


Figura 34: Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).

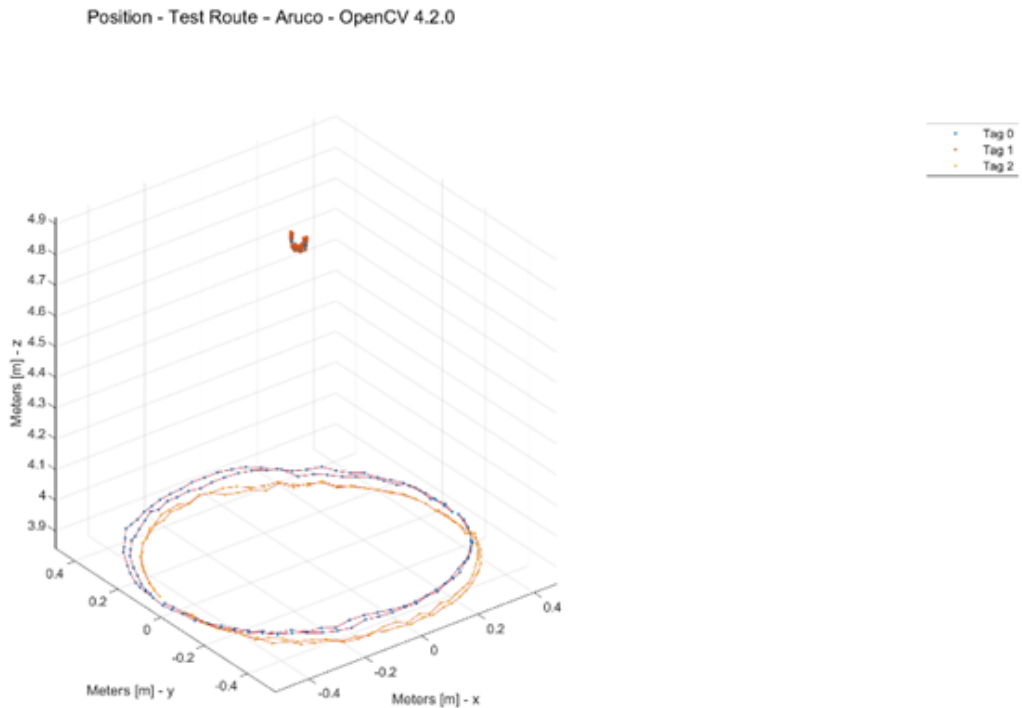


Figura 35: Gráfico de posição em 3D dos marcadores relativamente à câmara.

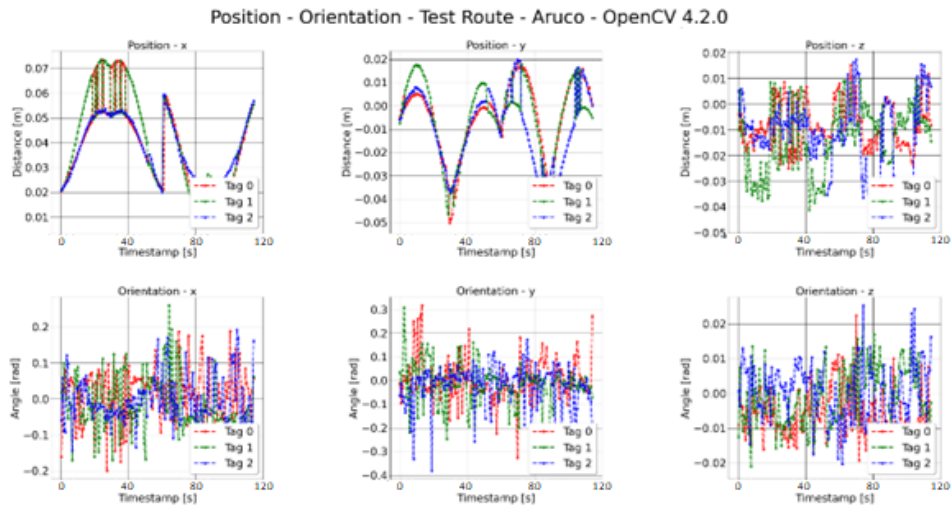


Figura 36: Gráfico de erro de posição e orientação, por eixo, de cada marcador relativamente à câmara.

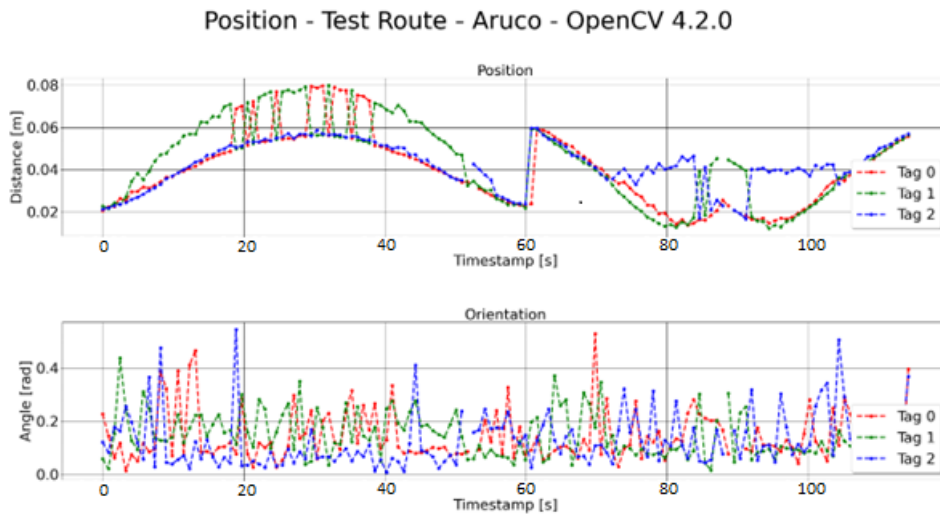


Figura 37: Gráfico de erro de posição e orientação por marcador.

câmara relativamente aos marcadores. Neste caso é de referir alguma preocupação relativamente à orientação, pois o erro é mais elevado.

4.2.3 *Conclusões da segunda fase de testes*

Os testes realizados em simulação e com recurso ao robô, de forma a permitir acesso ao valor real da pose da câmara e, assim, estimar o erro ao longo do tempo, permitiram validar a deteção dos marcadores e a estimação da pose da câmara com base nestes resultados. Permitiu ainda obter uma ordem de grandeza para os erros de deteção de cada marcador e o seu impacto na estimativa da pose da câmara.

4.3 TERCEIRA FASE DE TESTES

Os testes realizados nesta fase foram direcionados para percursos com o utilizador que, embora em ambiente laboratorial (Figura 38), procurassem ser semelhantes ao ambiente final (Figura 43), pelo menos em algumas características, tais como a distância entre marcadores e altura dos marcadores, este conjunto de elementos causa consequências no desempenho do algoritmo de localização.

Esta fase de testes teve por base a utilização dos óculos de RA (Epson® Moverio BT-35E) com a deteção de marcadores ArUco baseada em OpenCV®, tal como descrito anteriormente. Assim, o principal objetivo nesta fase era realizar um percurso em tempo real, recolhendo a pose do utilizador, i.e., a pose dos óculos de RA.

Foram colocados ao longo do laboratório doze marcadores ArUco (DICT_6X6_250), do marcador 0 até ao marcador 11, como mostra a Figura 38. Adicionalmente foi colocada a Intel® RealSense™ T265 junto aos óculos, para efeitos de depuração, embora esses dados não sejam apresentados neste relatório por limitações de espaço.

A Figura 39 e na Figura 43 ilustra o percurso definido para os testes. Testes estes que ditaram a distância entre os marcadores (apresentado na Figura 40) e a altura dos marcadores ao chão que se encontra no LRAFI. O ambiente real, no laboratório, foi ajustado de forma a reduzir o nível de luminosidade, procurando aproximar o ambiente laboratorial do ambiente real, resultando numa iluminação com cerca de 350 lux. O referencial de trabalho/mundo está colocado no canto superior direito do laboratório, ao nível do chão como mostra a Figura 39. Neste teste, o percurso é realizado por um ser humano, ao longo de uma linha reta, marcada a vermelho

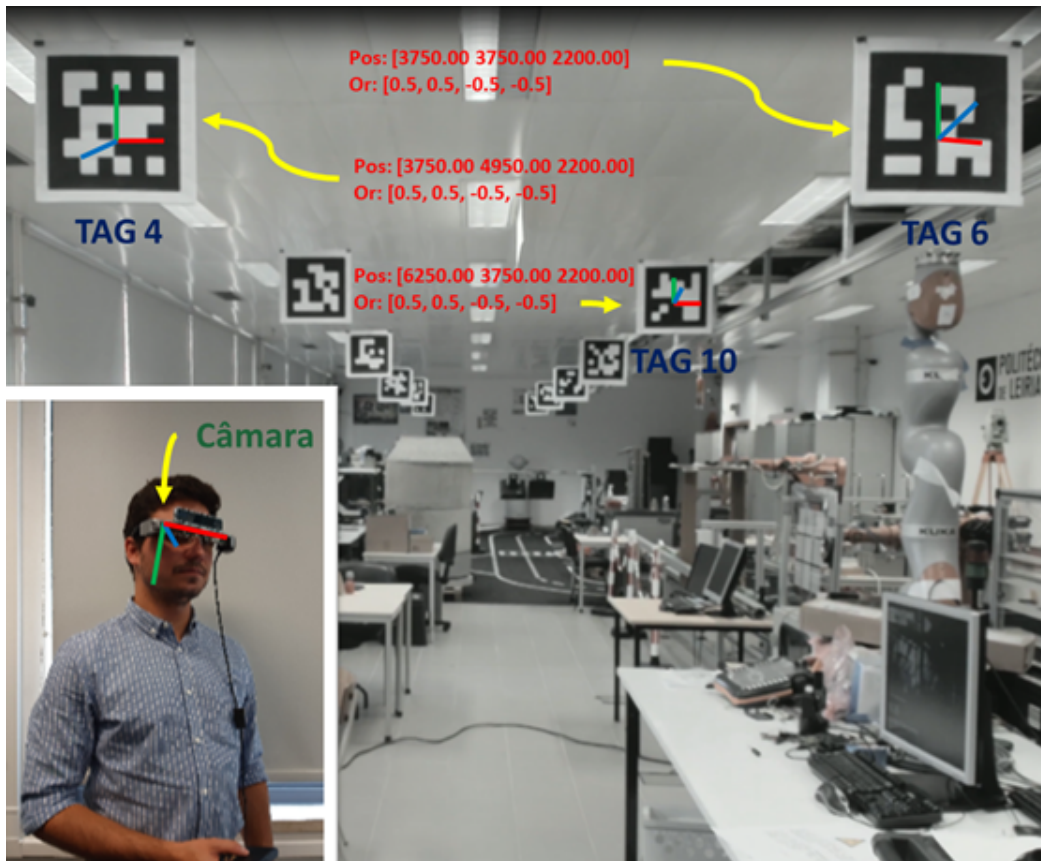


Figura 38: Ambiente real para testes com o ser humano equipado com os óculos de RA.

na Figura 39 iniciando o percurso na seta castanha e terminando o mesmo antes de intercepar a segunda linha vermelha que se encontra perpendicular à mesma, ao longo de 14m de distância, aproveitando apenas a linha reta com maior distancia.

Na Figura 41 encontram-se os resultados do percurso realizado pelo utilizador dos óculos de RA, sendo este orientado pelas instruções apresentadas nos óculos de RA.

4.3.1 Conclusões terceira fase de testes

Com os resultados obtidos nestes testes podem retirar-se várias conclusões. No percurso realizado pelo utilizador dos óculos de RA é notório que existem alguns desvios do utilizador ao longo desse percurso, devido aos objetos que se encontram ao longo do laboratório. Apesar de não termos um *ground-truth* de forma a comparar os dados apresentados na Figura 48 e na Figura 49, pode-se dizer que o erro ao longo do percurso, qualitativamente/quantitativamente, não será superior ao erro máximo definido para este projeto nos 3 eixos de posição, podemos ver que o percurso em termos de distância percorrida é cerca de 14m (eixo x). Em termos de desvios

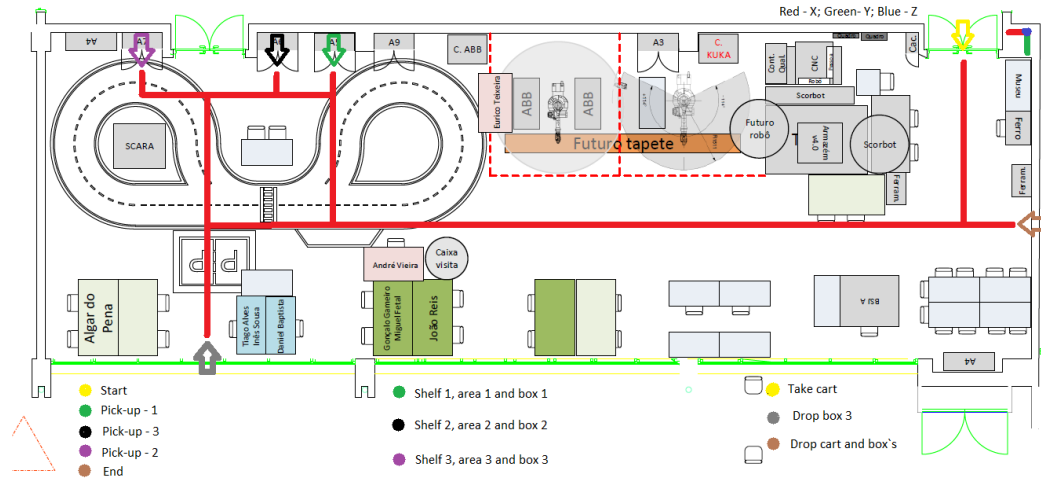


Figura 39: Vista de topo do ambiente real com o percurso a realizar pelo ser humano no laboratório (as cores dos círculos, acima da imagem, corresponde a legenda da setas na imagem).

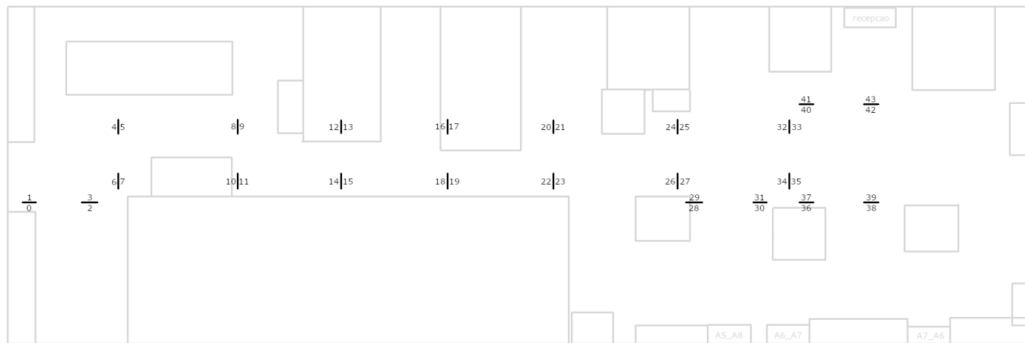


Figura 40: Vista de topo do ambiente real com os marcadores espalhados no percurso a realizar pelo ser humano no laboratório.

do percurso (eixo y) são notáveis alguns desvios dos objetos mas, além disso, é aqui que se encontram erros de maior dimensão. O esperado seria uma linha de pontos sensivelmente retilínea, existindo apenas pequenos desvios devido ao andar do utilizador. Em termos de altura (eixo z) o utilizador que realizou estes testes tem cerca de 1,75m, o que significa que, ao longo do percurso, tem-se uma altura média dos óculos de 1,6m, como se verificou. Em termos de orientação, pode dizer-se que “continha” um *ground-truth* pois, como o utilizador estava a ser orientado pelos óculos de RA, onde é apresentada a projeção das setas no ecrã, indicando assim a orientação do percurso (ver Figura 42) com base na posição e orientação enviada pela localização. Pode assim dizer-se que a orientação apresentada nos óculos foi coerente com o percurso idealizado.

Comparando os resultados nesta fase de testes com a anterior, podemos dizer que o erro em termos de posição e orientação aumentou. Na segunda fase de teste,

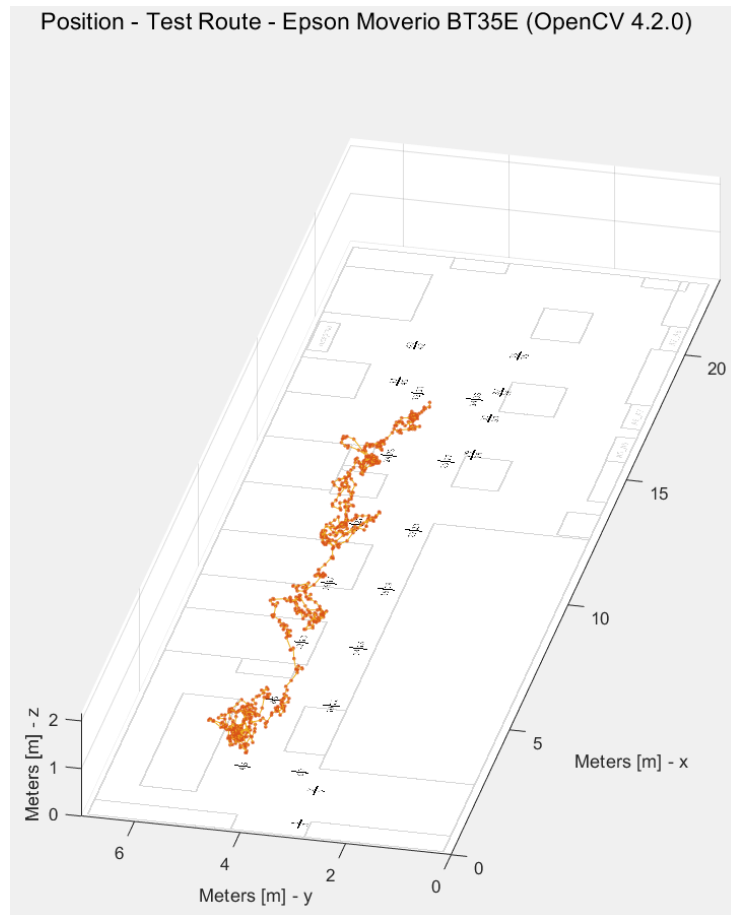


Figura 41: Gráfico de posição em 3D (teste com o utilizador).

tivemos valores de erros máximos de posição de cerca de 0.015m e orientação de cerca de 0.08 rads (5°) nos testes em ambiente simulado, enquanto que nos testes em ambiente real esse valor aumentou, tendo um erro máximo de posição de 0,08m e de orientação de 0,4 rads (22°) como se pode observar na Figura 37 e na Figura 36.

4.4 TESTE E DEMONSTRAÇÃO DO SISTEMA EM AMBIENTE REAL

A abordagem nesta fase de testes utiliza duas estratégias. Por um lado, recorre-se aos dados determinados pelo sensor Intel® RealSense™ T265 como *ground-truth*, por outro lado, recorre-se às imagens da câmara do dispositivo de RA (Epson® Moverio BT-35E) e a algoritmos de odometria visual implementados sobre a unidade de processamento (Huawei® EML-L29), complementados pela unidade inercial dos óculos de RA

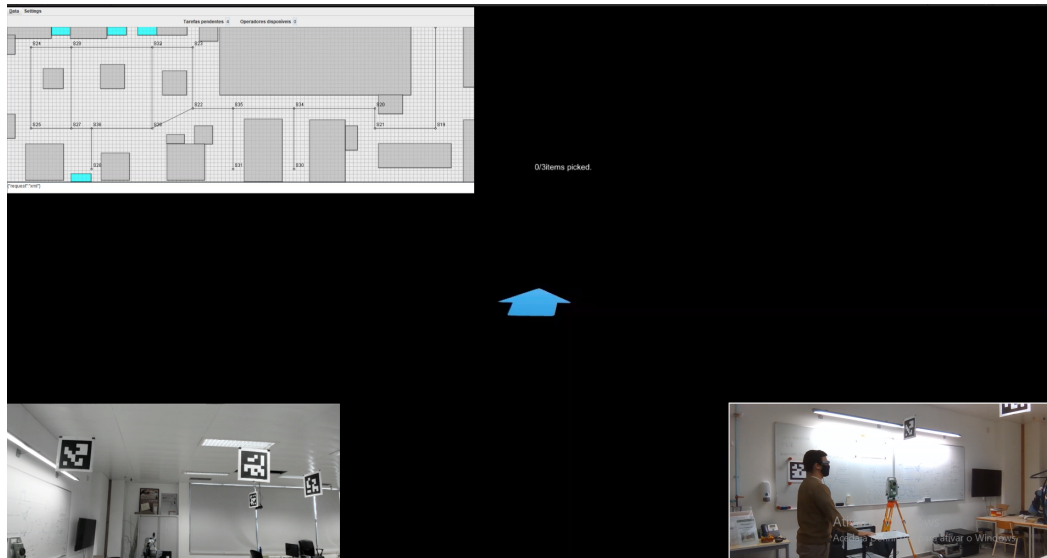


Figura 42: Vista da projeção das setas no ecrã dos óculos, com os marcadores ao longo dos corredores do laboratório e projeção gráfica algoritmo de otimização de rotas.

Estes testes, realizados num armazém da empresa Matceramica, tiveram em conta a adaptação da operação típica dos operadores ao sistema desenvolvido.

Na Figura 44 pode observar-se parte do ambiente onde foram realizados os testes, estando visível no computador a informação que é exibida ao utilizador nos óculos pela componente da realidade aumentada (desenvolvida por outros elementos da equipa no projeto ARWARE).

Tendo em conta o mapa do armazém, realizado por outros elementos do projeto ARWARE e ilustrada na Figura 45, foram realizados vários percursos, marcados a vermelho nessa mesma Figura. Previamente, recorrendo a um teodolito, foi feito um levantamento da pose de todos os marcadores, incluindo essa informação no mapa do armazém. Além dos marcadores já referidos anteriormente, foram também colocados marcadores mais pequenos, com um dicionário disponível pela ArUco, mas diferente dos de maior dimensão, com o objetivo de utilizar marcadores com número de bits distinto (biblioteca DICT 6X6 250 - marcadores maiores e DICT 4X4 250 - marcadores mais pequenos), como se pode ver na Figura 46. Estes marcadores foram adicionados de modo a aumentar a precisão na fase em que o operador se dirige às prateleiras, dado que os marcadores colocados ao longo dos corredores deixam de ser visíveis nessa altura.

A pose estimada ao longo do teste realizado está ilustrada na Figura 47, na Figura 48 e na Figura 49. Analisando o resultado obtido, pode verificar-se que, ao longo do percurso realizado, no eixo x, onde foi percorrida uma distância por volta de 7.5m perpendicularmente com o conjunto das prateleiras, é notório um ligeiro desvio da

4.4 TESTE E DEMONSTRAÇÃO DO SISTEMA EM AMBIENTE REAL



Figura 43: Vista com os marcadores ao longo dos corredores do armazém da empresa Matceramica.

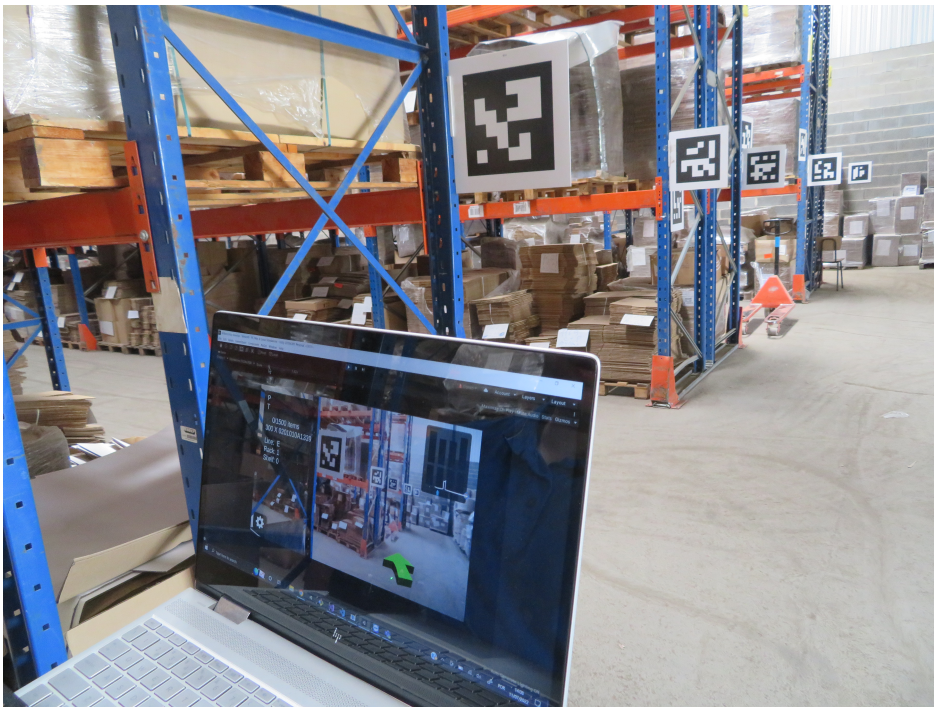


Figura 44: Teste em ambiente real, armazém da Matceramica.

TESTES E RESULTADOS

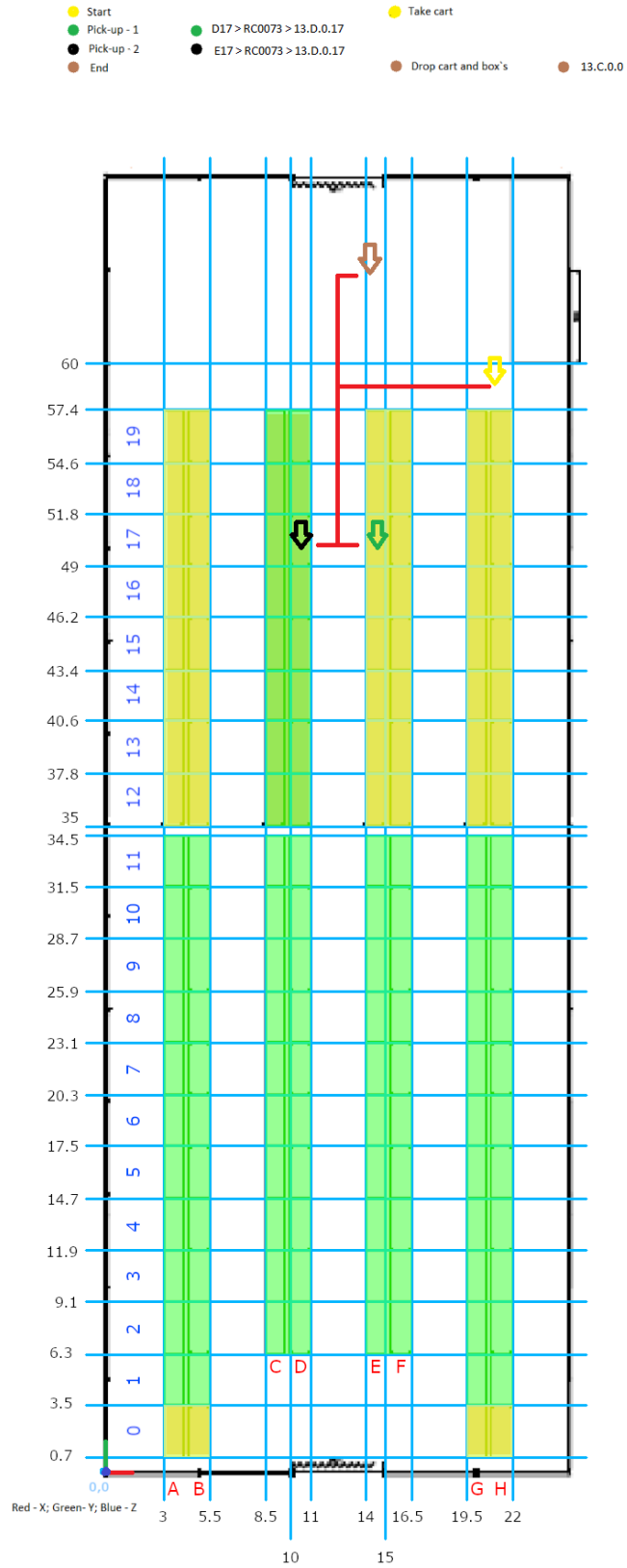


Figura 45: Mapa 2D (vista de topo) do ambiente real com o percurso a realizar pelo ser humano.



Figura 46: Vista com os marcadores ao longo das prateleiras do armazém da empresa Matceramica.

rota, que na realidade o utilizador não realizou. Já no eixo y (no corredor que contém prateleiras/marcadores de ambos os lados) a distância percorrida junto às prateleiras foi sensivelmente 10m, onde se pode afirmar que, os resultados obtidos, apesar de terem um ligeira diferença com a rota realizada, o algoritmo apresenta estimativas da pose aceitável. No entanto, quando o utilizador retorna pelo mesmo percurso, após passar o corredor onde se encontram os marcadores, e este deixar de visualizar os mesmos, o algoritmo apresenta resultados de apenas mais 3m aproximadamente do que na realidade o utilizador andou (6.5m). Tal acontece pelo facto de existir falta de atualização da informação e o algoritmo mantém-se na mesma pose. Por fim, em termos de altura, ou seja, o eixo z, temos variações muito baixas, tal como aconteceu nos testes realizados em ambiente laboratorial. Em termos de orientação temos resultados idênticos aos anteriores, na terceira fase de testes, pois mais uma vez a orientação do utilizador é feita pelos óculos de RA, tendo em consideração que a pose é enviada pela localização, pode dizer-se que a orientação apresentados nos óculos foi coerente com o percurso idealizado.

Ao realizar os testes, uma das primeira constatações foi o facto do nível de luminosidade ser bastante baixo, além de poder variar substancialmente ao longo do dia e ao longo do ano, dado que o nível de luz artificial disponível no armazém é

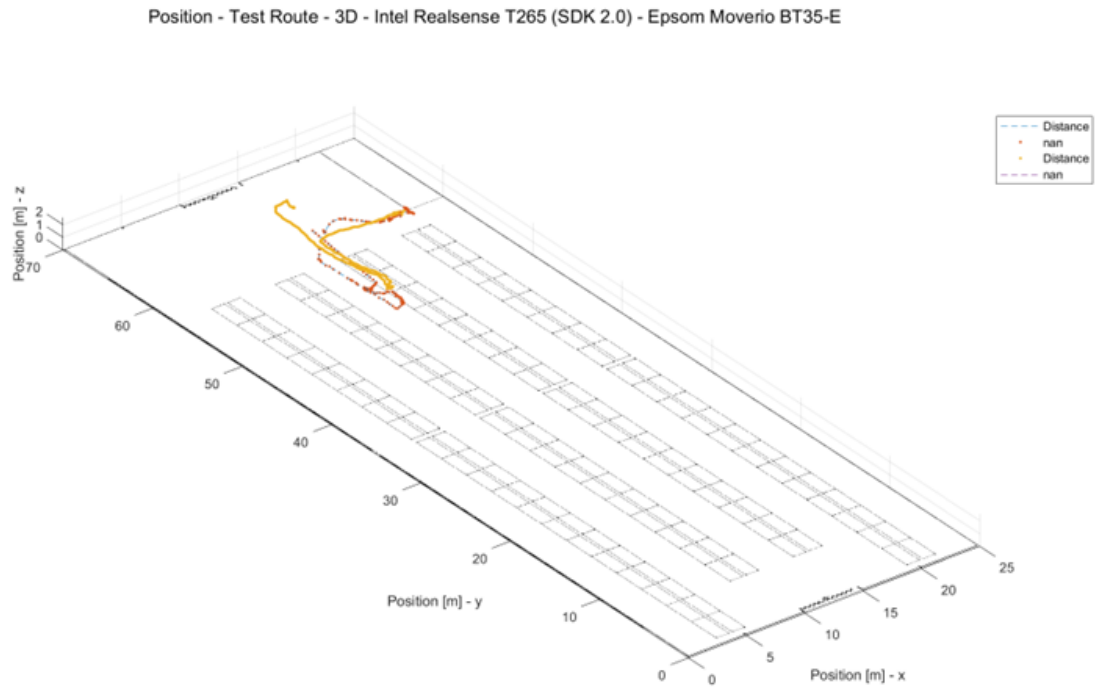


Figura 47: Visão geral da posição estimada ao longo do percurso realizado no armazém da Matceramica.

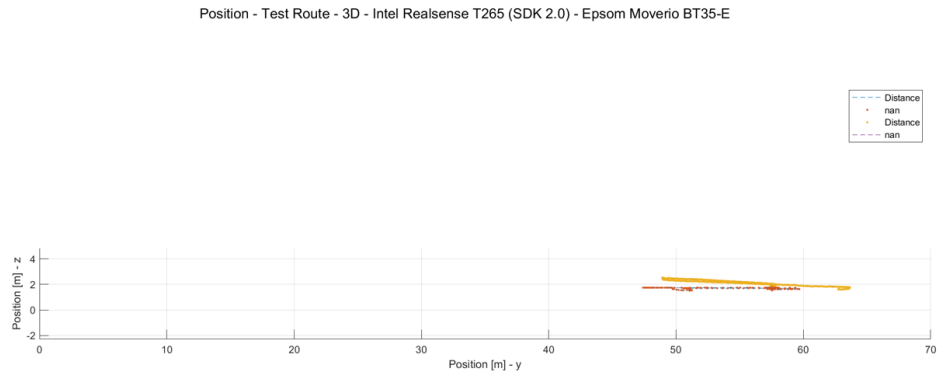


Figura 48: Visão eixo z e y da posição estimada com o algoritmo ao longo do percurso realizado no armazém da Matceramica.

4.4 TESTE E DEMONSTRAÇÃO DO SISTEMA EM AMBIENTE REAL

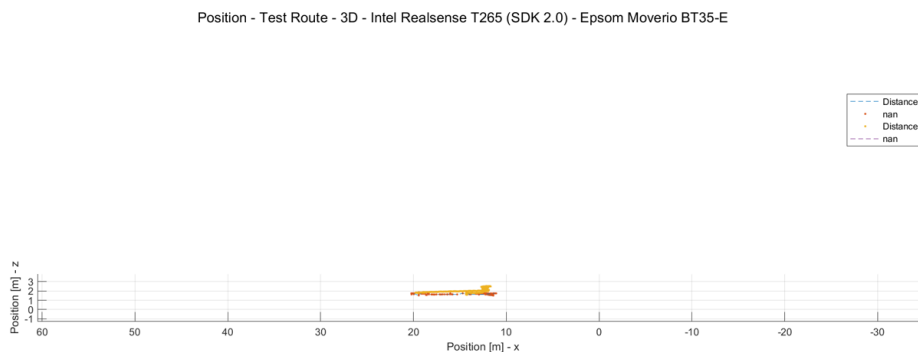


Figura 49: Visão eixo z e x da posição estimada ao longo do percurso realizado no armazém da Matceramica.

bastante reduzido. Este facto leva a que seja necessário um maior tempo de exposição aquando da aquisição de imagens, o que leva não só a uma diminuição da taxa de aquisição de imagens, como também à perda de nitidez na presença de movimento (*motion blur*). A luz natural tem como principal problema o facto de se alterar consoante a hora do dia, ao passo que a luz artificial, apresenta problemas como a deteção da variação da luminosidade das luzes pela câmara, devido à frequência da rede. Para além disso, como foi anteriormente previsto, o hardware utilizado poderia ter alguns problemas em termos de captações de imagens. Esta limitação faz com que marcadores mais afastados possam não ser detetados ou tenham um erro superior em termos de localização estimada, o que influencia negativamente o erro de estimação da pose. Estas influências negativas no algoritmo são reduzidas pelos filtros implementados, melhorando toda a componente de localização.

O conjunto de testes realizados nesta fase permitiram avaliar o impacto das alterações descritas anteriormente, bem como melhorar a interligação entre a componente de LF e RA.

Para chegar ao algoritmo e parametrização utilizada nos testes agora apresentados, foram ainda testadas várias versões do algoritmo desenvolvido, tendo em conta a possibilidade de parametrização de vários dos passos (parâmetros de deteção dos códigos QR, critérios de paragem do ICP, entre outros), de forma a procurar reduzir o erro na estimativa da pose dos óculos, bem como o desempenho em geral.

Nesta fase, concluiu-se que o algoritmo era capaz de estimar a pose dos óculos a uma média de 21 vezes por segundo, embora ainda como um erro médio de

estimativa que não se coaduna com a projeção 3D de informação por parte da RA. A origem mais provável deste erro será a dificuldade na deteção dos marcadores, devido à baixa luminosidade que se encontrava neste espaço, comparativamente com os testes realizados em laboratório.

De salientar ainda que estes testes incluíram todos os componentes envolvidos no sistema global, nomeadamente a RA, o Planeador, o ERP, o modelador de Armazém e a Localização Aproximada, não obstante esta última ser simulada, dada a ausência da instalação do hardware necessária para o efeito no armazém.

4.5 LIMITAÇÕES DO SISTEMA

Ao longo do projeto foi identificado um conjunto de restrições técnicas, que a seguir se descrevem:

- Ambientes com pouca luz causam uma grande dificuldade. Dado que a principal fonte de informação advém do processamento das imagens capturadas por uma câmara, é essencial ter condições de iluminação adequadas;
- A falta de textura no ambiente, também se torna uma dificuldade. O algoritmo desenvolvido, quando processa uma imagem capturada, aproveita a maioria da informação disponibilizada pela mesma, assim sendo, quanto maior for o número de elementos em destaque maior será o sucesso;
- Idealmente o processamento de toda a aplicação deveria ser feita num *smartphone*, ou num outro dispositivo de pequenas dimensões. No entanto, com o avanço dos desenvolvimentos observou-se que é necessário um nível de processamento bastante elevado, o que levou à necessidade de utilizar um dispositivo com maiores dimensões, como por exemplo um computador;
- O tamanho do “conjunto de equipamentos” criado de forma a cumprir os requisitos, é considerável, embora o peso não seja notório (comparativamente com os óculos da Hololens da Microsoft®), pois é distribuído entre os óculos de RA (transportado pelo utilizador na cara) e um processador móvel (transportado pelo utilizador numa mochila).

Futuramente para diminuir as limitações do sistema desenvolvido, pode-se recorrer a um conjunto de tecnologias e fontes de dados convenientemente integradas, através de diferentes algoritmos e equipamentos de processamento.

CONCLUSÕES E TRABALHO FUTURO

A RA é uma tecnologia recente e que tem um potencial de expansão grande, e que certamente irá revolucionar a indústria. Hoje em dia existe um número considerável de empresas a produzir equipamentos de RA, uma boa parte das quais com o objetivo de melhorar a eficiência dos processos industriais como, por exemplo, os armazéns de logística industriais.

Neste relatório foi apresentado e detalhado um sistema de localização para ambientes interiores, com o principal objetivo de criar uma aplicação de localização interior para aplicação a uns óculos de RA, estimando a pose do utilizador ao longo do percurso. O sistema de localização baseou-se numa câmara, acelerómetros e giroscópios (ambos de 3 eixos), e na distribuição de marcadores ArUco em localizações previamente definidas.

Apesar disso, o algoritmo de localização contém algumas limitações técnicas a considerar como, por exemplo: ambientes com pouca luz são uma grande dificuldade, dado que a principal fonte de informação advém do processamento das imagens capturadas por uma câmara, é essencial ter condições de iluminação adequadas; a falta de textura no ambiente no ambiente também se tornam uma dificuldade, dado que, o algoritmo desenvolvido quando processa uma imagem capturada, aproveita a deteção dos marcadores disponibilizada pela mesma, assim sendo, quanto maior for o número de marcadores detetados, menor será o erro da estimativa da pose; disponibilidade de poderosos processadores móveis, idealmente o processamento de toda a aplicação deveria ser feito num *smartphone*, mas com o avanço dos desenvolvimentos observou-se que é necessário um processador móvel com alta capacidade. Note-se ainda que o projeto foi desenvolvido desde início num computador, e só mais tarde se pretendia passar para um *smartphone*, ou para um outro dispositivo de menores dimensões, sendo que neste momento utiliza-se uma junção de equipamentos sendo que todo o dispositivo criado tenha um tamanho considerável; apesar do seu tamanho, o peso não é notório, pois é distribuído entre os óculos de RA (transportado pelo utilizador na cara) e um processador móvel (transportado pelo utilizador numa mochila); a capacidade de armazenamento de energia (bateria). Devido ao consumo de energia do dispositivo móvel ser relevante, ainda se têm de

alimentar os restantes dispositivos, aumentando assim a necessidade de ter acesso a uma capacidade energética elevada.

Na fase de testes e resultados, foram realizadas diferentes etapas, numa primeira etapa, foram testados em contexto laboratorial, alguns equipamentos e SDK em rotinas/percursos idênticos aos armazéns industriais de logística, de forma a validar os algoritmos de determinação da pose de uma forma qualitativa, orientando assim os desenvolvimentos seguintes. A segunda etapa de testes foi focada no teste no algoritmo de localização interior com base em marcadores, onde foram criados dois ambientes de testes, em simulação, obtendo aqui erros máximos de 0.015m na posição e 0.08 rad (5 graus) na orientação. Em ambiente real, mas ainda em laboratório, o erro máximo aumentou, sendo que em termos de posição este foi de 0.08m e de orientação de 0.4 rad (22°). Na terceira a etapa de testes, onde foi realizado um percurso em ambiente real (ainda em laboratório) em conjunto com os óculos de RA, estando estes a orientar o utilizador num percurso definido, aqui pode-se afirmar que apesar de não existir um *ground-truth*, o erro máximo de posição não supera o valor admitido neste projeto, já relativamente à orientação esta foi analisada com base na seta projetado nos óculos, tendo estar a transmitir informação ao utilizador da orientação correta do percurso, estas contêm umas ligeiras oscilações. Na quarta etapa, onde o algoritmo foi testado em ambiente real (no armazém de logística da Matceramica), onde os erros não foram superiores aos da etapa anterior, existindo ligeiras dificuldades na deteção dos marcadores, comparativamente com as etapas anteriores, conseqüentemente a pose tem um erro médio aproximado de 0.5m e uma orientação com oscilações elevadas.

Face aos resultados, verifica-se a necessidade de trabalho de aperfeiçoamento futuro, sendo necessário reduzir o erro de posição, podendo recorrer a métodos de localização interior como, por exemplo, VIO ou SLAM. Além disso, apesar da orientação ter um erro associado, o seu principal problema é ao nível da oscilação, problema este que afeta bastante o utilizador dos óculos de RA pois, para além de o deixar confuso, pode perturbar a sua visão. Além disso, a melhoria da performance do algoritmo é algo a realizar, para reduzir o erro ainda existente e aumentar a frequência de atualização da estimativa da pose. Atualmente a estimativa da pose é realizada com uma média de 21 vezes por segundo, o que, apesar de ser suficiente para a projeção nos óculos de RA, futuramente ainda se pode migrar para outro *hardware* como, por exemplo, um *smartphone*, melhorando também assim a portabilidade de todo o sistema.

BIBLIOGRAFIA

- Abawi, D.F., J. Bienwald e R. Dorner (2004). “Accuracy in optical tracking with fiducial markers: an accuracy function for ARToolKit”. Em: *Third IEEE and ACM International Symposium on Mixed and Augmented Reality*, pp. 260–261. DOI: [10.1109/ISMAR.2004.8](https://doi.org/10.1109/ISMAR.2004.8).
- Aqel, Mohammad O.A. et al. (2016). “Review of visual odometry: types, approaches, challenges, and applications”. Em: *SpringerPlus* 5.1. ISSN: 21931801. DOI: [10.1186/s40064-016-3573-7](https://doi.org/10.1186/s40064-016-3573-7).
- Azuma, Ronald T. (ago. de 1997). “A Survey of Augmented Reality”. Em: *Presence: Teleoperators and Virtual Environments* 6.4, pp. 355–385. DOI: [10.1162/pres.1997.6.4.355](https://doi.org/10.1162/pres.1997.6.4.355). eprint: <https://direct.mit.edu/pvar/article-pdf/6/4/355/1623026/pres.1997.6.4.355.pdf>.
- Bolte, Nils-Ole e Daniel Christopher Goll (2020). “Potential analysis of track-and-trace systems in the outbound logistics of a Swedish retailer”. Tese de mestrado. Jönköping University, Jönköping International Business School, p. 88.
- Bradski, G. (2000). “The OpenCV Library”. Em: *Dr. Dobb’s Journal of Software Tools*.
- Breunig, Markus M et al. (2000). “LOF: identifying density-based local outliers”. Em: *Proceedings of the 2000 ACM SIGMOD international conference on Management of data*, pp. 93–104.
- Campbell, Mike et al. (2017). “The State of Industrial Augmented Reality 2017”. Em: *Ptc*, p. 10.
- Carlos, Manuel e Barbosa Figueiredo (2016). “Daniela Andreia Costa Miranda Otimização do sistema de gestão de armazenamento e inventário numa empresa de retalho”. Em.
- Caudell, T.P. e D.W. Mizell (2003). “Augmented reality: an application of heads-up display technology to manual manufacturing processes”. Em: 659–669 vol.2. DOI: [10.1109/hicss.1992.183317](https://doi.org/10.1109/hicss.1992.183317).
- Chen, Pan, Jianga Shang e Fuqiang Gu (2019). “Learning RSSI feature via ranking model for Wi-Fi fingerprinting localization”. Em: *IEEE Transactions on Vehicular Technology* 69.2, pp. 1695–1705.

- Cirulis, Arnis e Egils Ginters (2013). “Augmented reality in logistics”. Em: *Procedia Computer Science* 26, pp. 14–20. ISSN: 18770509. DOI: [10.1016/j.procs.2013.12.003](https://doi.org/10.1016/j.procs.2013.12.003).
- Cohen, L. et al. (2018). “Augmented and Virtual Reality in Operations”. Em: *Capgemini Research Institute*, pp. 1–32. URL: <https://www.capgemini.com/research/augmented-and-virtual-reality-in-operations/>.
- Comuniello, Antonella et al. (2019). “Ultrasound Time of Flight based positioning using the Bluetooth Low Energy protocol”. Em: *2019 IEEE International Symposium on Measurements & Networking (M&N)*, pp. 1–6. DOI: [10.1109/IWMN.2019.8805003](https://doi.org/10.1109/IWMN.2019.8805003).
- Deliottee (2015). “The warehouse of tomorrow, today Ubiquitous Computing, Internet of Things, Machine Learning”. Em.
- DeTone, Daniel, Tomasz Malisiewicz e Andrew Rabinovich (2018). “Self-improving visual odometry”. Em: *arXiv preprint arXiv:1812.03245*.
- Diete, Alexander et al. (set. de 2016). “Exploring a multi-sensor picking process in the future warehouse”. Em: pp. 1755–1758. DOI: [10.1145/2968219.2968270](https://doi.org/10.1145/2968219.2968270).
- Duda, Alexander e Udo Frese (2018). “Accurate Detection and Localization of Checkerboard Corners for Calibration.” Em: *BMVC*, p. 126.
- Exploration, Surface (2022). “Bionic Design of Multi-Toe Quadruped Robot for Planetary Surface Exploration”. Em.
- Fiala, M. (2005). “Comparing ARTag and ARToolkit Plus fiducial marker systems”. Em: *IEEE International Workshop on Haptic Audio Visual Environments and their Applications*, 6 pp.-. DOI: [10.1109/HAVE.2005.1545669](https://doi.org/10.1109/HAVE.2005.1545669).
- Fraga-Lamas, Paula et al. (fev. de 2018). “A Review on Industrial Augmented Reality Systems for the Industry 4.0 Shipyard”. Em: *IEEE Access* 6, pp. 13358–13375. DOI: [10.1109/ACCESS.2018.2808326](https://doi.org/10.1109/ACCESS.2018.2808326).
- François, Rémy (1989). “Engineering training”. Em: *Education + Training* 31.1, p. 28. ISSN: 00400912. DOI: [10.1108/eb055187](https://doi.org/10.1108/eb055187).
- Fuentes-Pacheco, Jorge, Jose Ascencio e J. Rendon-Mancha (nov. de 2015). “Visual Simultaneous Localization and Mapping: A Survey”. Em: *Artificial Intelligence Review* 43. DOI: [10.1007/s10462-012-9365-8](https://doi.org/10.1007/s10462-012-9365-8).
- Garrido-Jurado, S. et al. (2014). “Automatic generation and detection of highly reliable fiducial markers under occlusion”. Em: *Pattern Recognition* 47.6, pp. 2280–2292. ISSN: 0031-3203. DOI: <https://doi.org/10.1016/j.patcog.2014.01.005>.
- Guerreiro, Rui Filipe Cardoso, Pedro Manuel Quintas Aguiar e Universidade Técnica de Lisboa (2004). “Modelos 3D densos a partir de imagens com sobreposição

- parcial: factorização com dados desconhecidos”. Tese de doutoramento. Instituto Superior Técnico.
- Horaud, Radu e Fadi Dornaika (jun. de 1995). “Hand-Eye Calibration”. Em: *I. J. Robotic Res.* 14, pp. 195–210. DOI: [10.1177/027836499501400301](https://doi.org/10.1177/027836499501400301).
- Jacobs, Trent (jul. de 2018). “AR Headsets Give Oil and Gas Sector the Quicker Fix”. Em: *Journal of Petroleum Technology* 70.07, pp. 32–34. ISSN: 0149-2136. DOI: [10.2118/0718-0032-JPT](https://doi.org/10.2118/0718-0032-JPT). eprint: <https://onepetro.org/JPT/article-pdf/70/07/32/2211773/spe-0718-0032-jpt.pdf>.
- Janin, A.L., D.W. Mizell e T.P. Caudell (1993). “Calibration of head-mounted displays for augmented reality applications”. Em: *Proceedings of IEEE Virtual Reality Annual International Symposium*, pp. 246–255. DOI: [10.1109/VRAIS.1993.380772](https://doi.org/10.1109/VRAIS.1993.380772).
- Jha, Gauri, Pawan Singh e Lavanya Sharma (2019). “Recent advancements of augmented reality in real time applications”. Em: *International Journal of Recent Technology and Engineering* 8.2 Special Issue 7, pp. 537–542. ISSN: 22773878. DOI: [10.35940/ijrte.B10100.0782S719](https://doi.org/10.35940/ijrte.B10100.0782S719).
- Kim, Pyojin et al. (2022). “A benchmark comparison of four off-the-shelf proprietary visual-inertial odometry systems”. Em: *Sensors* 22.24, p. 9873.
- Kohn, Vanessa e David Harborth (2018). “Augmented reality - A game changing technology for manufacturing processes?” Em: *26th European Conference on Information Systems: Beyond Digitization - Facets of Socio-Technical Change, ECIS 2018* September.
- Ledwich, Luke e Stefan Williams (2004). “Reduced SIFT features for image retrieval and indoor localisation”. Em: *Australian conference on robotics and automation*. Vol. 322. CiteSeer Australia, p. 3.
- Li, Fan et al. (set. de 2012). “A Reliable and accurate indoor localization method using phone inertial sensors”. Em: pp. 421–430. DOI: [10.1145/2370216.2370280](https://doi.org/10.1145/2370216.2370280).
- Li, Peiliang et al. (out. de 2017). “Monocular Visual-Inertial State Estimation for Mobile Augmented Reality”. Em: pp. 11–21. DOI: [10.1109/ISMAR.2017.18](https://doi.org/10.1109/ISMAR.2017.18).
- Liu, Jian et al. (2010). “Route optimization of stacker in automated storage and retrieval system based on genetic algorithms”. Em: *Shenyang Jianzhu Daxue Xuebao (Ziran Kexue Ban)/Journal of Shenyang Jianzhu University (Natural Science)* 26.5, pp. 15–17. ISSN: 16712021.
- Machado, Jéssica Milene Segundo (2021). “Middleware para Armazéns Inteligentes”. Tese de doutoramento.
- Mama, Mouchili (2008). “Mathematical Modelling of The Global Positioning System Tracking Signals”. Em: May.

- Martins, Pedro Filipe et al. (2020). “Monocular Camera Calibration for Autonomous Driving — a comparative study”. Em: *2020 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, pp. 306–311. DOI: [10.1109/ICARSC49921.2020.9096104](https://doi.org/10.1109/ICARSC49921.2020.9096104).
- Mautz, Rainer (2012). “Indoor Positioning Technologies”. Em: *ETH Zurich, Department of Civil, Environmental and Geomatic Engineering, Institute of Geodesy and Photogrammetry*, p. 10.
- Mirabdollah, M Hossein e Bärbel Mertsching (2015). “Fast techniques for monocular visual odometry”. Em: *Pattern Recognition: 37th German Conference, GCPR 2015, Aachen, Germany, October 7-10, 2015, Proceedings 37*. Springer, pp. 297–307.
- Monteiro, Francisco e Santos Br (2017). “SmartLocator - Indoor Human Localization using a Smartphone”. Em.
- More, Even e Even More (2020). “Augmented Reality drives efficiency in logistics”. Em: pp. 1–2.
- Ozturkcan, Selcen (2021). “Service innovation: Using augmented reality in the IKEA Place app”. Em: *Journal of Information Technology Teaching Cases* 11.1, pp. 8–13. DOI: [10.1177/2043886920947110](https://doi.org/10.1177/2043886920947110). eprint: <https://doi.org/10.1177/2043886920947110>.
- Paavilainen, Janne et al. (2017). “The Pokémon go experience: A location-based augmented reality mobile game goes mainstream”. Em: *Conference on Human Factors in Computing Systems - Proceedings 2017-May*. Figure 1, pp. 2493–2498. DOI: [10.1145/3025453.3025871](https://doi.org/10.1145/3025453.3025871).
- Pradhan, Chittaranjan e Chandrashekhar N. Bhende (2019). “Online load frequency control in wind integrated power systems using modified Jaya optimization”. Em: *Engineering Applications of Artificial Intelligence* 77, pp. 212–228. ISSN: 0952-1976. DOI: <https://doi.org/10.1016/j.engappai.2018.10.003>.
- Reljić, Vule et al. (2021). “Augmented Reality Applications in Industry 4.0 Environment”. Em: *Applied Sciences* 11.12. ISSN: 2076-3417. DOI: [10.3390/app11125592](https://doi.org/10.3390/app11125592).
- Renaudin, Valérie, Aigul KHAMITOVA e Miguel Ortiz (set. de 2018). *IPIN 2018 - Ninth International Conference On Indoor Positioning And Indoor Navigation*. ISBN: 978-2-85782-748-1.
- Riedlinger, Urs e Leif Oppermann (2018). “On the Accuracy of Google Tango in Comparison to ARCore”. Em.
- Sandra, CJ et al. (2021). “INDOOR NAVIGATION SYSTEM USING AUGMENTED REALITY”. Em.

- Santana, Andre et al. (jan. de 2008). “Localization of a Mobile Robot based in Odometry and Natural Landmarks using Extended Kalman Filter.” Em: pp. 187–193. DOI: [10.13140/2.1.2081.8882](https://doi.org/10.13140/2.1.2081.8882).
- Sarupuri, Bhuvaneshwari (2016). “Enhancing depth cues with AR visualization for forklift operation assistance in warehouse.” Em: URL: <http://ir.canterbury.ac.nz:80/handle/10092/12038>.
- Sauer, F. et al. (2001). “Augmented reality visualization of ultrasound images: system description, calibration, and features”. Em: *Proceedings IEEE and ACM International Symposium on Augmented Reality*, pp. 30–39. DOI: [10.1109/ISAR.2001.970513](https://doi.org/10.1109/ISAR.2001.970513).
- Sivakumar, S. e Raman Venkatesan (jul. de 2015). “Meta-heuristic approaches for minimizing error in localization of wireless sensor networks”. Em: *Applied Soft Computing* 36. DOI: [10.1016/j.asoc.2015.05.053](https://doi.org/10.1016/j.asoc.2015.05.053).
- Skurowski, Przemysław et al. (2022). “Performance of QR Code Detectors near Nyquist Limits”. Em: *Sensors* 22.19. ISSN: 1424-8220. DOI: [10.3390/s22197230](https://doi.org/10.3390/s22197230). URL: <https://www.mdpi.com/1424-8220/22/19/7230>.
- Stoltz, Marie H el ene et al. (2017). “Augmented Reality in Warehouse Operations: Opportunities and Barriers”. Em: *IFAC-PapersOnLine* 50.1, pp. 12979–12984. ISSN: 24058963. DOI: [10.1016/j.ifacol.2017.08.1807](https://doi.org/10.1016/j.ifacol.2017.08.1807).
- Study, Case (2021). “Augmented Reality With AR, Mitsubishi Electric enhances the support experience and improves repair processes.” Em.
- Sutherland, Ivan E. (1968). “A head-mounted three dimensional display”. Em: *Proceedings of the December 9-11, 1968, fall joint computer conference, part I on - AFIPS '68 (Fall, part I)*. ACM Press. DOI: [10.1145/1476589.1476686](https://doi.org/10.1145/1476589.1476686).
- Tejesh, B. Sai Subrahmanya e S. Neeraja (2018). “Warehouse inventory management system using IoT and open source framework”. Em: *Alexandria Engineering Journal* 57.4, pp. 3817–3823. ISSN: 1110-0168. DOI: <https://doi.org/10.1016/j.aej.2018.02.003>.
- T ordal, Sondre Sanden e Geir Hovland (2017). “Relative vessel motion tracking using sensor fusion, ArUco markers, and MRU sensors”. Em.
- Torres-Sospedra, Joaqu n et al. (2018). “Off-Line Evaluation of Mobile-Centric Indoor Positioning Systems: The Experiences from the 2017 IPIN Competition”. Em: *Sensors* 18.2. ISSN: 1424-8220. DOI: [10.3390/s18020487](https://doi.org/10.3390/s18020487).
- Tsai, Shu-pei (mar. de 2019). “Augmented reality enhancing place satisfaction for heritage tourism marketing”. Em: *Current Issues in Tourism* 23, pp. 1–6. DOI: [10.1080/13683500.2019.1598950](https://doi.org/10.1080/13683500.2019.1598950).

- Tully, Jeffrey et al. (2015). “Recording medical students’ encounters with standardized patients using Google Glass: providing end-of-life clinical education”. Em: *Academic medicine* 90.3, pp. 314–316.
- VAISALA (2020). “GMP Warehouse Mapping Step-by-Step Guidelines for Validating Life Science Storage Facilities”. Em: p. 9. URL: https://www.vaisala.com/sites/default/files/documents/VIM-GLO-CMS-GMP-Warehouse-Mapping-White-Paper-B211170EN-B_1.pdf.
- Walko, Christian (2018). “Integration of augmented-reality-glasses into a helicopter simulator with front projection”. Em: *Deutscher Luft-und Raumfahrtkongress (DLRK)*.
- Weeks Joanna K e Pakpoor, Jina e Park (2021). “Aproveitando realidade aumentada e tomografia computadorizada para ensinar anatomia de cabeça e pescoço para estudantes de medicina do primeiro ano”. Em: *radiologia acadêmica* 28. Ed. por Elsevier.
- Williams, Mark Kenny (2019). “Augmented Reality Supported Batch Picking System”. Em: p. 65.
- Xu, Xiang-Bin e Xin-Jian Zhou (2010). “The Picking System Optimization Based on Genetic Algorithm and Implemented by eM-Plant”. Em: *2010 Second International Conference on Computer Modeling and Simulation*. Vol. 4, pp. 143–147. DOI: [10.1109/ICCMS.2010.169](https://doi.org/10.1109/ICCMS.2010.169).
- Yang, Zhice et al. (2015). “Wearables Can Afford”. Em: pp. 317–330. DOI: [10.1145/2742647.2742648](https://doi.org/10.1145/2742647.2742648).

ANEXOS

ANÁLISE COMPARATIVA DE ÓCULOS DE REALIDADE AUMENTADA

Tabela 5: Análise comparativa de modelos de óculos de realidade aumentada (-"significa que não foi possível apurar).

Empresa	Modelo	Preço [€]	Peso [g]	<i>Tracking</i>	Software anfitrião
Glassup	F4	NA	65	Não	Linux
SONY	Smart Eyeglass Developer Edition	699 ¹	77	Não	Android 4.4
ODG	R-7	2750,00 ²	181,43	-	Reticle OS on Android 4.4.4

Continua na próxima página

1 <https://www.aniwaa.com/product/vr-ar/sony-smarteyeglass/>

2 <https://www.osterhoutgroup.com/r-7-smartglasses>

Tabela 5 – continuação da página anterior

Empresa	Modelo	Preço [€]	Peso [g]	Tracking	Software anfitrião
ODG	R-9	1800,00 ³	181,43	Ultra wide-angle fisheye camera for enhanced environmental tracking and positioning	Reticle OS on Android™ Nougat™
ODG	R-8	1000,00 ⁴	141,74	6 DoF Tracking	Reticle OS on Android™ Nougat™
DAQRI	Smart Glasses (Ainda não disponível)	4995	335	Unidade de processamento de visão dedicada para rastreamento de 6 DoF	-
Microsoft HoloLens	Commercial Suite	5000	579	-	-
Microsoft HoloLens	Development Edition	3000	579	Não	Windows 10
Magic Leap	One	2295,00	345	Head tracking, 6 DoF	-
<i>Continua na próxima página</i>					

³ <https://www.aniwaa.com/product/vr-ar/odg-r-9/>

⁴ <https://www.aniwaa.com/product/vr-ar/odg-r-8/>

Tabela 5 – continuação da página anterior

Empresa	Modelo	Preço [€]	Peso [g]	Tracking	Software anfitrião
Epson	Moverio BT-2000	2029,52	290 (óculos) + 265 (controlador)	3 DoF	Android 4.04
Epson	Moverio BT-2200	2184,32	270 (óculos) + 265 (controlador)	3 DoF	Android 4.05
Epson	Moverio BT-350	922,50	119 (óculos) + 129 (controlador)	Não	Android 5.1
Epson	Moverio BT-35E	799,50	119 (óculos) + 45 (controlador)	Não	Android 4.04
Realwear	HMT-1	2000,00	380	Deteta marcos bluetooth	Android 6.0
Meta	Meta 2	1758,00	500	Head tracking	Windows 10
Vuzix	M300	1699,99	140	Head tracking	Android 6 e iOS
Vuzix	Blade	1099,00	-	Head tracking	Android e iOS
<i>Continua na próxima página</i>					

Tabela 5 – continuação da página anterior

Empresa	Modelo	Preço [€]	Peso [g]	Tracking	Software anfitrião
Vuzix	M100	1079,99	371	Não	Android ICS 4.04 API15
EverySight	Raptor	799,00	98	-	Android
Recon	Jet Pro	700,00	-	Sim	-
Epson	Moverio BT-200	699,00	87,9	Não	Android 4.04
Optinvent	ORA-2	699,00	90	-	Android AOSP 4.4.2
Epson	Moverio BT-300	685,80	69 (óculos) + 129 (controlador)	Não	Android 5.1
Laforge	Beta	590,00	-	-	-
Google Glass	-	-	36	-	-
Meta	Meta 1	-	-	-	-
Realwear	hmt-1z1	-	-	Deteta marcos bluetooth	Android 6.1
ODG	R-7HL	?	226,79	Não	Reticle OS on Android 4.4.4
DreamWorld	DreamGlass	619,00	240	Hand gesture recognition	Android e iOS

RESULTADOS DOS TESTES COM MARCADORES EM AMBIENTE SIMULAÇÃO (TRAJETÓRIAS)

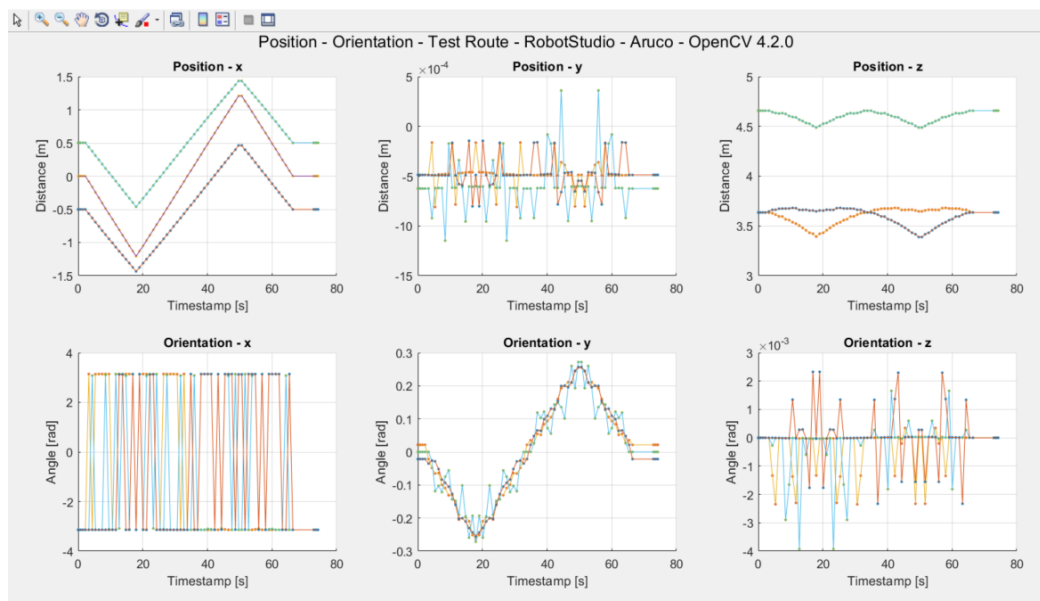


Figura 50: Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).

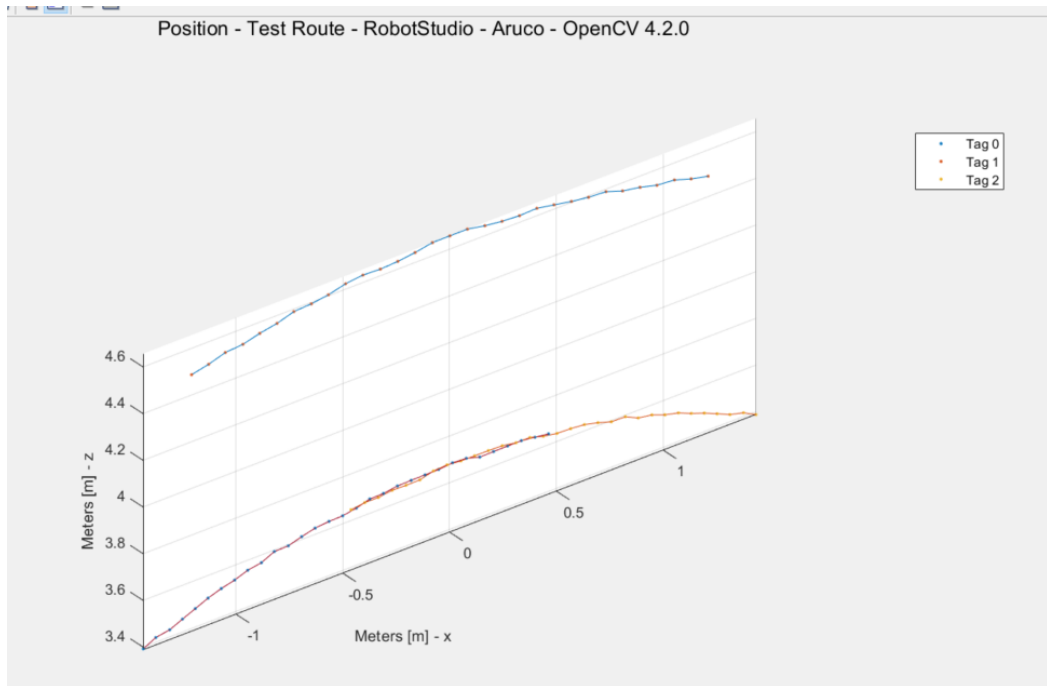


Figura 51: Gráfico de posição em 3D dos marcadores relativamente à câmara.

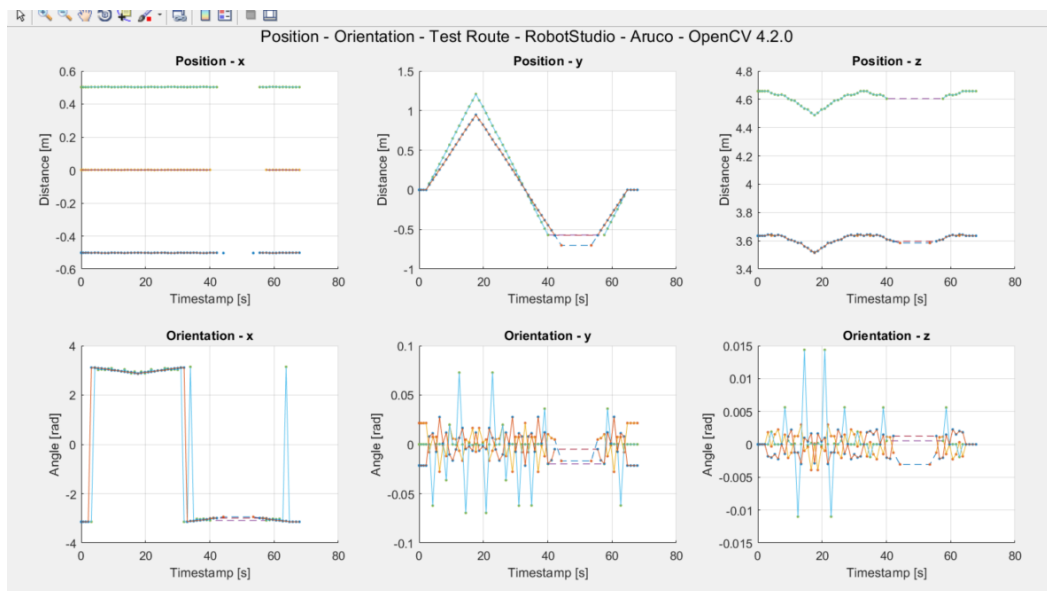


Figura 52: Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).

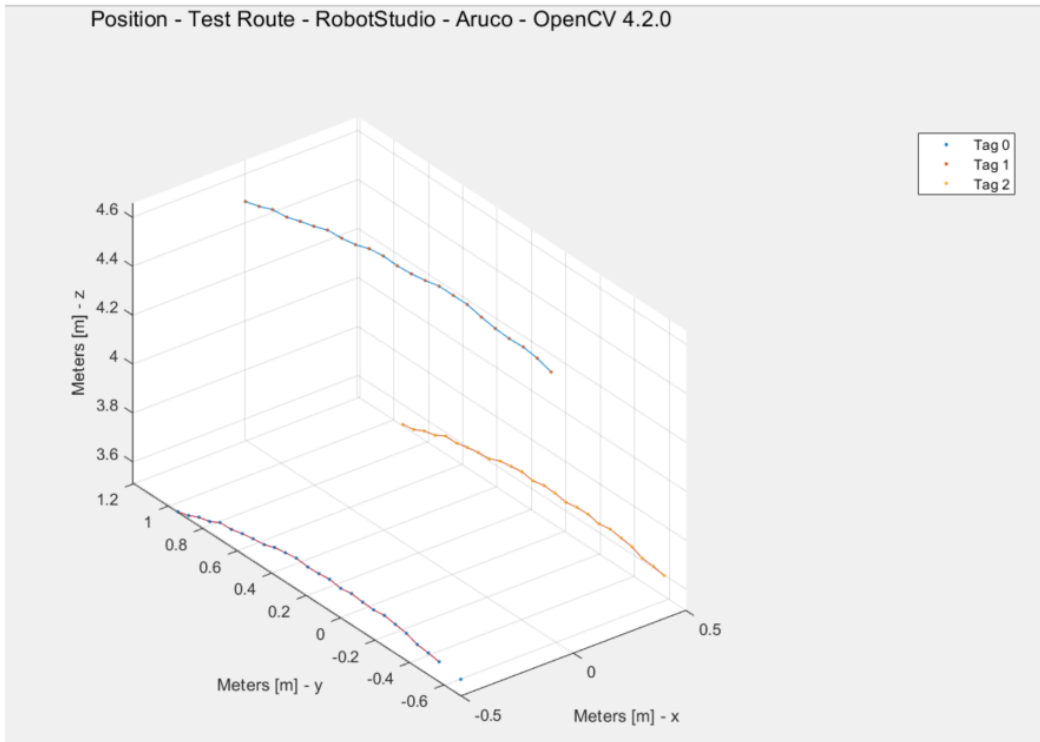


Figura 53: Gráfico de posição em 3D dos marcadores relativamente à câmara.

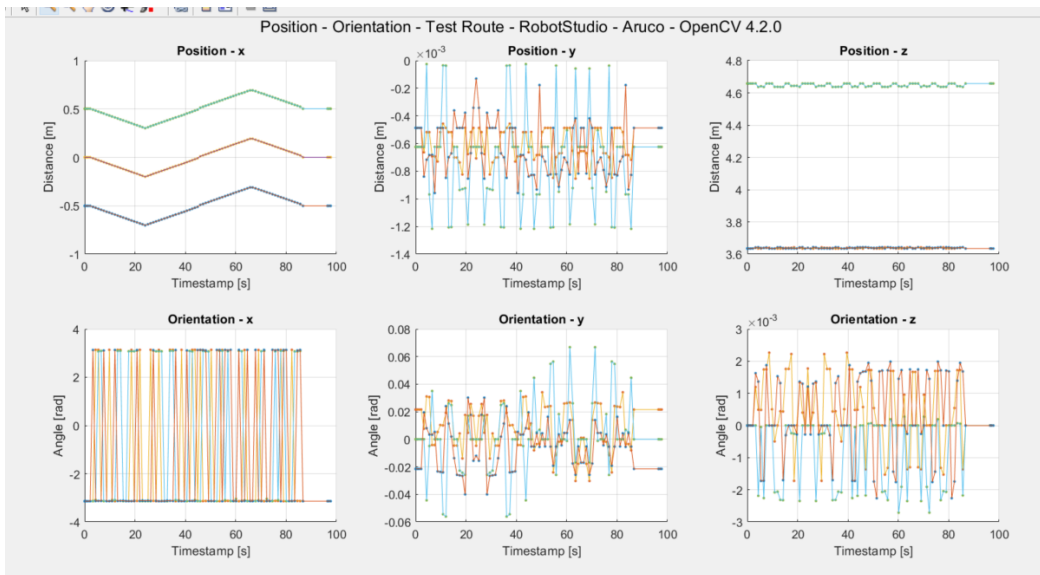


Figura 54: Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).

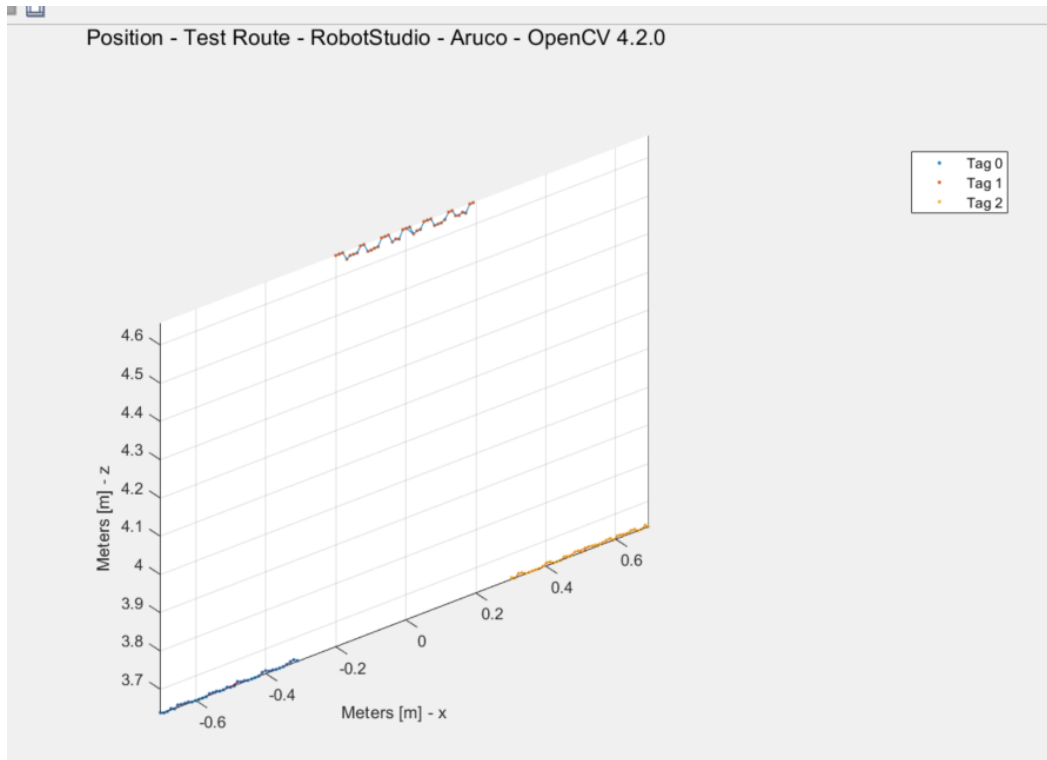


Figura 55: Gráfico de posição em 3D dos marcadores relativamente à câmara.

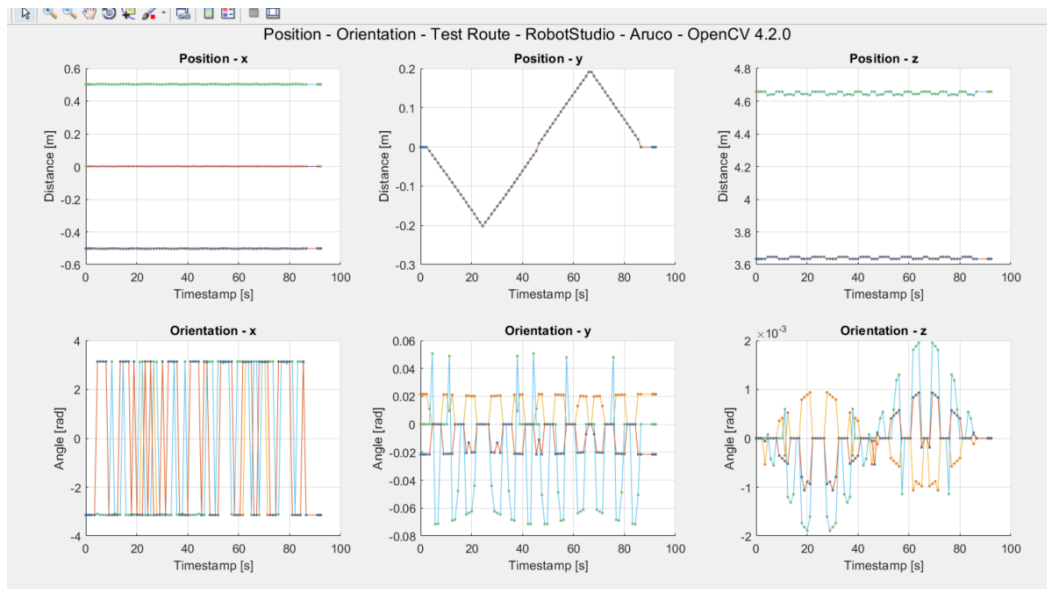


Figura 56: Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).

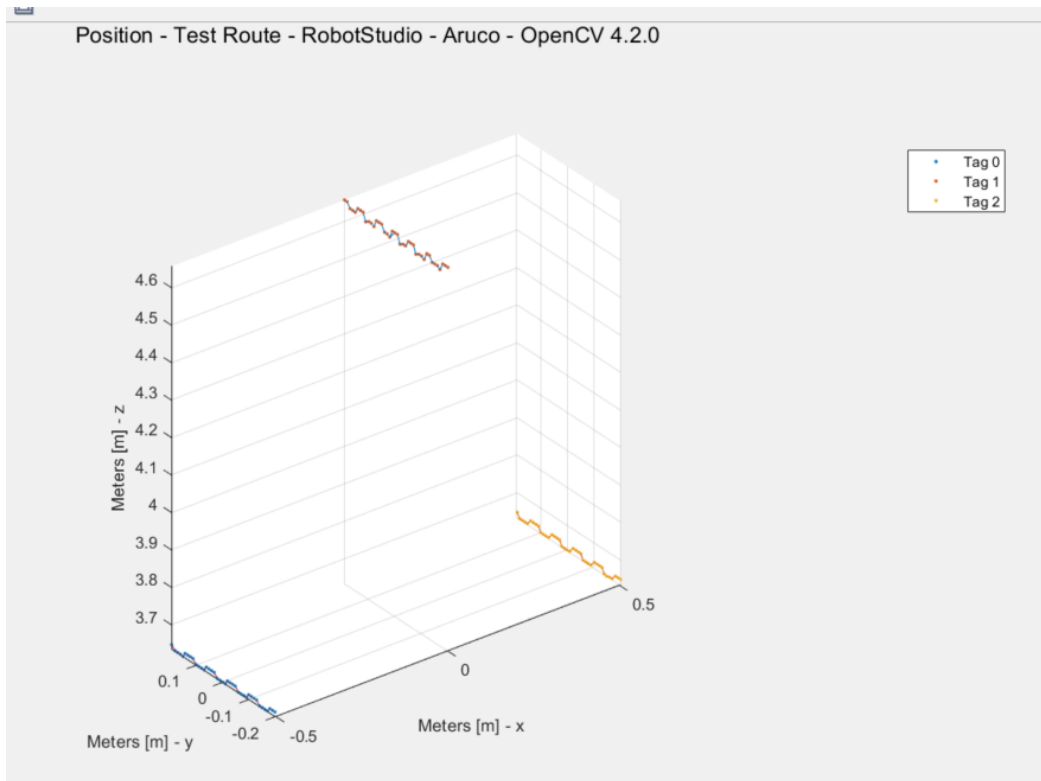


Figura 57: Gráfico de posição em 3D dos marcadores relativamente à câmara.

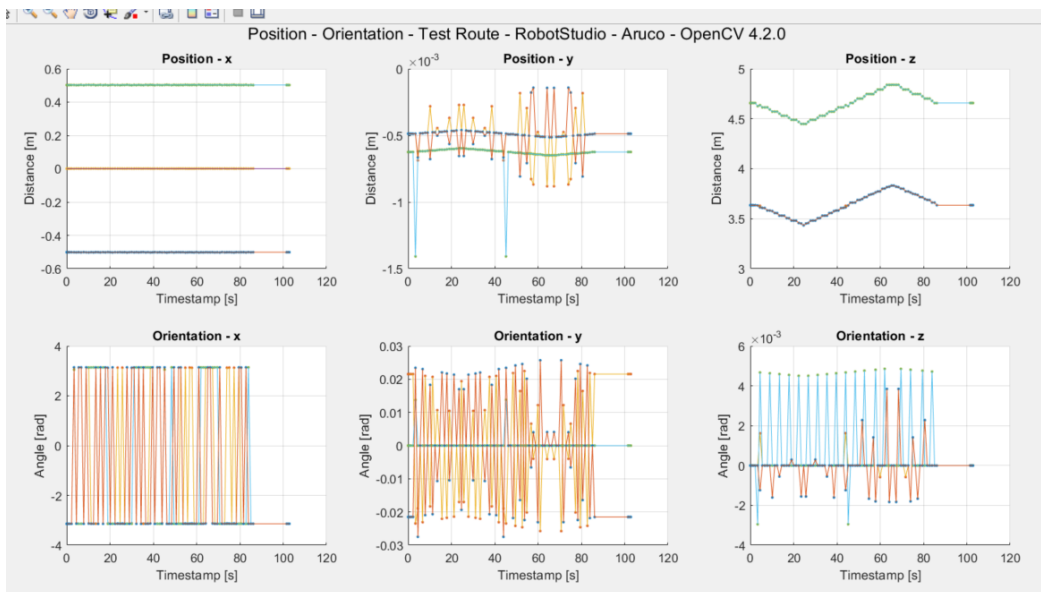


Figura 58: Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).

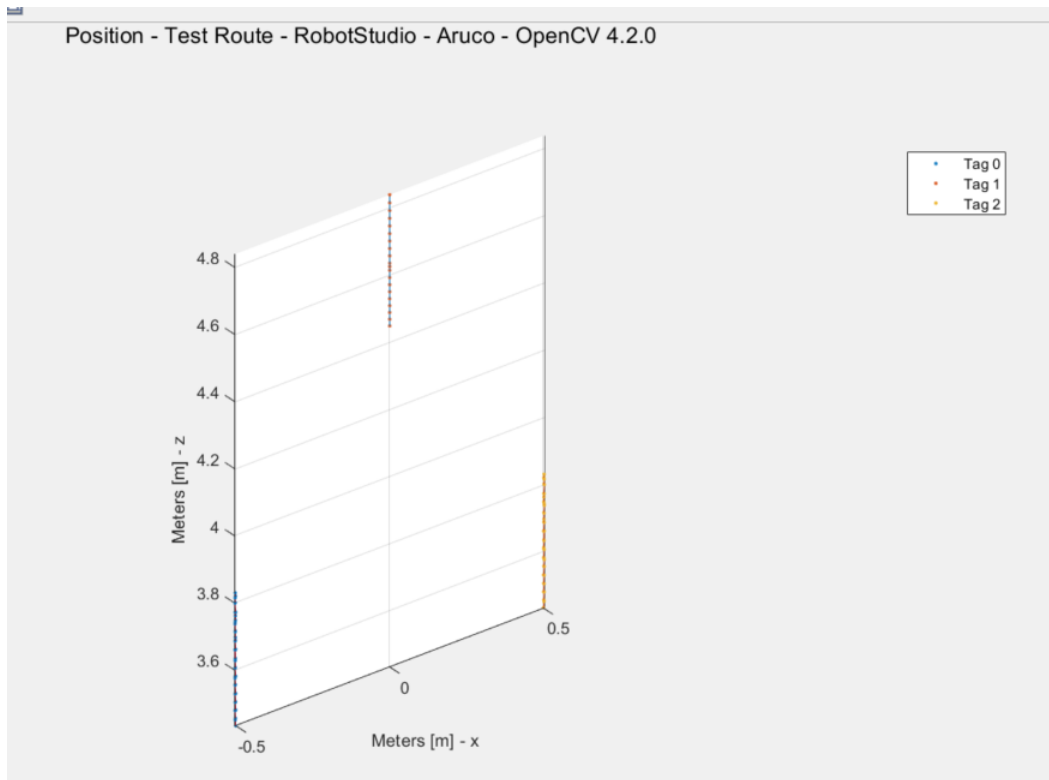


Figura 59: Gráfico de posição em 3D dos marcadores relativamente à câmara.

TESTES DE MARCADORES EM AMBIENTE REAL (TRAJETÓRIAS)

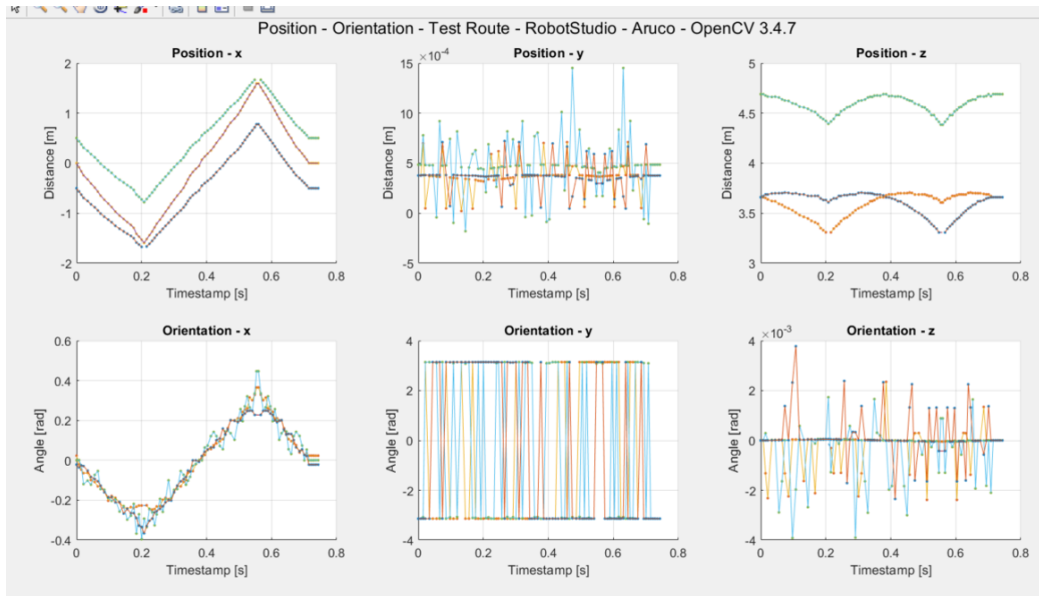


Figura 60: Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).

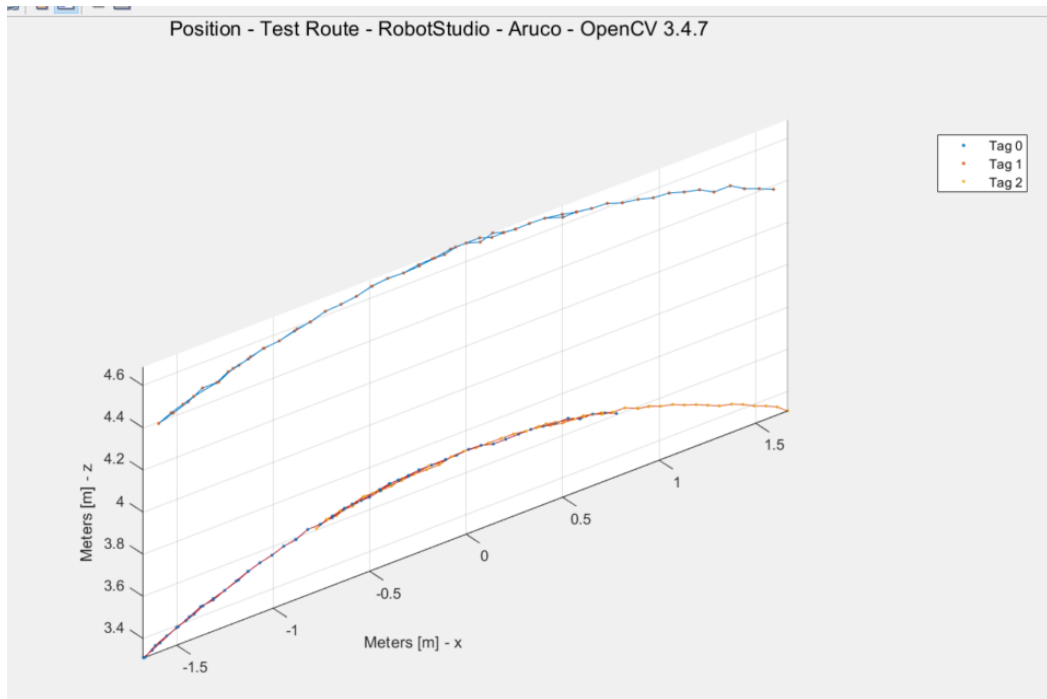


Figura 61: Gráfico de posição em 3D dos marcadores relativamente à câmara.

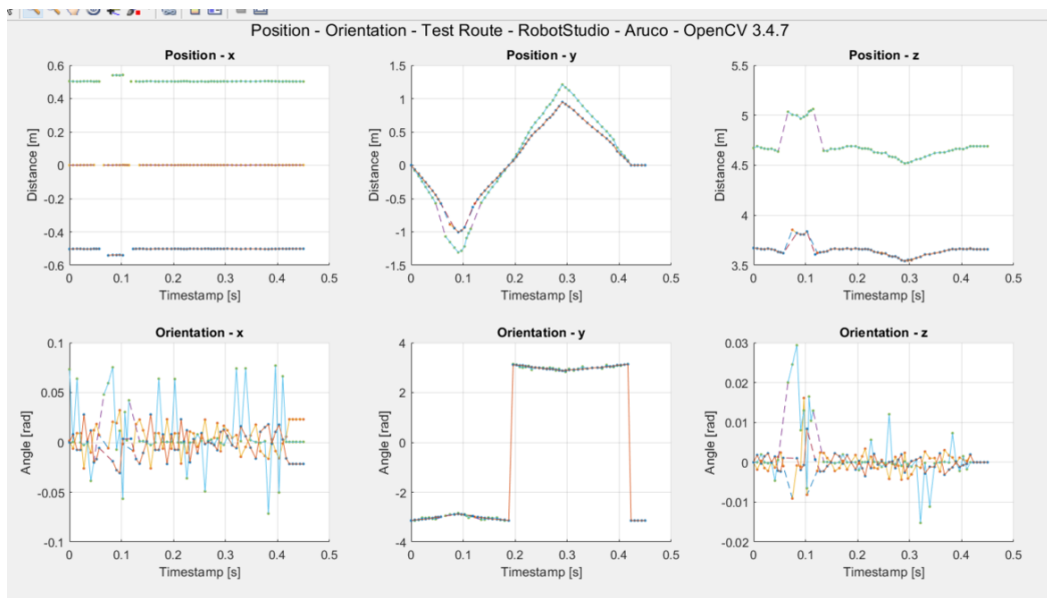


Figura 62: Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).

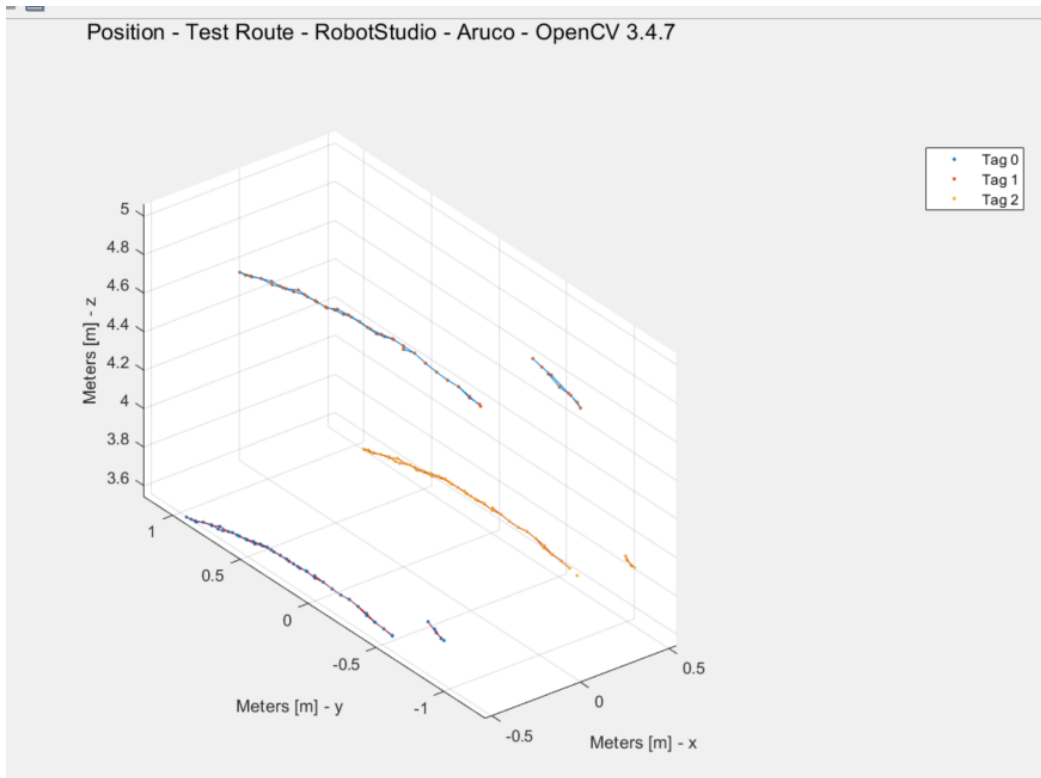


Figura 63: Gráfico de posição em 3D dos marcadores relativamente à câmara.

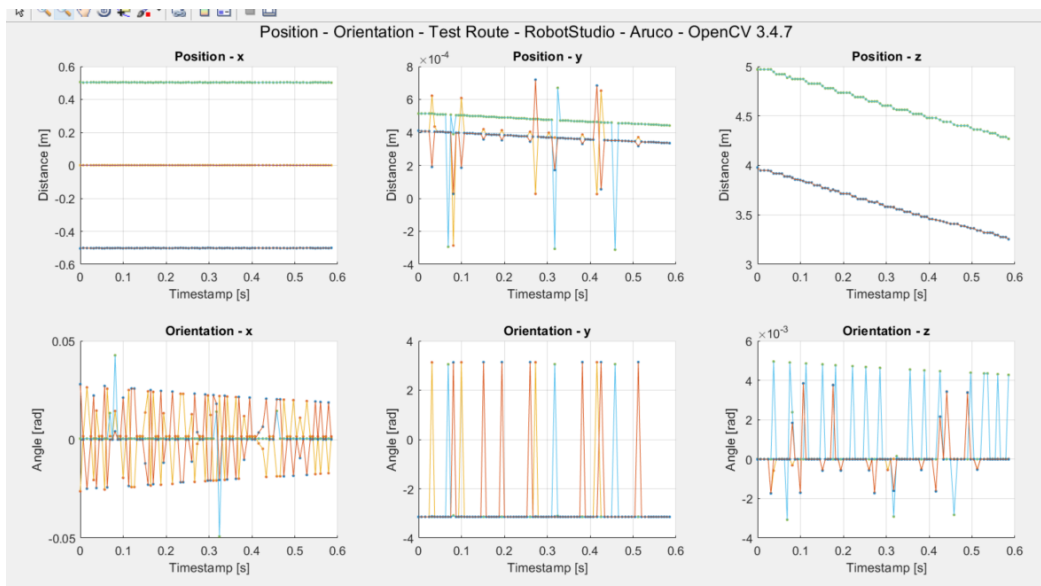


Figura 64: Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).

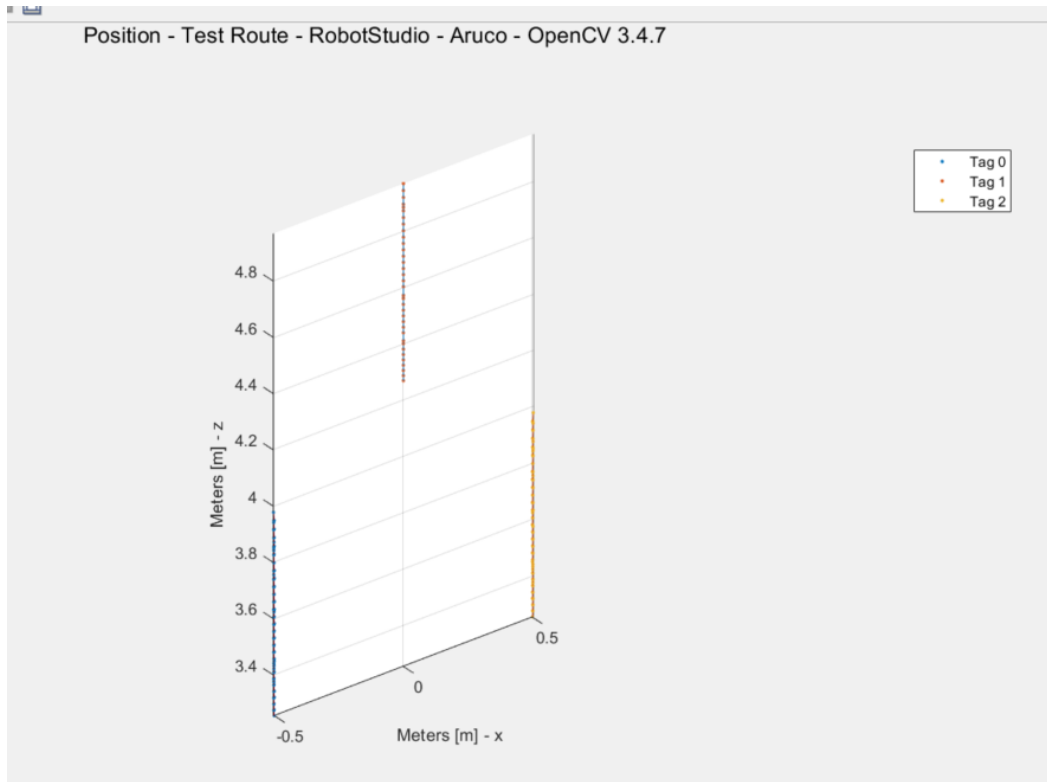


Figura 65: Gráfico de posição em 3D dos marcadores relativamente à câmara.

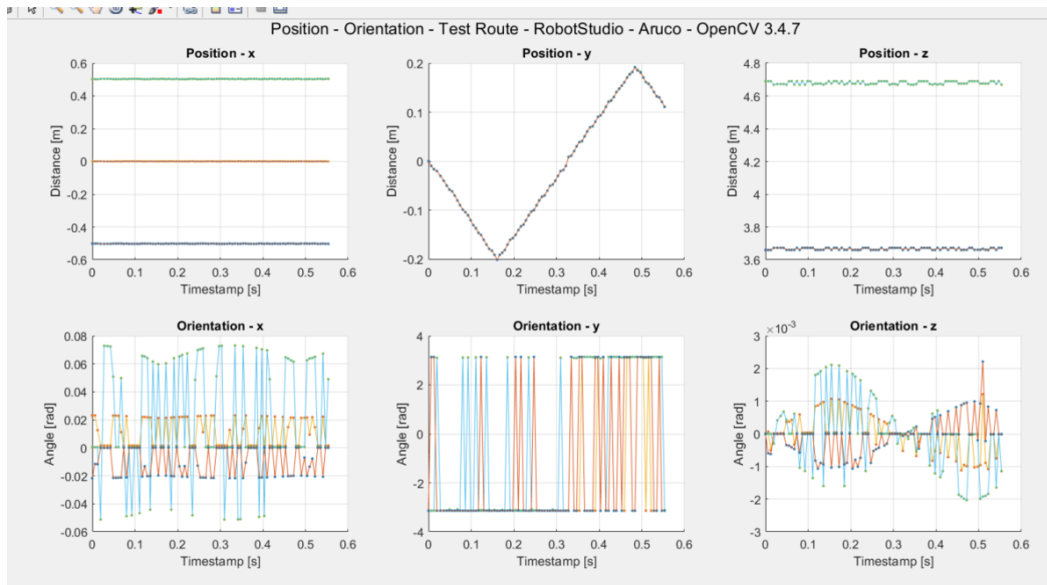


Figura 66: Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).

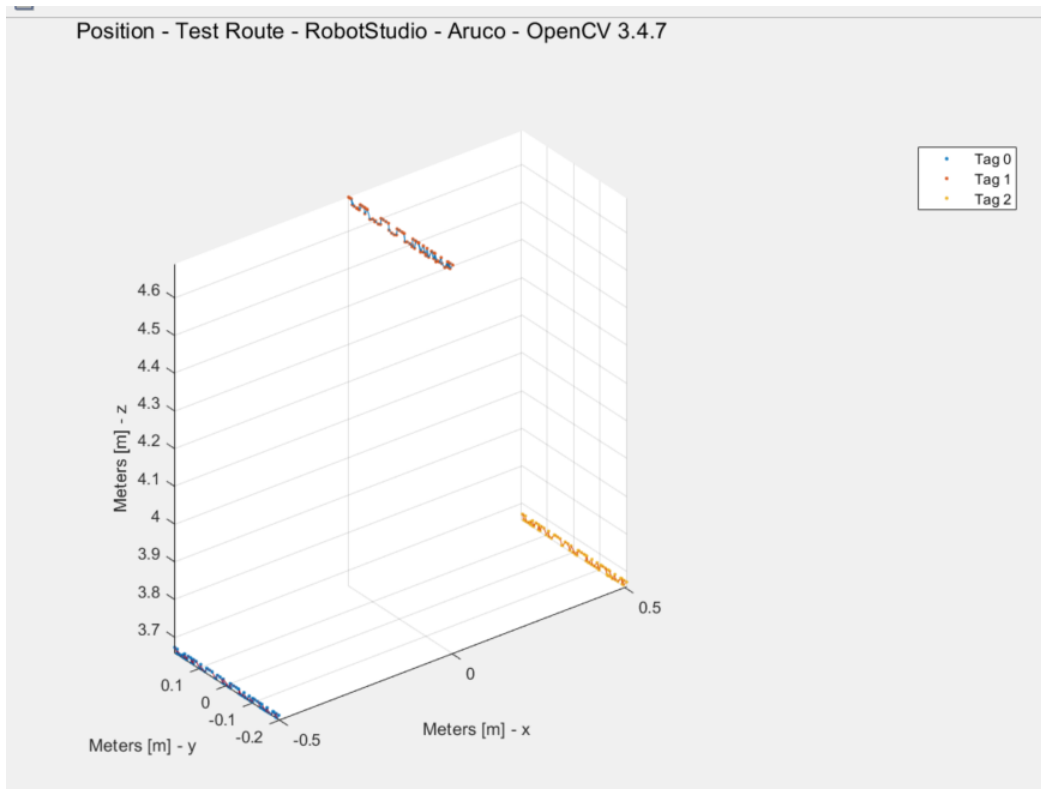


Figura 67: Gráfico de posição em 3D dos marcadores relativamente à câmara.

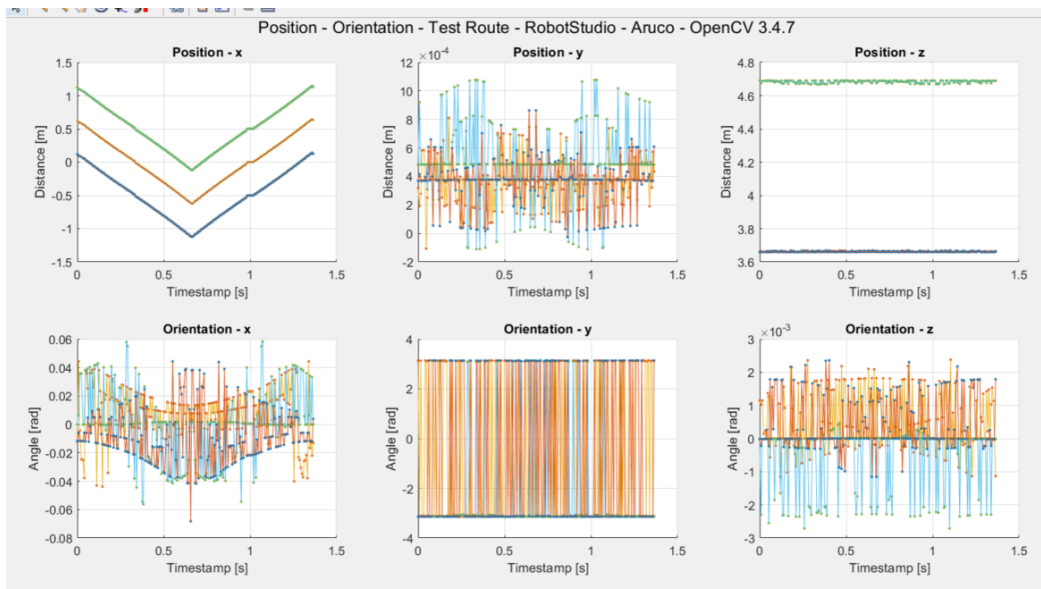


Figura 68: Gráfico de posição e orientação por eixo (A linha azul representa a TAG0, a linha laranja representa a TAG1 e a linha amarela representa a TAG2).

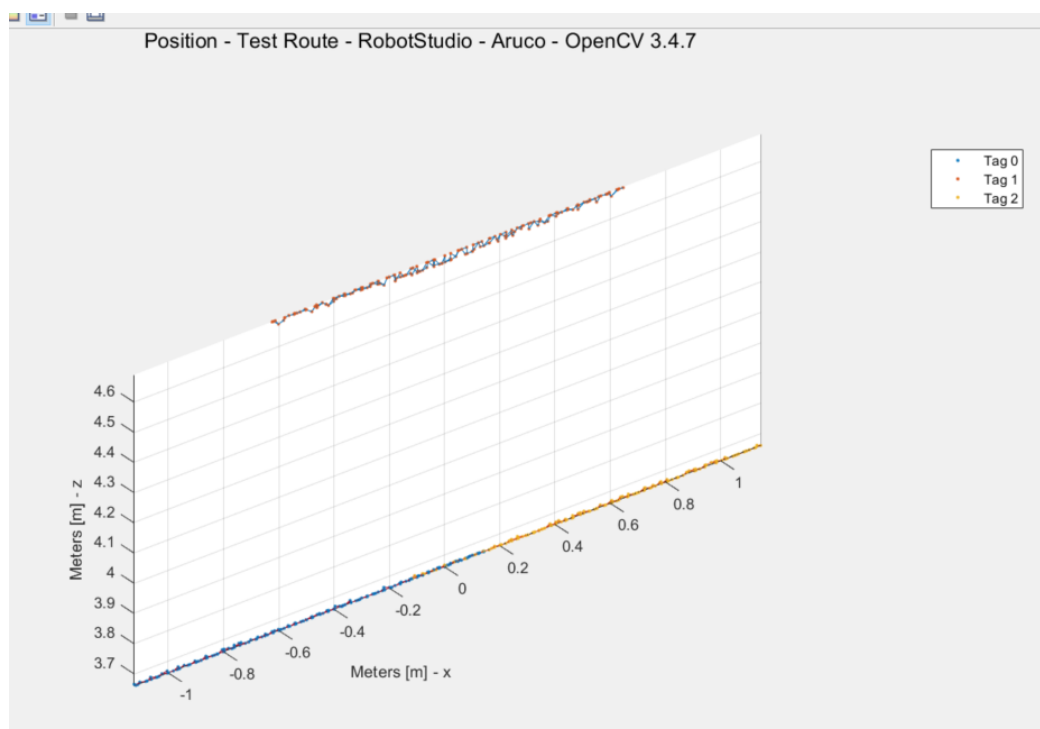


Figura 69: Gráfico de posição em 3D dos marcadores relativamente à câmara.

DECLARAÇÃO

Declaro, sob compromisso de honra, que o trabalho apresentado neste projeto, com o título “*Sistema de Localização Interior para Operações Logísticas com Realidade Aumentada em Armazéns*”, é original e foi realizado por Jaime Miguel Conceição Silva (2182458) sob orientação de Professor Doutor Hugo Filipe Costelha de Castro (hugo.costelha@ipleiria.pt) e Professor Doutor Carlos Fernando Couceiro de Sousa Neves (carlos.neves@ipleiria.pt).

Leiria, Setembro de 2023

Jaime Miguel Conceição Silva