



# ESCOLA NAVAL



*talant de biefaire*

Matilde de Macedo Correia Vieira

**Integração de um sistema de apoio à aterragem de Veículos Aéreos Não Tripulados (UAVs) em navios**

Dissertação para obtenção do Grau de Mestre em Ciências Militares Navais, na especialidade de Engenharia Naval Ramo de Armas e Eletrónica



Alfeite

2023





# ESCOLA NAVAL

talant de bi-faire



Matilde de Macedo Correia Vieira

*Integração de um sistema de apoio à aterragem de  
Veículos Aéreos Não Tripulados (UAVs) em navios*

Dissertação para obtenção do Grau de Mestre em  
Ciências Militares Navais, na especialidade de Engenharia Naval Ramo  
de Armas e Eletrónica

**Orientação de:** Bruno Duarte Damas

Nuno Alexandre Antunes Martins Pessanha Santos

*A Aluna Mestranda,*

*O Orientador,*

---

Matilde Vieira



---

Bruno Duarte Damas

Alfeite

2023



“A ship in port is safe, but that’s not what ships are built for.”

Grace Hopper



# Agradecimentos

Antes de mais, gostaria de expressar a minha sincera gratidão a todas as pessoas que se cruzaram e me influenciaram de forma positiva ao longo de toda a minha jornada académica.

Em primeiro lugar, gostaria de agradecer aos meus dois orientadores, Engenheiro Bruno Damas e Engenheiro Nuno Pessanha, por todos os conhecimentos que me transmitiram, pelo apoio incondicional e compreensão que me proporcionaram ao longo deste ano letivo.

Quero agradecer à guarnição da Célula de Experimentação Operacional de Veículos Não Tripulados (CEOV) pela maneira como me acolheram, tanto ao longo do exercício *Guerriers de La Méditerranée*, como também nas suas instalações.

Gostaria de expressar a minha gratidão ao Pelotão Intelligence Surveillance and Reconnaissance (ISR) do Batalhão de Fuzileiros nº2 (BF2) do Destacamento de Apoio de Combate (DAC), pelo dinamismo e prontidão com que se disponibilizaram a planear e executar os voos a bordo da fragata NRP Corte-Real.

À guarnição do Centro de Investigação da Academia da Força Aérea (CI-AFA) gostaria de agradecer pela forma atenciosa e cativante, que se disponibilizaram a receber-me no vosso espaço.

À minha família e amigos, pelo apoio incondicional, tanto nos momentos bons como nos menos positivos, estou sinceramente grata e orgulhosa de vos ter sempre ao meu lado. Sem a vossa paciência e apoio inesgotável não estaria agora a terminar o meu percurso académico.

Por fim, dirijo os meus agradecimentos aos meus camaradas de curso "CALM Manuel Armando Ferraz", que no decorrer destes cinco anos foram um pilar na minha formação. E que nunca nos esqueçamos que: "Na eternidade do amanhã, nos conheçam por destemidos".



# Resumo

Os Veículos Aéreos Não Tripulados (VANTs) são utilizados nos dias de hoje em inúmeras aplicações, abrangendo vigilância operacional, reconhecimento em tempo real, atividades humanitárias, entre outras. Os veículos aéreos não tripulados podem ser operados manualmente por um piloto ou autonomamente, com trajetórias pré-definidas. A aterragem de veículo aéreo não tripulado em navios apresenta uma dificuldade acrescida devido à área limitada disponível a bordo. Adicionalmente, existem áreas navegáveis em que o sinal *Global Positioning System* (GPS) pode ter falhas intermitentes de sinal ou sofrer empastelamento, aumentando o desafio de aterragem dos veículos aéreos não tripulados. Consequentemente, torna-se fundamental auxiliar e automatizar a fase de aterragem de veículo aéreo não tripulado em navios, uma vez que é durante este período que ocorre o maior número de acidentes. A presente dissertação explora a visão computacional, aplicando duas câmaras por forma a localizar o veículo, apresentando-se como uma abordagem viável em situações em que ocorra falha de sinal GPS. Foram obtidos três conjuntos de dados reais, com um aparato experimental dedicado para o efeito, através da captura de imagens em três cenários distintos. Um dos cenários envolveu a aquisição de imagens a bordo de um navio da Marinha Portuguesa, evidenciando a diversidade dos dados obtidos. A calibração das câmaras foi efetuada de acordo com o método de Zhang e a determinação dos parâmetros intrínsecos e extrínsecos pelo método *Perspective-n-Point*. O alinhamento temporal entre as câmaras e os dados de telemetria do veículo aéreo não tripulado, permitem o sincronismo dos dados e uma maior precisão temporal. Os *datasets* adquiridos foram devidamente anotados para a análise do desempenho nos algoritmos implementados. O *dataset* disponibilizado é composto por cerca de vinte mil imagens anotadas através de *bounding boxes*. De forma a demonstrar a aplicação prática do *dataset*, os dados foram submetidos a duas aplicações distintas, o Filtro de Kalman Estendido e a Triangulação *Stereo*.

**Palavras-chave:** VANT, Aterragem autónoma, Visão estereoscópica, PnP, Filtro de Kalman, Geometria epipolar



# Abstract

Unmanned Aerial Vehicles (UAVs) have a wide range of applications, including operational surveillance, real-time reconnaissance, and humanitarian activities, among many others equally significant. Unmanned Aerial Vehicles can be operated manually by a pilot or autonomously on predefined routes. The limited area available on board makes landing unmanned aerial vehicles on ships an additional challenge. Furthermore, in some navigable areas, GPS (Global Positioning System) signals may experience intermittent signal failures, causing the challenge of landing the unmanned aerial Vehicle to increase. Therefore, assisting and automating the landing phase of unmanned aerial vehicles aboard ships is essential. This dissertation explores a feasible way of locating a unmanned aerial vehicle, with the help of computer vision aided by two cameras, in case of GPS (Global Positioning System) signal failure or jamming. A dedicated experimental apparatus captured images in three distinct scenarios, resulting in three real datasets. One of the scenarios involved acquiring images on board a Portuguese Navy ship, highlighting the diversity of data obtained. Camera calibration, defining the intrinsic and extrinsic parameters, was carried out according to the Zhang and Perspective-n-Point method. Time history and unmanned aerial vehicle telemetry data alignment allowed data synchronization and improved temporal accuracy. The available dataset consists of about 20,000 images annotated through bounding boxes. To demonstrate the practical application of the dataset, the data were submitted to two different procedures, the Kalman Filter and Epipolar Geometry.

**Keywords:** UAV, Autonomous landing, Stereo vision, PnP, Kalman filter, Epipolar geometry



# Índice

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Contexto Geral . . . . .	1
1.2	Motivação e Relevância para a Marinha Portuguesa . . . . .	6
1.3	Formulação do Problema . . . . .	7
1.4	Objetivos . . . . .	8
1.5	Estrutura da Dissertação . . . . .	9
<b>2</b>	<b>Revisão de Literatura e Enquadramento Teórico</b>	<b>11</b>
2.1	Classificação de UAVs . . . . .	11
2.2	Aterragem de UAVs a bordo de navios . . . . .	13
2.3	Deteção e seguimento de aeronaves e UAVs . . . . .	24
2.3.1	Deteção de aeronaves e UAVs . . . . .	24
2.3.2	Seguimento de aeronaves e UAVs . . . . .	27
2.4	Datasets . . . . .	28
<b>3</b>	<b>Metodologia</b>	<b>33</b>
3.1	Implementação das câmaras a bordo: disposição, geometria, referen- ciais e parâmetros . . . . .	33
3.1.1	Disposição das câmaras . . . . .	34
3.1.2	Geometria e Referenciais: Formação das Imagens . . . . .	35
3.1.3	Parâmetros extrínsecos e intrínsecos das câmaras . . . . .	37
3.2	Calibração das câmaras . . . . .	41
3.2.1	Parâmetros intrínsecos: método de Zhang . . . . .	41
3.2.2	Parâmetros extrínsecos: método <i>Perspective-n-Point</i> . . . . .	42
3.3	Deteção do UAV . . . . .	45
3.4	Estimação de posição e velocidade do UAV: geometria epipolar e filtro de Kalman estendido . . . . .	46
3.4.1	Geometria epipolar e triangulação <i>stereo</i> . . . . .	46
3.4.2	Filtro de Kalman estendido . . . . .	51

<b>4</b>	<b>Aquisição de dados</b>	<b>59</b>
4.1	Aquisição de <i>datasets</i> em diferentes cenários . . . . .	59
4.2	Descrição do hardware utilizado e aparato experimental . . . . .	61
4.3	Aquisição de dados 1 . . . . .	66
4.4	Aquisição de dados 2 . . . . .	70
4.5	Aquisição de dados 3 . . . . .	72
<b>5</b>	<b>Tratamento de dados</b>	<b>75</b>
5.1	Calibração das câmaras . . . . .	75
5.2	Alinhamento dos dados para avaliação de algoritmos de detecção e estimação . . . . .	82
5.3	Anotação do dataset com geração de <i>bounding boxes</i> para o UAV . .	85
5.4	Descrição do formato de dados utilizados para o dataset e disponibi- lização de endereço web para acesso ao público . . . . .	88
5.4.1	Organização da aquisição de dados 1 . . . . .	88
5.4.2	Organização da aquisição de dados 2 . . . . .	89
5.4.3	Organização da aquisição de dados 3 . . . . .	90
<b>6</b>	<b>Resultados Experimentais</b>	<b>91</b>
6.1	Teste experimental dos dados sintéticos . . . . .	91
6.1.1	Triangulação <i>stereo</i> para estimativa de posição 3D do UAV e respetiva velocidade . . . . .	93
6.1.2	Filtro de Kalman estendido para estimativa de posição 3D do UAV e respetiva velocidade . . . . .	97
6.2	Aplicação prática dos dados reais . . . . .	102
6.2.1	Aplicação prática dos dados da aquisição de dados 2 . . . . .	103
6.2.2	Aplicação prática dos dados da aquisição de dados 3 . . . . .	107
6.3	Simulação de <i>outliers</i> . . . . .	111
6.4	Discussão dos resultados . . . . .	113
<b>7</b>	<b>Conclusões e Trabalho Futuro</b>	<b>117</b>
7.1	Conclusões . . . . .	117
7.2	Trabalho Futuro . . . . .	118
	<b>Bibliografia</b>	<b>121</b>
	<b>Apêndices</b>	<b>129</b>
	<b>Anexos</b>	<b>129</b>

I	Desenvolvimento das derivadas parciais do Jacobiano	129
II	Desenvolvimento da fórmula de Haversine	131
III	Definição de RMSE e MAE	133



# Lista de Figuras

1.1	Zona de Jurisdição Portuguesa. . . . .	2
1.2	Aplicações de UAVs. . . . .	3
1.3	UAV utilizado na distribuição de abastecimento médico em Ruanda. . . . .	3
1.4	Entrega de encomendas e comida ao domicílio utilizando UAVs pela Amazon. . . . .	4
1.5	UAVs da Força Aérea Portuguesa que auxiliam no combate a incêndios. . . . .	4
1.6	Resumo das causas que podem influenciar a aterragem com sucesso a bordo de um navio. . . . .	6
2.1	Tipos de veículos aéreos não tripulados . . . . .	12
2.2	Sistema de Navegação Inercial (INS) . . . . .	15
2.3	Aterragem através do sistema GNSS . . . . .	15
2.4	Image Based Visual Servoing (IBVS) . . . . .	16
2.5	Position Based Visual Servoing . . . . .	17
2.6	Cenários possíveis para a aterragem de UAVs . . . . .	17
2.7	NRP Vasco da Gama da Classe Vasco da Gama . . . . .	18
2.8	NRP Sines da Classe Viana do Castelo . . . . .	18
2.9	Pista de aterragem de um navio da Marinha dos Estados Unidos da América . . . . .	18
2.10	Rede de abordagem utilizada numa lancha da Marinha Portuguesa. . . . .	18
2.11	Marcador 2D (Nguyen, K. Kim, Y. Lee e Park 2017). . . . .	20
2.12	Fluxograma do sistema integrado de aterragem a bordo . . . . .	20
2.13	Aterragem através do sistema TLS . . . . .	22
2.14	Visão binocular <i>stereo</i> . . . . .	23
2.15	UAV com câmara binocular . . . . .	24
2.16	Deteção de UAV através de Rede Neuronal Artificial (RNA) . . . . .	25
2.17	<i>Dataset</i> simulado no Ubuntu . . . . .	29
2.18	<i>Dataset</i> de deteção de um objeto (esquerda) e de vários objetos simultâneos (direita) (Y. Ma, Q. Li, Chu, Y. Zhou e C. Xu 2021). . . . .	29
2.19	<i>Dataset</i> real de aterragem de UAV com estimativa da altitude . . . . .	30

2.20	<i>Dataset</i> real de aterragem de UAV numa plataforma de simulação do balanço do navio. . . . .	30
2.21	<i>Dataset</i> real de um UAV em aproximação. . . . .	31
3.1	Disposição das câmaras a bordo de um navio. . . . .	34
3.2	Visão de objetos através de uma câmara. . . . .	35
3.3	Referenciais estabelecidos: referencial do mundo (W), da câmara (C) e da imagem (I). . . . .	37
3.4	Distorções radiais e tangenciais. . . . .	39
3.5	Marcação de pontos no marcador H. . . . .	44
3.6	Geometria Epipolar aplicada a bordo de um navio. . . . .	47
3.7	Geometria Epipolar detalha. . . . .	47
3.8	Projeção de um ponto numa linha epipolar. . . . .	49
3.9	Filtro de Kalman. . . . .	52
4.1	<i>Datasets</i> em diferentes cenários: cenário 1 (esquerda), cenário 2 (centro) e cenário 3 (direita). . . . .	60
4.2	Estruturas de suporte e câmaras IP. . . . .	62
4.3	Recetores GPS utilizados nos ensaios experimentais. . . . .	63
4.4	marcador H bilateral e as suas dimensões. . . . .	64
4.5	UAV <i>Tarot X6 Hexacopter</i> utilizado no primeiro <i>dataset</i> . . . . .	65
4.6	UAV <i>Autel EVO II Dual 640T Quadcopter</i> utilizado no segundo <i>dataset</i> . . . . .	65
4.7	UAV <i>DJI Matrice 600 Pro</i> utilizado no terceiro <i>dataset</i> . . . . .	66
4.8	Imagens do Exercício <i>Guerriers de La Méditerranée</i> . . . . .	66
4.9	Aterragem vertical do <i>dataset</i> . . . . .	67
4.10	Aterragem horizontal do <i>dataset</i> . . . . .	67
4.11	Aproximações do UAV da Aquisição de dados 1. . . . .	68
4.12	Imagens dos ensaios a bordo no NRP Corte-Real. . . . .	70
4.13	Aproximações do UAV da Aquisição de dados 2. . . . .	70
4.14	Imagens dos ensaios em colaboração com o CIAFA. . . . .	72
4.15	Aproximações do UAV da Aquisição de dados 3. . . . .	72
5.1	Ensaio para aquisição de imagens para calibração das câmaras. . . . .	76
5.2	Fluxograma de procedimentos de calibração das câmaras. . . . .	76
5.3	Imagens com o padrão xadrez reconhecido com as câmaras sem zoom. . . . .	77
5.4	Imagem antes da correção (esquerda) e imagem após a correção (direita) sem zoom. . . . .	77

5.5	Imagens com o padrão xadrez reconhecido com as câmaras com 21,13% zoom. . . . .	77
5.6	Imagem antes da correção (esquerda) e imagem após a correção (direita) com 21,13% zoom. . . . .	78
5.7	Imagens com o padrão xadrez reconhecido com as câmaras com 61,03% zoom. . . . .	78
5.8	Imagem antes da correção (esquerda) e imagem após a correção (direita) com 61,03% zoom. . . . .	78
5.9	Pontos bidimensionais da aquisição de dados 2. . . . .	80
5.10	Pontos bidimensionais da aquisição de dados 3. . . . .	81
5.11	Aterragem do UAV da aquisição 1 da câmara 1 (esquerda) e da câmara 2 (direita). . . . .	82
5.12	Aterragem do UAV da aquisição 2 da câmara 1 (esquerda) e da câmara 2 (direita). . . . .	83
5.13	Aterragem do UAV da aquisição 3 da câmara 1 (esquerda) e da câmara 2 (direita). . . . .	84
5.14	Imagens da aquisição de dados 1 com <i>bounding box</i> do UAV capturada pela câmara 1 (esquerda) e câmara 2 (direita). . . . .	86
5.15	Imagens da aquisição de dados 2 com <i>bounding box</i> do UAV capturada pela câmara 1 (esquerda) e câmara 2 (direita) . . . . .	86
5.16	Imagens da aquisição de dados 3 com <i>bounding box</i> do UAV capturada pela câmara 1 (esquerda) e câmara 2 (direita) . . . . .	87
6.1	Posição do UAV das câmaras 1 (esquerda) e 2 (direita) dos dados sintéticos. . . . .	92
6.2	Coordenadas na imagem $u_1, v_1, u_2$ e $v_2$ em função do tempo dos dados sintéticos . . . . .	93
6.3	Posição 2D e 3D do UAV através da triangulação <i>stereo</i> dos dados sintéticos sem ruído. . . . .	94
6.4	Estabilidade da posição 3D dos dados sintéticos sem ruído. . . . .	94
6.5	Evolução do ruído na triangulação <i>stereo</i> . . . . .	95
6.6	Velocidades do UAV em função do tempo dos dados sintéticos através da triangulação <i>stereo</i> . . . . .	96
6.7	Erros da posição 3D do UAV ao longo do tempo na triangulação <i>stereo</i> . . . . .	96
6.8	Inicialização do FKE no ponto do marcador H (0, 0, 0). . . . .	97
6.9	Inicialização do FKE num ponto distante. . . . .	98
6.10	Inicialização do FKE num ponto calculado através da triangulação <i>stereo</i> . . . . .	99

6.11	Aumento da variância do ruído do processo. . . . .	99
6.12	Evolução do ruído no filtro de Kalman estendido. . . . .	100
6.13	Aumento da variância do ruído da medição. . . . .	100
6.14	Velocidades do UAV em função do tempo dos dados sintéticos através do filtro de Kalman estendido. . . . .	101
6.15	Erros da posição 3D do UAV ao longo do tempo no FKE. . . . .	101
6.16	Posição do UAV das imagens capturadas pela câmara 1 (esquerda) e 2 (direita) na aquisição de dados 2. . . . .	103
6.17	Coordenadas na imagem $u_1, v_1, u_2$ e $v_2$ em função do tempo da aquisição de dados 2 . . . . .	104
6.18	Velocidades do UAV em função do tempo da aquisição de dados 2. . . . .	105
6.19	Posição 3D do UAV através do filtro de Kalman da aquisição de dados 2. . . . .	105
6.20	Posição 3D do UAV através do FKE da aquisição de dados 2. . . . .	106
6.21	Posição 3D do UAV através da telemetria da aquisição de dados 2. . . . .	106
6.22	Posição do UAV das câmaras 1 (esquerda) e 2 (direita) na aquisição de dados 3. . . . .	107
6.23	Coordenadas na imagem $u_1, v_1, u_2$ e $v_2$ em função do tempo da aquisição de dados 3 . . . . .	108
6.24	Velocidades do UAV em função do tempo da aquisição de dados 3. . . . .	109
6.25	Posição 3D do UAV através da triangulação <i>stereo</i> da aquisição de dados 3. . . . .	109
6.26	Posição 3D do UAV através do filtro de Kalman da aquisição de dados 3. . . . .	110
6.27	Posição 3D do UAV através da telemetria da aquisição de dados 3. . . . .	110
6.28	Triangulação <i>stereo</i> com <i>outliers</i> . . . . .	111
6.29	Filtro de Kalman estendido com <i>outliers</i> . . . . .	112
6.30	Filtro de Kalman estendido com <i>outliers</i> e maior incerteza. . . . .	113
6.31	Trajetórias do UAV através da triangulação <i>stereo</i> , filtro de Kalman estendido e <i>ground truth</i> do UAV. . . . .	113
6.32	Triangulação <i>stereo</i> e filtro de Kalman estendido perante diferentes níveis de ruído. . . . .	114
6.33	Trajetórias do UAV através da triangulação <i>stereo</i> , filtro de Kalman estendido e <i>ground truth</i> do UAV. . . . .	114
6.34	Posições 3D do UAV através da triangulação <i>stereo</i> , filtro de Kalman estendido e <i>ground truth</i> do UAV com outliers. . . . .	115

6.35 Erros da posição 3D do UAV ao longo do tempo no FKE e triangulação <i>stereo</i> . . . . .	116
---	-----



# Lista de Tabelas

2.1	Classes de veículos aéreos não tripulados consoante o peso . . . . .	12
4.1	Características Principais das Câmaras . . . . .	61
4.2	Características <i>6-Port 10/100Mbps Desktop Switch with 4-Port PoE+</i> . . . . .	62
4.3	Características do navegador GPS . . . . .	63
4.4	Características do UAV da Aquisição de Dados 1. . . . .	64
4.5	Características do UAV da Aquisição de Dados 2 . . . . .	65
4.6	Características do UAV da Aquisição de Dados 3 . . . . .	66
4.7	Parâmetros das câmaras do Ensaio 1 ao 10 da Aquisição de dados 3 . . . . .	68
4.8	Parâmetros das câmaras do Ensaio 11 da Aquisição de dados 1 . . . . .	69
4.9	Parâmetros das câmaras do Ensaio 1 da Aquisição de dados 2 . . . . .	71
4.10	Parâmetros das câmaras do Ensaio 2 ao 5 da Aquisição de dados 2 . . . . .	71
4.11	Parâmetros das câmaras durante todos os Ensaios da Aquisição de dados 3 . . . . .	73
5.1	Parâmetros intrínsecos das câmaras: Erro de reprojeção, matrizes intrínsecas e coeficientes de distorção de cada câmara. . . . .	79
5.2	Matrizes dos parâmetros extrínsecos das duas câmaras. . . . .	81
5.3	Momento de aterragem da aquisição de dados 1. . . . .	83
5.4	Momento de aterragem da aquisição de dados 2. . . . .	83
5.5	Momento de aterragem da aquisição de dados 3 antes da correção. . . . .	84
5.6	Momento de aterragem da aquisição de dados 3 após a correção. . . . .	85



# Lista de Equações

3.1	Transformação de coordenadas cartesianas para coordenadas homogéneas. . . . .	36
3.2	Coordenadas homogéneas do Referencial W, C e I. . . . .	36
3.3	Coordenadas cartesianas do Referencial W, C e I. . . . .	37
3.4	Matriz homogénea dos parâmetros extrínsecos. . . . .	38
3.5	Matriz homogénea das coordenadas da câmara. . . . .	38
3.6	Equação dos parâmetros extrínsecos. . . . .	38
3.7	Formação da imagem bidimensional. . . . .	38
3.8	Definição das variáveis $u$ e $v$ . . . . .	39
3.9	Coefficientes de distorção. . . . .	39
3.10	Distorção radial. . . . .	40
3.11	Distorção tangencial. . . . .	40
3.12	Definição das variáveis $x'$ e $y'$ . . . . .	40
3.13	Desenvolvimento das variáveis $x'$ e $y'$ . . . . .	40
3.14	Desenvolvimento das variáveis $u$ e $v$ . . . . .	40
3.15	Definição da matriz de $u$ e $v$ . . . . .	40
3.16	Ponto 3D observado por cada câmara. . . . .	48
3.17	Desenvolvimento da equação do plano. . . . .	48
3.18	Vetor de translação da câmara 1 para a câmara 2. . . . .	48
3.19	Desenvolvimento da matriz essencial. . . . .	49
3.20	Desenvolvimento das equações de projeção. . . . .	49
3.21	Matriz fundamental. . . . .	49
3.22	Coordenadas dos pontos projetados. . . . .	50
3.23	Desenvolvimento das coordenadas dos pontos projetados. . . . .	50
3.24	Velocidade em cada eixo através da triangulação <i>stereo</i> . . . . .	51
3.25	Definição do estado do UAV. . . . .	51
3.26	Definição do centro das <i>bounding boxes</i> . . . . .	52
3.27	Desenvolvimento da incerteza do estado de atualização. . . . .	53
3.28	Estado atual do UAV. . . . .	53
3.29	Definição da matriz de controlo. . . . .	53

3.30	Desenvolvimento do estado do UAV e definição da matriz de transição de estado. . . . .	54
3.31	Matriz da covariância da incerteza do estado atual. . . . .	54
3.32	Desenvolvimento da matriz da covariância da incerteza na dinâmica. .	54
3.33	Desenvolvimento das coordenadas da imagem. . . . .	55
3.34	Definição das coordenadas da câmara. . . . .	55
3.35	Equação final das coordenadas da imagem. . . . .	55
3.36	Matriz do filtro de Kalman estendido. . . . .	55
3.37	Definição das variáveis da matriz do EKF. . . . .	56
3.38	Cálculo do Jacobiano. . . . .	56
3.39	Cálculo da primeira matriz do Jacobiano. . . . .	56
3.40	Cálculo da segunda matriz do Jacobiano. . . . .	56
3.41	Equação do ganho de Kalman. . . . .	57
3.42	Definição do estado após a atuação atualização e a incerteza associada.	57
I.1	Cálculo da primeira matriz do Jacobiano. . . . .	129
I.2	Desenvolvimento da primeira matriz do Jacobiano. . . . .	129
I.3	Valores da variáveis A e B da primeira matriz do Jacobiano. . . . .	129
II.1	Fórmula trigonométrica de Haversine. . . . .	131
II.2	Desenvolvimento da fórmula de Haversine. . . . .	131
III.1	Fórmula do RMSE . . . . .	133
III.2	Fórmula do MAE . . . . .	133

# Lista de Abreviaturas, Siglas e Acrónimos

<b>2D</b>	Bidimensional
<b>3D</b>	Tridimensional
<b>ASPOF</b>	Aspirante a Oficial
<b>BF2</b>	Batalhão de Fuzileiros n <sup>o</sup> 2
<b>CAD</b>	<i>Computer-Aided Design</i>
<b>CBR</b>	<i>Constant Bit Rate</i>
<b>CEOV</b>	Célula de Experimentação Operacional de Veículos Não Tripulados
<b>CIAFA</b>	Centro de Investigação da Academia da Força Aérea
<b>CINAV</b>	Centro de Investigação Naval
<b>COCO</b>	<i>Common Objects in Context</i>
<b>DAC</b>	Destacamento de Apoio de Combate
<b>DL</b>	<i>Deep Learning</i>
<b>DLT</b>	<i>Direct Line Transform</i>
<b>DT-YOLO</b>	<i>Detection and Tracking You Only Look Once</i>
<b>EKF</b>	<i>Extended Kalman Filter</i>
<b>EPnP</b>	<i>Efficient Perspective-n-Point</i>
<b>FK</b>	Filtro de Kalman
<b>FKE</b>	Filtro de Kalman Estendido
<b>FPS</b>	<i>Frames Per Second</i>
<b>GCS</b>	<i>Ground Control Station</i>
<b>GNSS</b>	<i>Global Navigation Satellite System</i>
<b>GPS</b>	<i>Global Positioning System</i>
<b>IBVS</b>	<i>Image-Based Visual Servoing</i>
<b>IMU</b>	<i>Inertial Measurement Unit</i>
<b>INS</b>	<i>Inertial Navigation System</i>
<b>IP</b>	<i>Internet Protocol</i>
<b>ISR</b>	<i>Intelligence Surveillance and Reconnaissance</i>
<b>IST-ID</b>	Instituto Superior Técnico - Investigação e Desenvolvimento
<b>JSON</b>	<i>JavaScript Object Notation</i>

<b>KF</b>	<i>Kalman Filter</i>
<b>LPI</b>	<i>Landing Period Indicator</i>
<b>MAE</b>	<i>Mean Absolute Error</i>
<b>MATLAB</b>	<i>Matrix Laboratory</i>
<b>MAV</b>	<i>Micro Aerial Vehical</i>
<b>NN</b>	<i>Neural Network</i>
<b>NPO</b>	Navio Patrulha Oceânico
<b>NRP</b>	Navio da República Portuguesa
<b>P3P</b>	<i>Perspective-Three-Point</i>
<b>PBVS</b>	<i>Position-Based Visual Servoing</i>
<b>PnP</b>	<i>Perspective-n-Point</i>
<b>PoE</b>	<i>Power Over Ethernet</i>
<b>PTZ</b>	<i>Pan Tilt Zoom</i>
<b>RANSAC</b>	<i>Random Sample Consensus</i>
<b>RBPF</b>	<i>Rao-Blackwellized Particle Filter</i>
<b>RGB</b>	<i>Red Green Blue</i>
<b>RGB-D</b>	<i>Red Green Blue-Depth</i>
<b>RMSE</b>	<i>Root Mean Square Error</i>
<b>RNA</b>	Rede Neuronal Artificial
<b>ROI</b>	<i>Region Of Interest</i>
<b>SANT</b>	Sistemas Aéreos Não-tripulados
<b>SAR</b>	<i>Search And Rescue</i>
<b>SLAD</b>	<i>Safe Landing Areas Detection</i>
<b>SLAM</b>	<i>Simultaneous Localization and Mapping</i>
<b>SPA</b>	<i>Signal Processing Algorithm</i>
<b>SRR</b>	<i>Search and Rescue Region</i>
<b>SSD</b>	<i>Single Shot Detector</i>
<b>TLS</b>	<i>Terrestrial Laser Scanning</i>
<b>ToF</b>	<i>Time of Fly</i>
<b>UAV</b>	<i>Unmanned Aerial Vehicle</i>
<b>UBiF</b>	<i>Unscented Bingham Filter</i>
<b>UGV</b>	<i>Unmanned Ground Vehicle</i>
<b>UKF</b>	<i>Unscented Kalman Filter</i>
<b>UPF</b>	<i>Unscented Particle Filter</i>
<b>UPnP</b>	<i>Uncalibrated Perspective-n-Point</i>
<b>USRP</b>	<i>Universal Software Radio Peripheral</i>
<b>USV</b>	<i>Unmanned Surface Vehicle</i>

<b>VANT</b>	Veículo Aéreo Não Tripulado
<b>VBR</b>	<i>Variable Bit Rate</i>
<b>VTOL</b>	<i>Vertical Take-Off and Landing</i>
<b>Wi-Fi</b>	<i>Wireless Fidelity</i>
<b>YOLO</b>	<i>You Only Look Once</i>
<b>ZEE</b>	Zona Económica Exclusiva



# Capítulo 1

## Introdução

A aterragem de um UAV de forma autónoma tem sido um desafio durante a última década devido à sua complexidade e de todos os parâmetros que nela estão envolvidos. Para efetuar a aterragem com sucesso, tem de se considerar também todos os fatores externos, como por exemplo as condições meteorológicas, o meio onde este sobrevoa, bem como, a plataforma onde se pretende aterrar.

A introdução aborda na Secção 1.1 o contexto geral em que a presente dissertação assenta. A Secção 1.2 abrange a motivação e relevância para a Marinha Portuguesa. O problema particular investigado na presente dissertação é definido na Secção 1.3. Na Secção 1.4 estão descritos os objetivos que se pretendem alcançar. E por fim, na Secção 1.5 descreve-se a estrutura da presente dissertação.

### 1.1 Contexto Geral

Um *Unmanned Aerial Vehicle* (UAV) é uma aeronave que opera sem nenhum piloto a bordo, podendo ter ou não a influência de um ser humano, dependendo do seu nível de automatização. Os UAV podem operar em variadas situações como o reconhecimento aéreo em tempo real, vigilância tanto a nível operacional como em atividades civis, ações humanitárias, distribuição de equipamentos médicos ou até mesmo de resgate de emergência (Nex et al. 2022).

De acordo com a Direção-Geral de Recursos Naturais, Portugal possui uma vasta Zona Económica Exclusiva (ZEE), a vigésima maior a nível mundial e quinta a nível europeu, que se divide em três subáreas: "subárea do continente (287 521 quilómetros quadrados), subárea dos Açores (930 687 quilómetros quadrados) e subárea da Madeira (442 248 quilómetros quadrados)", tornando-se quase impossível vigiar a mesma com os recursos que temos disponíveis. A área da ZEE é cerca

de dezanove vezes superior à área de Portugal Continental. Segundo a Convenção Internacional sobre Busca e Salvamento Marítimo de 1979, acresce a esta zona anteriormente mencionada área territorial de aproximadamente noventa e três mil quilómetros quadrados, bem como uma Região de Busca e Salvamento (SRR) de cerca de seis milhões de quilómetros quadrados (DGRM 2018), como é ilustrada na Figura 1.1.

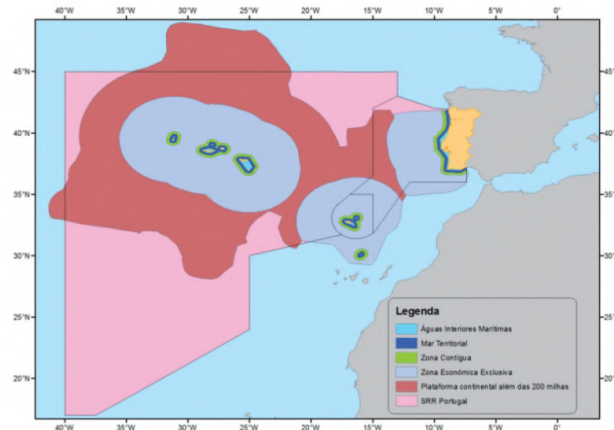


FIGURA 1.1: Zona de Jurisdição Portuguesa (DGRM 2018).

A vasta extensão é quase inatingível na sua totalidade, dando origem ao conceito de veículos autónomos não tripulados, a fim de se conseguir abranger uma área ainda maior de proteção e vigilância.

Os UAVs são recursos financeiramente mais económicos, para além de não necessitarem de uma tripulação a bordo, não colocando em risco desnecessário a vida de uma guarnição. A utilização de UAVs em missões de busca e salvamento, *Search and Rescue* (SAR), bem como em reconhecimento de outras unidades navais ou locais remotos requer a capacidade de descolagem, voo e aterragem com sucesso (Gonçalves e Damas 2022). Deste modo, a integração de um sistema de apoio à aterragem de UAVs em navios será um ponto marcante no desenvolvimento de novas tecnologias e estudos.

Os UAVs, por sua vez, podem ser controlados remotamente por um ou mais pilotos simultaneamente, desde que tenham acesso e o UAV esteja dentro do seu raio de alcance de controlo. Atualmente, também é possível que o UAV realize voos de forma autónoma, dispensando a necessidade de intervenção humana.

Diversos UAVs são configurados pelo operador antes do lançamento, e podem ser programadas as rotas de voo, a origem de onde descola e aterra. Nesses

casos, o UAV realiza operações de forma autónoma, sem a necessidade de intervenção do piloto ao longo da missão (Hassanalian e Abdelkefi 2017). Nos dias que decorrem, os UAVs são utilizados para exercer todo o tipo de funções, desde vigilância, a reconhecimento de zonas, mas também já estão inseridos na área da logística, como se observa através da Figura 1.2.

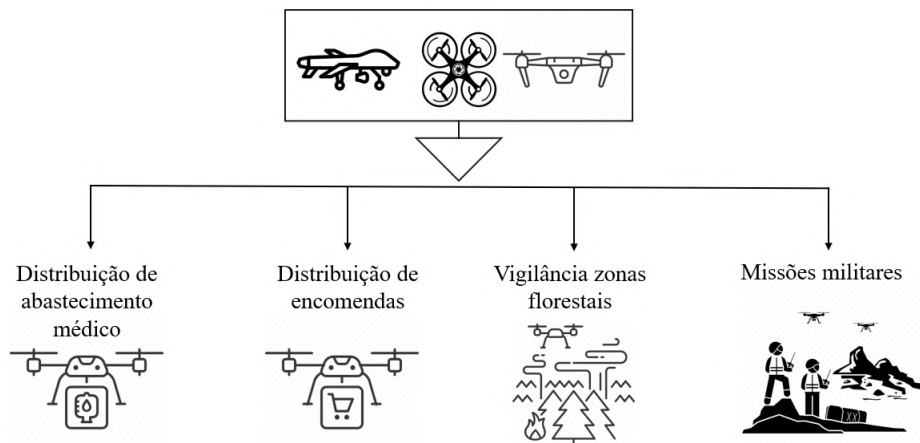


FIGURA 1.2: Aplicações de UAVs.

Um destes exemplos, é o caso dos UAVs que estão a ser empenhados na distribuição de sangue e outros medicamentos na África Oriental. Um experimento real foi proposto em Ruanda em 2016, conseguindo resultados incríveis e passíveis de serem adaptados a muitos outros países. Os médicos requisitavam abastecimento médico através de uma aplicação, e mesmo estando em zonas mais rurais e de difíceis acessos, os UAVs Zipline de asa fixa, entregavam-nos no destino (Ackerman e Strickland 2018), como ilustra a Figura 1.3.



FIGURA 1.3: UAV empregue na distribuição de abastecimento médico em Ruanda (Guardian 2018).

Na vertente do comércio, empresas como a Amazon e Third Space Auto, criaram parcelas de entregas de encomendas de forma ainda mais expedita e ecológica, empregando UAVs nas entregas e comida ao domicílio, como se observa na Figura 1.4. A Amazon desenvolveu uma nova divisão com o nome Amazon Prime

Air, com o objetivo de entregar de forma mais rápida e precisa as encomendas efetuadas *online*, tendo já testado o seu serviço em várias cidades dos Estados Unidos (Jung e H. Kim 2018).

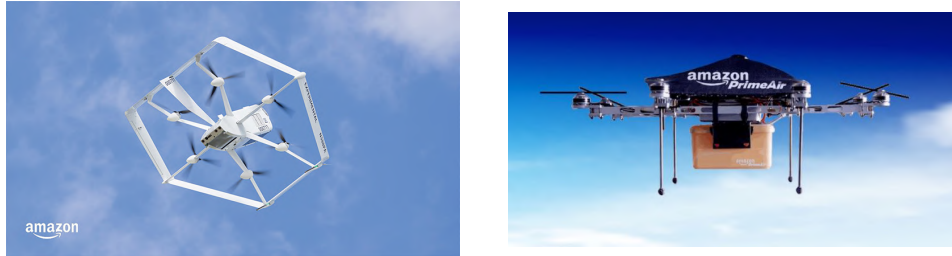


FIGURA 1.4: Entrega de encomendas e comida ao domicílio utilizando UAVs pela Amazon (Oitzman 2022).

Em Portugal, desde 2020, foram atribuídos à Força Aérea Portuguesa os "primeiros 12 drones vocacionados para vigilância, observação e coordenação", destacados para o Dispositivo Especial de Combate a Incêndios Rurais, como demonstra a Figura 1.5. A sua grande vantagem é a deteção inicial de incêndios, auxiliar e monitorizar as áreas em perigo, e desta forma assessorar as tomadas de decisão no Comando de Operações de Combate (República Portuguesa 2020).



FIGURA 1.5: UAVs da Força Aérea Portuguesa que auxiliam no combate a incêndios (República Portuguesa 2020).

No ramo militar internacional, também já existem vários UAVs a desempenharem papéis cruciais tanto em tempo de paz como em tempo de guerra. OS UAVs podem desempenhar funções de vigilância, reconhecimento de zonas e do próprio inimigo, como também têm a capacidade de carregar material explosivo e munições, conseguindo identificar um alvo e disparar (Konert e Balcerzak 2021).

A necessidade de auxílio no mar, para cobrir a extensa área marítima existente, como referido anteriormente, requer a participação de UAVs, evitando assim que o navio tenha de se deslocar até ao local de assistência. Desde fiscalizações a navios pesqueiros, vigilância de navios suspeitos de tráfico, ou até mesmo navios

que necessitem de auxílio imediato, é muito mais rápido empenhar um UAV para uma resposta eficiente. Desta forma, emerge a necessidade de conseguir controlar os UAVs remotamente, bem como garantir a sua descolagem e aterragem sem incidentes.

Surge então, a exigência de se conseguir controlar um UAV a partir de um navio. Desde conseguir controlar uma plataforma móvel sujeita a oscilações, como é o caso do convés de um navio, local que será o ponto de partida e chegada do UAV. Antes do UAV descolar, é possível delinear toda a rota que se pretende que ele efetue, se a qualquer momento da missão o UAV perder o sinal com o piloto, tem a capacidade de concluir a missão e retornar ao ponto de onde descolou. No caso de se perder o sinal GPS, ou as comunicações com o piloto, ocorre então, a possibilidade de através de câmaras a bordo do navio, identificar a sua posição e auxiliar o piloto a efetuar a aterragem, ou até o próprio UAV conseguir receber os dados e aterrar de forma autónoma.

O primeiro obstáculo para a aterragem dos UAVs em navios é a oscilação da própria plataforma, devido às condições meteorológicas e marítimas (Riola, Diaz e Giron-Sierra 2011). Outro desafio que torna a aterragem de UAVs de forma autónoma em navios, é a restrição do local de aterragem, deixando pouca margem de erro ao UAV e ao algoritmo de aterragem, caso ocorra um erro de dez metros, existe o risco de o UAV cair ao mar. A perda ou empastelamento de sinal GPS (*Global Positioning System*) também se torna uma agravante no que se refere à aterragem de UAV em navios. Muitas vezes a navegar em zonas mais remotas, o sinal GPS é escasso e degradado, ocorrendo várias falhas momentâneas, o que torna a fase de aterragem instável. No entanto, através de visão computacional, depende das câmaras que observam o UAV e não do sinal GPS (Xia, Mantegh e W. Xie 2021).

O horário de operação do UAV constitui um dos problemas primordiais no âmbito da utilização de visão computacional a bordo de navios. Caso ocorra uma falha do sinal GPS, a aterragem do UAV com sucesso depende das boas condições de visibilidade e do horário de operação, entre outros ilustrados na Figura 1.6, o cenário torna-se mais complexo (Good, Aboutalib, Thai, Yamaoka, C. Kim, Wilkinson e Findlay 2016). Considerando que a baixa luminosidade possa apresentar-se como uma barreira, principalmente quando o UAV efetua voos noturnos, os métodos de visão precisam de ser adaptados para visão noturna ou térmica.

Em concordância com os desafios apresentados anteriormente, a presente dissertação pretende auxiliar a aterragem de UAVs a bordo de navios. Contudo, existem vários tipos de abordagem possível, como é o caso da visão monocular e

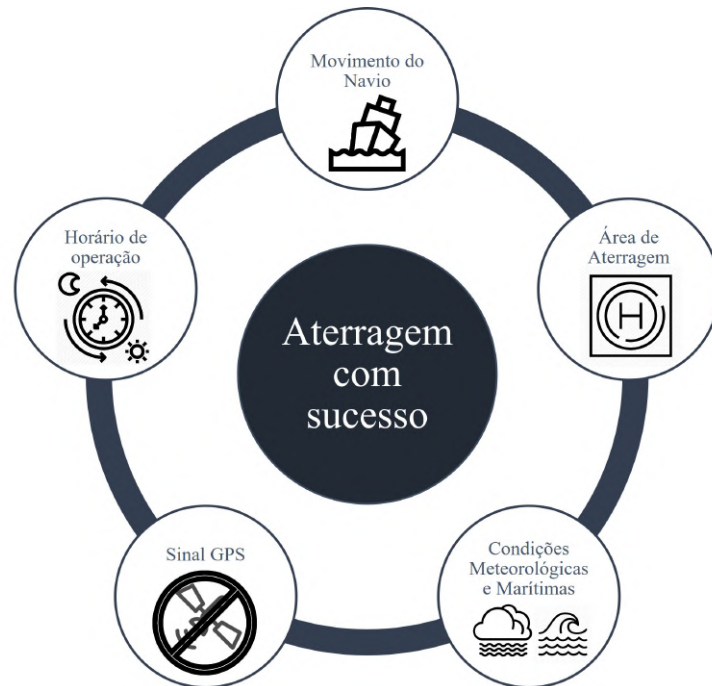


FIGURA 1.6: Resumo das causas que podem influenciar a aterragem com sucesso a bordo de um navio.

da visão estereoscópica. Como se pretende estimar a localização 3D da aeronave, propôs-se implementar câmaras a bordo, dispostas perto do convés onde se pretende que o UAV aterre. De seguida, e sabendo as suas localizações a bordo, é possível prever através de visão *stereo* a localização do UAV, auxiliando desta forma na sua aterragem. Este método vem auxiliar principalmente em casos de falha de sinal GPS, atuando como auxílio ao próprio piloto do UAV.

Através de duas câmaras RGB (*Red Green Blue*), que irão utilizar o mesmo princípio de visão dos olhos humanos, acompanhando o UAV ao longo da sua trajetória descendente. Inicialmente, durante os ensaios descritos ao longo dos próximos capítulos, as câmaras irão adquirir imagens de UAVs em aproximação, simulando a aterragem dos mesmos.

## 1.2 Motivação e Relevância para a Marinha Portuguesa

A relevância e emprego operacional de UAVs na Marinha Portuguesa têm sido crescente ao longo dos anos. Os UAVs desempenham um papel fundamental na "fiscalização dos espaços marítimos, a deteção e combate de ilícitos, busca e

salvamento, a coordenação do plano de resgate em massa para os navios-cruzeiros e o combate à poluição"(Marcelino 2021).

O voo de um UAV baseia-se nas seguintes fases: fase de descolagem, fase de sobrevoo e fase de aterragem (Azevedo 2022). Sendo que é ao redor desta última fase que a presente dissertação se irá basear.

A aterragem de um UAV num navio torna-se numa missão desafiante, por vários motivos, como o facto de os navios da Marinha Portuguesa não terem uma pista de aterragem dedicada para este efeito e possuírem uma área bastante limitada o que dificulta a execução da mesma. Atualmente é utilizada nos navios uma rede de abordagem ao longo dessa área de aterragem permitindo assim que o UAV colida contra esta e acabe por ficar retido.

A principal motivação para a elaboração desta dissertação será conseguir automatizar o máximo possível a fase de aterragem de um UAV a bordo de um navio, e demonstrar que através de duas câmaras e conhecendo as suas localizações a bordo, se consiga obter a localização do UAV. Desta forma, torna-se um conceito aplicável principalmente quando existe empastelamento ou falha no sinal GPS, dependendo então apenas das câmaras em utilização.

O projeto VOAMASIS teve início em fevereiro de 2019, tendo terminado em janeiro de 2023, e é financiado por Portugal 2020 e Fundação para a Ciência e Tecnologia. As parcerias deste projeto são o CINAV (Centro de Investigação Naval), ISR (*Institute for Systems and Robotics*) e CIAFA (Centro de Investigação da Academia da Força Aérea). As entidades beneficiárias deste projeto são o IST-ID (Associação do Instituto Superior Técnico para a Investigação e Desenvolvimento), que é o líder do projeto, o Estado Maior da Armada e da Força Aérea. O projeto VOAMASIS conta também com o apoio financeiro da União Europeia.

Os resultados do projeto VOAMASIS terão impacto em operações nacionais estratégicas, como por exemplo em ambientes florestais ou marítimos. A base deste programa tem foco noutro projeto denominado SEAGULL, cujo objetivo é investigar e desenvolver estudos para operações que envolvam vigilância, segurança da vida humana e ambiental, usufruindo da tecnologia de UAVs.

## 1.3 Formulação do Problema

A presente dissertação visa mitigar as dificuldades de aterragem dos UAVs, localizando o UAV na sua fase final do voo. Deste modo, é apresentado um sistema

integrado que se destina a fornecer dados durante as operações de aterragem dos UAVs em navios.

Uma vez que, durante a aproximação de um UAV ao navio, a câmara do próprio UAV pode perder o navio do seu campo de visão, recorrer câmaras localizadas a bordo do navio poderá ser uma vantagem. Na medida em que, através de duas câmaras é possível acompanhar continuamente a trajetória do UAV, através de visão *stereo*, desde o momento em que este entra no cone de aproximação à pista de aterragem.

Através de duas câmaras localizadas a bordo do navio, sabendo as suas disposições exatas, define-se um referencial. Efetuando a aquisição de dados em cenários distintos é possível criar um dataset robusto, no que concerne ao comportamento que o UAV poderá realizar. É fundamental proceder-se à calibração das câmaras previamente, de forma a obter os parâmetros intrínsecos e extrínsecos.

A posição do UAV obtida através de cada câmara tem em consideração um referencial fixo a bordo do navio, como por exemplo um alvo. Conhecendo também o posicionamento de uma câmara em relação à outra, é possível gerar um sistema de referencial. Deste modo, é viável reconstruir tridimensionalmente o cenário envolvente bem como a aproximação do UAV.

No final, o *dataset* elaborado é submetido para a análise prática através de dois métodos distintos: o Filtro de Kalman Estendido e a Triangulação *Stereo*, baseada em Geometria Epipolar.

## 1.4 Objetivos

O objetivo da presente dissertação é a integração de um sistema de apoio à aterragem de UAVs em navios, desta forma o projeto foi segmentado em fases distintas descritas posteriormente.

No contexto de auxílio na manobra de aterragem do UAV a bordo, propõe-se a utilização de duas câmaras instaladas a bordo do navio. A utilização de duas câmaras ao invés de uma, tem por objetivo a obtenção de posições tridimensionais (3D) por meio da análise de imagens bidimensionais (2D). Estas câmaras tem como objetivo capturar imagens sincronizadas que serão utilizadas para desenvolver e testar algoritmos de estimação de posição, e posteriormente, a sua velocidade e orientação de um UAV.

De forma a viabilizar este processo, é necessário efetuar a aquisição de vários conjuntos de dados reais, que irão formar os *datasets*. Os *datasets* consistem nas imagens capturadas pelas câmaras sincronizadas, em concordância com o *ground-truth* do UAV. Todos os dados referidos serão fundamentais para o desenvolvimento e validação do algoritmo de estimação de posição, permitindo a sua avaliação precisa do desempenho técnico.

Uma vez adquiridos os *datasets*, procede-se à fase seguinte: teste e validação do conceito proposto. Nesta fase, são utilizados algoritmos conhecidos, como visão *stereo* e o filtro de Kalman estendido, para processar as imagens e estimar a posição do UAV.

Assim sendo, a combinação das câmaras instaladas a bordo, dos *datasets* reais e dos algoritmos de estimação de posição permitirá o desenvolvimento, teste e validação de um sistema capaz de fornecer informações precisas acerca da posição, velocidade e orientação do UAV durante a sua fase de aterragem. E deste modo, contribuir para a eficiência das operações de voo destas aeronaves especialmente em ambientes desafiadores, como é o caso de um navio em pleno mar.

## 1.5 Estrutura da Dissertação

A presente dissertação é composta por diversos capítulos, em que cada um contribui para a exploração aprofundada dos elementos-chave, que são os seguintes:

- **Capítulo 2 - Revisão de Literatura e Enquadramento Teórico:** Este capítulo foca os temas desenvolvidos por outros autores, como a classificação de UAVs, aterragem de UAVs a bordo de navios, deteção e seguimento de aeronaves e UAVs e *datasets* desenvolvidos relacionados com o presente tema;
- **Capítulo 3 - Metodologia:** O capítulo descreve a implementação das câmaras a bordo, os cálculos dos parâmetros intrínsecos e extrínsecos das câmaras, a calibração das câmaras, a estimação de posição e velocidade do UAV;
- **Capítulo 4 - Aquisição de Dados:** Este capítulo aborda a descrição geral da aquisição de *datasets* em diferentes cenários, começando pelo hardware utilizado e o restante aparato experimental, e das três aquisições de dados;
- **Capítulo 5 - Tratamento de dados:** O capítulo concentra no cálculo da calibração das câmaras e as suas matrizes intrínsecas e extrínsecas, bem como,

do alinhamento dos dados para a avaliação de algoritmos de detecção e estimação. Foca também a anotação do *dataset* com geração de *bounding boxes* para os UAVs e descreve o formato de dados do *dataset*;

- **Capítulo 6 - Resultados Experimentais:** Este capítulo explora o uso de triangulação *stereo* e de um filtro de Kalman estendido para a estimativa da posição 3D do UAV e a sua velocidade. Discute-se os resultados entre ambos os métodos e a telemetria obtida pelo UAV;
- **Capítulo 7 - Conclusões:** O último capítulo compara os resultados esperados com os obtidos, onde se reflete sobre as divergências dos métodos e aborda trabalhos a desenvolver futuramente na área em análise.

# Capítulo 2

## Revisão de Literatura e Enquadramento Teórico

Neste Capítulo, analisa-se os conceitos já desenvolvidos por outros autores de forma a suportar a dissertação em questão e o trabalho com ela relacionado. A Secção 2.1 aborda a classificação de UAVs. A Secção 2.2 apresenta estudos já efetuados de aterragens de UAVs. A Secção 2.3 foca a deteção e seguimento de aeronaves e UAVs. E por fim, a Secção 2.4 que aprofunda *datasets* já desenvolvidos por outros autores.

### 2.1 Classificação de UAVs

Ao longo dos anos verificou-se um desenvolvimento exponencial de vários tipos de aeronaves não tripuladas, com empregos, alcances, pesos, tipo de asas, velocidades distintas entre outras características discriminatórias.

Para além da maior distinção entre UAVs, já referenciada no Capítulo 1, acerca de os UAVs serem pilotados de forma autónoma ou com recurso a piloto. A Figura 2.1, relata os diversos tipos de aeronaves não tripuladas existentes até ao momento.

Outra distinção dos UAVs é o seu modo de propulsão, existem três principais categorias: propulsão de combustão, propulsão elétrica e propulsão híbrida, na qual é efetuada a junção de ambas. De salientar que a utilização de propulsão por combustão é o método mais poluente devido à queima de combustíveis fósseis, sendo que os UAVs cada vez mais estão a evoluir tanto para a propulsão híbrida como elétrica.

Para além da propulsão elétrica ser mais ecológica e apresentar um desenvolvimento mais sustentável, consegue oferecer ao UAV uma maior diversidade de

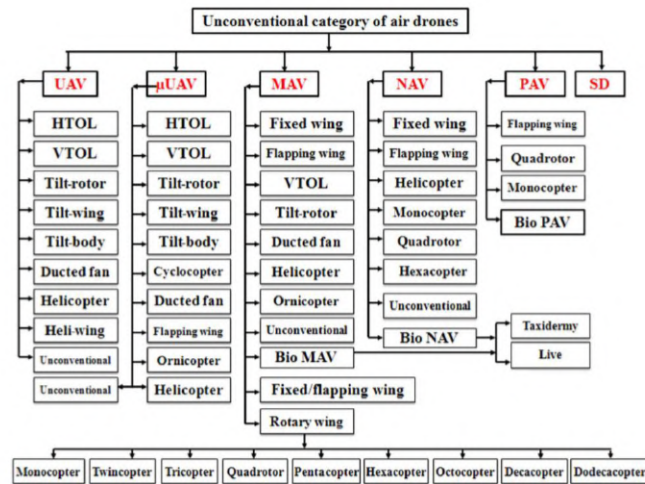


FIGURA 2.1: Tipos de veículos aéreos não tripulados (Hassanalian e Abdelkefi 2017).

aerodinâmica, como o caso dos MAV (*Micro Aerial Vehicle*). No entanto, atualmente, os motores elétricos ainda tem de evoluir bastante a nível de capacidade de potência e peso que conseguem carregar, bem como a própria dissipação de energia e autonomia no ar, em relação aos motores de combustão (B. Zhang, Song e F. Zhao 2022).

Os UAVs também poderão ser classificados de acordo com o seu peso, tendo em conta estudos dos diferentes autores, como pode demonstrar a Tabela 2.1.

TABELA 2.1: Classes de veículos aéreos não tripulados consoante o peso (adaptado de Hassanalian e Abdelkefi 2017).

Autores	Classe	Intervalo de pesos
Brooke-Holland	Classe I (a)	$W \leq 200g$
	Classe I (b)	$200g < W \leq 2Kg$
	Classe I (c)	$2Kg < W \leq 20Kg$
	Classe I (d)	$20Kg < W \leq 150Kg$
	Classe II	$150Kg < W \leq 600Kg$
	Classe III	$W > 600Kg$
Arjomandi, Agostino, Mammone, Nelson e Zhou	Super heavy	$W > 2000Kg$
	Heavy	$200Kg < W \leq 2000Kg$
	Medium	$50Kg < W \leq 200Kg$
	Light	$5Kg < W \leq 50Kg$
	Micro	$W \leq 5Kg$
Weibel e Hansman	Micro	$W < 2lbs$
	Mini	$2lbs < W \leq 30lbs$
	Tactical	$30lbs < W \leq 1000lbs$
	Medium and high altitude	$1000lbs < W \leq 30,000lbs$
	Heavy	$W > 30,000lbs$

## 2.2 **Aterragem de UAVs a bordo de navios**

O grande desafio de ter um UAV a sobrevoar os oceanos ou empenhado em missões dedicadas que necessitem que este aterre num navio, é a fase final a que este é submetido. Enquanto a fase de sobrevoar pode ser previamente atribuída com um plano de trajetória já delineado, a fase de aterragem coloca outros desafios. Na fase de aterragem acrescem os fatores como o balanço, cabeceio, arfagem, guinada, abatimento e avanço do próprio navio, adicionados às condições meteorológicas e oceanográficas que já se faziam sentir quando o UAV estava em fase de voo. Deste modo, a aterragem em navios de UAVs torna-se um desafio perante as inúmeras variáveis que se pretendem controlar.

De acordo com o estudo efetuado por Kevin Williams, os acidentes sofridos por UAVs devem-se a vários fatores, como por exemplo, falhas na manutenção, que ronda entre os 10 % a 18 %. Os acidentes causados pelo envolvimento humano e os seus erros ao operar o próprio UAV abrangem cerca de 48 % a 68 %. Já os acidentes causados pela própria falha da aeronave, ou do sistema da mesma, circunscrevem entre 32 % a 68 % (Williams 2004). Evidencia-se, assim, que a componente humana apresenta um impacto significativo na operação do UAV. Quanto maior for o apoio dado ao piloto, bem como a crescente automação nos sistemas de aterragem, menor será a taxa de acidentes por parte dos UAVs.

A aterragem de um UAV que recorre à visão é dividido em três partes: a primeira que se baseia na deteção do alvo e dos seus pontos de referência, na recolha dos dados onde pretende aterrar, nas condições meteorológicas e oceânicas. A segunda parte é a orientação do voo UAV, é nesta fase que se efetuam as correções, tendo em conta uma posição relativa, tanto da altitude como da posição a que o alvo se encontra. E a terceira, e última fase, é a aterragem autónoma, na qual o UAV tem como função adaptar a sua altitude e velocidade, tendo em conta o tipo de alvo e cenário em que se encontram (Xin, Tang, Gai e H. Liu 2022).

A distinção de método passivo e ativo no modo de aquisição de imagem é crucial. Os métodos passivos são aqueles em que as câmaras responsáveis pela aquisição de imagem estão localizadas a bordo do navio, na superfície terrestre ou em outro local que não esteja no UAV. Já por método ativo entende-se que a aquisição de imagem está a ser executada pelo UAV ou por uma câmara que a ele esteja conectada na mesma infraestrutura.

A aterragem em navios baseada exclusivamente no sinal GPS é considerado um método bastante limitativo devido à sua incapacidade de adquirir com precisão a

altitude do UAV, bem como, o balanço do próprio navio. Neste sentido, é necessário combinar o uso do GPS com pelo menos mais dois instrumentos: o sistema de navegação do navio e um sistema com capacidade de medir a altitude a que se encontra o UAV ao longo da aterragem.

Para garantir uma aterragem com sucesso, é fundamental obter informações precisas sobre a posição GPS, a velocidade e orientação do navio. No entanto, nem sempre esses dados estão prontamente disponíveis. Embora seja possível realizar *visual servoing* a partir de uma câmara instalada no UAV, surgem adversidades, como a perda de visibilidade do navio durante a manobra de aterragem e o balanço do mesmo.

De forma a superar os dois problemas identificados anteriormente é ter câmaras a bordo do navio, alinhadas devidamente com o convés de aterragem. Assim, é possível oferecer uma visão contínua, atualizada e precisa da posição em que o UAV se encontra, uma vez que é uma alternativa mais precisa.

Além disso, depender exclusivamente do sinal GPS poderá não ser desejável, tendo em conta que, em alto mar o sinal GPS pode sofrer perdas, obstruções físicas, interferência eletromagnética, condições atmosféricas adversas e até mesmo problemas no hardware do equipamento ou na receção do sinal podem afetar a fiabilidade do GPS. Deste modo, é necessário considerar alternativas complementares para garantir uma aterragem bem sucedida. A combinação de câmaras a bordo, juntamente com outras tecnologias de estimação de posição, pode oferecer uma abordagem mais robusta e precisa para enfrentar esses desafios.

A aterragem de UAVs em navio de forma autónoma tem vindo a ser estudada ao longo dos anos, como por exemplo, a aterragem de UAVs em navios através do Sistema de Navegação Inercial (INS) ou do Sistema de Navegação Global por Satélite (GNSS), o mais utilizado tem sido a combinação em simultâneo dos dois sistemas INS/GNSS (G. Xu, Y. Zhang, Ji, Cheng e Tian 2009).

O INS (*Inertial Navigation System*) é um sistema de navegação inercial que determina a posição, orientação e velocidade de um objeto, neste caso de um UAV. Este sistema consegue também fornecer dados como o rumo, inclinação e rotação do próprio navio. Normalmente, através de sensores inerciais, mais precisamente três giroscópios e um acelerómetro de três eixos. Estes sensores compõem a Unidade de Medida Universal (IMU), que consegue fornecer dados como o rumo, inclinação e rotação do UAV. Com isso, o INS é capaz de fornecer um total de seis graus de liberdade, abrangendo os movimentos nos três eixos de rotação (balanço, guinada

e cabeceio) e três eixos de translação (aceleração linear em x, y e z), segundo a Figura 2.2 (Christ e Wernli 2014):

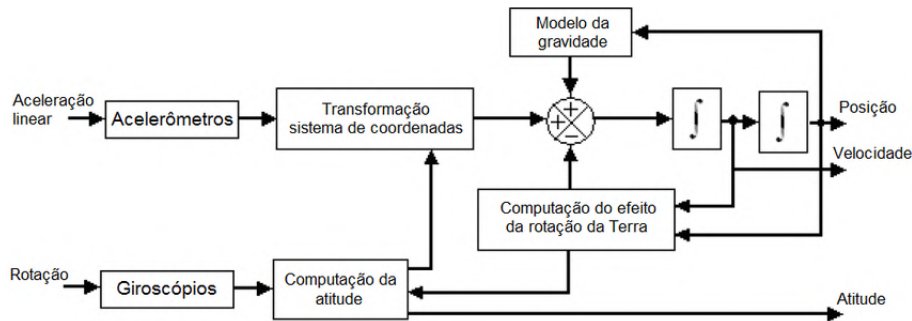


FIGURA 2.2: Sistema de Navegação Inercial (INS) (Silva, A. Morais, J. Morais e Gedraite 2014).

O GNSS é um sistema que consegue aceder a vários satélites simultaneamente, este recebe dados de posição, navegação e sincronização temporal tanto a nível global como de uma região específica, como representa a Figura 2.3. O desempenho do sistema GNSS é assente em quatro tópicos fundamentais: a precisão, a integridade, a continuidade e a disponibilidade (Zhu, Bétaille e Berbineau 2018).

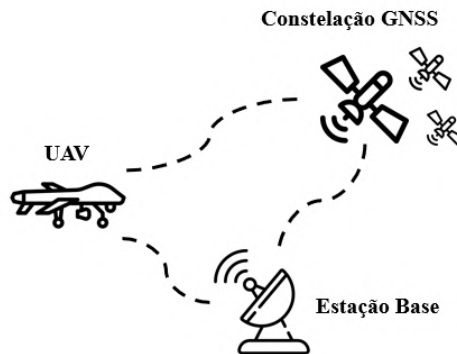


FIGURA 2.3: Aterragem através do sistema GNSS.

O GNSS distingue-se do GPS, na medida em que o GPS é um Sistema de Posição Global, criado pelos norte americanos, enquanto que um sistema GNSS poderá utilizar o GPS assim como outros sistemas de posição como o Galileo (sistema Europeu), Glonass (sistema Russo), BeiDou (sistema Chinês), entre outros, todos em simultâneo.

De acordo com os autores W. Kong et al. (2014) utilizam o método de aterragem baseado na visão, utilizam apenas uma câmara e uma unidade de processamento de imagens. Neste estudo foram utilizados os seguintes sistemas: Sistema

de Navegação, através do qual extraíram os movimentos do navio e a trajetória pretendida pelo mesmo; o Sistema de Orientação que tinha por base planear a missão e as tomadas de decisão, substituindo assim um piloto humano; e por fim, os Sistemas IBVS (*Imaged-Based Visual Servoing*) e PBSV (*Pose-Based Visual Servoing*) que são sistemas de servo monitorização baseadas na imagem e posição, respetivamente (Hassanalian e Abdelkefi 2017).

O *Visual Servoing* é uma abordagem de controlo que recorre a informações visuais de forma a obter o movimento de objetos. Ao contrário dos sensores tradicionais, como *encoders* ou sensores inerciais, o *Visual Servoing* recorre a dados visuais obtidos através de câmaras para estimar, controlar e orientar em relação a um objeto. Esta abordagem requer aquisição de dados visuais contínuos em tempo real, posteriormente os dados serão processados e é extraída a informação desejada, como por exemplo o próprio UAV e pontos de referência (J. Wu, Z. Jin, A. Liu, Yu e F. Yang 2022). Dois exemplos de *Visual Servoing* são os seguintes: o *Image-Based Visual Servoing* (IBVS) e o *Position-Based Visual Servoing* (PBVS).

O Sistema IBVS utiliza imagens bidimensionais (2D) para extrair informações de uma aeronave durante a fase de aterragem. Embora não seja possível através deste sistema adquirir diretamente a profundidade, compensa-se a limitação através de processamento de imagem. O IBVS consegue fornecer dados sobre a trajetória, velocidade e posição da aeronave. Posteriormente, calcula o erro tendo em consideração as informações visuais que obteve do UAV comparando com o local onde idealmente se deveria encontrar (Miranda-Moya, Castañeda, Gordillo e H. Wang 2017), como ilustra a Figura 2.4. O sistema IBVS pode ser integrado com outros sistemas de controlo de tráfego aéreo e deteção de obstáculos. É um sistema indicado para ensaios com diferentes cenários, pois consegue adotar uma forma mais flexível de adquirir dados.

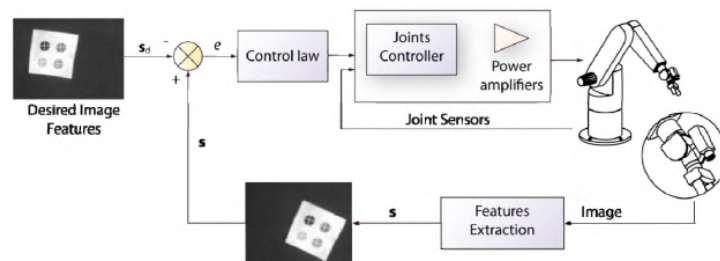


FIGURA 2.4: Image Based Visual Servoing (IBVS) (Garcia, Corrales, Pomares e Torres 2009).

O sistema IBVS também foi utilizado para aterragem de UAVs quer em alvos estáticos quer em alvos móveis, pelos autores G.Cho et al. (2022). Através de

imagens obtidas pela câmara do UAV, método ativo, e com os dados de velocidade e movimentação do próprio navio, combinando todos os dados com o Filtro de Kalman, adicionaram um ganho IBVS de forma a compensar os desequilíbrios. Na última parte do estudo aplicaram o sistema IBVS em situação real, onde desenvolveram um método de aproximação do navio e a aterragem no mesmo, onde obtiveram um erro de 0,2 metros com o navio a navegar a mais de 3 nós (Cho, Choi, Bae e Oh 2022).

O Sistema PBSV consegue através de posição tridimensional (3D) controlar o erro entre a posição onde se encontra o UAV e a posição desejada,, de acordo com a Figura 2.5. Este sistema é valorizado em relação ao anterior, na medida em que consegue ter dados acerca da posição muito mais precisos (Miranda-Moya, Castañeda, Gordillo e H. Wang 2017).

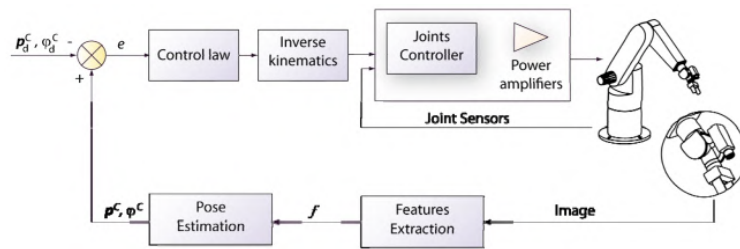


FIGURA 2.5: Position Based Visual Servoing (Garcia, Corrales, Pomares e Torres 2009).

A aterragem em navios pode ser abordada em três vertentes diferentes, ilustradas na Figura 2.6: aterragem numa plataforma estática, que é o caso de um navio fundeado ou atracado; a aterragem numa plataforma em movimento, como por exemplo, um navio com seguimento e a aterragem numa área desconhecida, tal como a aterragem de um UAV noutra navio desconhecendo as suas dimensões (Xin, Tang, Gai e H. Liu 2022).

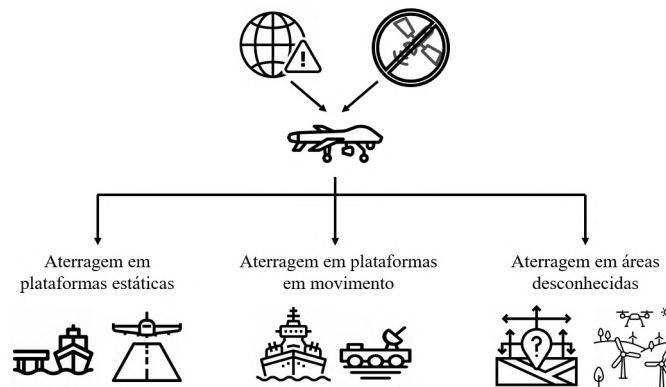


FIGURA 2.6: Cenários possíveis para a aterragem de UAVs.

Em relação à aterragem numa plataforma previamente conhecida e estudada, em casos tanto de plataforma estática como em movimento, temos a pista de aterragem de helicópteros. Estas pistas de aterragem são normalmente circulares, e têm uma letra pintada no chão, como é o exemplo das fragatas da Marinha Portuguesa da Classe Vasco da Gama e Classe Bartolomeu Dias, e também dos navios Patrulhas Oceânicos da Classe Viana do Castelo, como se consegue verificar nas Figuras 2.7 e 2.8, respetivamente.



FIGURA 2.7: NRP Vasco da Gama da Classe Vasco da Gama (Portuguesa 2023b).



FIGURA 2.8: NRP Sines da Classe Viana do Castelo (Portuguesa 2023a).

É também possível a aterragem de UAV numa pista de aterragem a bordo do navio, que seria a forma ideal para UAV de asa fixa. No entanto, é importante salientar que a Marinha Portuguesa não possui na sua esquadra este tipo específico, colmatando esta falha com uma rede de abordagem, na qual o UAV é recolhido na fase de aterragem, evitando que este colida contra o próprio navio, como se pode verificar na Figura 2.10. A Figura 2.9 é um exemplo de uma dessas pista de aterragem, a bordo de um navio da Marinha dos Estados Unidos da América.



FIGURA 2.9: Pista de aterragem de um navio da Marinha dos Estados Unidos da América (Institute 2023).



FIGURA 2.10: Rede de abordagem utilizada numa lancha da Marinha Portuguesa (Nuno Santos 2020b).

No que se refere à aterragem de UAVs em **plataformas estáticas**, esse é o caso, por exemplo, de navios atracados ou fundeados com condições meteorológicas

e oceanográficas bastante favoráveis à sua estabilidade. Existem diversos estudos que abordam este método e técnicas similares.

Exemplo disso, é o modelo estudado por M. Rebelo et al (2021), onde foi proposta uma estrutura *Unmanned Ground Vehicle* (UGV) no qual o UGV iria permitir que o UAV aterrasse ou se aproximasse o suficiente, de modo a conseguir descarregar uma carga. Sabendo inicialmente as coordenadas de cada veículo, do UAV e do UGV, e a linha reta que distância os dois, obteve-se o modelo cinemático tanto do UGV como do próprio UAV, bem como a formação desejável em que estes se devem posicionar para que a trajetória de aproximação seja efetuada com sucesso. O UGV utiliza um controlador em cascata cinemático e um compensador dinâmico enquanto que o UAV utilizada um controlador não linear. Neste estudo foram utilizadas quatro câmaras OptiTrack, que conseguem seguir alvos em movimento simultâneos e transformá-los em cenários 3D em tempo real, tornando este estudo num método passivo. Todas as comunicações do presente estudo, quer entre UAV, UGV, computador e as câmaras, foram efetuadas através de uma rede *Wireless Fidelity* (Wi-Fi) (Rabelo, Brandão e Sarcinelli-Controller 2021).

De salientar, que a cooperação entre o UGV e o UAV, num sistema de resolução OptiTrack, apresenta desafios exigentes, de forma que o torna quase impossível de replicar a bordo de um navio. Sendo de realçar que as condições de aterragem a bordo de navios é substancialmente diferente, especialmente na relação cooperativa entre o UAV e o UGV, não sendo possível aplicar no navio as mesmas ações que se aplicariam num UGV.

Já no estudo de P. Nguyen et al. (2017) é proposto um modelo de aterragem de UAV em qualquer fase do dia sem recorrer ao sinal GPS, evitando um dos problemas de hoje em dia, a perda de sinal. Utilizam uma câmara montada no próprio UAV, tornando este estudo um método ativo, de forma a captar imagens, utilizando um processador de baixa potência de modo a aumentar a velocidade de processamento. Através da informação horária e da luminosidade do ambiente em que está a ocorrer o ensaio, o algoritmo é selecionado para o modo diurno ou noturno, de forma automática consoante ambos os parâmetros. Deste modo, foi utilizado um marcador 2D que se colocou no solo como um alvo, de forma a ser identificado pela câmara presente no UAV, mesmo que não se consiga ver o padrão total, o algoritmo foi ponderado para o reconhecer parcialmente (Nguyen, K. Kim, Y. Lee e Park 2017). O marcador 2D está representado na Figura 2.11.

No que diz respeito aos UAVs que aterram em **plataformas em movimento**, como o caso dos navios que se encontram com seguimento, temos o exemplo

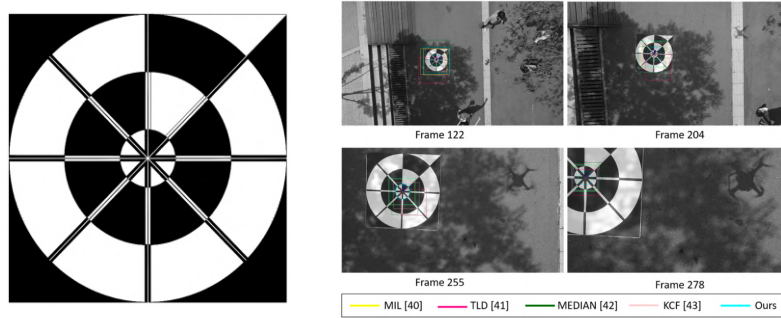


FIGURA 2.11: Marcador 2D (Nguyen, K. Kim, Y. Lee e Park 2017).

do estudo efetuado por Y. Meng et al. (2019). Através do sistema de visão instalado no próprio UAV, tornando-o num método ativo de aquisição de imagem, e dos infravermelhos localizados no convés do navio é possível obter, em tempo real, as coordenadas das imagens captadas. Para além disso, os dados recolhidos são fundidos com os dados obtido pela IMU e altímetro, conseguindo assim estimativas de posição do UAV ainda mais precisas (Meng, W. Wang, H. Han e Ban 2019). Com base no artigo, os autores apresentaram o seguinte fluxograma de acordo com o seu estudo, presente na Figura 2.12.

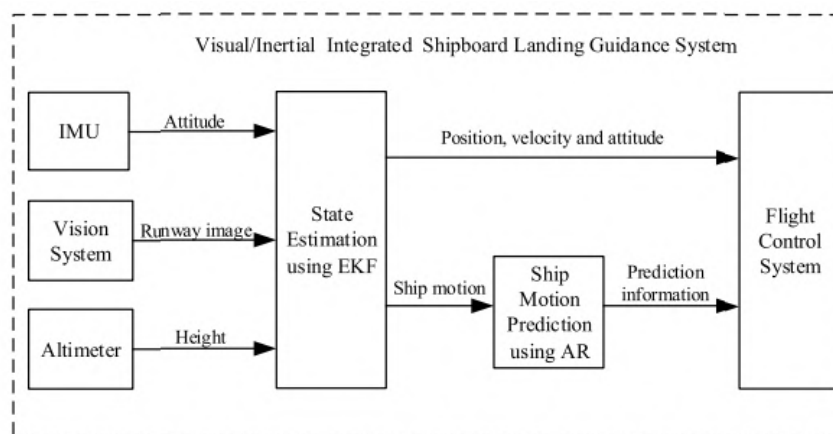


FIGURA 2.12: Fluxograma do sistema integrado de aterragem a bordo (Meng, W. Wang, H. Han e Ban 2019).

O método que os autores utilizaram, previstos para um navio a navegar até aos 20 nós, permite tendo em conta o movimento anterior do navio e o atual, a uma distância de 450 metros fazer ajustes no algoritmo antes do UAV aterrar no convés do navio. Conta com um erro de previsão em termos de altitude de 0,3 metros e 4 segundos antes de aterrar o UAV.

De acordo com o estudo elaborados pelos autores N. Santos et al. (2019), foram desenvolvidos algoritmos tanto pelo método passivo como ativo, empregando

visão computacional. Pelo método ativo, utilizaram a câmara do UAV bem como marcadores na rede de abordagem localizada no navio, através de Filtros de Kalman (KF) e *Efficient Perspective-n-Point* (EPnP) para o cálculo da abordagem do UAV à rede. Já no método passivo foi utilizada apenas uma câmara a bordo do navio, em simulação com um modelo CAD e a trajetória calculada através de uma estação terrestre (GCS) e transmitida posteriormente ao UAV. Neste último caso foi utilizado um Filtro de Partículas para estimar a posição do UAV, um Filtro de *Kalman Unscented* (UKF) de forma a filtrar o movimento e um Filtro de *Bingham Unscented* (UBiF) para filtrar o movimento rotacional (N. Santos, Lobo e Bernardino 2019).

Segundo os autores V. Saj et al. (2022), num estudo que efetuaram em setembro de 2022, através de visão computacional conseguiram criar um algoritmo de aterragem de UAVs de forma autónoma. Utilizaram uma plataforma com um máximo de 6 graus de liberdade, de forma a simular o balanço de um navio, e através da câmara do UAV, recorrendo a um método ativo. Os autores numa fase final compararam os dados com um controlador PID e realizaram vários ensaios em diferentes condições meteorológicas (Saj, B. Lee, Kalathil e Benedict 2022).

Os autores S. Abujoub et al. (2020) propuseram um algoritmo de previsão de sinal (SPA) para a aterragem de UAV em navios em movimento. O SPA é um algoritmo que consegue prever os movimentos que o navio irá realizar, com base na estimativa e previsão. Este algoritmo é limitativo no estudo em questão, devido a necessitar de recolha de dados inicialmente. No estudo deles, para além do algoritmo de previsão, utilizaram um sistema capaz de compensar os movimentos do UAV de forma a aterrar autonomamente. Este sistema de LPI (*Landing Period Indicator*) determina logo quais são os parâmetros estáticos e dinâmicos, e efetua a relação entre os movimentos normais do navio com os movimentos em que se encontram condições seguras para o UAV aterrar (Abdujoub, McPhee e Irani 2020).

Em relação à aterragem em **áreas desconhecidas**, a utilização de alvos visuais artificiais como pontos de referência terrestres, que apresentem características adequadas para uma aterragem do UAV de forma segura, como superfícies planas e estáveis, desempenha um papel fundamental. Neste tipo de aterragem é fundamental a existência de algoritmos complexos e funcionais que consigam associar padrões, plataformas ou paisagens de forma expedita e calcular automaticamente a trajetória de aterragem. Neste tipo de aterragem também engloba a eventual perda de sinal GPS. Normalmente, é necessário recorrer a estas aterragens quando existem situações de emergência, ou o próprio UAV se depare com problemas que necessitem que aterre, antes da sua perda ou colisão com a superfície terrestre.

O estudo proposto por S. Alam e J. Oluoch (2022) sustenta-se na revisão de todos os métodos possíveis de aterragem. Salientam a criação de algoritmos que processam e detetam SLAD (*Safe Landing Areas Detection*), de maneira a identificar o local mais seguro através da visão. As áreas de aterragem desconhecidas poderão ser também estáticas ou dinâmicas (Alam e Oluoch 2021).

O sistema TLS (*Terrestrial Laser Scanning*), resumidamente, efetua o varrimento terrestre através de um laser. Os dados da superfície em que se pretende aterrar, podem ser recolhidos pelo próprio UAV (Wojcik, Klapa, Mitka e Piech 2019), de acordo com a Figura 2.13. A grande vantagem neste caso seria a sua empregabilidade para a aterragem em superfícies desconhecidas.

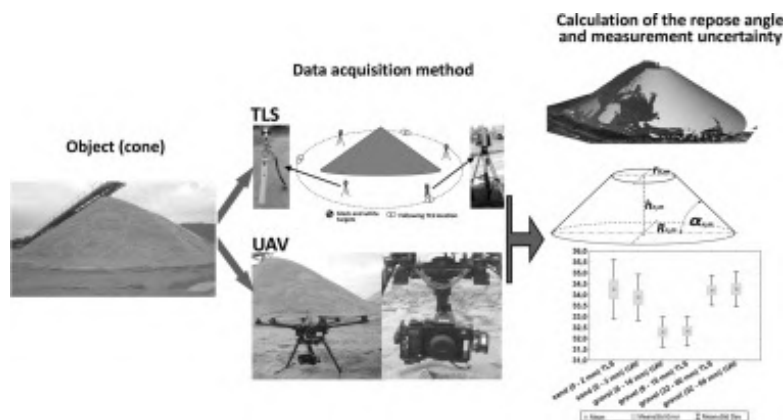


FIGURA 2.13: Aterragem através do sistema TLS (Wojcik, Klapa, Mitka e Piech 2019).

Em 2018, os autores T. Yang et al. (2018) sugeriram um sistema de aterragem autónoma, baseado na visão computacional, de UAVs em situações de emergência e em ambientes desconhecidos. Neste estudo foi utilizado um algoritmo de localização e um mapeamento em simultâneo (SLAM), assim sendo o UAV consegue construir um terreno tridimensional. Porém antes da aterragem o UAV, através da câmara que tem a bordo, método ativo, examina e constrói um cenário que prevê que seja seguro para se aproximar. Numa fase seguinte, o algoritmo calcula uma trajetória de forma a aterrar autonomamente na área que considera segura. Numa fase de aproximação o UAV irá receber dados de ajuste tanto da posição como velocidade (T. Yang, P. Li, H. Zhang, J. Li e Z. Li 2018).

No estudo elaborado por F. Liu et al. (2022) sugeriram um sistema de identificação de áreas desconhecidas para a aterragem de UAVs. Neste caso, foi utilizado a câmara do UAV, método ativo, através da qual se retirou um conjunto de dados visuais, que posteriormente, deram origem à criação de uma rede de aprendizagem profunda de reconhecimento e busca de áreas seguras. Em simultâneo, foi utilizado

o sistema LIDAR 3D, que gera um mapa 3D do ambiente que o rodeia e calcula a área que considera ser mais segura para a aproximação (F. Liu, Shan, Xiong e Fang 2022).

Os autores M. Ikura et al. (2020) propuseram também um sistema similar ao do estudo anterior. No entanto, estes recorreram ao uso de uma luz laser para fazer o varrimento das áreas abrangentes, simultaneamente com a imagem captada pela câmara do UAV. Após terem adquirido um panorama geral do terreno, procederam à procura de uma região de interesse (ROI) para a aterragem segura do UAV (Ikura, Miyashita e Ishikawa 2020).

Todos os estudos anteriormente referidos, utilizaram apenas uma câmara para a aquisição de imagem, sendo que os projetos a seguir apresentados recorreram a duas câmaras. Ao utilizar duas câmaras ao invés de uma, é possível com o cruzamento dos dados obtidos por ambas, calcular a profundidade a que se encontra o UAV. De uma forma geral, com duas imagens 2D conseguimos identificar a posição 3D do objeto em questão a ser estudado, como ilustrada na Figura 2.14.

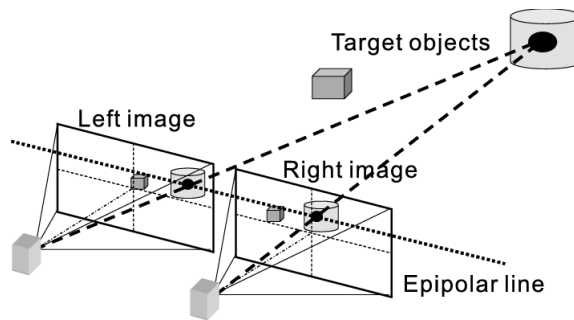


FIGURA 2.14: Visão binocular *stereo* (Hariyama, Yokoyama e Kamayama 2008).

Uma câmara convencional RGB não é capaz de fornecer diretamente informações sobre a profundidade de um objeto, e conseqüentemente, não é capaz de conceber informações sobre a posição 3D do objeto em relação a um referencial externo fixo. No entanto, as câmaras RGB-D, embora sejam mais dispendiosas, permitem obter essa informação com recurso ao *Time of Fly* (ToF). Por outro lado, a utilização de duas câmaras RGB também pode disponibilizar informações de profundidade, através da técnica de visão binocular *stereo* apresentada na figura 2.15.

O princípio ToF é uma técnica que calcula a distância entre um sensor e um objeto. A distância é medida tendo em conta o tempo que leva um sinal a ser

emitido, refletido pelo objeto e retornar ao sensor. O sensor é capaz de fornecer imagens 3D através de todos os sinais recebidos (Foix, Alenya e Torras 2014).

No estudo em que os autores X. Pan et al. (2008) estiveram envolvidos, foi proposto um método de aquisição de imagens combinado, com base em duas câmaras monoculares, alinhadas no solo, que recolhiam imagens em simultâneo. Deste modo, é possível calcular a visão *stereo* da posição do UAV, ou seja, através de duas imagens 2D captadas no mesmo momento, e sabendo a posição de cada uma, conseguimos obter a posição 3D, através da profundidade que é posteriormente calculada (Pan, D. Ma, L. Jin e Z. Jiang 2008).

Outra abordagem, da visão *stereo*, será o próprio UAV conseguir sobrevoar com duas câmaras a bordo, ou uma câmara binocular, ilustrada na Figura 2.15. Um exemplo deste caso, é o estudo efetuado pelos autores Y. Ma et al. (2022), no qual foi proposto um algoritmo de deteção e controlo, este baseia-se nos mesmos módulos do artigo anterior (Dolgoplov, Kazantsev e Bezuhliy 2016).

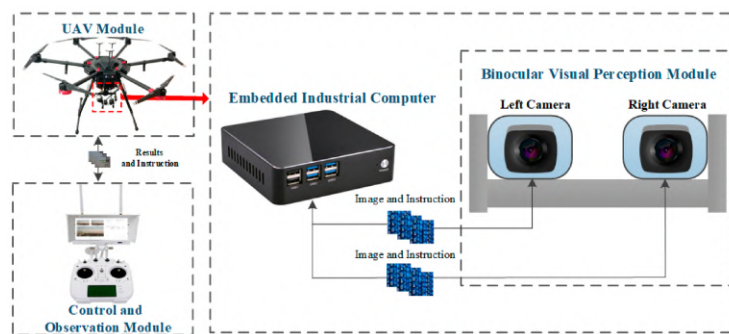


FIGURA 2.15: UAV com uma câmara binocular (Y. Ma, Q. Li, Chu, Y. Zhou e C. Xu 2021).

## 2.3 Deteção e seguimento de aeronaves e UAVs

O desenvolvimento de novas áreas de investigação de inteligência artificial tem tido um crescimento ao longo dos últimos anos, desta forma, surge a também a necessidade de expandir o setor das aeronaves e UAVs. Neste Secção pretende-se desenvolver os setores de deteção e seguimento de aeronaves e UAVs.

### 2.3.1 Deteção de aeronaves e UAVs

A deteção de aeronaves e UAVs consiste em determinar a sua localização e categorizar cada objeto encontrado. Usualmente, em trabalhos de deteção de objetos, foca-se em três fases distintas: a seleção da área de interesse, a extração

das características e procede-se para a classificação do objeto identificado (Z. Zhao, Zheng, S. Xu e X. Wu 2019). Normalmente, os objetos detetados são representados por uma *bounding box*, que consiste num retângulo que envolve um certo objeto específico numa imagem, no presente caso, o UAV.

O progresso das redes neuronais profundas, durante a última década, tem sido fundamental para viabilizar o desenvolvimento de redes de deteção de objetos com um desempenho mais preciso e rápido em comparação com métodos anteriormente existentes.

Uma Rede Neuronal Artificial (RNA) ou *Neural Network* (NN) sustenta-se no funcionamento do sistema nervoso de um ser vivo, como por exemplo o cérebro humano. Este consiste no funcionamento de várias unidades interligados, formando os neurónios, este processo é capaz de reconhecer padrões. As RNA poderão ter o formato da Figura 2.16:

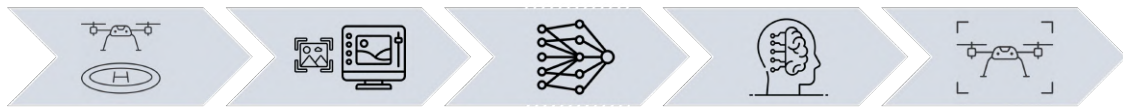


FIGURA 2.16: Deteção de UAV através de Rede Neuronal Artificial.

A utilização de redes neuronais como um sistema de aprendizagem máquina pode ser utilizado para a localização e deteção de objetos, como por exemplo a deteção de um UAV. Geralmente, subdivide-se em três etapas, a primeira é o pré-processamento de dados em que as imagens captadas são todas padronizadas, de forma a ter um conjunto de dados homogéneos na entrada da rede. De seguida, as imagens são analisadas para extrair dados relevantes que auxiliam a identificar a existência de um UAV na imagem, dados como: cores, formas, padrões visuais distintos, entre outros. São atribuídos pesos aos diferentes parâmetros consoante a importância que se pretenda dar a cada um, desta forma a rede neuronal aprende a associar os parâmetros específicos de um UAV, e tem a capacidade de prever em novos *inputs* se existe ou não um UAV.

*Deep Learning* são redes neuronais mais profundas, estas conseguem adquirir, aprender, prever, associar padrões, realizar tarefas e cálculos a uma velocidade muito superior à capacidade de um ser humano (B. Zhang, Song e F. Zhao 2022).

A aterragem de UAVs em navios pode apresentar duas abordagens diferentes no que se refere à deteção de objetos. Na primeira abordagem, o próprio UAV é

equipado com uma câmara que obtém dados do navio durante a sua aproximação. Na segunda abordagem, as câmaras são instaladas a bordo do navio e são responsáveis pela deteção da posição do UAV durante a sua aproximação.

A deteção do UAV através de câmaras a bordo do navio, permite estabelecer uma referência fixa no convés do navio, considerando os movimentos que o navio executa. Com base nessa referência, é possível calcular a distância, velocidade e posição do UAV em relação ao convés.

A deteção do navio por parte do UAV, requer que este acompanhe continuamente o navio. Este método apresenta uma maior complexidade, em questões como perdas de visão do navio durante a manobra de aterragem, e o movimento do próprio navio, que através de uma câmara a bordo do UAV não é tão perceptível.

Começando pela **deteção do UAV a partir de bordo de um navio**, em 2019, os autores A. Barisic et al. (2019) apresentaram uma proposta para a deteção de UAVs através de visão computacional. Neste estudo, adotaram o uso de uma rede neuronal convolucional YOLO (*You Only Look Once*), na qual aplicavam um problema de regressão e criavam em torno do UAV detetado em cada imagem captada uma *bounding box* (Barisic, Car e Bogdan 2019).

Em 2020, o autor Nuno Santos (2020), utilizou um método de deteção de ROI (*Region Of Interest*) nas imagens captadas onde poderá estar a sobrevoar o UAV. Para este efeito, utilizou rede neuronal YOLOv3 e SSD, comparando ambas, na deteção do UAV reduzindo a amostragem das imagens captadas idealmente utilizado para UAVs de pequenas dimensões (Nuno Santos 2020a).

A deteção de um UAV proposta pelos autores W. Nie et al. (2021), foi baseada no sinal de *Wi-Fi* e simultâneo com deteção de sinal por rádio-frequência (Nie, Z. Han, M. Zhou, L. Xie e Q. Jiang 2021). Inicialmente, captam o sinal utilizando USRP (*Universal Software Radio Peripheral*), os dados são processados por uma análise multi-resolução, que irão ajudar a prever os próximos dados de entrada. A deteção do UAV, de acordo com o artigo, é realizada através de um pré-processamento do sinal, utilizando transformadas para extrair informações de tempo e frequência. Consecutivamente, efetua-se a seleção de um limiar de deteção para distinguir as componentes de um sinal, e calcula-se o desvio padrão para medir a variação entre o sinal pré-processado. Caso o desvio padrão exceda um determinado valor, indica que se está na presença de um UAV, efetuado através da análise do sinal *Wi-Fi*.

A **detecção do navio a partir do UAV**, é proposta pelos autores A. Dolgoplov et al. (2016), criando um algoritmo com base em dois controladores cascata, um dedicado à forma do navio e outro para as características extraídas automáticas. Ao processarem as imagens RGB extraídas pela câmara do UAV, calcula-se o mapa Gaussiano das mesmas. Posteriormente, realiza-se a convolução de cada canal separadamente, resultando em mapas de resposta, somando-os obtém-se a a segmentação aproximada do navio que o UAV está a visualizar. Por fim, agrupam-se de acordo com um *cluster* hierárquico as elipses que contornam os navios detetados (Dolgoplov, Kazantsev e Bezuhliy 2016).

No estudo dos autores C. Pires et al. (2022), foi efetuada através de um modelo em cascata que se dividia em duas etapas: a detecção do navio e a segmentação. Durante a detecção identificaram com *bounding box* áreas de possível existência de navios, em simultâneo a segmentação efetuava *segmentation mask*, concentrando o processamento apenas numa região (Pires, Damas e Bernardino 2022). Os modelos foram treinados com imagens aéreas obtidas por um UAV, com detecção e segmentação em tempo real.

À semelhança do artigo anterior, os autores M. Ribeiro et al. (2022), apresentaram um método de detecção de navios num cenário de vigilância marítima. Através de um conjunto de dados sintéticos, efetuou-se o treino de redes de detecção e segmentação através de um modelo em cascata (Ribeiro, Damas e Bernardino 2022).

Os autores X. Zhang et al. (2021) propuseram uma detecção, neste caso de navios, através de imagens captadas pelo UAV, na qual utilizaram uma rede neuronal profunda DT-YOLO. Esta rede neuronal é a melhoria proposta pelos autores em relação à RNA de YOLOv3 (X. Zhang, Shi, Z. Wu e J. Liu 2020).

#### 2.3.2 Seguimento de aeronaves e UAVs

O seguimento de aeronaves e UAVS pode ser dividido em duas partes: o seguimento do UAV a partir de uma superfície estável ou dinâmica; e o seguimento de um navio por parte do próprio UAV.

O **seguimento do navio a partir do UAV**, caso do estudo realizado pelos autores T. Trong et al. (2021), em que foi proposto um método de seguimento de UAVs utilizando o algoritmo Deep SORT. Este algoritmo estima os trajetos futuros que o UAV irá desempenhar, tendo em conta os Filtros de Kalman (Trong, Hai, Van, Thai, Chi e Quang 2021).

Os autores Y. Li et al. (2022), apresentaram um método de seguimento, comparado com o anterior, baseado em Deep SORT, mas este foi combinado com a rede neuronal YOLOv5 (Y. Li, Yuan, Y. Wang e B. Zhang 2022).

O **seguimento do UAV a partir de bordo de um navio** pode ser feito recorrendo a um modelo de estima de pose baseado no filtro de partículas RBPF (*Rao-Blackwellized Particle Filter*). Para o seguimento do UAV, utilizou uma rede neuronal baseada em YOLO, a qual tinha como função aplicar um *autoencoder* às imagens captadas em tempo real de entrada, e compará-las com as características extraídas, adicionalmente, calculou também os pesos das partículas (Ferreira 2021).

Os autores N. Santos et al. (2020), propuseram um método de seguimento de UAVs baseado numa UPF (*Unscented Particle Filter*). O filtro, inicialmente, entra na fase de deteção, gerando hipóteses de localizações do UAV, caso este não seja identificado naquele *frame*, a localização atribuída é a sua última posição. Este artigo foca-se especialmente no filtro de movimento rotacional (N. Santos, Lobo e Bernardino 2020).

## 2.4 Datasets

Um *dataset* é um conjunto de dados que contém registos de uma certa situação ou ambiente (Documentation 2010). No âmbito desta dissertação, os *dataset* adquiridos são guardados sequencialmente, de forma a obter o movimento que o UAV efetua durante a sua aterragem.

Os *datasets* desempenham um papel fundamental no treino e validação de algoritmos em diversas áreas, como na visão computacional e *machine learning*. Os conjuntos de dados representam as amostras representativas de um estudo em questão, permitindo que os algoritmos sejam treinados de forma a reconhecer padrões e tomar decisões com base nesses padrões. Possibilitam também a avaliação, eficácia e robustez de algoritmos tendo em conta o seu desempenho, em cenários e condições variados. Adicionalmente, os *datasets* permitem identificar e mitigar possíveis desequilíbrios nos algoritmos, quando existem disparidades entre resultados obtidos.

Existem diversos estudos sobre a aterragem de UAVs, porém muitos são baseados em conjuntos de dados simulados, não recorrendo a dados reais. Vários autores optam por simulações computacionais como fonte de dados para os seus estudos, por exemplo, o estudo efetuado pelos autores W. Li et al. (2022). Os autores propuseram uma simulação de aterragem de um UAV num USV. A simulação

efetuada no Ubuntu, onde definiram que a ondulação não deveria exceder os 10 graus de oscilação do navio (W. Li, Ge e G. Ye 2022), como ilustrada na Figura 2.17.

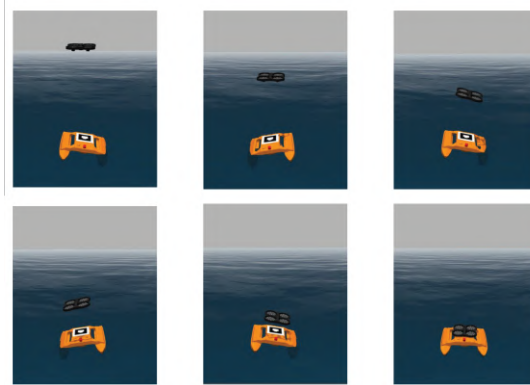


FIGURA 2.17: *Dataset* simulado no Ubuntu (W. Li, Ge e G. Ye 2022).

Nos estudos seguintes, os *datasets* foram obtidos através de voos reais por parte do UAV, em que este efetuou aproximações ao navio.

Em 2021, os autores Y. Ma et al. (2021), criaram um *dataset* com base numa câmara binocular, cerca de 800 imagens no total onde obtiveram um conjunto de dados da câmara da direita em simultâneo com um conjunto de dados da câmara da esquerda, da mesma forma que se pretende obter nesta dissertação. No estudo foram obtidos *datasets* durante várias alturas do dia, em que mudam os parâmetros adquiridos como a visibilidade e condições meteorológicas, bem como as condições de voo do próprio UAV (Y. Ma, Q. Li, Chu, Y. Zhou e C. Xu 2021). Uma das vantagens deste estudo, é a capacidade de detetar mais do que um objeto em cada imagem captada, como é observado na Figura 2.18.

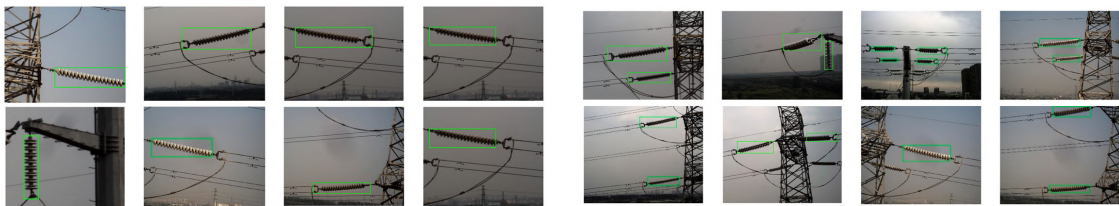


FIGURA 2.18: *Dataset* de deteção de um objeto (esquerda) e de vários objetos simultâneos (direita) (Y. Ma, Q. Li, Chu, Y. Zhou e C. Xu 2021).

No estudo dos autores X. Zhang et al. (2021) foi proposto um *dataset* visual-inercial adquirido de um ambiente real. Ao utilizarem uma câmara, que recolhe vídeos, e uma IMU, que recolhe dados do INS, formam um *dataset* com dados visuais e inerciais de aterragem do UAV. Neste caso, a velocidade de aquisição de

imagens foi de 17 fps (X. Zhang, He, Z. Ma, Jun e K. Yang 2021). O *dataset* neste estudo conta com um total de 2 400 imagens e dados do IMU em simultâneo, como se observa na Figura 2.19.

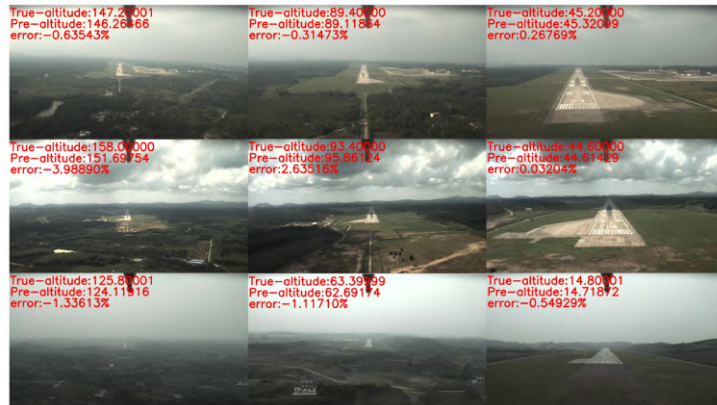


FIGURA 2.19: *Dataset* real de aterragem de UAV com estimativa da altitude (X. Zhang, He, Z. Ma, Jun e K. Yang 2021).

Já o estudo elaborado pelos autores V. Saj et al. (2022), utilizou inicialmente simulação para prever a aterragem do UAV, e só posteriormente, é que recorreu a um *dataset* real. Neste estudo, como a aquisição de imagens foi num edifício coberto, os autores instalaram uma ventoinha de forma a simular vento, ilustrada na Figura 2.20. A aquisição de imagens efetuada pelo UAV em simultâneo com a telemetria do mesmo, geraram um *dataset* que simula de forma aproximadamente real a abordagem do UAV a um navio (Saj, B. Lee, Kalathil e Benedict 2022).



FIGURA 2.20: *Dataset* real de aterragem de UAV numa plataforma de simulação do balanço do navio (Saj, B. Lee, Kalathil e Benedict 2022).

Em 2020, o autor Nuno Santos, propôs um *dataset* real recolhido a bordo de um navio, com uma câmara no convés que conseguia visualizar o UAV a aterrar, que tinha capacidade de extrair imagens com a velocidade de 30 fps. Anteriormente, através de simulação computacional criou um *dataset* com cerca de 335 mil imagens,

e aproximadamente um quarto do conjunto total foi utilizado para validação do estudo (Nuno Santos 2020a), ilustradas na Figura 2.21.



FIGURA 2.21: *Dataset* real de aterragem de UAV em aproximação a um navio (Nuno Santos 2020a).



# Capítulo 3

## Metodologia

Durante este Capítulo, apresenta-se a metodologia para a detecção em tempo real de UAVs numa pista a bordo de um navio. Utilizando duas câmaras estrategicamente posicionadas, é possível efetuar a detecção, em tempo real, durante a fase de aproximação e aterragem no convés. A presente abordagem permite a visão contínua e precisa do UAV em movimento, aferindo a monitorização e estimação da sua posição e velocidade, tendo em consideração um referencial fixo no navio e a localização de cada câmara e marcador H em relação ao referencial definido.

A Secção 3.1 desenvolve a implementação das câmaras a bordo do navio. A Secção 3.2 aborda a calibração das câmaras. A Secção 3.3 foca a detecção do UAV. E por fim, a Secção 3.4 analisa diferentes abordagens para a estimação da posição e velocidade de um UAV, como a geometria epipolar e o filtro de Kalman estendido.

### **3.1 Implementação das câmaras a bordo: disposição, geometria, referenciais e parâmetros**

A implementação de câmaras a bordo é uma prática para a obtenção de informações visuais em tempo real do UAV na fase de aterragem, bem como, pode ser implementada como apoio na tomada de decisões.

De forma a se conseguir obter a localização do UAV em tempo real é necessário considerar a disposição das câmaras a bordo. Outro fator é a geometria envolvida no processo bem como os referenciais definidos. Por fim, tem de se considerar os parâmetros de cada câmara de forma a poder estimar posteriormente a localização do UAV a partir da sua detecção nas imagens das câmaras.

### 3.1.1 Disposição das câmaras

As câmaras dispostas a bordo de um dos navios da Marinha Portuguesa terão a função de captar imagens, em tempo real, do UAV em aproximação. As imagens captadas permitirão determinar a posição e velocidade do UAV durante a fase de aproximação e, consecutivamente, de aterragem.

Considerando os navios existentes da Marinha Portuguesa, nos quais a presente dissertação possa ser aplicada, incluem-se as lanchas, os patrulhas, os patrulhas oceânicas, as corvetas e as fragatas. A boca máxima das lanchas é de 5,9 metros, nos patrulhas é de 7,7 a 9,0 metros, nos patrulhas oceânicas é de 12,9 metros, nas corvetas é de 10,3 metros e nas fragatas é de 14,2 metros.

De acordo com as especificações anteriores, foi proposto as câmaras terem uma distância lateral entre 6 metros e 14 metros, tendo em consideração a distância máxima lateral possível de implementar a bordo. As câmaras serão dispostas no convés do navio, no mesmo plano horizontal, e devidamente alinhadas consoante a Figura 3.1. É fundamental também ambas as câmaras capturem o marcador H que se localizará no chão do convés, o que permitirá obter os parâmetros extrínsecos de ambas as câmaras e definir um referencial tridimensional.

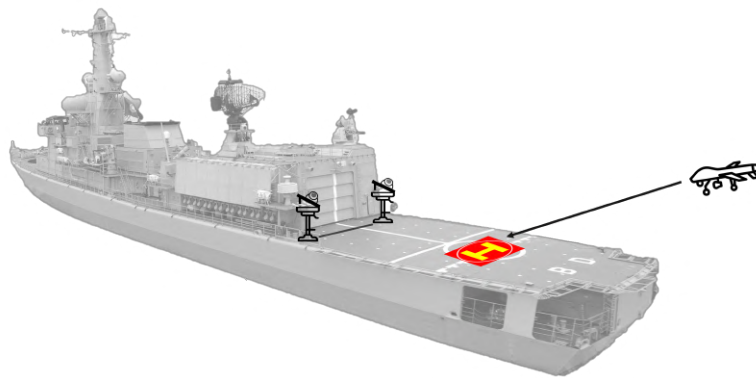


FIGURA 3.1: Disposição das câmaras a bordo de um navio.

Ambas as câmaras são dispostas numa estrutura, que tem a capacidade de efetuar o *pan* dos  $0^\circ$  aos  $360^\circ$  e *tilt* dos  $0^\circ$  aos  $90^\circ$ , de forma manual. O proposto é definir no início de cada ensaio o *pan* e *tilt* adequados das estruturas, de forma a captar o melhor ângulo para a aquisição de imagens.

Estabelece-se a relação de que quanto maior for a distância focal num sistema de duas câmaras, maior será o seu zoom ótico. Resulta numa ampliação com melhor resolução e a capacidade de detetar objetos com maior detalhe e a uma distância superior, como ilustrado na Figura 3.2.

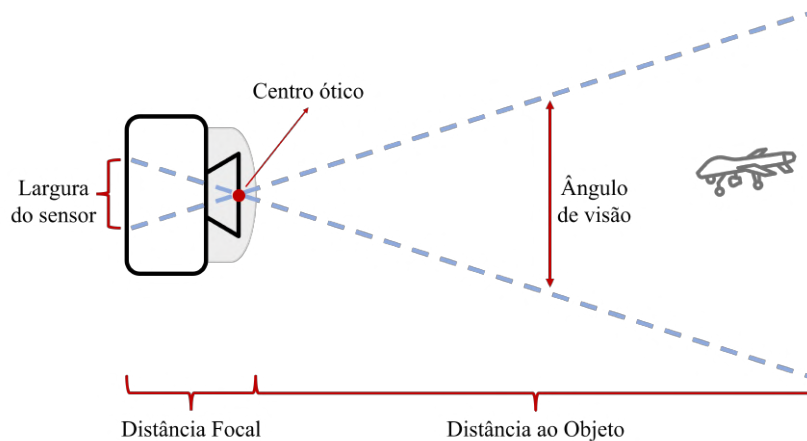


FIGURA 3.2: Visão de objetos através de uma câmara.

O aumento da distância focal implica uma ampliação das imagens observadas por cada câmara individualmente. De forma a manter a nitidez e o foco adequado, quando se executa captura de imagens para os ensaios, é fundamental afastar o objeto da câmara que o observa. Para obter o ajuste preciso entre a câmara e o objeto, é necessário estabelecer o equilíbrio entre a distância do UAV e ambas as câmaras, que ocorre quanto maior for a distância focal. A perspectiva entre as câmaras irá ser mais pronunciada quando o UAV se encontrar mais próximo, pois a diferença será maior entre imagens. Desta forma, quanto maior for a distância focal menor será o ângulo de visão e vice-versa.

De acordo com a visão estereoscópica, quanto maior for a distância entre a posição em que cada câmara se localiza, tendo em conta uma linha base na qual ambas estão alinhadas, maior e mais precisa será a percepção de profundidade do UAV, bem como a sensação tridimensional de que o UAV se encontra mais próximo do navio. No entanto, a distância entre as duas câmaras fica limitada à largura máxima do convés em que as mesmas serão instaladas.

### 3.1.2 Geometria e Referenciais: Formação das Imagens

As coordenadas cartesianas são utilizadas para se conseguir descrever um ponto num plano ou espaço, como por exemplo localizar o UAV numa imagem capturada pela câmara ou o UAV no espaço, respetivamente. No plano cartesiano os pontos podem ser localizados tendo em conta os eixos  $x$  e  $y$ , caso seja um plano bidimensional, acrescentando o eixo  $z$  para a representação tridimensional.

A transformação de coordenadas cartesianas para coordenadas homogêneas, permite uma maior flexibilidade na sua utilização em transformações geométricas

complexas, como por exemplo rotações, translações e projeções. A representação em coordenadas homogêneas, à semelhanças das coordenadas cartesianas, tem três variáveis que correspondem à posição vertical, horizontal e de profundidade no espaço. Ao introduzir uma variável adicional neste contexto, denominada fator de escala, é possível representar as propriedades intrínsecas das transformações e projeções. Geralmente, é atribuído ao fator de escala o valor de uma unidade.

Deste modo os pontos podem ser transformados de coordenadas cartesianas para coordenadas homogêneas, de acordo com a Equação 3.1:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \rightarrow \begin{bmatrix} \lambda x \\ \lambda y \\ \lambda z \\ \lambda \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}, \text{ com } \lambda \neq 0 \quad (3.1)$$

De salientar que, pontos em coordenadas cartesianas pertencem a  $R^N$ , já em coordenadas homogêneas pertencem a  $R^{N+1}$ . Em coordenadas homogêneas, os vetores iguais à parte de um fator de escala representam o mesmo ponto em coordenadas cartesianas. Para converter coordenadas homogêneas para cartesianas, é necessário efetuar a normalização dos vetores, de forma a que o último componente iguale o valor unitário.

Os referenciais em coordenadas homogêneas estabelecidos, na presente dissertação, visam a estabelecer correlações espaciais entre pontos em diferentes contextos, assim criaram-se três referenciais distintos. O referencial do mundo (W), que representa pontos tridimensionais tendo em consideração o espaço. O sistema de referências da câmara (C) é estabelecido de forma a abranger os pontos tridimensionais que estão intrinsecamente ligados à posição da câmara. O referencial da imagem (I) é definido para representar os pontos bidimensionais capturados por cada câmara. Os sistemas referenciais podem se observar através da Figura 3.3.

As coordenadas homogêneas do referencial W, C e I representam-se, respectivamente, de acordo com a Equação 3.2:

$$\bar{P}_W = \begin{bmatrix} x_W \\ y_W \\ z_W \\ 1 \end{bmatrix}, \bar{P}_C = \begin{bmatrix} x_C \\ y_C \\ z_C \\ 1 \end{bmatrix} \text{ e } \bar{P}_I = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad (3.2)$$

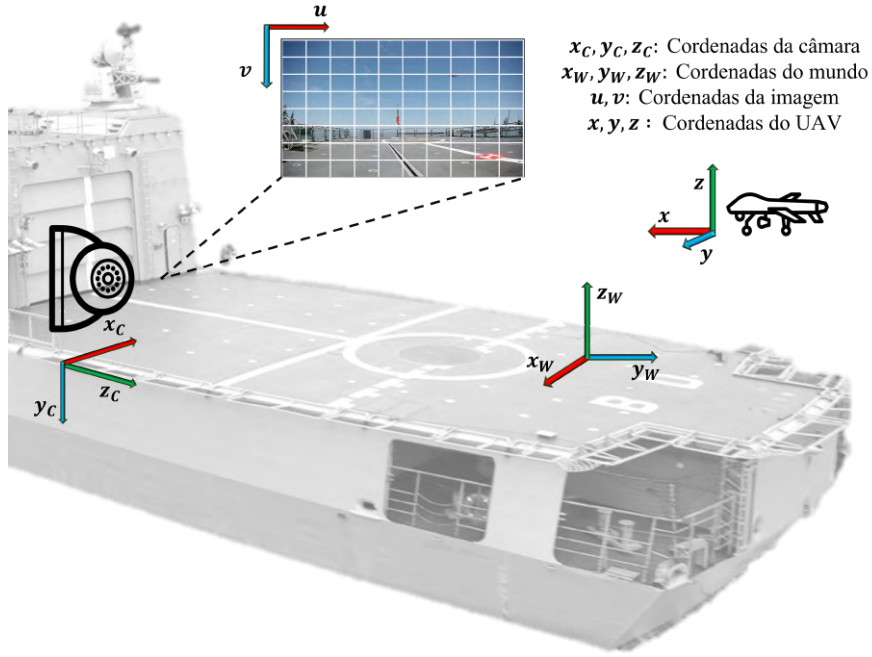


FIGURA 3.3: Referenciais estabelecidos: referencial do mundo (W), da câmera (C), da imagem (I) e do UAV.

Em coordenadas cartesianas, com a devida transformação caso necessitem de ser normalizados, tem-se os referenciais do mundo (W), da câmera (C) e da imagem (I), respetivamente, como demonstra a Equação 3.3:

$$P_W = \begin{bmatrix} x_W \\ y_W \\ z_W \end{bmatrix}, P_C = \begin{bmatrix} x_C \\ y_C \\ z_C \end{bmatrix} \text{ e } P_I = \begin{bmatrix} u \\ v \end{bmatrix} \quad (3.3)$$

### 3.1.3 Parâmetros extrínsecos e intrínsecos das câmaras

Os **parâmetros extrínsecos** de uma câmara referem-se às variáveis que descrevem a orientação (rotação e translação) tridimensional da câmara em relação a um sistema de coordenadas de referência. Os parâmetros extrínsecos permitem a reconstrução geométrica tridimensional do cenário captado através de imagens bidimensionais, bem como a calibração espacial e a análise de movimentos através de visão computacional, de acordo com a Figura 3.3.

Através da equação  $\bar{P}_c = C_{T_W} \bar{P}_W$  consegue-se obter a matriz homogénea dos parâmetros extrínsecos, que é dada pela Equação 3.4:

$$C_{\bar{T}_W} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & R_{3 \times 3} & \cdot & T_{3 \times 1} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.4)$$

A matriz homogénea das coordenadas da câmara é um elemento da representação geométrica para o sistema de visão computacional da câmara, e tendo em consideração as coordenadas definidas anteriormente e a ilustração da Figura 3.2, deduz se a Equação 3.5:

$$\bar{P}_C = C_{\bar{T}_W} \bar{P}_W \quad (3.5)$$

A partir da relação anterior entre coordenadas homogéneas nos referenciais do mundo e da câmara, de acordo com a Figura 3.3, é possível obter a relação correspondente seguinte entre coordenadas cartesianas através da Equação 3.6:

$$P_C = R_{3 \times 3} P_W + T_{3 \times 1} \quad (3.6)$$

A formação da imagem bidimensional, é composta pelos eixos  $u$  e  $v$ , como retratados na figura 3.2. Tendo em conta que existem duas câmaras a capturar duas imagens diferentes, irão existir dois valores distintos para  $u$  e dois valores distintos para  $v$ . Considerando um ponto aleatório numa imagem captura obtém-se a Equação 3.7, em que  $f$  representa a distância focal,  $u_o$  e  $v_o$  os centro óticos das imagens, respetivamente.

$$\bar{P}_I = \begin{bmatrix} f & \gamma & u_o & 0 \\ 0 & f & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \bar{P}_C = \begin{bmatrix} fx_C + \gamma y_C + u_o z_C \\ fy_C + v_o z_C \\ z_C \end{bmatrix} \quad (3.7)$$

A equação seguinte descreve a relação entre  $P_I$  e  $P_C$ , que depende dos parâmetros intrínsecos das câmaras. Considera-se  $\gamma = 0$  nos cálculos da Equação 3.8, devido aos valores obtidos na Secção 5.1.

$$\begin{aligned} u &= f \frac{x_C}{z_C} + \gamma \frac{y_C}{z_C} + u_o \\ v &= f \frac{y_C}{z_C} + v_o \end{aligned} \quad (3.8)$$

Os **parâmetros intrínsecos** de uma câmara englobam conceitos como o comprimento focal e a posição ao centro ótico. O comprimento focal indica a convergência dos raios de luz incidentes na câmara, o centro ótico expressa a posição ótica central em relação ao plano da imagem e os coeficientes de distorção radial identificam as distorções geométricas da lente.

Devido à distorção das imagens captadas por ambas as câmaras, para além dos parâmetros intrínsecos e extrínsecos a serem considerados, surgem os coeficientes de distorção. Os coeficientes de distorção medem o grau de distorção nas imagens obtidas, tendo em conta o modelo e as lentes das mesma.

As distorções causadas pelas lentes podem ser: *barrel* quando corresponde a uma distorção radial negativa ou *pincushion* quando ocorre uma distorção radial positiva. A distorção tangencial é uma distorção ótica, tal como a radial, porém esta afeta as imagens capturas por uma câmara devido a imperfeições na lente ou montagem imprecisa da mesma. Normalmente, os cantos das imagens capturadas poderão ficar comprimidos ou esticados tendo em conta o eixo ótico da câmara. Ambas as distorções encontram-se ilustradas na Figura 3.4.

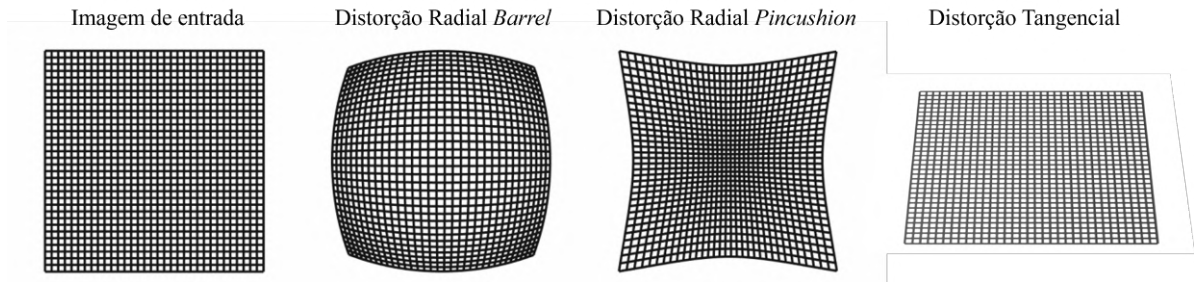


FIGURA 3.4: Distorções radiais e tangenciais (adaptada de Sadekar 2020).

Os coeficientes de distorção, referidos anteriormente, podem ser representados pela matriz  $k$  da Equação 3.9:

$$k = \begin{bmatrix} k_1 & k_2 & p_1 & p_2 & k_3 \end{bmatrix} \quad (3.9)$$

A distorção radial e tangencial do sistema de câmaras pode ser definida pelas Equações 3.10 e 3.11 respetivamente, tanto no eixo de  $x$  como de  $y$ , e considerando os coeficientes de distorção:

$$\begin{aligned} x_{radial} &= x_C(1 + k_1r^2 + k_2r^4 + k_3r^6) \\ y_{radial} &= y_C(1 + k_1r^2 + k_2r^4 + k_3r^6) \end{aligned} \quad (3.10)$$

$$\begin{aligned} x_{tangencial} &= x_C + (2 + p_1x_Cy_C + p_2(r^2 + 2x_C^2)) \\ y_{tangencial} &= y_C + (p_1(r^2 + 2y_C^2) + 2p_2x_Cy_C) \end{aligned} \quad (3.11)$$

De forma a concluir as características óticas e geométricas de ambas as câmaras, para além dos parâmetros intrínsecos, considera-se também a distorção da lente. Assim apresentam-se as Equações 3.12, 3.13 e 3.14:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \frac{x_C}{z_C} \\ \frac{y_C}{z_C} \end{bmatrix} \quad (3.12)$$

$$\begin{bmatrix} x'' \\ y'' \end{bmatrix} = \begin{bmatrix} x'(1 + k_1r^2 + k_2r^4) \\ y'(1 + k_1r^2 + k_2r^4) \end{bmatrix}, \text{ com } r^2 = x'^2 + y'^2 \quad (3.13)$$

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f_x & 0 \\ 0 & f_y \end{bmatrix} \begin{bmatrix} x'' \\ y'' \end{bmatrix} + \begin{bmatrix} u_o \\ v_o \end{bmatrix} \quad (3.14)$$

Considerando que  $f_x = f_y$  as equações anteriormente desenvolvidas, descubra-se a Equação 3.15:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{f}{z_C}(1 + k_1r^2 + k_2r^4) \begin{bmatrix} x_C \\ y_C \end{bmatrix} + \begin{bmatrix} u_o \\ v_o \end{bmatrix}, \text{ com } r^2 = \left(\frac{x_C}{z_C}\right)^2 + \left(\frac{y_C}{z_C}\right)^2 \quad (3.15)$$

## 3.2 Calibração das câmaras

A calibração das câmaras é um processo que estima os seus parâmetros, estabelecendo uma relação entre um objeto 3D que se encontra num cenário com pontos de referência tridimensional  $(x, y, z)$  num espaço de coordenadas bidimensional  $(u, v)$ . Desta forma estabelece-se uma relação entre o cenário real e as imagens captadas pelas câmaras (Y.-J. Zhang 2023).

### 3.2.1 Parâmetros intrínsecos: método de Zhang

Desde as décadas de 1940 e 1950, que autores como Edgar Graham e Harry Carslaw desenvolviam técnicas de calibração inicial de câmaras. O método de calibração de câmaras através de um tabuleiro de xadrez foi publicado por Zhang Zhengyou em 2000, de forma simplificada tornou-se um dos métodos mais utilizados na visão computacional.

O método de Zhang propõe uma técnica de calibração de câmaras com base nas imagens captadas pela mesma. A utilização de um tabuleiro de xadrez com a dimensões pré-definidas e conhecidas, e a sua movimentação perante a câmara em diferentes ângulos e posições permite obter a matriz dos parâmetros intrínsecos, os coeficientes de distorção (Zhengyou Zhang 2000).

Ao recorrer-se apenas a imagens capturadas pelas câmaras simplifica que existe a necessidade de ter conhecimentos sobre geometria tridimensional complexa. O modelo de Zhang permite para além da movimentação do tabuleiro de xadrez, a movimentação da própria câmara e mesmo assim efetuar a calibração com sucesso. No caso da presente dissertação, a calibração de cada câmara será efetuada independentemente.

O modelo de Zhang pode subdividir-se em seis etapas, sendo que numa primeira fase efetua-se a captura de imagens de cada câmara, com diferentes valores de zoom, em que o tabuleiro se movimenta perante a câmara em orientações diversas.

De seguida recorre-se à deteção dos cantos do padrão de xadrez em cada imagem, o autor Zhang Zhengyou propôs um algoritmo que permite identificar pontos de referência no padrão, o qual é abordado na calibração das câmaras presentes na dissertação. A terceira fase do método é a extração de pontos bidimensionais e corresponder a cada um o seu ponto tridimensional. As coordenadas dos pontos tridimensionais são conhecidos à priori, e sendo o tabuleiro de xadrez padrão, é aplicado de forma expedita no algoritmo.

A estimação dos parâmetros iniciais tendo em conta a fase anterior, irá relacionar os pontos bidimensionais com os tridimensionais e estima os parâmetros iniciais intrínsecos da câmara, como por exemplo a distância focal e o centro ótico. Com a distância focal e o centro ótico é possível formar a matriz dos parâmetros intrínsecos.

A etapa seguinte do método é refinar não linearmente os parâmetros obtidos anteriormente recorrendo à máxima verosimilhança. Durante esta fase é efetuado o ajuste dos pontos bidimensionais com os tridimensionais, de modo a que o erro de projeção diminua.

Por fim, recorre-se à avaliação dos resultados obtidos, onde se verifica a precisão e a qualidade da calibração da câmara obtida. É possível com as imagens iniciais, e os parâmetros calculados e otimizados, processar as imagens e corrigi-las, para que, de forma visual, se realize o efeito dos parâmetros na imagem final sem distorção.

### 3.2.2 Parâmetros extrínsecos: método *Perspective-n-Point*

O método *Perspective-n-Point* (PnP) proposto por Martin Fischler em 1980, tem o objetivo de estimar uma posição ou orientação tridimensional de um objeto relativamente a uma câmara. Através da relação de pontos bidimensionais obtidos pelas imagens capturadas e os mesmos pontos mas em coordenadas tridimensionais, é possível obter a posição de um objeto.

O algoritmo do método PnP tem em consideração os pontos 2D das imagens e as suas correspondências tridimensionais, a geometria da câmara, os parâmetros intrínsecos, como a distância focal e o centro ótico.

Existem diversas abordagens do método PnP, como por exemplo o algoritmo EPnP (*Efficient Perspective-n-Point*), que é um método mais robusto a ruídos utilizando um número elevado de correspondências para estimar posições mais precisas. O método *Direct Line Transform* (DLT) é um método mais sensível a erros e ruídos, o que torna o algoritmo mais lento, funciona para qualquer número de correspondências, e estima posições menos precisas. O algoritmo *Uncalibrated Perspective-n-Point* (UPnP) é utilizado em situações em que são desconhecidos os parâmetros intrínsecos das câmaras, apesar de o tornar mais flexível, os valores de estimativa são menos precisos. O método *Perspective-Three-Point* (P3P) consegue resolver a estimação da posição com apenas três correspondências de pontos, é um

dos algoritmos mais rápidos para ser implementado em casos de aquisição em tempo real, pode também gerar várias soluções possíveis de estimação da posição.

Para o efeito da presente dissertação, como estas medidas pretendem ser efetuadas à priori, podendo ser efetuadas até mesmo antes do UAV descolar, ou o navio largar para o mar, desde que as estruturas se mantenham na mesma posição, optou-se pelo algoritmo PnP Ransac.

O algoritmo PnP Ransac utiliza *Random Sample Consensus* (RANSAC) de forma a obter uma estimativa da posição mais precisa, ou seja, dos pontos todos que são propostos o algoritmo escolhe aleatoriamente alguns desse pontos e calcula a estimação da posição. Desta forma, elimina erros de possíveis pontos *outliers* ou ruidosos, conseguindo uma robustez na estimação da posição mais precisa. O algoritmo atua subdividido em oito etapas principais.

A primeira etapa corresponde à seleção aleatória de correspondências de pontos 2D-3D e assim formam um modelo base da posição.

A fase seguinte baseia-se na estimação de posição, o qual aprovisiona a posição estimada inicial do UAV perante o cenário.

A terceira etapa é a avaliação das correspondências, que através da posição estimada inicial anteriormente e das correspondências 2D-3D que não foram selecionadas na primeira etapa irão avaliar se a estimação está de acordo com os valores obtidos.

A quarta fase calcula o erro residual para cada correspondência, é avaliada a discrepância entre os pontos bidimensionais projetados e os pontos bidimensionais reais da imagem.

Com base no erro, a fase seguinte, define um limite sobre quais as correspondências que se consideram dentro do expectável e as correspondências que são considerados *outliers*.

A sexta fase, após se definir o limite das correspondências, aperfeiçoa a posição estimada com apenas os pontos 2D-3D que se encontram dentro do limite.

A próxima etapa, aprimora a posição estimada através da repetição do processo desde a primeira fase até à sexta, e varia a escolha das correspondências que se encontram dentro do limite.

Por fim, a oitava etapa emite o valor final da posição tridimensional estimada, com consistência no resultados final, tendo em conta todas as correspondências.

De forma a obter os parâmetros extrínsecos das câmaras colocadas no convés do navio, decidiu-se recorrer a um marcador visual com uma figura geométrica em forma da letra "H". O marcador visual, colocado no convés de forma visível por ambas as câmaras, permitirá estimar a localização das mesmas no convés, e consequentemente, obter os seus parâmetros extrínsecos.

Define-se um sistema de coordenadas do mundo no centro do marcador visual, de forma a que o vetor do eixo do  $x$  aponte para ré do navio, o vetor do eixo do  $y$  para estibordo e o vetor do eixo do  $z$  para cima. Mede-se a posição dos doze pontos facilmente identificáveis nas imagens das câmaras, de acordo com a Figura 3.5, em relação ao referencial definido.

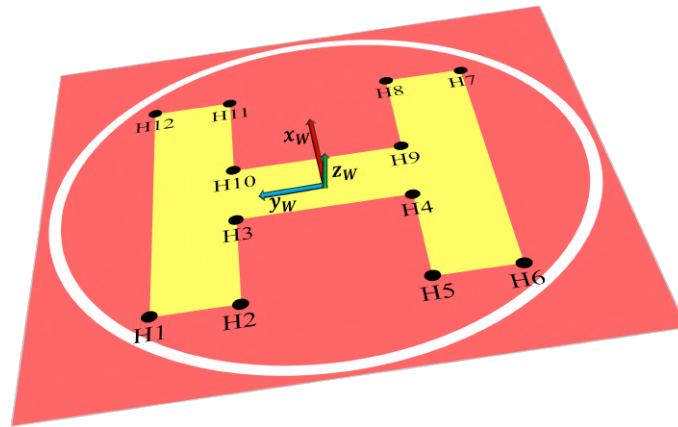


FIGURA 3.5: Marcação de pontos no marcador H.

Para o funcionamento correto do método PnP são necessários pelo menos três pontos para efetuar a relação 2D-3D, e quantos mais correspondências existirem menor será o erro e maior a precisão na estimação. Os pontos tridimensionais são medidos tendo em conta as dimensões reais do marcador H, em metros, e dá-se a denominação a cada ponto da seguinte forma, em metros:

$$\begin{aligned}
 H_1 &= (x_1; y_1; z_1) & H_7 &= (x_7; y_7; z_7) \\
 H_2 &= (x_2; y_2; z_2) & H_8 &= (x_8; y_8; z_8) \\
 H_3 &= (x_3; y_3; z_3) & H_9 &= (x_9; y_9; z_9) \\
 H_4 &= (x_4; y_4; z_4) & H_{10} &= (x_{10}; y_{10}; z_{10}) \\
 H_5 &= (x_5; y_5; z_5) & H_{11} &= (x_{11}; y_{11}; z_{11}) \\
 H_6 &= (x_6; y_6; z_6) & H_{12} &= (x_{12}; y_{12}; z_{12})
 \end{aligned}$$

Os pontos bidimensionais de imagens de ambas as câmaras serão representados conforme a representação, respetivamente para a câmara 1 e câmara 2, tendo em consideração que a medida  $u$  são os píxeis medidos da esquerda para a direita a começar do canto superior esquerdo de cada imagem e a medida  $v$  na direção descendente do mesmo referencial inicial.

## 3.3 Detecção do UAV

No contexto da detecção do UAV que se encontra na fase de aproximação, é imperativo recorrer a uma Rede Neuronal (RN) com a capacidade de operar em tempo real. A RN tem de conseguir identificar de forma eficaz e precisa a presença de um UAV no cenário real. A RN dever ter também a capacidade de analisar continuamente as imagens que são captadas pelas câmaras a bordo.

Para além de uma RN ter de as características mencionadas, é fundamental recorrer a uma arquitetura de rede viável. É crucial que a RN tenha dados suficientes para treino, validação e teste.

A diversidade de dados de treino que devem ser capturadas pelas câmaras, devem conter aproximações do UAV em cenários distintos sobre condições de iluminação, ângulos e trajetórias diversas. Quanto maior for a variedade de dados recolhidos, maior será a robustez da RN.

Nas anotações efetuadas sobre as imagens recolhidas, tem de se considerar as variações de tamanho das *bounding boxes*, as diversas formas de UAVs, ou até mesmo, considerar as obstruções visuais durante as aquisições de imagens.

A abordagem da família *You Only Look Once* (YOLO) surge como uma solução viável devido à sua capacidade de executar deteções de objetos de forma expedita em tempo real (Redmon, Divvala, Girshick e Farhadi 2017).

A arquitetura YOLO divide cada imagem numa grade e faz a associação de cada célula à deteção do objeto a identificar. Ao subdividir as imagens em secções torna o processo de deteção mais ágil, e com capacidade de identificar o objeto mesmo em tamanhos diferentes (Terven e Cordova-Esparza 2023).

Enquanto que a deteção estabelece o cenário inicial, a estimação da posição e velocidade do UAV foca a trajetória ao longo do tempo até ao momento de aterragem. Embora a deteção de UAVs seja um elemento crucial nos sistemas de apoio à aterragem de UAVs, é necessário reconhecer que a presente dissertação direciona o foco na estimação da posição e velocidade.

## 3.4 Estimação de posição e velocidade do UAV: geometria epipolar e filtro de Kalman estendido

A estimação de posição e velocidade é uma das áreas de visão computacional que se baseia no seguimento preciso de um certo objeto, o UAV na presente dissertação, que se encontre em movimento, de forma a conseguir localizá-lo tridimensionalmente num espaço. Existem diversas possibilidades de calcular a estimativa da pose de um UAV, no entanto, na presente dissertação irão ser analisados dois métodos, na Subsecção 3.4.1 aborda-se a utilização de geometria epipolar e triangulação *stereo* e na Subsecção 3.4.2 recorre-se ao filtro de Kalman estendido.

### 3.4.1 Geometria epipolar e triangulação *stereo*

O estudo envolvente na área da geometria epipolar, retrocede desde meados do século XIX, porém só a partir do século XX é que existiu um avanço significativo no setor de geometria projetiva. Os autores Gerg Hevesy, Theodor Scheimpflug, Zhengyou Zhang entre outros igualmente importantes, desenvolveram o campo da geometria epipolar atualizando-a para novos métodos.

O método da triangulação *stereo* tem vantagens na sua utilização pois depende exclusivamente das correspondências de pontos tridimensionais com pontos bidimensionais adquiridos pelas imagens de cada câmara através da geometria epipolar. O método reduz a procura de correspondências 2D-3D para um linha epipolar, tornando o algoritmo computacional mais rápido, no entanto mais sensível à ocorrência de erros de posição (Z. Zhang, Deriche, Faugeras e Luong 1995). À semelhança do filtro de Kalman, este método também depende do conhecimento dos parâmetros extrínsecos e intrínsecos das câmaras.

Este método permite estabelecer relações espaciais entre as câmaras e o cenário tridimensional que elas estão a captar. Com base na representação visual apresentada na Figura 3.6, é ilustrado o processo de correlação 2D-3D. A figura 3.8 ilustra de forma mais detalhada a geometria epipolar aplicada, recorrendo às designações atribuídas na Figura 3.7.

O ponto X pode ser considerado o ponto central tridimensional de um UAV em aproximação ao navio, por exemplo, ou qualquer outro objeto presente no cenário. Os  $x_1$  e  $x_2$  são a projecção bidimensional do ponto X nas imagens 1 e 2, da

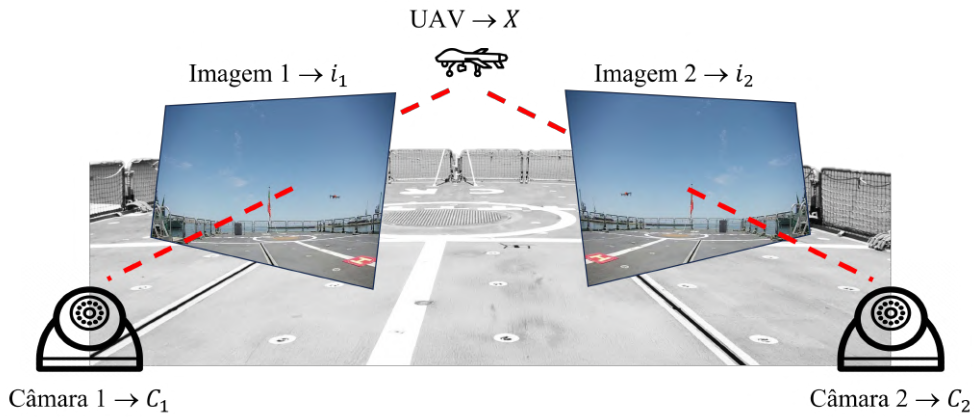


FIGURA 3.6: Geometria Epipolar aplicada a bordo de um navio.

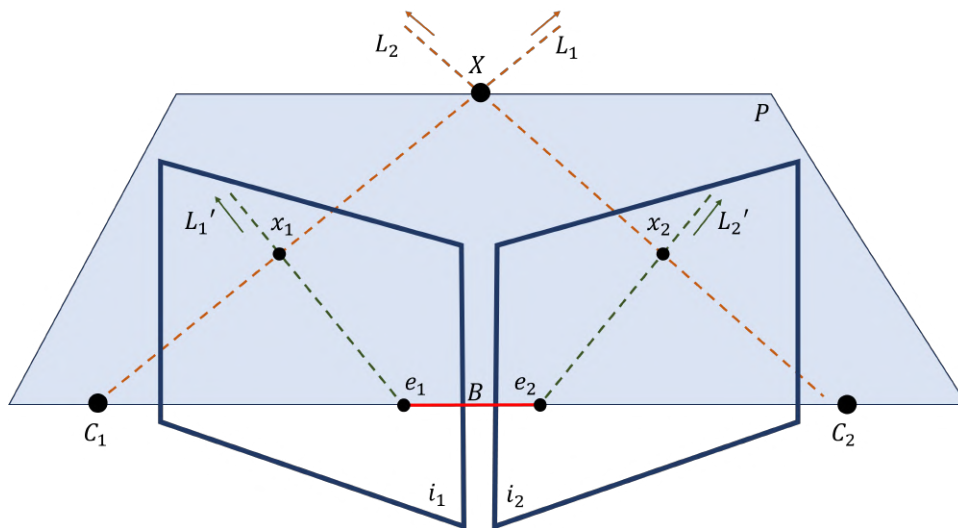


FIGURA 3.7: Geometria Epipolar detalhada.

câmera 1 e 2 respectivamente. Os vetores diretores  $L_1$  e  $L_2$  correspondem às linha epipolares da câmera 1 e 2 em relação ao objeto. O plano é denominado por  $P$  e a linha base por  $B$ , que é a linha que une os dois epipolos ( $e_1$  e  $e_2$ ).

Um epipolo ( $e$ ) é a projeção do centro ótico de uma câmera na vista da outra câmera, e como existem duas câmaras existem dois epipolos. O  $e_1$  é a projeção do centro ótico da câmera 2 na imagem 1, e por sua vez, o  $e_2$  é a projeção do centro ótico da câmera 1 na imagem 2. As linhas epipolares, como por exemplo,  $L_1'$  e  $L_2'$  intersectam o plano  $P$  com os planos das imagens  $i_1$  e  $i_2$ . Por sua vez, os pontos correspondentes localizam-se sobre as linhas epipolares.

Na aplicação da geometria epipolar o primeiro passo a efetuar é a calibração das câmaras, que à semelhança do método anterior, envolve os cálculos dos parâmetros intrínsecos de ambas as câmaras.

O vetor de translação  $T = (C1 - C2)$ , considera o movimento da posição onde se encontra a câmara 1 em deslocação para a câmara 2. Cada câmara tem a sua matriz de rotação ( $R$ ) da câmara. Um ponto  $X$  tridimensional no espaço, tem a relação entre os pontos observados através de cada câmara, de acordo com a Equação 3.16:

$$X_1 = R(X_2 - T) \quad (3.16)$$

A equação do plano epipolar ( $P$ ) pode ser obtida através dos vetores de translação ( $T$ ), bem como do vetor  $X_1$  e da sua translação para a câmara 2,  $X_1 - T$ , desenvolvendo a equação do plano de acordo com a Equação 3.17 obtém-se:

$$(X_1 - T)^t \times X_1 = 0 \Leftrightarrow (R^T X_2)^T T \times X_1 = 0 \quad (3.17)$$

Sabendo que  $[T_x, T_y, T_z]$  é o vetor de translação da câmara 1 para a câmara 2, é possível deduzir a Equação 3.18:

$$T_x X_1 = S X_1, \text{ com } S = \begin{bmatrix} 0 & -T_x & T_y \\ T_z & 0 & -T_x \\ -T_y & T_x & 0 \end{bmatrix} \quad (3.18)$$

A matriz  $S$  representa os vetores tridimensionais, como por exemplo o caso da translação ( $T$ ) de uma câmara para a outra, é originada a partir de um vetor e da sua transporta negativa e iguala-se à matriz original. O resultado das relações anteriores resulta numa matriz designada por *skew-symmetric*.

A próxima fase é o cálculo da matriz Essencial ( $E$ ), é uma componente com dimensões  $3 \times 3$  que visa a relacionar as posições das câmaras num sistema de visão *stereo*, quer da câmara 1 em relação à câmara 2, como vice-versa. A matriz Essencial é utilizada para calcular a relação entre pontos correspondentes em duas imagens diferentes, captadas por câmaras distintas que podem estar em movimento. A matriz tem "rank" de 2, sabe-se que  $R$  é a matriz de rotação. Desta forma estabelece-se a ligação entre os parâmetros extrínsecos do sistema *stereo* e as restrições epipolar, obtendo-se a Equação 3.19:

$$X_2^T E X_1 = 0, \text{ com } E = RS \quad (3.19)$$

Atribuindo a  $X_1$  e  $X_2$  as suas equações de projeções e dividindo pelas respectivas profundidades relativas, normalizando os pontos obtém-se a Equação 3.20:

$$x_2^T E x_1 = 0 \quad (3.20)$$

Em que a parcela  $E x_1$  expressa uma linha projetiva no plano da imagem 2  $L_2$ , que cruza o epipolo 2 ( $e_2$ ) e passa pelo ponto  $x_2$ :  $L_2 = E x_1$ .

Na fase seguinte calcula-se a matriz Fundamental ( $F$ ), contém informações da matriz Essencial e dos parâmetros intrínsecos de cada câmara, em píxeis. A matriz  $F$  pode ser representada de acordo com a Equação 3.21:

$$F = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \quad (3.21)$$

Em que a primeira linha representa a relação dos parâmetros da câmara 1 com a projeção da imagem 1 na imagem 2. A segunda linha representa a relação dos parâmetros da câmara 2 com a projeção da imagem 2 na imagem 1. A terceira linha representa as linha epipolares da imagem 1 que corresponde a pontos na imagem 2. Assim, é possível mapear um ponto numa imagem para um linha epipolar na outra imagem, como ilustra a Figura 3.8.

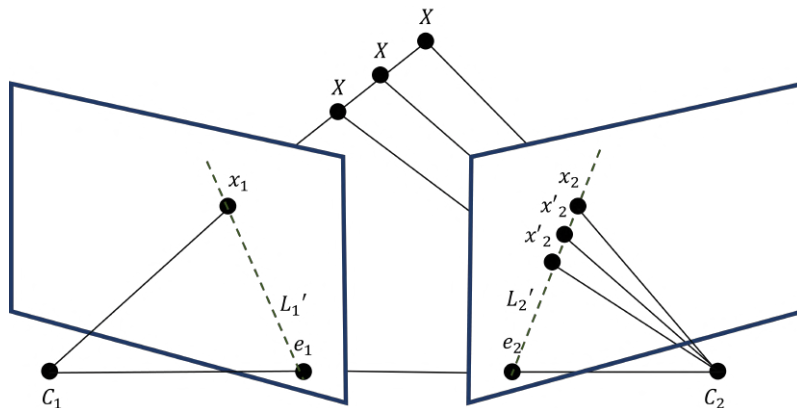


FIGURA 3.8: Projeção de um ponto numa linha epipolar.

As matrizes dos parâmetros intrínsecos das câmaras representadas por  $K_1$  e  $K_2$ , possibilitam o cálculo das coordenadas dos pontos projetados,  $x'_1$  e  $x'_2$  da seguinte forma para a imagem 1 e imagem 2, respetivamente de acordo com a Equação 3.22:

$$x_1 = K'_1 x'_1 \quad x_2 = K'_2 x'_2 \quad (3.22)$$

Tendo em consideração a Equação 3.23 e que a parcela  $F'x'_1 = L'_2$  que representa a equação da linha projetiva, dá-se que:

$$x_2'^T F x'_1 = 0 \quad \text{com} \quad F = K_2^{-T} E K_1^{-1} \quad (3.23)$$

A triangulação *stereo* consiste num processo composto por diversas fases. À semelhança do filtro de Kalman, o método de triangulação recorre à calibração das câmaras, utilizando os parâmetros intrínsecos e extrínsecos para os cálculos do estado do UAV.

A fase seguinte é a identificação de correspondências, em que para cada ponto da câmara 1 necessita de um ponto correspondente na câmara 2. A correlação é efetuada através de técnicas de correspondência *stereo*.

Após serem recolhidas as correspondências necessárias, é aplicada a geometria epipolar. Na presente etapa, simplifica-se o processo de correspondência da posição do UAV, de duas dimensões para uma dimensão, ao associar a uma das câmaras uma linha epipolar.

No passo seguinte calcula-se a disparidade, que é a diferença horizontal entre as posições dos pontos correspondentes nas imagens de cada câmara. Através dos valores da disparidade recorre-se aos valores da distância entre ambas as câmaras (linha base ( $B$ ) na Figura 3.7) e os respetivos parâmetros de forma a calcular a profundidade do UAV.

Uma vez que, a profundidade de cada imagem já foi definida, procede-se à transformação das coordenadas  $(x, y, z)$  do ponto correspondente para um sistema de coordenadas do mundo tridimensional.

Através das posições tridimensionais calculadas com coordenadas em  $x$ ,  $y$  e  $z$  anteriormente, é possível determinar a velocidade em cada eixo. A velocidade em

cada eixo é calculada pela diferença entre as posições consecutivas dividindo pelo intervalo de tempo entre as posições respectivas, na Equação 3.24.

$$\begin{aligned}v_x(k) &= \frac{x(k)-x(k-1)}{T} \\v_y(k) &= \frac{y(k)-y(k-1)}{T} \\v_z(k) &= \frac{z(k)-z(k-1)}{T}\end{aligned}\tag{3.24}$$

### 3.4.2 Filtro de Kalman estendido

O Filtro de Kalman (FK) representa um método estatístico que tem o objetivo de estimar o estado de um objeto de um sistema dinâmico. É um método matemático que aplica equações e dados de entrada sucessivos, como é o caso das imagens que adquirem a posição espacial do UAV e os dados da sua telemetria, e consegue estimar rapidamente o seu valor verdadeiro, como a posição, velocidade, distância, entre outros, do UAV, quando os valores medidos apresentam erros imprevisto, variações aleatórias ou ruído (Kalman 1960).

De forma a compreender a evolução das variáveis recorre-se a um sistema que consiga analisar a evolução das mesmas ao longo do tempo. O filtro de Kalman assenta em duas fases principais que são: a fase da predição e a fase da atualização, de acordo com a Figura 3.9.

Na fase da predição é estimado o estado do objeto com base no modelo dinâmico e as variáveis de entrada do sistema. Na fase da atualização é corrigida a estimativa do estado do objeto obtida na fase de predição, de forma a tornar o estado mais coerente com as observações obtidas nesse instante.

O filtro de Kalman começa por se assumir que o estado do UAV ( $x$ ), tendo em conta que  $P_x$ ,  $P_y$  e  $P_z$  representam a posição tridimensional em  $x$ ,  $y$  e  $z$  respetivamente, e  $v_x$ ,  $v_y$  e  $v_z$  representam a velocidade nos três eixos, de acordo com a Equação 3.25.

$$x = \begin{bmatrix} P_x \\ P_y \\ P_z \\ V_x \\ V_y \\ V_z \end{bmatrix}\tag{3.25}$$

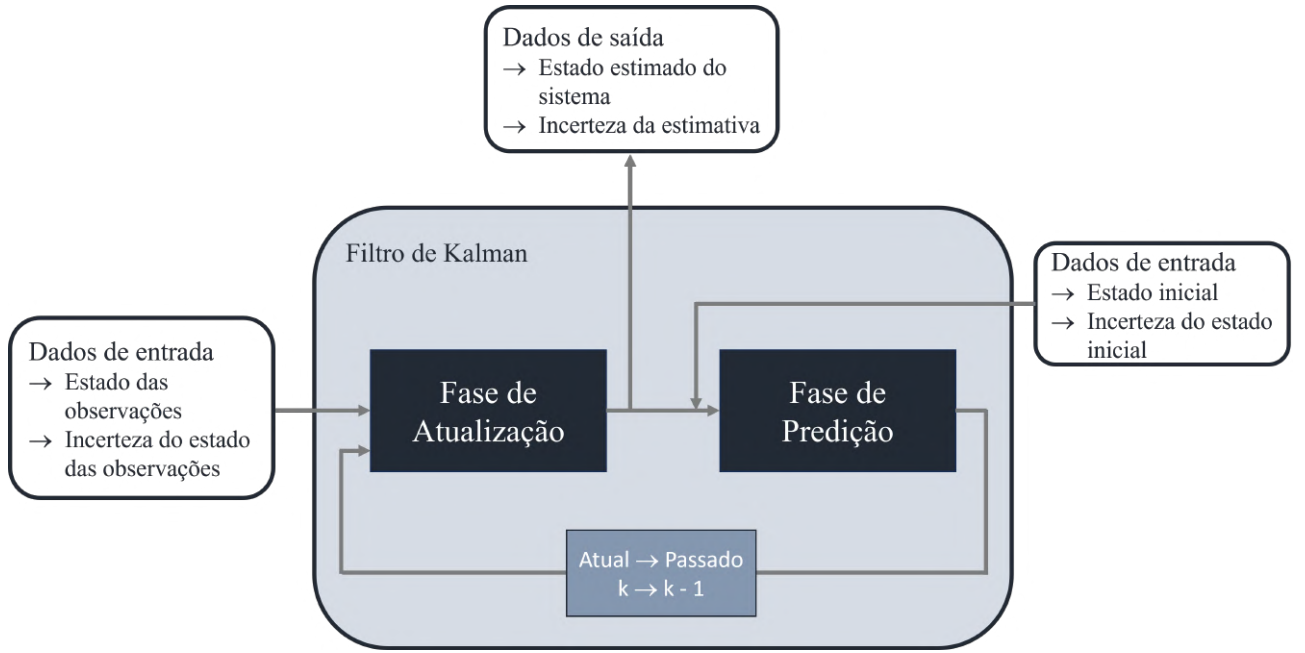


FIGURA 3.9: Filtro de Kalman.

Tendo em conta as *bounding boxes* das imagens recolhidas por ambas as câmaras tem-se  $z$  que é a observação do FK, em que  $u_1$  e  $v_1$  são as coordenadas em píxeis do centro do UAV na câmara 1 e  $u_2$  e  $v_2$  são as coordenadas em píxeis do centro do UAV na câmara 2, de acordo com a Equação 3.26.

$$z = \begin{bmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \end{bmatrix} \quad (3.26)$$

O sistema pode ser descrito através das equações da cinemática para cada uma das três dimensões existentes, de acordo com a Equação 3.27:

$$\begin{cases} v_x(k) = v_x(k-1) + \Delta T a_x(k-1) \\ p_x(k) = p_x(k-1) + \Delta T v_x(k-1) + \frac{1}{2} \Delta T^2 a_x(k-1) \\ v_y(k) = v_y(k-1) + \Delta T a_y(k-1) \\ p_y(k) = p_y(k-1) + \Delta T v_y(k-1) + \frac{1}{2} \Delta T^2 a_y(k-1) \\ v_z(k) = v_z(k-1) + \Delta T a_z(k-1) \\ p_z(k) = p_z(k-1) + \Delta T v_z(k-1) + \frac{1}{2} \Delta T^2 a_z(k-1) \end{cases} \quad (3.27)$$

A fase de predição estima o estado atual do UAV, através da relação de equações da posição ou velocidade do UAV no instante anterior e a posição ou velocidade do UAV no instante atual. Desta fase, surge ainda a matriz de covariância do erro, que contém a incerteza associada a cada estimativa do estado do UAV.

Na fase de predição aplica-se um modelo dinâmico em que se assume uma relação linear com a velocidade constante, de forma a que as variáveis envolvidas evoluem de forma previsível ao longo do tempo. Tendo em consideração o estado anterior do UAV  $X_{k-1}$  e o estado da entrada do sistema  $U_{k-1}$ , consegue-se obter o estado após a atuação que se representa por  $X_k$ , de acordo com a Equação 3.28:

$$\bar{X}_k = F_k X_{k-1} + B_k U_{k-1} \quad (3.28)$$

Em que se assume que  $F_k$  é a matriz de transição de estado que relaciona o estado anterior  $X_{k-1}$  com o estado atual  $X_k$ , e  $B_k$  é a matriz de controlo que relaciona o vetor de controlo  $U_{k-1}$  com o estado atual. A matriz  $B_k$  pode ser representada na Equação 3.29:

$$B_k = \begin{bmatrix} \frac{1}{2}\Delta T^2 & 0 & 0 \\ 0 & \frac{1}{2}\Delta T^2 & 0 \\ 0 & 0 & \frac{1}{2}\Delta T^2 \\ \Delta T & 0 & 0 \\ 0 & \Delta T & 0 \\ 0 & 0 & \Delta T \end{bmatrix} \quad (3.29)$$

No presente caso, despreza-se a parcela  $B_k X_{k-1}$ , pois assume-se que a aceleração  $a = \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix}$  é uma perturbação gaussiana com média nula, o que significa que em média a aceleração tem valor zero. Desta forma, o estado após a atuação é representado de acordo com a Equação 3.30:

$$\bar{X}_k = F_k X_{k-1}, \text{ com } F_k = \begin{bmatrix} 1 & 0 & 0 & \Delta T & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta T & 0 \\ 0 & 0 & 1 & 0 & 0 & \Delta T \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.30)$$

Isto significa que o estado do UAV passa a ser mais complexo tendo em consideração o modelo dinâmico anterior. Ainda na fase de predição é também calculada a incerteza do estado de  $X$  que é representada por uma matriz da covariância da incerteza,  $\bar{\Sigma}_k$ , de acordo com a Equação 3.31.

$$\bar{\Sigma}_k = F_k \Sigma_{k-1} F_k^T + Q_k \quad (3.31)$$

Em que a variável  $Q_k$  representa a matriz da covariância da incerteza na dinâmica e pode ser dada pela Equação 3.32:

$$Q_k = V(B_k a_{k-1}) = B_k V(a) B_k^T = B_k \sigma_a^2 I_{3 \times 3} B_k^T = \sigma_a^2 B_k B_k^T$$

$$= \sigma_a^2 \begin{bmatrix} \frac{1}{4} \Delta T^4 & 0 & 0 & \frac{1}{2} \Delta T^3 & 0 & 0 \\ 0 & \frac{1}{4} \Delta T^4 & 0 & 0 & \frac{1}{2} \Delta T^3 & 0 \\ 0 & 0 & \frac{1}{4} \Delta T^4 & 0 & 0 & \frac{1}{2} \Delta T^3 \\ \frac{1}{2} \Delta T^3 & 0 & 0 & \Delta T^2 & 0 & 0 \\ 0 & \frac{1}{2} \Delta T^3 & 0 & 0 & \Delta T^2 & 0 \\ 0 & 0 & \frac{1}{2} \Delta T^3 & 0 & 0 & \Delta T^2 \end{bmatrix} \quad (3.32)$$

A fase de atualização corrige os valores obtidos na fase de predição, considerando as observações, no presente caso, se consideram as posições do UAV obtidas pelas câmaras. Nesta fase é proposto um modelo de observação para cada câmara específica, que tem o objetivo de estabelecer uma relação matemática entre as informações obtidas em cada imagem com as coordenadas do mundo. Desta forma tem se em consideração a Equação 3.33:

$$P_I = h(P_w), \text{ com } (P_w) = h_{INT} \circ h_{EXT} (P_w) \quad (3.33)$$

Na presente etapa ocorre a transformação de coordenadas do mundo  $P_w$  para coordenadas relativas tendo em conta os parâmetros extrínsecos de cada câmara, em que  $R_{3 \times 3}$  é a matriz de rotação dos eixos x, y e z e  $T_{3 \times 1}$  é a matriz de translação dos três eixos, de acordo com a Equação 3.34:

$$P_c = h_{EXT} (P_w) = C_{T_w} \overline{P_w} = R_{3 \times 3} P_w + T_{3 \times 1} \quad (3.34)$$

Um ponto no referencial da câmara (C) é transformado para um ponto no referencial da imagem (I), de acordo com a equação anterior. Para transformação são utilizados os parâmetros intrínsecos e coeficientes de distorção de cada câmara específica, obtém-se a Equação 3.35:

$$P_I = h_{INT} (P_c) \Rightarrow \begin{bmatrix} u \\ v \end{bmatrix} = \frac{f}{z_c} (1 + k_1 r^2 + k_2 r^4) \begin{bmatrix} x_c \\ y_c \end{bmatrix} + \begin{bmatrix} u_o \\ v_o \end{bmatrix}, \quad r^2 = \left( \frac{x_c}{z_c} \right)^2 + \left( \frac{y_c}{z_c} \right)^2 \quad (3.35)$$

Tendo em consideração que a função  $h()$ , deduzida anteriormente, é uma equação não linear e não depende da velocidade do UAV, é necessário recorrer ao Filtro de Kalman Estendido (FKE).

O FKE é baseado no filtro de Kalman clássico porém permite lidar com sistemas não lineares. Utiliza aproximações lineares de Taylor, e a linearização é executada através de derivadas parciais, obtendo a matriz Jacobiana, que descreve a relação entre as variáveis do sistema. O EKF combina a fase de predição e atualização de forma a estimar o estado do UAV de um sistema dinâmico não linear.

Através da matriz do filtro de Kalman estendido  $H_k$ , sabendo que  $\widehat{X}_{k-1}$  é a estimação mais recente do estado de  $X$ , tem se a Equação 3.36 e 3.37:

$$H_k = \left. \frac{\partial h()}{\partial x} \right|_{x = \widehat{X}_{k-1}} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & 0 & 0 & 0 \\ h_{21} & h_{22} & h_{23} & 0 & 0 & 0 \end{bmatrix}, \text{ onde :} \quad (3.36)$$

$$\begin{aligned} h_{11} &= \frac{\partial u}{\partial x_w} & h_{12} &= \frac{\partial u}{\partial y_w} & h_{13} &= \frac{\partial u}{\partial z_w} \\ h_{21} &= \frac{\partial v}{\partial x_w} & h_{22} &= \frac{\partial v}{\partial y_w} & h_{23} &= \frac{\partial v}{\partial z_w} \end{aligned} \quad (3.37)$$

O cálculo do Jacobiano da função  $\left. \frac{\partial h(\cdot)}{\partial x} \right|_{x = \bar{X}_{k-1}}$  é efetuado através da derivada da função composta, que é dada pela Equação 3.38:

$$\left. \frac{\partial h(\cdot)}{\partial x} \right|_{x = \bar{X}_{k-1}} = \left. \frac{\partial h(\cdot)_{INT}}{\partial P_C} \right|_{x = h_{EXT}(\bar{X}_{k-1})} \times \left. \frac{\partial h(\cdot)_{EXT}}{\partial x} \right|_{x = \bar{X}_{k-1}} \quad (3.38)$$

Em que a primeira parcela do cálculo do Jacobiano será uma matriz com as dimensões  $2 \times 3$ , que é representada de acordo com a Equação 3.39. O cálculo das derivadas parciais do Jacobiano encontram-se no Anexo I.

$$\left. \frac{\partial h(\cdot)_{INT}}{\partial P_C} \right|_{x = h_{EXT}(\bar{X}_{k-1})} = \begin{bmatrix} \frac{du}{dx_C} & \frac{du}{dy_C} & \frac{du}{dz_C} \\ \frac{dv}{dx_C} & \frac{dv}{dy_C} & \frac{dv}{dz_C} \end{bmatrix} \quad (3.39)$$

A segunda parcela do cálculo do Jacobiano é uma matriz com dimensões  $3 \times 6$  que corresponde à Equação 3.40:

$$\left. \frac{\partial h(\cdot)_{EXT}}{\partial x} \right|_{x = \bar{X}_{k-1}} = \begin{bmatrix} \cdot & \cdot & \cdot & 0 & 0 & 0 \\ \cdot & R_{3 \times 3} & \cdot & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & 0 & 0 & 0 \end{bmatrix} \quad (3.40)$$

Após a multiplicação das matrizes Jacobianas anteriores, obtém-se o valor do Jacobiano final de  $H_k$ , que é uma matriz com dimensões  $4 \times 6$ . O Jacobiano final corresponde à justaposição dos Jacobianos obtidos para a estimativa atual da posição do UAV, calculados para cada câmara. A operação permite calcular a relação dos movimentos das câmaras com as variáveis independentes, e por consequente, obter a posição estimada do UAV.

Na fase de atualização tem-se em consideração o estado calculado na fase de predição,  $\bar{X}_k$ . O modelo de incerteza incluem-se as incertezas dos valores medidos ou obtidos através dos sensores (câmaras). A incerteza é expressa pela matriz de covariância  $R_k$ , expressa em píxeis, pela diagonal da matriz. Os valores obtidos refletem a falta de precisão e confiabilidade.

O ganho de Kalman é também calculado nesta etapa tendo em consideração que  $R_k$  é a matriz da covariância do sensor de cada câmara, assim obtém-se a Equação 3.41:

$$K_k = \bar{\Sigma}_k H_k^T (H_k \bar{\Sigma}_k H_k^T + R_k)^{-1} \quad (3.41)$$

A Equação 3.42 atualiza a estimativa do estado  $\hat{X}_k$  e a incerteza correspondente  $\Sigma_k$ , considerando a incorporação da observação  $Z_k$ :

$$\begin{aligned} \hat{X}_k &= \bar{X}_k + K_k (Z_k - H_k \bar{X}_k) \\ \Sigma_k &= \bar{\Sigma}_k - K_k H_k \bar{X}_k \end{aligned} \quad (3.42)$$

Em que  $Z_k$  representa a matriz do estado do UAV a cada instante. Desta forma chama-se inovação do estado do UAV à parcela:  $Z_k - H_k \bar{X}_k$ .



# Capítulo 4

## Aquisição de dados

A aquisição de dados é um procedimento crucial que possibilita a recolha sistemática de informações para análise futura. A Secção 4.1 descreve os procedimentos adotados por forma a adquirir os conjuntos de dados com a precisão necessária para a correta validação dos resultados obtidos. A Secção 4.2 descreve o *hardware* e o aparato experimental que foi utilizado em cada aquisição. A Secção 4.3 desenvolve mais detalhadamente a primeira aquisição de dados. A Secção 4.4 aborda a aquisição de dados 2, e a Secção 4.5 a aquisição de dados 3.

### 4.1 Aquisição de *datasets* em diferentes cenários

A aquisição de dados foi elaborada em três contextos diferentes, porém com o mesmo objetivo, ilustrados na Figura 4.1.

Na primeira aquisição de dados, o foco foi conseguir estabelecer métodos de trabalho mais eficientes, tanto na montagem como na própria aquisição de imagens e localizações das câmaras. A aquisição efetuada numa área terrestre ampla, teve o intuito de identificar áreas passíveis para melhorias nas aquisições subsequentes.

A segunda aquisição de dados foi efetuada a bordo de uma fragata com o marcador H presente, ao contrário da aquisição anterior. O marcador H, com as dimensões 3D conhecidas, acrescentou a possibilidade do cálculo dos parâmetros extrínsecos. Os dados desta aquisição conseguem ser implementados num filtro de Kalman estendido e na triangulação *stereo*, contribuindo para a tomada de decisão mais precisa e informada no que remete à posição do UAV perante o navio.

A terceira aquisição foi efetuada novamente em zonas amplas terrestres, com o mesmo objetivo da segunda aquisição. Os novos dados adquiridos colmataram algumas falhas que decorreram na obtenção da segunda aquisição de imagens, o

que contribuiu para o enriquecimento do *dataset* final. Nesta aquisição também se recorreu ao marcador H com as dimensões previamente medidas.



FIGURA 4.1: *Datasets* em diferentes cenários: cenário 1 (esquerda), cenário 2 (centro) e cenário 3 (direita).

Foram desenvolvidas duas aplicações distintas para extrair o máximo de *frames* possíveis de ambas as câmaras. A primeira foi desenvolvida em linguagem Python e a segunda foi desenvolvida em linguagem C++, tendo obtido uma taxa de aquisição de 12 *frames* por segundo (fps) e de 25 fps respetivamente. Como quanto maior o número de aquisições por segundo mais robusto se tornava a aquisição, a linguagem C++ desempenhava um trabalho mais rápido, daí se ter optado pela segunda alternativa.

Desta forma, foi criado um código em linguagem C++, com o compilador do Visual Studio Community 2022, para a extrações dos *frames* de ambas as câmaras em simultâneo. O código tem como principal função aceder a ambas as câmaras através do endereço IP, e efetuar um loop em que se retira um *frame* e escreve num ficheiro de texto uma linha com as informações do ano, dia, mês, seguido das informações da hora, minuto, segundo e milissegundo a que o *frame* foi capturado.

Foi elaborada uma *checklist* para cada cenário, de procedimentos e montagem, que contém todos os materiais necessários bem como os passos a serem cumpridos antes de cada ensaio, englobando os procedimentos necessários antes da aproximação e aterragem do UAV.

De forma a anotar os dados de cada ensaio no decorrer das aquisições de dados, foi elaborado um ficheiro Excel. O ficheiro Excel proposto serve como ficha de ensaio para a anotação de dados como *pan*, *tilt* e *zoom* de cada câmara, distância ao solo de cada câmara, distância entre câmaras, posição GPS de cada câmara e do marcador H, altitude do local, entre outros parâmetros.

## 4.2 Descrição do hardware utilizado e aparato experimental

O aparato experimental utilizado para efetuar a aquisição de imagem, foi constituído por: duas câmaras *Internet Protocol* (IP), um *switch*, um GPS, um computador, uma fonte de alimentação, três cabos de rede, duas estruturas criadas para simular o *pan* e *tilt* das câmaras, um marcador H. Cada aquisição foi efetuada com um UAV diferente e em unidades militares distintas.

As duas estruturas criadas, onde assentavam as câmaras IP, tem a capacidade de rodar 360° no eixo horizontal (*pan*), e no eixo vertical dos 0° aos 90° (*tilt*), sendo que este último só consegue alterar os ângulos de 5° em 5°, de acordo com a Figura 4.2. As estruturas foram desenvolvidas com o propósito específico de aquisição de imagem para a dissertação. O suporte metálico permite que a estrutura efetue um *pan* de 360°. A base de madeira tem as dimensões de 0,4 metros por 0,4 metros, sendo que as estão centradas no meio da base. Em cima de cada estrutura foi também afixada uma bússola de forma a conseguir diminuir o erro de alinhamento entre ambas, como ilustra a Figura 4.2. As estruturas tem também a capacidade de variar a sua altura desde os 0,6 metros até aos 0,81 metros.

Nesta dissertação foram utilizadas duas câmaras da marca Dahua, que pertencem ao CINA, com as características principais descritas de acordo com a Tabela 4.1 (Technology 2023):

TABELA 4.1: Características Principais das Câmaras

<b>Modelo:</b>	DH-IPC-HDBW8232EP-Z-SL
<b>Peso:</b>	3,02Kg
<b>Dimensões:</b>	254mm*254mm*257mm
<b>Temperatura operacional:</b>	-40° C até +60° C
<b>S/N:</b>	5F05BA7PAGA0E3D
<b>Fabricado:</b>	China

Para além das características anteriores cada câmara tem a capacidade de efetuar *pan* e *tilt*, dos 0° aos 360°, no entanto, no presente modelo de câmara não é possível efetuar alterações parâmetros através de *software*. O *Bit Rate* das câmaras pode ser optado por dois controlos diferentes: *Constant Bit Rate* (CBR) e *Variable Bit Rate* (VBR). As câmaras ainda contam com as seguintes propriedades:

- 1/1.8" 2Megapixel progressive scan CMOS;
- H.265 & H.264 triple-stream encoding;

- 50/60fps, 1080P(1920×1080);
- WDR(120dB), Day/Night(ICR), 3DNR, AWB, AGC, BLC;
- Multiple network monitoring: Web viewer, CMS(DSS/PSS) & DMSS;
- 4.1mm 16.4mm motorized lens;
- Max. IR LEDs Length 50m;
- Micro SD memory, IP67, IK10, POE+;
- 316L stainless steel housing.



FIGURA 4.2: Estruturas de suporte e câmaras IP.

Posteriormente, de forma a captar imagens de ambas as câmaras e sincronizá-las, utilizou-se um *switch* com 6 portas de 10/100Mbps e quatro delas são portas de *PoE+*. O *switch* utilizado era da marca Tp-link, e padece das especificidades de acordo com a Tabela 4.2 (Tp-link 2023):

TABELA 4.2: Características *6-Port 10/100Mbps Desktop Switch with 4-Port PoE+*

<b>Modelo:</b>	TL-SF1006P
<b>Peso:</b>	0,39Kg
<b>Dimensões:</b>	158mm*101mm*25mm
<b>Temperatura operacional:</b>	0° C até +40° C
<b>Fonte de Alimentação:</b>	74W (220V/50Hz com 67W PD conectados)
<b>Interface (RJ45):</b>	Standard 802.3, PoE Ports 1–4, Power Supply 67W

Os recetores de GPS, da marca Garmin, ilustradas na Figura 4.3, que foram responsáveis pela aquisição das posições das câmaras e do marcador H durante os ensaios, contam com as características da Tabela 4.3 (Garmin 2023):

## 4.2. Descrição do hardware utilizado e aparato experimental

TABELA 4.3: Características do navegador GPS

<b>Modelo:</b>	eTrex Vista HCx
<b>Peso:</b>	0,156Kg
<b>Dimensões:</b>	56mm*107mm*30mm
<b>Temperatura operacional:</b>	-40° C até +60° C
<b>Fonte de Alimentação:</b>	2 baterias AA
<b>Resolução:</b>	176 x 220 pixels
<b>Erro de Posição:</b>	-4m



FIGURA 4.3: Recetores GPS utilizados nos ensaios experimentais.

Foi desenvolvido um marcador H para os as aquisições de dados 2 e 3. O marcador H tem as seguintes dimensões gerais: 80 centímetros de largura, 80 centímetros de comprimento e 0,2 centímetros de espessura. O marcador H é caracterizado pela letra H, que tem as seguintes dimensões: 48 centímetros verticais, 48 centímetros horizontais e uma largura de 12 centímetros para cada linha. A circunferência tem um diâmetro de 76,6 centímetros e uma espessura de 0,5 centímetros.

O marcador H pode ser utilizado bilateralmente, sendo que um lado é caracterizado pela cor cinzenta para o fundo e branca para a letra e circunferência. O outro tem o fundo vermelho, com a letra a amarelo e a circunferência a branco, de acordo com a Figura 4.4.

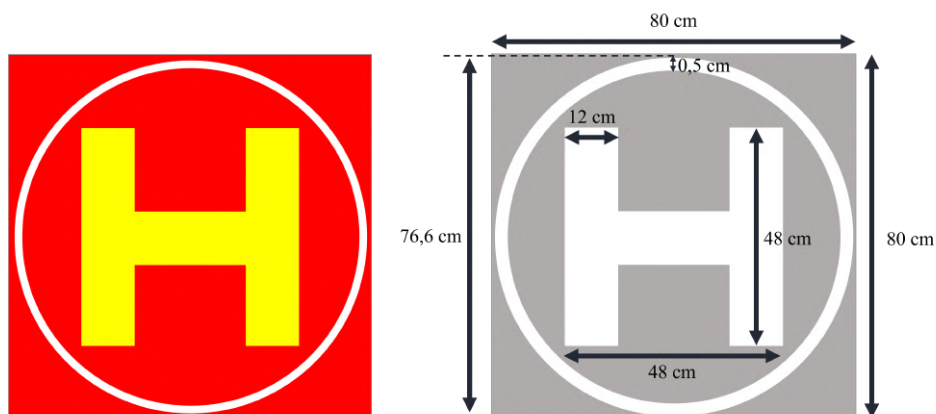


FIGURA 4.4: marcador H bilateral e as suas dimensões.

Cada aquisição de dados foi conduzida com recurso a UAVs que apresentavam características diferenciadas. Consequentemente, foram empregados três UAVs distintos, um para cada operação de aquisição de dados em cenários variados. A inclusão de variedade de UAVs contribuiu para a consolidação de um *dataset* mais robusto.

O primeiro UAV utilizado para aquisição de um *dataset* inicial, ilustrado na Figura 4.5, pertencia ao CEOV, e era um *hexacopter* da marca Tarot, de acordo com o representado na Tabela 4.4 (Tarot 2011):

TABELA 4.4: Características do UAV da Aquisição de Dados 1.

<b>Modelo:</b>	Tarot X6 Hex-Copter FPV TL6X001
<b>Peso:</b>	2.7Kg
<b>Dimensões entre eixos:</b>	960mm
<b>Dimensões de um braço:</b>	316mm
<b>Dimensões do corpo principal:</b>	328mm
<b>Dimensões de cada Hélice:</b>	18polegadas
<b>Altura do UAV:</b>	395mm
<b>Temperatura operacional:</b>	-10° C até +40° C
<b>Fonte de Alimentação:</b>	Baterias de 22.2V 10000-20000mAh
<b>Máximo peso para descolar:</b>	9Kg
<b>Autonomia de voo:</b>	40 minutos (sem <i>payload</i> )

## 4.2. Descrição do hardware utilizado e aparato experimental



FIGURA 4.5: UAV *Tarot X6 Hexacopter* utilizado no primeiro *dataset*.

Para a aquisição do segundo *dataset* foi utilizado um UAV *quadcopter* da marca Autel, ilustrado na Figura 4.6, pertencia ao Pelotão ISR, e é descrito com as características de acordo com a Tabela 4.5 (Autelpilot 2023):

TABELA 4.5: Características do UAV da Aquisição de Dados 2

<b>Modelo:</b>	EVO II Dual 640T
<b>Peso:</b>	1.15Kg
<b>Dimensões entre eixos:</b>	457mm
<b>Dimensões de um braço:</b>	113mm
<b>Dimensões do corpo principal:</b>	230mm
<b>Dimensões de cada Hélice:</b>	5.9polegadas
<b>Altura do UAV:</b>	108mm
<b>Temperatura operacional:</b>	-10° C até +40° C
<b>Fonte de Alimentação:</b>	Baterias de 11.55V 7100mAh
<b>Máximo peso para descolar:</b>	2Kg
<b>Autonomia de voo:</b>	35 minutos (sem <i>payload</i> )

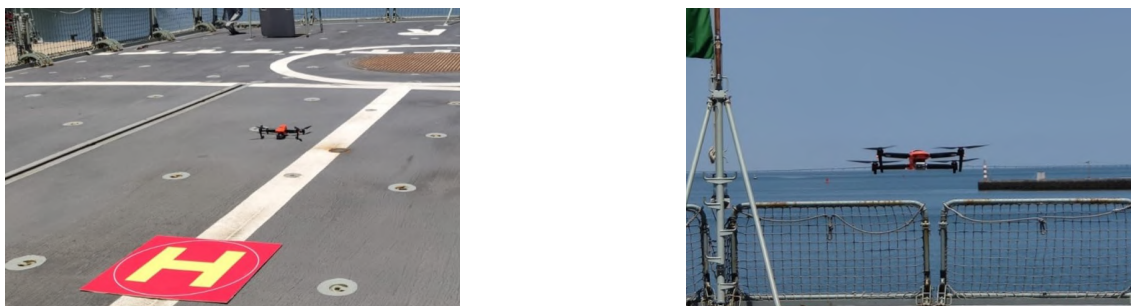


FIGURA 4.6: UAV *Autel EVO II Dual 640T Quadcopter* utilizado no segundo *dataset*.

Por fim, na aquisição do terceiro *dataset* foi utilizado um UAV *hexacopter* da marca DJI, ilustrado na Figura 4.7, que pertencia ao CIAFA, e conta com as características de acordo com a Tabela 4.6 (DJI 2023).

TABELA 4.6: Características do UAV da Aquisição de Dados 3

<b>Modelo:</b>	Matrice 600 Pro
<b>Peso:</b>	9.5Kg
<b>Dimensões entre eixos:</b>	1668mm
<b>Dimensões de um braço:</b>	267.5mm
<b>Dimensões do corpo principal:</b>	1133mm
<b>Dimensões de cada Hélice:</b>	21polegadas
<b>Altura do UAV:</b>	727mm
<b>Temperatura operacional:</b>	-10° C até +40° C
<b>Fonte de Alimentação:</b>	Baterias de 22.2V 4500mAh
<b>Máximo peso para descolar:</b>	15.5Kg
<b>Autonomia de voo:</b>	38 minutos (sem <i>payload</i> )



FIGURA 4.7: UAV *DJI Matrice 600 Pro* utilizado no terceiro *dataset*.

### 4.3 Aquisição de dados 1

O primeiro *dataset* foi recolhido em Mafra, na Escola Prática de Infantaria. A área onde decorreu o exercício foi desenvolvida no âmbito de treino e formação para o combate em Áreas Edificadas, bem como o treino de novas tecnologias, como o caso do exercício em que o *dataset* foi recolhido, ilustrado na Figura 4.8. O exercício proporcionado pelas Forças Armadas Portuguesas, denominado *Guerriers de La Méditerranée*, no âmbito da Iniciativa Defesa 5+5, teve como principal objetivo o treino conjunto entre as Operações Especiais de Portugal, Espanha, França, Itália, Malta, Marrocos, Tunísia, Mauritânia, Líbia e Argélia.



FIGURA 4.8: Imagens do Exercício *Guerriers de La Méditerranée*.

### 4.3. Aquisição de dados 1

No exercício participou também a Célula de Experimentação Operacional de Veículos Não Tripulados (CEOV), a qual se disponibilizou para efetuar vários voos dedicados, independentes do Exercício, para a obtenção do *dataset* desenvolvido neste dissertação.

Durante a recolha do primeiro conjunto de dados, o UAV efetuou aterragens verticais, pois trata-se de um UAV VTOL, como ilustra a Figura 4.9. Enquanto que, os conjuntos de dados subsequentes foram adquiridos pelo mesmo UAV, porém este simulou a aproximação e aterragem de um UAV de asa fixa, ilustrada na Figura 4.10.

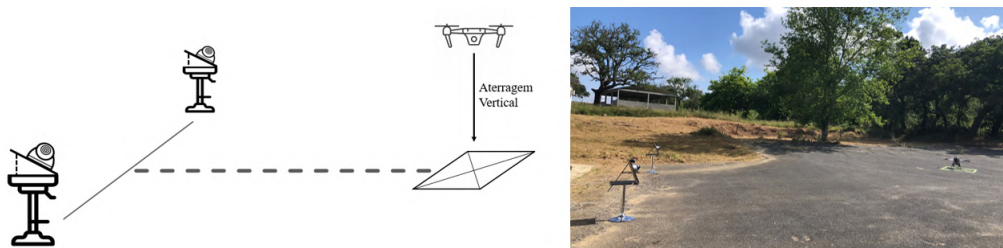


FIGURA 4.9: Aterragem vertical do *dataset*.

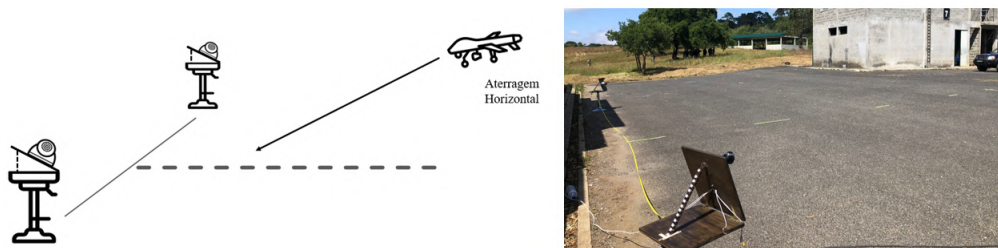


FIGURA 4.10: Aterragem horizontal do *dataset*.

Em acréscimo aos dados obtidos pelas duas câmaras IP, obteve-se também a telemetria do próprio UAV do primeiro *dataset*, contanto com os seguintes dados: data, tempo, *TimeUS*, intensidade do sinal, velocidade angular no eixo do x, y e z, aceleração linear nos eixos do x, y e z, altitude e altitude geométrica.

Os ensaios são delimitados consoante as aproximações do UAV, deste modo efetuaram-se doze ensaios diferentes. A imagem geral de todas as aproximações que o UAV efetuou encontra-se na Figura 4.11.

Os parâmetros das câmaras do primeiro ao décimo ensaio mantiveram-se constantes, de acordo com a Tabela 4.7:



FIGURA 4.11: Aproximações do UAV da Aquisição de dados 1.

TABELA 4.7: Parâmetros das câmaras do Ensaio 1 ao 10 da Aquisição de dados 3

<b>Modo de Codificação:</b>	H.265
<b>Resolução:</b>	1920*1080(1080P)
<b>Velocidade de Projeção:</b>	50
<b>Controlo de Taxa de Bits:</b>	VBR
<b>Taxa Máxima de Bits:</b>	1280 Kb/S
<b>Intervalo de Frames:</b>	50

O ensaio 1 serviu para testar se o sistema estava montado corretamente e se a comunicação entre ambas as câmaras e o computador funcionava sem problemas. O UAV aterrou de forma lenta e verticalmente, era controlado à distância pelo piloto, no entanto encontrava-se sempre um coordenador no terreno a dar informações visuais ao piloto. O ensaio 1 contém no total 1 873 imagens, sendo que 922 pertencem à câmara 1 e 951 a câmara 2.

Após o UAV trocar de baterias, no ensaio 2 foram capturadas imagens de ambas câmaras do UAV a descolar verticalmente de forma lenta com auxílio no local do observador que transmitia ao piloto à distância e afastou-se do local. No total o ensaio 2 tem 798 imagens, 399 imagens capturada por cada câmara.

No ensaio 3 o UAV aproxima-se lentamente do local de aterragem mas o ensaio foi interrompido antes de captar o momento de aterragem. No total tem 4 636 imagens, sendo que 2 298 imagens foram obtidas pela câmara 1 e 2 338 imagens pela câmara 2.

O ensaio 4 realiza uma aproximação gradual em direção ao local de aterragem, onde executa uma aterragem controlada. Após aterrar o UAV descola novamente e retorna ao ponto de partida inicial. O ensaio abrange 2 894 imagens no

### 4.3. Aquisição de dados 1

---

total, em que 1 441 imagens foram capturadas pela câmara 1 e 1 453 imagens pela câmara 2.

No quinto ensaio, o UAV efetua uma aproximação progressiva, no entanto não executa a aterragem final, conta com 714 imagens no total, 358 imagens obtidas pela câmara 1 e 356 imagens pela câmara 2.

O ensaio 6 decorre de forma semelhante e é composto por um total de 625 imagens capturadas, sendo que 310 imagens adquiridas pelas câmara 1 e 315 imagens pela câmara 2.

O ensaio 7 o UAV efetua a aproximação moderada e procede à aterragem. No total o ensaio tem 1087 imagens, em que 541 imagens pertencem à câmara 1 e 546 imagens à câmara 2.

No ensaio 8 o UAV, à semelhança do ensaio 5 e 6, efetua a aproximação mas não conclui a aterragem. O ensaio é composto por 808 imagens, sendo que 399 imagens foram capturadas pela câmara 1 e 409 pela câmara 2.

O ensaio 9 é também equivalente ao ensaio 8, no total é caracterizado por 560 imagens, em que 276 imagens são da câmara 1 e 284 imagens da câmara 2.

O UAV no ensaio 10 descola do local de aterragem, afasta-se, efetua a aproximação e simula a aterragem e volta a afastar-se e efetua novamente uma aproximação. No ensaio foram capturas 3 016 imagens, nas quais 1 521 imagens pela câmara 1 e 1 495 imagens pela câmara 2.

No ensaio 11 existiu uma alteração dos parâmetros de ambas as câmaras, de forma a obter um maior número de imagens capturadas por segundo, passando a ser os valores da Tabela 4.8:

TABELA 4.8: Parâmetros das câmaras do Ensaio 11 da Aquisição de dados 1

<b>Modo de Codificação:</b>	H.265
<b>Resolução:</b>	1280*720(720P)
<b>Velocidade de Projeção:</b>	50
<b>Controlo de Taxa de Bits:</b>	VBR
<b>Taxa Máxima de Bits:</b>	1024 Kb/S
<b>Intervalo de Frames:</b>	50

No total de todos os ensaios da presente aquisição perfaz 23 864 imagens, nas quais foram capturadas 11 930 imagens pela câmara 1 e 11 934 imagens pela câmara 2.

## 4.4 Aquisição de dados 2

O segundo *dataset* foi recolhido em colaboração do Batalhão de Fuzileiro, especificamente com o Pelotão *Intelligence Surveillance and Reconnaissance* (ISR) do Batalhão de Fuzileiros nº2 (BF2) do Destacamento de Apoio de Combate (DAC), e com a guarnição da fragata NRP Corte-Real. O pelotão ISR disponibilizou-se a efetuar os voos na Base Naval do Alfeite, onde a fragata se encontrava atracada. Os ensaios a bordo da fragata para além do treino interno do Pelotão ISR, teve como objetivo a aquisição de dados para a presente dissertação.

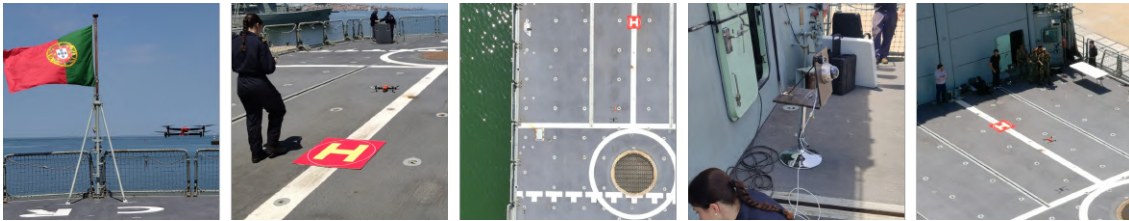


FIGURA 4.12: Imagens do Exercício dos ensaios a bordo no NRP Corte-Real.

Durante a recolha de dados, o UAV do pelotão ISR efetuou três aproximações ao navio, uma por bombordo, outra por estibordo, e por fim a última com a proa da fragata. As simulações de aterragem foram na sua maioria na diagonal, simulando a aterragem de um UAV de asa-fixa, como ilustra a Figura 3.1 do Capítulo 3.

A aquisição de dados efetuada a bordo do NRP Corte-Real, em colaboração com o Pelotão ISR, dividiu-se essencialmente em três ensaios, com aproximações distintas. A Figura 4.13 representa as trajetórias adotadas pelo UAV durante as aproximações:

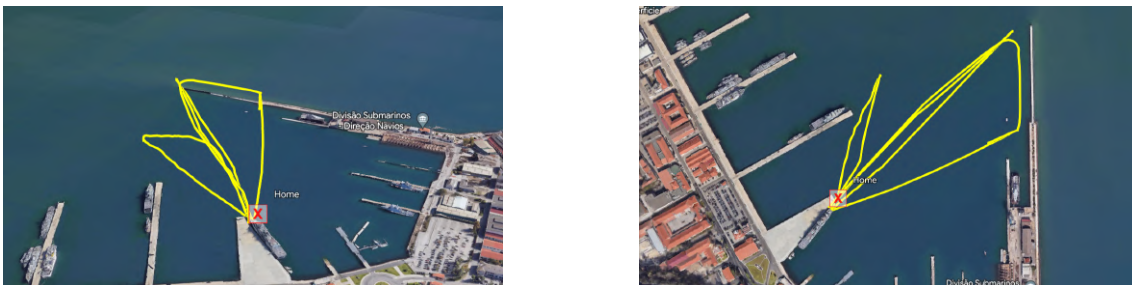


FIGURA 4.13: Aproximações do UAV da Aquisição de dados 2.

Os parâmetros das câmaras no primeiro ensaio efetuado a bordo do navio foram definidos de acordo com a Tabela 4.9:

TABELA 4.9: Parâmetros das câmaras do Ensaio 1 da Aquisição de dados 2

<b>Modo de Codificação:</b>	H.265
<b>Resolução:</b>	1920*1080(1080P)
<b>Velocidade de Projeção:</b>	50
<b>Controlo de Taxa de Bits:</b>	VBR
<b>Taxa Máxima de Bits:</b>	1280 Kb/S
<b>Intervalo de Frames:</b>	50

O ensaio 1 foi o teste de funcionamento de todo o equipamento montado e verificação de comunicações com ambas as câmaras e sincronismo das mesmas. Durante a captura de imagens neste ensaio ainda se estavam a efetuar as medições de distância entre câmaras e marcador H, e as respetivas posições GPS.

Nos restantes ensaios, do segundo ao quinto, as propriedades de ambas as câmaras foram alteradas de acordo com a Tabela 4.10 de forma aumentar o número de *frames* captados por segundo:

TABELA 4.10: Parâmetros das câmaras do Ensaio 2 ao 5 da Aquisição de dados 2

<b>Modo de Codificação:</b>	H.265
<b>Resolução:</b>	1280*720(720P)
<b>Velocidade de Projeção:</b>	50
<b>Controlo de Taxa de Bits:</b>	VBR
<b>Taxa Máxima de Bits:</b>	1024 Kb/S
<b>Intervalo de Frames:</b>	50

O ensaio 2 foi um teste após mudar a resolução das imagens capturadas, não contém imagens do UAV a sobrevoar nem a aterrar, apenas capturou o cenário e o marcador H. Este ensaio é composto por 2 174 imagens, em que 1 085 pertencem à câmara 1 e 1 089 à câmara 2.

No ensaio 3 o UAV aproximou-se e simulou a aterragem, no entanto, o não chega a aterrar no convés. A partir deste ensaio a câmara 2 começou a captar imagens de forma irregular. Desta modo, o ensaio tem um total de 6 691 imagens, das quais 3 018 imagens foram capturadas pela câmara 1 e 3 673 imagens pela câmara 2.

O ensaio 4 captura a aproximação do UAV apenas com a câmara 1, pois a câmara 2 encontrava-se danificada e a frequência de aquisição de imagem diminuiu bruscamente. Desta forma, no total foram capturadas 7 305, em que 5 320 imagens eram da câmara 1 e 1 985 imagens da câmara 2.

No ensaio 5, devido às restrições da câmara 2 e da passagem de pessoas em frente à câmara 1, a câmara 1 capturou a aproximação do UAV com dificuldade. No total o ensaio tem 6 275 imagens, nas quais 5 724 imagens foram adquiridas através da câmara 1 e 551 imagens da câmara 2.

Em todos os ensaios da presente aquisição capturaram-se no total 24 398 imagens, nas quais 16 129 imagens foram adquiridas pela câmara 1 e 8 269 imagens pela câmara 2.

## 4.5 Aquisição de dados 3

O terceiro dataset foi recolhido em colaboração com o Centro de Investigação da Academia da Força Aérea (CIAFA) na Base Aérea nº1, que se localiza em Sintra. À semelhança da aquisição anterior, a recolha de dados no âmbito da dissertação foi obtida durante o treino interno da guarnição do CIAFA, de acordo com a Figura 4.14.



FIGURA 4.14: Imagens dos ensaios em colaboração com o CIAFA.

A aquisição de dados foi efetuada numa zona ampla e sem obstruções, em que o UAV efetuou inúmeras aproximações de diversos ângulos, no total foram efetuados catorze ensaios distintos. A Figura 4.15 demonstra o local onde foram os ensaios e os voos do UAV:

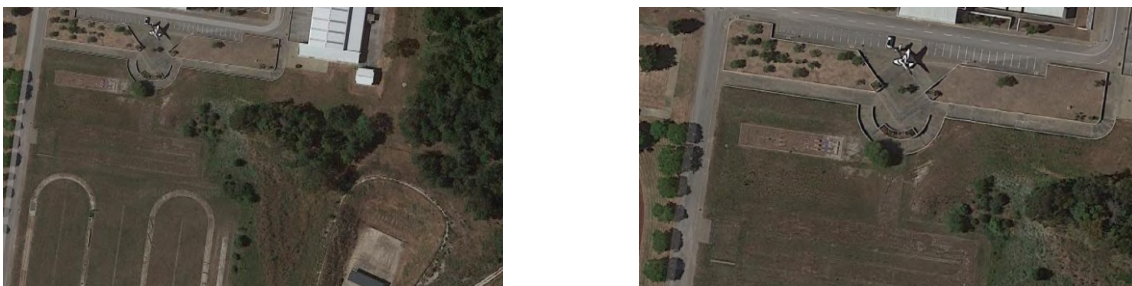


FIGURA 4.15: Aproximações do UAV da Aquisição de dados 3.

As configurações de ambas as câmaras permaneceram as mesmas ao longo dos quatorze ensaios efetuados, de acordo com a Tabela 4.11:

TABELA 4.11: Parâmetros das câmaras durante todos os Ensaios da Aquisição de dados 3

<b>Modo de Codificação:</b>	H.265
<b>Resolução:</b>	1280*720(720P)
<b>Velocidade de Projeção:</b>	50
<b>Controlo de Taxa de Bits:</b>	VBR
<b>Taxa Máxima de Bits:</b>	1024 Kb/S
<b>Intervalo de Frames:</b>	50

O ensaio 1 foi o teste de verificação, em que ainda se estavam a efetuar medições de distâncias e posições GPS. Neste ensaio o UAV permaneceu localizado em cima do marcador H. O ensaio contém no total 535 imagens, nas quais 274 imagens foram capturadas pela câmara 1 e 261 imagens pela câmara 2.

O UAV no ensaio 2 descola perto do local onde o marcador H está posicionado e afasta-se. Neste ensaio foram captadas 2 319 imagens, em que 1 151 imagens pertencem à câmara 1 e 1 168 imagens à câmara 2.

Durante todo o ensaio 3 é possível observar o UAV a sobrevoar o espaço aéreo envolvente no cenário que as câmaras capturam. No total foram adquiridas 3 893 imagens, das quais 1 936 imagens foram obtidas através da câmara 1 e 1 957 imagens da câmara 2.

O ensaio 4 capturou imagens do UAV em aproximação gradual e posterior aterragem, no entanto o UAV fez ligeiras correções na sua altitude na fase final do voo. Foram capturadas 5 465 imagens no total, em que 2 727 imagens são da câmara 1 e 2 738 imagens da câmara 2.

Ao longo do ensaio 5, à semelhança do ensaio 3, o UAV sobrevoou e efetuou várias alterações laterais de posição. O ensaio é composto por 6 216 imagens no total, das quais 3 112 imagens foram capturadas pela câmara 1 e 3 104 imagens pela câmara 2.

No ensaio 6 o UAV aproxima-se e efetua a aterragem, porém avança ligeiramente para fora do campo de visão das câmaras, que deixam de o capturar por um instante. No entanto, o UAV corrige a sua posição e aterra visivelmente em ambas as câmaras. No total foram capturadas 3 466 imagens, em que 1 724 imagens são da câmara 1 e 1 742 imagens da câmara 2.

O ensaio 7 capturou o UAV a sobrevoar lateralmente, bem como efetuar séries de aproximação e afastamento, sem aterrar. Contém no total 1 899 imagens, em que 950 imagens foram adquiridas pela câmara 1 e 949 imagens pela câmara 2.

No ensaio 8 o UAV distancia-se para uma posição de espera no ar, aproxima-se gradualmente e efetua uma aterragem perto do marcador H. No total foram capturadas 5 618 imagens, das quais 2 837 imagens da câmara 1 e 2 781 imagens da câmara 2.

Durante o ensaio 9 o UAV aproxima-se e aterra, no entanto a aterragem não se consegue captar com a câmara 2 pois aterra fora do campo de visão. Assim apenas é possível verificar a última fase do voo através da câmara 1. O ensaio é constituído por 6 311 imagens, nas quais 3 144 imagens são da câmara 1 e 3 167 imagens da câmara 2.

No ensaio 10 o UAV descola do terreno e mantém-se a sobrevoar no campo de visão de ambas as câmaras. Foram capturadas no total 2 283 imagens, em que 1 208 imagens pertencem à câmara 1 e 1 075 imagens à câmara 2.

Durante o ensaio 11 o UAV afastou-se para a posição de partido do próximo ensaio. No total foram capturadas 1 595 imagens do UAV em afastamento, das quais 791 imagens foram capturadas pela câmara 1 e 804 imagens pela câmara 2.

O ensaio 12 captou o UAV em aproximação e consecutivamente em aterragem perto do marcador H. O ensaio contém no total 4 820 imagens, em que 2 399 imagens foram adquiridas através da câmara 1 e 2 421 imagens da câmara 2.

Ao longo do ensaio 13 o UAV aproximou-se e efetuou a aterragem perto do marcador H, voltou a descolar e a afastou-se para o ponto de espera aéreo. É composto por um total de 7 222 imagens, em que 3 616 imagens são da câmara 1 e 3 606 imagens da câmara 2.

No ensaio 14 o UAV sobrevoa o campo de visão de ambas as câmaras, efetua uma aproximação oscilante e posteriormente aterra próximo do marcador H. No total é caracterizado por 6 928 imagens, das quais 3 449 imagens são da câmara 1 e 3 479 imagens da câmara 2.

O número total de imagens capturadas em todos os ensaios da presente aquisição é de 58 570 imagens, das quais 29 318 imagens foram obtidas pela câmara 1 e 29 252 imagens pela câmara 2.

# Capítulo 5

## Tratamento de dados

O presente Capítulo visa a explorar os métodos aplicados durante o processamento de tratamento de dados. A Secção 5.1 descreve através do método de Zhang como é que as câmaras foram calibradas. A Secção 5.2 aborda o alinhamento dos dados de forma a implementá-los em algoritmos de deteção e estimação. A Secção 5.3 expõe as anotações do *dataset*. Por fim, a Secção 5.4 especifica o formato dos dados do *dataset* e a sua disponibilidade para acesso público.

### 5.1 Calibração das câmaras

A calibração de ambas as câmaras foi efetuada através do método de calibração de Zhang. O método proposto por Zhengyou Zhang (1999), tem como objetivo determinar os parâmetros intrínsecos de uma câmara. Através da observação de um padrão xadrez, na Figura 5.1, em várias posições e ângulos diferentes, são extraídos pontos de interesse no padrão. Com os pontos de interesse fazem-se correspondências no espaço tridimensional, através de técnicas de geometria obtém-se os parâmetros intrínsecos das câmaras (Zhengyou Zhang 1999).

O tabuleiro utilizado tem as dimensões de 24 centímetros de comprimento por 17 centímetros de largura, e cada quadrado tem 2,4 centímetros comprimento e largura. O tabuleiro utilizado foi retirado da biblioteca OpenCV e tem a designação de 9x6 (OpenCV 2023).

Cada câmara foi montada numa estrutura com um ângulo perpendicular ao chão, com a altura de 60 centímetros do chão para efetuar a aquisição de imagens do tabuleiro, de acordo com a Figura 5.1. Cada câmara tirou um conjunto de imagens, com as resoluções 1920x1080 píxeis e 1280x720 píxeis, com e sem zoom ótico. Obtiveram-se assim doze conjuntos de dados, cada um com cerca de 30 imagens, perfazendo um total de 338 imagens.

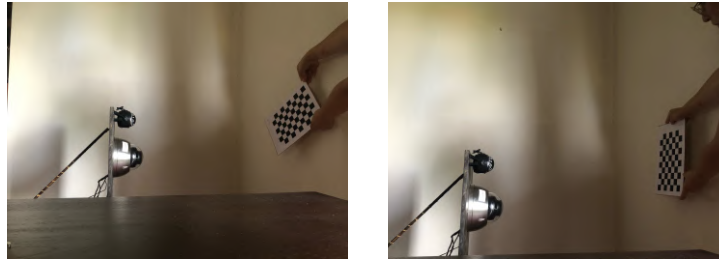


FIGURA 5.1: Ensaios para aquisição de imagens para calibração das câmaras.

Para a calibração das câmaras efetuaram testes sem zoom ótico, e com dois zooms óticos diferentes para cada câmara e resolução. Tendo como referência que o zoom mínimo corresponde a 0 e o zoom máximo corresponde a 852, o primeiro zoom utilizado foi de 180 que corresponde a 21,13% e o segundo zoom foi de 520 que corresponde a 61,03% do zoom total que a câmara consegue efetuar.

Para a obtenção dos parâmetros intrínsecos das câmaras foi elaborado um algoritmo em linguagem C++, no Microsoft Visual Studio Community 2022, em que se recorreu à biblioteca OpenCV. A calibração foi realizada individualmente para cada câmara, onde se procedeu consoante o fluxograma da Figura 5.2.

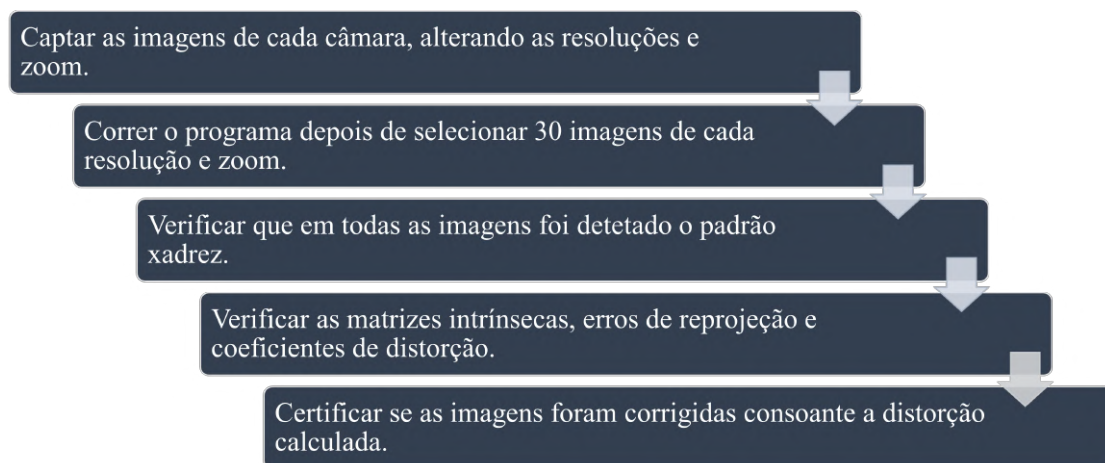


FIGURA 5.2: Fluxograma de procedimentos de calibração das câmaras.

O código em C++, começa por definir variáveis como o nome dos arquivos das imagens e o tamanho do tabuleiro de xadrez, de seguida cria um vetor de pontos tridimensional e um vetor de pontos bidimensional de forma a armazenar os pontos encontrados em cada imagem. Os cantos do tabuleiro são encontrados através da função 'cv::findChessboardCorners()' e através da função 'cv::cornerSubPix()' é refinado a sua localização. Dentro dos cantos já definidos é efetua-se a busca do padrão

## 5.1. Calibração das câmaras

---

de xadrez, caso se encontre o padrão na imagem, o valor dos pontos tridimensionais são guardados na matriz  $Q$ . O padrão encontrado em cada imagem encontra-se ilustrado na Figura 5.3.

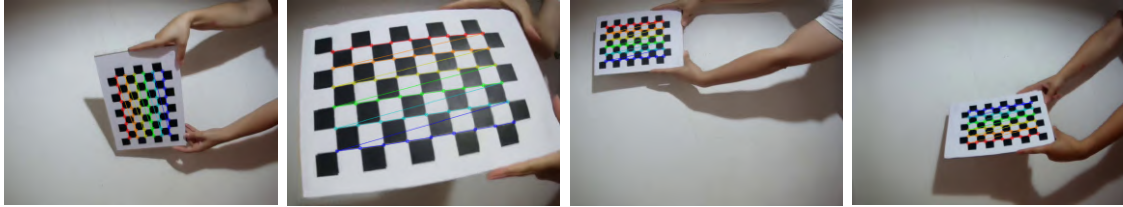


FIGURA 5.3: Imagens com o padrão xadrez reconhecido com as câmaras sem zoom.

Após terem sido reconhecidos os padrões em todas as imagens, calcula-se a matriz intrínseca, o erro de projeção e os coeficientes de distorção. O código em linguagem C++ referido anteriormente tem a capacidade de mostrar as imagens com as correções da distorção calculadas anteriormente, como demonstrado na Figura 5.4.

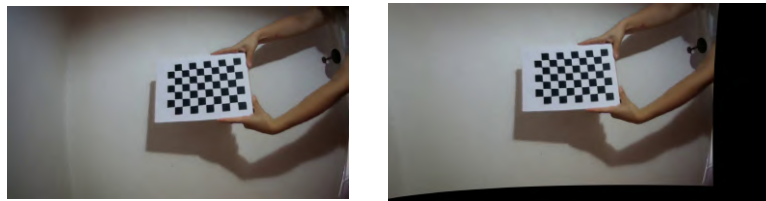


FIGURA 5.4: Imagem antes da correção (esquerda) e imagem após a correção (direita) sem zoom.

A calibração foi efetuada para as câmaras sem zoom e com zoom. Os resultados anteriores descritos foram obtidos com as câmaras sem estarem com zoom. Os resultados seguintes foram para as câmaras com zoom em 180, ou seja, nos 21,13% presentes na Figura 5.5.

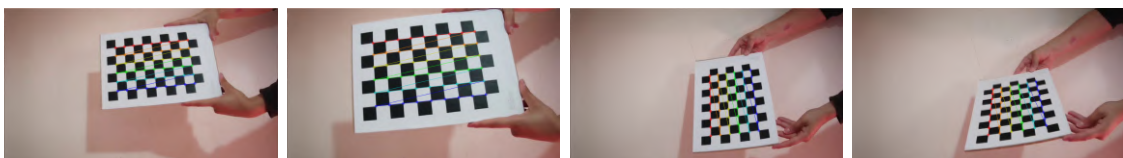


FIGURA 5.5: Imagens com o padrão xadrez reconhecido com as câmaras com 21,13% zoom.

É possível verificar a imagem antes e após a correção (direita) com 21,13% zoom, de acordo com a Figura 5.6.



FIGURA 5.6: Imagem antes da correção (esquerda) e imagem após a correção (direita) com 21,13% zoom.

No ensaios com o nível de zoom em 520, que corresponde a 61,03% da capacidade total de ampliação das câmaras, como se pode observar na Figura 5.7.

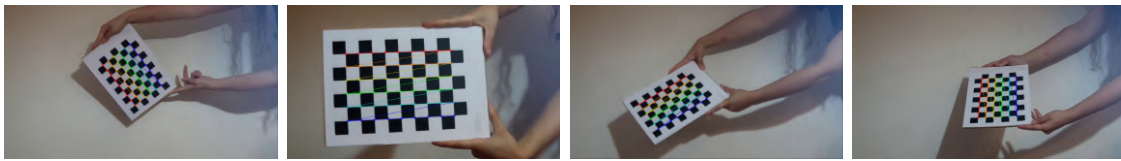


FIGURA 5.7: Imagens com o padrão xadrez reconhecido com as câmaras com 61,03% zoom.

No que diz respeito às correções, que se observam na Figura 5.8, é possível verificar que são atenuam a distorção que a imagem inicial sofreu tendo em conta a distorção da lente.

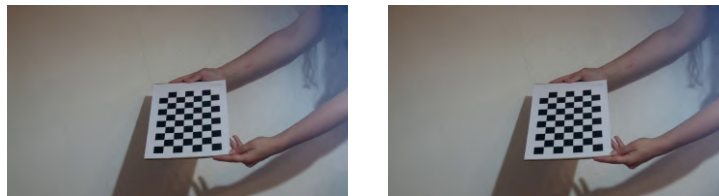


FIGURA 5.8: Imagem antes da correção (esquerda) e imagem após a correção (direita) com 61,03% zoom.

É possível concluir que à medida que o zoom ótico diminui a distância focal também diminui de forma a obter um ampliação menor do cenário em questão. Os valores das matrizes tendo em conta o zoom e resolução encontram-se na Tabela 5.1.

TABELA 5.1: Parâmetros intrínsecos das câmaras: Erro de reprojeção, matrizes intrínsecas e coeficientes de distorção de cada câmara.

		Sem Zoom	Com 21,13% Zoom	Com 61,03% Zoom
Câmara 1	720p	Reprojection error = 0.488674 $K = \begin{bmatrix} 762.70117 & 0 & 639.5 \\ 0 & 762.70117 & 359.5 \\ 0 & 0 & 1 \end{bmatrix}$ $k = [-0.326955 \ 0.160193 \ 0 \ 0 \ 0]$	Reprojection error = 0.451012 $K = \begin{bmatrix} 1095.4264 & 0 & 639.5 \\ 0 & 1095.4264 & 359.5 \\ 0 & 0 & 1 \end{bmatrix}$ $k = [-0.334413 \ 0.259229 \ 0 \ 0 \ 0]$	Reprojection error = 0.412219 $K = \begin{bmatrix} 1764.3596 & 0 & 639.5 \\ 0 & 1764.3596 & 359.5 \\ 0 & 0 & 1 \end{bmatrix}$ $k = [-0.25729 \ -0.975039 \ 0 \ 0 \ 0]$
	1080p	Reprojection error = 0.81397 $K = \begin{bmatrix} 1214.3004 & 0 & 959.5 \\ 0 & 1214.3004 & 539.5 \\ 0 & 0 & 1 \end{bmatrix}$ $k = [-0.377662 \ 0.221886 \ 0 \ 0 \ 0]$	Reprojection error = 0.554948 $K = \begin{bmatrix} 1438.4447 & 0 & 959.5 \\ 0 & 1438.4447 & 539.5 \\ 0 & 0 & 1 \end{bmatrix}$ $k = [-0.274352 \ 0.282446 \ 0 \ 0 \ 0]$	Reprojection error = 0.496807 $K = \begin{bmatrix} 2640.0886 & 0 & 959.5 \\ 0 & 2640.0886 & 539.5 \\ 0 & 0 & 1 \end{bmatrix}$ $k = [-0.565936 \ 3.78786 \ 0 \ 0 \ 0]$
Câmara 2	720p	Reprojection error = 0.46182 $K = \begin{bmatrix} 774.46228 & 0 & 639.5 \\ 0 & 774.46228 & 359.5 \\ 0 & 0 & 1 \end{bmatrix}$ $k = [-0.347411 \ 0.152147 \ 0 \ 0 \ 0]$	Reprojection error = 0.523265 $K = \begin{bmatrix} 1092.3313 & 0 & 639.5 \\ 0 & 1092.3313 & 359.5 \\ 0 & 0 & 1 \end{bmatrix}$ $k = [-0.338718 \ -0.423347 \ 0 \ 0 \ 0]$	Reprojection error = 0.328422 $K = \begin{bmatrix} 1750.4392 & 0 & 639.5 \\ 0 & 1750.4392 & 359.5 \\ 0 & 0 & 1 \end{bmatrix}$ $k = [-0.489927 \ 2.1282 \ 0 \ 0 \ 0]$
	1080p	Reprojection error = 0.644841 $K = \begin{bmatrix} 1103.5454 & 0 & 959.5 \\ 0 & 1103.5454 & 539.5 \\ 0 & 0 & 1 \end{bmatrix}$ $k = [-0.311596 \ 0.142598 \ 0 \ 0 \ 0]$	Reprojection error = 0.513501 $K = \begin{bmatrix} 1690.7893 & 0 & 959.5 \\ 0 & 1690.7893 & 539.5 \\ 0 & 0 & 1 \end{bmatrix}$ $k = [-0.363301 \ 0.297577 \ 0 \ 0 \ 0]$	Reprojection error = 0.441015 $K = \begin{bmatrix} 2633.2617 & 0 & 959.5 \\ 0 & 2633.2617 & 539.5 \\ 0 & 0 & 1 \end{bmatrix}$ $k = [-0.410263 \ 3.85595 \ 0 \ 0 \ 0]$

No contexto de análise dos parâmetros extrínsecos, inicialmente mediram-se os pontos tridimensionais, considerando as dimensões reais do marcador H, em metros, pela ordem determinada na Subsecção 3.2.2 de acordo com a Figura 3.5. Assim sendo, os pontos encontram-se ordenados através da seguinte sequência:

$$\begin{aligned}
 H_1 &= (-0.24; 0.24; 0) & H_7 &= (0.24; -0.24; 0) \\
 H_2 &= (-0.24; 0.12; 0) & H_8 &= (0.24; -0.12; 0) \\
 H_3 &= (-0.06; 0.12; 0) & H_9 &= (0.06; -0.12; 0) \\
 H_4 &= (-0.06; -0.12; 0) & H_{10} &= (0.06; 0.12; 0) \\
 H_5 &= (-0.24; -0.12; 0) & H_{11} &= (0.24; 0.12; 0) \\
 H_6 &= (-0.24; -0.24; 0) & H_{12} &= (0.24; 0.24; 0)
 \end{aligned}$$

Os pontos bidimensionais da aquisição de dados 2 forma medidos em imagens de cada câmara no momento do ensaio a bordo do navio, de acordo com a Figura 5.9. Os resultados foram os seguintes, respetivamente para a câmara 1 e câmara 2, em píxeis:

$$\begin{aligned}
 h_{1,1} &= (1096; 628) & h_{1,2} &= (164; 600) \\
 h_{2,1} &= (1110; 623) & h_{2,2} &= (174; 603) \\
 h_{3,1} &= (1090; 619) & h_{3,2} &= (201; 601) \\
 h_{4,1} &= (1120; 612) & h_{4,2} &= (224; 607) \\
 h_{5,1} &= (1145; 615) & h_{5,2} &= (203; 610) \\
 h_{6,1} &= (1156; 611) & h_{6,2} &= (217; 614) \\
 h_{7,1} &= (1100; 602) & h_{7,2} &= (273; 605) \\
 h_{8,1} &= (1086; 605) & h_{8,2} &= (261; 602) \\
 h_{9,1} &= (1105; 609) & h_{9,2} &= (237; 605) \\
 h_{10,1} &= (1081; 615) & h_{10,2} &= (213; 599) \\
 h_{11,1} &= (1055; 611) & h_{11,2} &= (231; 596) \\
 h_{12,1} &= (1041; 615) & h_{12,2} &= (219; 593)
 \end{aligned}$$



FIGURA 5.9: Pontos bidimensionais da aquisição de dados 2.

Os pontos bidimensionais da aquisição 3 também foram registados nas imagens de cada câmara durante os ensaios realizados na Força Aérea Portuguesa, de acordo com a Figura 5.10. Os resultados correspondentes da câmara 1 e 2, em píxeis apresentam-se como:

$$\begin{aligned}
 h_{1,1} &= (1042; 625) & h_{1,2} &= (216; 629) \\
 h_{2,1} &= (1044; 624) & h_{2,2} &= (223; 632) \\
 h_{3,1} &= (1043; 623) & h_{3,2} &= (231; 632) \\
 h_{4,1} &= (1050; 621) & h_{4,2} &= (239; 634) \\
 h_{5,1} &= (1059; 622) & h_{5,2} &= (238; 635) \\
 h_{6,1} &= (1061; 621) & h_{6,2} &= (242; 636) \\
 h_{7,1} &= (1049; 619) & h_{7,2} &= (255; 636) \\
 h_{8,1} &= (1046; 620) & h_{8,2} &= (249; 634) \\
 h_{9,1} &= (1048; 620) & h_{9,2} &= (244; 634) \\
 h_{10,1} &= (1042; 622) & h_{10,2} &= (234; 631) \\
 h_{11,1} &= (1034; 622) & h_{11,2} &= (238; 631) \\
 h_{12,1} &= (1029; 622) & h_{12,2} &= (234; 629)
 \end{aligned}$$

Os pontos bidimensionais da aquisição de dados 3, são ilustrados de acordo com a Figura 5.10. Em que as duas primeiras imagens foram capturadas pela câmara 1 e as duas da direita foram capturadas pela câmara 2.



FIGURA 5.10: Pontos bidimensionais da aquisição de dados 3.

Uma vez que os parâmetros intrínsecos das câmaras foram calculados, é necessário descobrir os parâmetros extrínsecos. Deste modo, empregou-se o método PnP, sendo apenas aplicável nas aquisições de dados 2 e 3, das Secções 4.4 e 4.5, respetivamente. Deve-se ao facto de que foram as aquisições em que foi implementado o uso de um marcador H com dimensões conhecidas.

Através de um código gerado em Python, no Visual Studio Code, adquirem-se os parâmetros extrínsecos de cada câmara. O código começa por introduzir os parâmetros intrínsecos de cada câmara, bem como os respetivos coeficientes de distorção. Introduzem-se os pontos de referência do marcador H tridimensionais e os pontos de bidimensionais dos pontos de cada imagem.

Após ter a matriz dos parâmetros intrínsecos e extrínsecos, o código encontra a matriz dos parâmetros extrínsecos de acordo com a função `'cv2.solvePnP(Ransac())'`, descrita na Subsecção 3.2.2. Para a conversão do vetor de rotação numa matriz de rotação com dimensões  $3 \times 3$  recorre-se à função `'cv2.Rodrigues()'`.

Através da matriz de rotação obtida pela função `'cv2.Rodrigues()'` e a matriz de translação obtida pela função `'cv2.solvePnP(Ransac())'`, é possível extrair a matriz final dos parâmetros extrínsecos de cada câmara para cada aquisição diferente. Assim, de acordo com a Tabela 5.2 é possível verificar os valores das matrizes extrínsecas de ambas as aquisições:

TABELA 5.2: Matrizes dos parâmetros extrínsecos das duas câmaras.

		Aquisição de dados 2 (Navio)	Aquisição de dados 3 (AFA)
Câmara 1	720p	Matriz dos parâmetros extrínsecos	Matriz dos parâmetros extrínsecos
		$= \begin{bmatrix} -0.14870862 & -0.98405085 & -0.09762 & 2.7315897 \\ 0.18275966 & 0.06966779 & -0.98068614 & 1.51611248 \\ 0.971846 & -0.16367748 & 0.1694846 & 3.89044932 \end{bmatrix}$	$= \begin{bmatrix} -0.01425766 & -0.99730862 & -0.0719182 & 8.46568534 \\ 0.22213765 & 0.06696882 & -0.97271272 & 5.46736229 \\ 0.97491106 & -0.02984435 & 0.22058498 & 13.86630375 \end{bmatrix}$
Câmara 2	720p	Matriz dos parâmetros extrínsecos	Matriz dos parâmetros extrínsecos
		$= \begin{bmatrix} 0.29225881 & -0.95602766 & 0.02441103 & -2.52851655 \\ 0.15345378 & 0.0216857 & -0.98791784 & 1.46044934 \\ 0.94394742 & 0.29247366 & 0.1530439 & 3.98840853 \end{bmatrix}$	$= \begin{bmatrix} 0.03281279 & -0.99159796 & 0.12512718 & -6.8278819 \\ 0.30074559 & -0.1095963 & -0.94738627 & 4.63466513 \\ 0.95313977 & 0.06871783 & 0.29462255 & 11.26220236 \end{bmatrix}$

## 5.2 Alinhamento dos dados para avaliação de algoritmos de detecção e estimação

O alinhamento dos registos temporais desempenha uma papel crucial na avaliação dos algoritmos de detecção e estimação. É fundamental garantir o sincronismo dos registos temporais capturados por ambas as câmaras e pelos recetores GPS. A precisão temporal é um elemento determinante na coerência e na integridade dos dados analisados, tornando-se imperativo estabelecer uma correspondência temporal exata no sistema.

Para a sincronização entre os dados temporais de ambas as câmaras e o dos dados da telemetria do UAV, é necessário estabelecer um momento conhecido entre ambos os registos. Esse momento pode ser efetuado no momento exato em que o UAV efetua a aterragem. Assim sendo, é possível identificar um dos *frames* em que o UAV está a efetuar a aterragem.

Dentro do contexto, dada a ocorrência de três momentos de aquisições de dados heterogêneos e a utilização de três UAVs diferentes, é crucial observar cada aquisição isoladamente. Em cada um dos cenários, são extraídas duas imagens, uma de cada câmara e o respetivo *timestamp*, no momento em que o UAV aterra. Extrai-se também o *timestamp* da telemetria do UAV em que os dados demonstrem uma aceleração e altitude praticamente nulas. Após a recolha dos três dados temporais é possível determinar o *offset* da aquisição. Desta forma, é possível proceder à correção dos dados das câmaras com os dados do UAV, alcançando um sincronismo temporal do sistema.

Na primeira aquisição de dados, existiram variados ensaios em que o UAV aterrou na parada. Recorreu-se ao ensaio 7 para obter um dos momentos de aterragem de forma a efetuar a sincronização entre ambas as câmaras, e também com a telemetria do UAV. Nas imagens da Figura 5.11 é possível identificar o momento de aterragem em ambas as câmaras.



FIGURA 5.11: Aterragem do UAV da aquisição 1 da câmara 1 (esquerda) e da câmara 2 (direita).

## 5.2. Alinhamento dos dados para avaliação de algoritmos de detecção e estimação

Como se verifica, observando a Tabela 5.3 existe uma diferença de sincronismo entre a câmara 1 e 2 de 0,026 segundos. Entre a câmara 1 e os dados de telemetria do UAV existe uma assincronia de 0,016 segundos. E entre a câmara 2 e os dados de telemetria do UAV existe uma assincronia de 0,010 segundos.

TABELA 5.3: Momento de aterragem da aquisição de dados 1.

Momento de aterragem Aquisição de dados 1	Grupo Data-Hora	Tempo em Microsegundos
Câmara 1	31-05-2023 16:11:42.629	1685545902629000
Câmara 2	31-05-2023 16:11:42.655	1685545902655000
Telemetria UAV	31-05-2023 16:11:42.645	1685545902645000

Durante a aquisição de dados 2, efetuada a bordo do navio, é possível identificar um momento comum de aproximação ao convés do navio no ensaio 1, como se pode observar na Figura 5.12. Nos dados de telemetria do UAV considerou-se o momento em que a altitude e a aceleração eram menores, assim considerou-se que o momento em que o UAV aterrou, de acordo com a Tabela 5.4.

TABELA 5.4: Momento de aterragem da aquisição de dados 2.

Momento de aterragem Aquisição de dados 2	Grupo Data-Hora	Tempo em Microsegundos
Câmara 1	06-07-2023 14:14:31.077	1688649271077000
Câmara 2	06-07-2023 14:14:31.078	1688649271078000
Telemetria UAV	06-07-2023 14:19:48.000	1688645988000000



FIGURA 5.12: Aterragem do UAV da aquisição 2 da câmara 1 (esquerda) e da câmara 2 (direita).

A diferença de sincronismo entre a câmara 1 e a câmara 2, na aquisição de dados 2, é de 0,001 segundos. Entre a câmara 1 e os tempo da telemetria do UAV é de 317 segundos, ou seja, 5 minutos e 17 segundos. A diferença entre a câmara 2 e

a telemetria do UAV é de 316,922 segundos, ou seja, 5 minutos, 16 segundos e 922 milissegundos.

Na aquisição de dados 3, em diversos ensaios é possível verificar o UAV aterrar. Optou-se pelo ensaio 13, uma vez que é o ensaio em que se efetuou uma observação mais minuciosa e detalhada. Observando a Figura 5.13, de onde se retirou o *timestamp* de ambas imagens de cada câmara. Conjuntamente com os dados da telemetria do UAV pode-se analisar o momento em que o UAV efetuou a aterragem, de acordo com a Tabela 5.5.

TABELA 5.5: Momento de aterragem da aquisição de dados 3 antes da correção.

Momento de aterragem Aquisição de dados 3	Grupo Data-Hora	Tempo em Microssegundos
Câmara 1	24-07-2023 12:58:55.082	1690199935082000
Câmara 2	2023-07-24 12:58:51.429	1690199931429000
Telemetria UAV	24-07-2023 12:58:52.840	1690199932840000



FIGURA 5.13: Aterragem do UAV da aquisição 3 da câmara 1 (esquerda) e da câmara 2 (direita).

A diferença de tempo entre as imagens capturadas pela câmara 1 e pela câmara 2 é de 3,653 segundos. Entre a câmara 1 e o tempo da telemetria do UAV é de 2,242 segundos. E entre a câmara 2 e o tempo da telemetria do UAV é de 1,411 segundos.

Através dos dados obtidos por ambas as câmaras durante os diferentes cenários é possível verificar numa fase inicial que as câmaras obtêm ligeiras divergências entre dados de tempo de captação. Desta forma, é notável a assincronia dos *timestamps*, que mesmo corrigindo a diferença entre os mesmos irá gerar resultados não desejáveis perante os dois métodos utilizados.

Enquanto que na primeira e segunda aquisição de dados a diferença sincronismo entre câmaras encontra-se na ordem dos milissegundos, não se efetua nenhuma

### 5.3. Anotação do dataset com geração de *bounding boxes* para o UAV

---

alteração nos dados obtidos, no entanto é necessário obter a concordância das câmaras com a telemetria do UAV.

Na terceira aquisição, como as câmaras se encontram desfasadas em 3.653 segundo efetua-se o ajuste tendo como referencial o *timestamp* obtido pela câmara 1. Os resultados apresentados posteriormente terão em consideração o alinhamento no tempo, de acordo com a atualização descrita pela Tabela 5.6.

TABELA 5.6: Momento de aterragem da aquisição de dados 3 após a correção.

Momento de aterragem Aquisição de dados 3	Grupo Data-Hora	Tempo em Microsegundos
Câmara 1	24-07-2023 12:58:55.082	1690199935082000
Câmara 2	24-07-2023 12:58:55.082	1690199935082000
Telemetria UAV	24-07-2023 12:58:52.840	1690199932840000

Apesar da correção de *timestamps* entre ambas as câmaras e também com a telemetria do UAV em diferentes cenários, existe um erro na atribuição do grupo data-hora de cada imagem captada. O grupo data-hora associado a cada imagem corresponde ao tempo em que cada *frame* foi guardado no computador e não ao tempo real em que cada imagem foi captada. Deste modo, não é possível aferir a posição exata do UAV com o respetivo *timestamp*.

### 5.3 Anotação do dataset com geração de *bounding boxes* para o UAV

As anotações do dataset foram efetuadas através da plataforma Roboflow. A plataforma Roboflow permite que a anotação de imagens não processadas desenvolvam modelos de computação visual personalizados. Uma das suas funcionalidades é a capacidade de anotação de imagens, sendo possível criar mais do que uma *bounding box* por imagem e catalogá-la de acordo com o objeto que se observa. é possível também exportar os dados em vários formatos diferentes, como COCO, CreateML, Pascal VOC e YOLO. Desta forma, a plataforma Roboflow tem a capacidade de acordo com as necessidades de deteção e classificação de objetos.

Os dados de dois dos ensaios efetuado em Mafra, aquisição de dados 1 de acordo com a Secção 4.3, perfaz um total de 9 869 imagens. No ensaio 10 foram anotadas no total 3 016 imagens, em que 1 521 foram da câmara 1 e 1 496 imagens

da câmara 2. O ensaio 11 perfaz 6 853 imagens no total, das quais 3 465 imagens são da câmara 1 e 3 388 imagens da câmara 2.

As imagens foram anotadas manualmente, tendo em conta a posição em que o UAV se encontrava, como demonstra a Figura 5.14.



FIGURA 5.14: Imagens da aquisição de dados 1 com *bounding box* do UAV capturada pela câmara 1 (esquerda) e câmara 2 (direita).

As anotações efetuadas na aquisição de dados 2, foram todas do ensaio 3 e totalizam um conjunto de 2 118 imagens, em que 1 205 imagens pertencem à câmara 1 e 913 imagens à câmara 2. A Figura 5.15 ilustra duas imagens anotadas, uma de cada câmara.

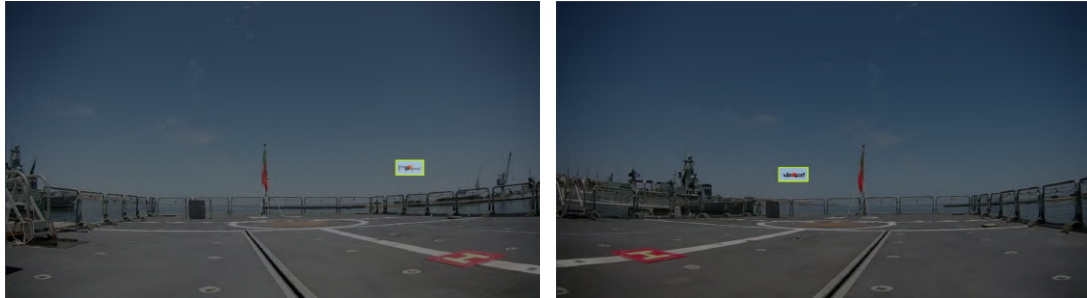


FIGURA 5.15: Imagens da aquisição de dados 2 com *bounding box* do UAV capturada pela câmara 1 (esquerda) e câmara 2 (direita)

Na aquisição de dados 3, foram anotadas no total 7138 imagens do ensaio 13. Sendo que, 3 541 imagens foram retiradas da câmara 1 e 3 597 imagens da câmara 2. Duas imagens, uma de cada câmara, que foram anotadas, tendo em conta a posição em que o UAV se encontrava, como se visualiza na Figura 5.16.

Consequentemente, abrangendo todas as imagens das três aquisições de dados que foram anotadas, obtém-se somando um total de 19 125 imagens.

Após terminar a anotação de todas as imagens das diversas aquisições de dados, as informações relativas às *bounding boxes*, considerando o fim para que



FIGURA 5.16: Imagens da aquisição de dados 3 com *bounding box* do UAV capturada pela câmara 1 (esquerda) e câmara 2 (direita)

serão necessárias, foram guardadas em três formatos distintos: COCO, YOLOv8 e Tensorflow *Object Detection*.

O formato *Common Objects in Context* (COCO) é configurado em *JavaScript Object Notation* (JSON), que consiste num conjunto de anotações que descrevem objetos presentes nas imagens, no presente caso o UAV. O formato é frequentemente utilizado na implementação de algoritmos para a visão computacional.

Os campos que o formato COCO gera diversos campos de anotação como: a identificação da anotação, a identificação de cada imagem anotada; a classe do objeto; as medidas das *bounding boxes* (xmin, ymin, largura e altura) e a área de cada *bounding box*.

O formato *You Only Look Once* versão 8 (YOLOv8) não é apenas um formato de anotação como o anterior, considera-se uma arquitetura de rede neuronal convolucional. O YOLOv8 é a versão mais recente da arquitetura YOLO, em que a eficiência e rapidez são as principais melhorias.

A arquitetura YOLO divide cada imagem em secções e cada célula pertencente prevê a existência de *bounding box* e a classe a que esta pertence. Após a deteção de cada *bounding box* é possível extrair campos de informação como:

- Identificação da anotação;
- Identificação da imagem;
- A classe do objeto;
- As coordenadas  $x$  e  $y$  que são o centro do objeto em relação à largura e altura, respetivamente;
- A largura e a altura da *bounding box* e a área de cada *bounding box*.

A arquitetura YOLOv8 e o formato TensorFlow *Object Detection* têm o mesmo objetivo, ambos lidam com modelos treinados de forma a realizar detecção de objetos para diversas aplicações. O TensorFlow *Object Detection* utiliza a biblioteca da Tensorflow da Google.

No TensorFlow *Object Detection*, após a anotação de cada *bounding box*, são extraídos os seguintes campos de informação: o nome de cada imagem; a largura e altura de cada *bounding box* e as medidas das *bounding boxes* (xmin, ymin, xmax e ymax).

Em resumo, a seleção dos formatos COCO e YOLOv8 para a anotação do *dataset* é sustentada no amplo espectro de aplicabilidade em diversas áreas, aliado ao fator de serem os métodos mais prevalentes na comunidade. O método do TensorFlow *Object Detection* foi escolhido pelas informações que fornecia, facilitando nos cálculos no Capítulo 6.

## 5.4 Descrição do formato de dados utilizados para o dataset e disponibilização de endereço web para acesso ao público

O *dataset* final é composto no geral por três *dataset* de cenários distintos: a aquisição de dados 1 efetuada na Escola Prática de Infantaria, aquisição de dados 2 efetuada a bordo da fragata NRP Corte Real e a aquisição de dados 3 efetuada em Base Aérea n<sup>o</sup>1.

O conjunto dos três *datasets* adquiridos e descritos nas Subsecções 5.4.1, 5.4.2 e 5.4.3, encontram-se disponíveis para consulta e potencial utilização para fins diversos, através do link seguinte: <https://outlook/EN-Dataset@marinha.pt>.

### 5.4.1 Organização da aquisição de dados 1

O *dataset* da aquisição de dados 1 é constituído uma *checklist* dedicada para o local específico de forma a efetuar a verificação de todos os passos e minimizar o erro.

No caso específico deste *dataset* existe um ficheiro Excel com a telemetria do UAV ao longo de todos os voos. O *dataset* tem pastas referentes para cada ensaio, que estão numerados de 1 a 11. Na pasta de cada ensaio encontra-se também um ficheiro texto com a descrição específica correspondente e um ficheiro Excel com as

características das câmaras, distâncias e medidas GPS retiradas durante o ensaio. Cada pasta do ensaio contém duas pastas com as imagens capturadas por cada câmara e a hora a que teve o início da aquisição de imagens.

Dentro da pasta de cada câmara encontram-se dois ficheiros texto: o ficheiro com o nome 'camera1.txt' contém a informação data-hora a que cada *frame* foi capturado no formato Ano-Mês-Dia\_Hora-Minutos-Segundos-Milissegundos; o segundo ficheiro com o nome 'cam1\_Frame\_TimeMicro' contém a conversão do grupo data-hora para microssegundos com o respetivo *frame*, os mesmo ficheiros existem para a câmara 2 apenas muda o nome.

O presente *dataset* contém uma pasta com as *bounding boxes* efetuadas no programa Roboflow, acessíveis em dois formatos diferentes: COCO e YOLOv8. É possível descarregar os dados das *bounding boxes* através do ficheiro Python que se encontra dentro de cada pasta respetiva.

## 5.4.2 Organização da aquisição de dados 2

O *dataset* da aquisição de dados 2 constituído por uma *checklist* dedicada para o local específico de forma a efetuar a verificação de todos os passos e minimizar o erro. No caso específico deste dataset existe um ficheiro Excel com a telemetria do UAV ao longo de todos os voos. O *dataset* tem pastas referentes para cada ensaio, que estão numerados de 1 a 5. Na pasta de cada ensaio encontra-se também um ficheiro texto com a descrição específica correspondente e um ficheiro Excel com as características das câmaras, distâncias e medidas GPS retiradas durante o ensaio. Cada pasta do ensaio contém duas pastas com as imagens capturadas por cada câmara e a hora a que teve o início da aquisição de imagens.

Dentro da pasta de cada câmara encontram-se dois ficheiros texto: o ficheiro com o nome 'camera1.txt' contém a informação data-hora a que cada *frame* foi capturado no formato Ano-Mês-Dia\_Hora-Minutos-Segundos-Milissegundos; o segundo ficheiro com o nome 'cam1\_Frame\_TimeMicro' contém a conversão do grupo data-hora para microssegundos com o respetivo *frame*, os mesmo ficheiros existem para a câmara 2 apenas muda o nome.

O *dataset* de aquisição de dados 2 inclui uma pasta com as *bounding boxes* anotadas no programa Roboflow. As anotações são disponibilizadas em três formatos: COCO, YOLOv8 e Tensorflow *Object Detection*. A obtenção dos dados das *bounding boxes* pode ser realizado através do ficheiro Python que se encontra em cada pasta correspondente.

### 5.4.3 Organização da aquisição de dados 3

O *dataset* da aquisição de dados 3 é constituído por uma ficha de ensaios gerais com as características das câmaras, distâncias e medidas GPS retiradas durante o ensaio. O *dataset* tem uma *checklist* dedicada para o local específico de forma a efetuar a verificação de todos os passos e minimizar o erro. No caso específico deste *dataset* existe uma pasta com a telemetria do UAV ao longo dos voos. O *dataset* tem pastas referentes para cada ensaio, que estão numerados de 1 a 14. Na pasta de cada ensaio encontra-se também um ficheiro texto com a descrição específica correspondente. Cada pasta do ensaio contém duas pastas com as imagens capturadas por cada câmara e a hora a que teve o início da aquisição de imagens.

Dentro da pasta de cada câmara encontram-se dois ficheiros texto: o ficheiro com o nome 'camera1.txt' contém a informação data-hora a que cada *frame* foi capturado no formato Ano-Mês-Dia\_Hora-Minutos-Segundos-Milissegundos; o segundo ficheiro com o nome 'cam1\_Frame\_TimeMicro' contém a conversão do grupo data-hora para microssegundos com o respetivo frame, os mesmo ficheiros existem para a câmara 2 apenas muda o nome.

O pasta 'Ensaio13' contém um ficheiro Excel com a telemetria do UAV específica daquela aquisição de dados tendo em conta o tempo que decorreram as aquisições de imagens.

O *dataset* de aquisição de dados 3 contém uma pasta com as *bounding boxes* efetuadas no programa Roboflow, acessíveis em três formatos diferentes: COCO, YOLOv8 e Tensorflow. É possível descarregar os dados das *bounding boxes* através do ficheiro Python que se encontra dentro de cada pasta respetiva.

# Capítulo 6

## Resultados Experimentais

O Capítulo 6 explora a utilização de tipos de dados distintos para a estimativa da posição do UAV e respetivas velocidades. Em que na Secção 6.1 a aplicação prática tem como base dados sintéticos com o objetivo de demonstrar a utilização de dois métodos para a estimativa da posição e velocidade do UAV. Enquanto que, na Secção 6.2 recorre-se aos dados obtidos através de aproximações reais do UAV. A Secção 6.3 descreve ambos os métodos quando são impostos *outliers* na trajetória. Na Secção 6.4 efetua-se discussão final dos resultados obtidos. Ao longo deste Capítulo, recorrem-se aos referenciais estabelecidos previamente na Figura 3.3 da Secção 3.1.2.

### 6.1 Teste experimental dos dados sintéticos

Por forma a verificar a implementação, foram testados os algoritmos propostos utilizando dados sintéticos. A utilização de dados sintéticos permite evidenciar o desempenho dos dois sistemas propostos, o filtro de Kalman estendido e a triangulação *stereo*.

Através da análise detalhada dos dados sintéticos pelos dois métodos, é possível determinar um que demonstre baixo erro na estimativa da posição e velocidade do UAV em tempo real. A comparação entre ambos os métodos permitirá avaliar qual dos métodos apresenta menos erros em cenários sintéticos.

A presente Secção reparte-se em dois momentos de análise, em que o primeiro recorre aos dados sintéticos de forma a demonstra a utilidade da triangulação *stereo* para a estimativa da posição tridimensional e a velocidade do UAV. Na continuação aborda-se o FKE e as estimativa da posição e velocidade do UAV calculadas.

Considera-se que ambas as câmaras estão sincronizadas temporalmente, garantindo a consistência. A sincronização também abrange os dados sintéticos de ambas as câmaras e a telemetria sintética do UAV.

Todos os cálculos de estimativa de posição do UAV com dados sintéticos, tiveram como base os resultados obtidos com o aparato experimental da base de dados 3, de acordo com a Secção 5.1. Embora os dados sejam sintéticos, a calibração das câmaras foi realizada com base em dados reais, considerando o marcador H real.

Os dados que foram propostos para a câmara 1 e para a câmara 2, simulam a posição do UAV através do centro das *bounding boxes*, numa imagem com resolução de 1280 por 720 píxeis. As posições do UAV simuladas consideram as coordenadas na imagem  $u$  e  $v$  de cada câmara, de acordo com a Figura 6.1:

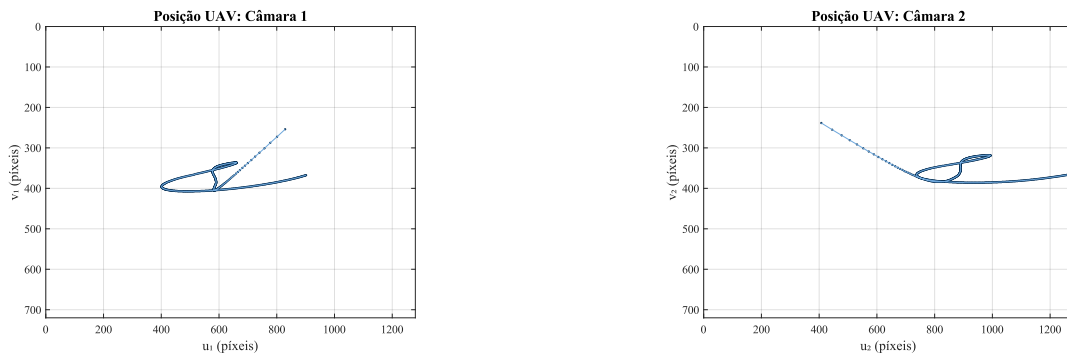


FIGURA 6.1: Posição do UAV das câmaras 1 (esquerda) e 2 (direita) dos dados sintéticos.

As coordenadas na imagem  $u_1$  e  $v_1$  são obtidas pelas posições simuladas do UAV na imagem da câmara 1 e as coordenadas na imagem  $u_2$  e  $v_2$  da câmara 2. As representações gráficas da Figura 6.2 descrevem a evolução das variáveis ao longo do tempo. Analisando os gráficos seguintes, verifica-se que ao longo do tempo o UAV efetua uma trajetória contínua, sem as limitações dos dados reais, pois não se verificam falhas na aquisição da posição do UAV ao longo da trajetória.

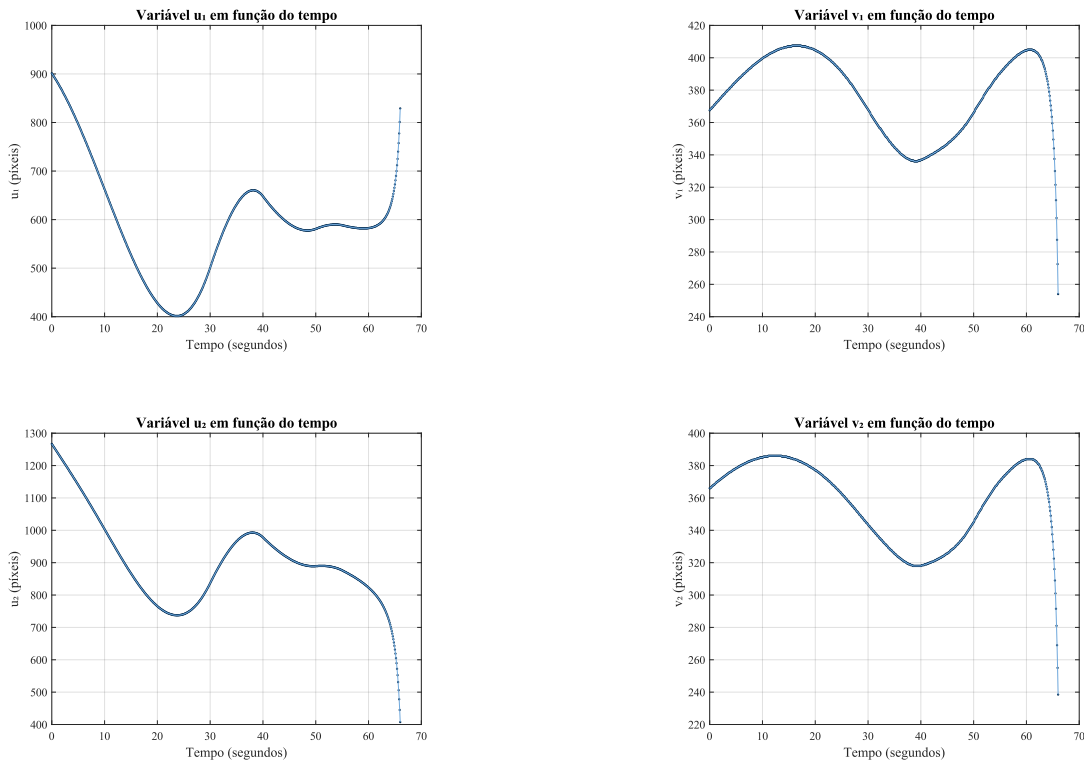


FIGURA 6.2: Coordenadas na imagem  $u_1$ ,  $v_1$ ,  $u_2$  e  $v_2$  em função do tempo dos dados sintéticos

Nos gráficos da Subsecção 6.1.1 e 6.1.2, as trajetórias a azul representam os valores dos dados sintéticos observados pelas câmaras, enquanto que as trajetórias a vermelho descrevem a trajetória efetua pelo *ground-truth* sintético do UAV.

### 6.1.1 Triangulação *stereo* para estimativa de posição 3D do UAV e respetiva velocidade

A triangulação *stereo* pode ser utilizada com o objetivo que estimar a posição e velocidade tridimensional de um UAV. A triangulação *stereo* recorre ao princípios da geometria epipolar de forma a conseguir extrair dados simplifcadas da localização do UAV bidimensional em imagens obtidas através das câmaras, como descrito anteriormente na Subsecção 3.4.1.

Os procedimentos de aquisição da posição do UAV, descritos na Secção 3.4, também são aplicáveis para a estimativa da velocidade a que este se desloca ao longo da aproximação. Desta forma, é possível calcular a posição do UAV, tendo em consideração as correspondências 2D-3D obtidas, como ilustra a Figura 6.3.

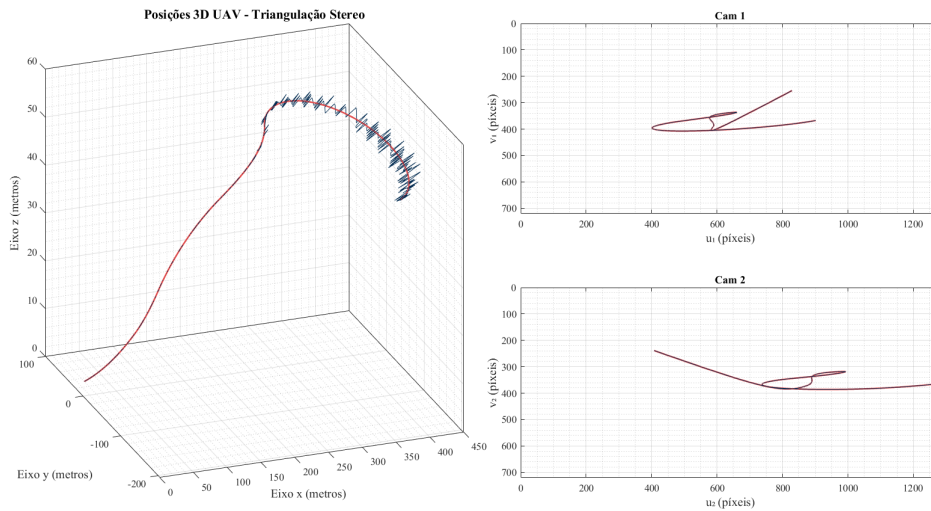


FIGURA 6.3: Posição 2D e 3D do UAV através da triangulação *stereo* dos dados sintéticos sem ruído.

Na ausência de ruído, o método de triangulação *stereo* demonstra a capacidade de seguir a trajetória com precisão entre os 100 e os 250 metros de distância do UAV ao marcador H, a uma altitude de aproximadamente 45 metros. Como se visualiza através do gráfico ilustrado na Figura 6.4:

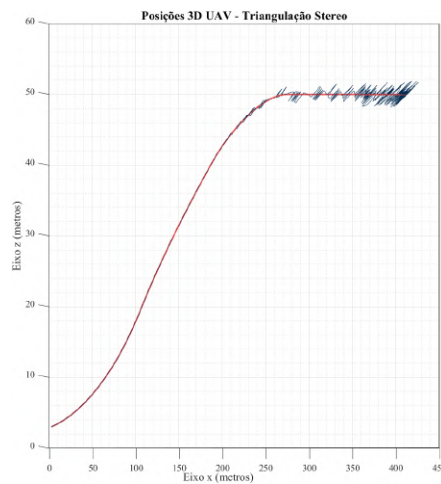


FIGURA 6.4: Estabilidade da posição 3D dos dados sintéticos sem ruído.

O método de triangulação *stereo* é sensível à presença de ruído. Nos dados sintéticos proposto é possível inserir ruído nas posições adquiridas em ambas as câmaras. Atribuiu-se um desvio de píxeis variáveis, e através da função `'randn()'` no MATLAB é gerada uma matriz aleatória com valores normalmente distribuídos, de acordo com a quantidade de ruído que se pretenda inserir.

De acordo com os gráficos ilustrados na Figura 6.5 é possível verificar que para níveis de ruídos diferentes, o cálculo da posição 3D do UAV também varia. À medida que o ruído aumenta, a exatidão das correspondências entre pontos nas imagens das câmaras reduz. Causando que as linhas epipolares definidas tenham mais incerteza na estimativa da posição 3D.

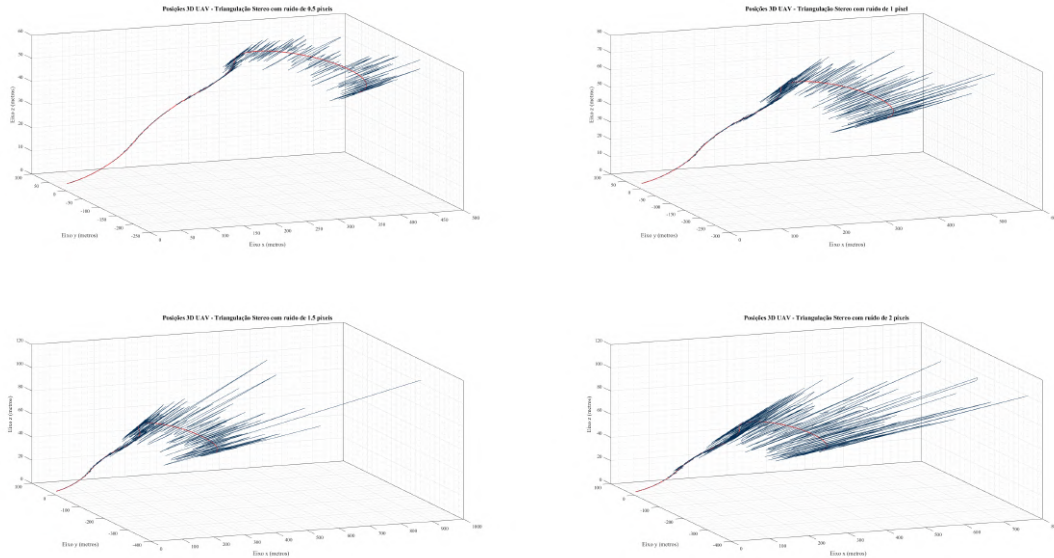


FIGURA 6.5: Evolução do ruído na triangulação *stereo*

Analisando o gráfico da Figura 6.5 observa-se que conforme a adição de ruído, iniciada em 0,5 píxeis, que incrementa gradualmente de meio em meio píxel até alcançar a variação de 2 píxeis, verifica-se que a posição estimada tem uma incerteza gradualmente crescente à medida que o ruído também aumenta. A partir de um erro de 1,5 píxeis existe um erro inicial de 500 metros o que não permite aterragem do UAV com sucesso.

Através das posições tridimensionais calculadas, é possível determinar de acordo com o referido anteriormente na Subsecção 3.4.1. Com os dados sintéticos calcula-se a diferença entre a posição atual e a posição anterior e divide-se pelo intervalo de tempo, que no caso é de 0.1 segundos. Assim sendo, a velocidade nos três eixos em função do tempo ilustra-se de acordo com o gráfico da Figura 6.6.

Com base na análise do gráfico observa-se que as velocidades iniciais apresentam valores elevados, o que corresponde à correção contínua da posição do UAV. A partir do instante  $t = 40$  segundos as oscilações na velocidade diminuem, a redução deve-se às menores variações na posição do UAV.

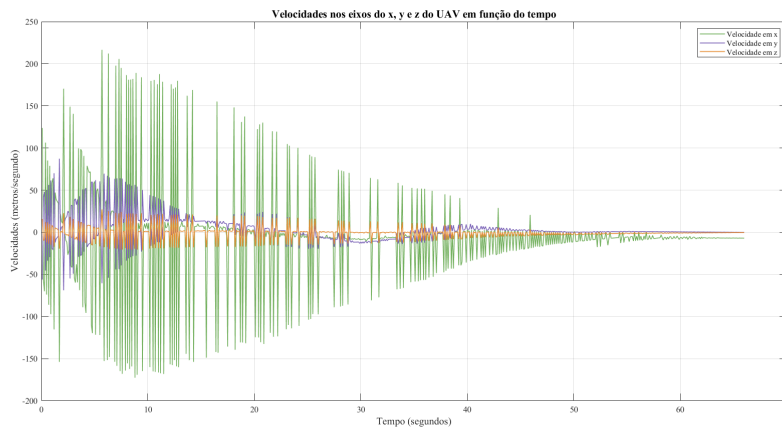


FIGURA 6.6: Velocidades do UAV em função do tempo dos dados sintéticos através da triangulação *stereo*.

Considerando os dados do *ground truth* do UAV e a trajetória 3D calculada através da triangulação *stereo*, é possível calcular o erro entre as trajetórias. Os erros foram calculados através de *Root Mean Square Error* (RMSE) e de *Mean Absolute Error* (MAE), de acordo com o Anexo III, conforme ilustrado no gráfico da Figura 6.7.

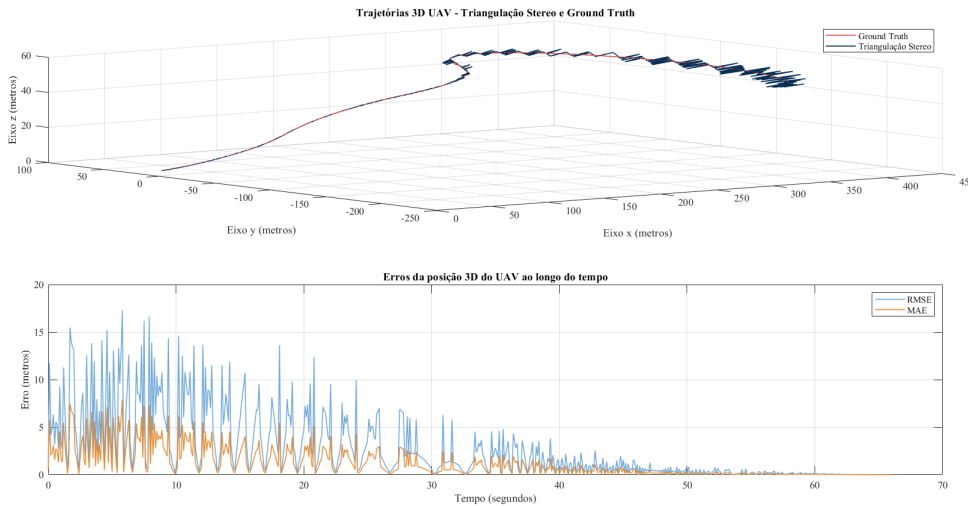


FIGURA 6.7: Erros da posição 3D do UAV ao longo do tempo na triangulação *stereo*.

Observando a Figura 6.7, conclui-se que quando o UAV se encontra longe do marcador, numa fase inicial da trajetória, qualquer erro ligeiro na detecção, mesmo que de um ou dois píxeis, corresponde a um erro de vários metros na estimativa 3D, e que tende a diminuir à medida que o UAV se aproxima do navio.

### 6.1.2 Filtro de Kalman estendido para estimativa de posição 3D do UAV e respetiva velocidade

No âmbito da estimativa da posição tridimensional e velocidade do UAV, recorre-se ao FKE. Como descrito na Subsecção 3.4.2, o FKE através das fase de predição e atualização consegue prever estimativas da posição do UAV em tempo real. O sistema à medida que recolhe novos dados acerca da posição do UAV, atualiza a sua estimativa de localização do UAV. O filtro de Kalman foi também aplicado para estimar a velocidade a que o UAV se aproxima do local de aterragem.

Ambas as estimativas, da posição e velocidade, têm em consideração a dinâmica do movimento e as incertezas correspondentes às medições efetuadas. Quanto mais precisa for a calibração de cada câmara e as medições efetuadas no local, maior será a precisão das estimativas efetuadas pelo FKE.

O filtro de Kalman estendido, contrariamente ao método anterior, necessita de um ponto de inicialização, ou seja, de um estado inicial antes de começar a incorporar observações reais. O estado inicial com que o filtro começa a calcular as posições é fundamental no desempenho do mesmo.

Ao começar o filtro com um ponto de inicialização conhecido, considerando por exemplo o marcador H, com as coordenada  $(0, 0, 0)$ , o filtro converge num menor espaço de tempo. O FKE consegue adquirir uma posição estimada próxima da posição real aos 200 metros de distância ao marcador H e cerca de 45 metros de altitude. Exemplificando graficamente de acordo com a Figura 6.8:

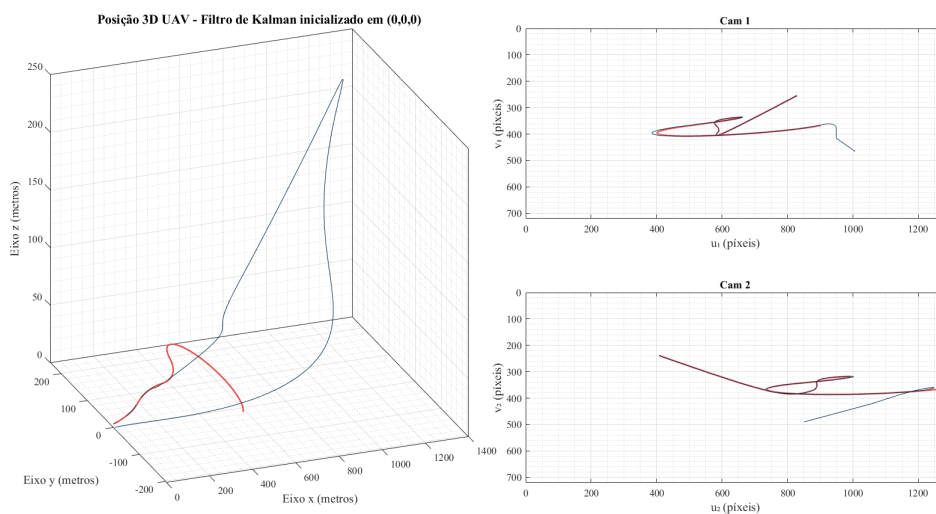


FIGURA 6.8: Inicialização do FKE no ponto do marcador H  $(0, 0, 0)$ .

Caso o FKE comece com um ponto muito distante, o filtro tem dificuldades em estimar a posição em casos extremos, como se encontra ilustrado na Figura 6.9. Acabando por só conseguir convergir um valor estimado na fase final da aterragem ou em casos drásticos o FKE acaba por divergir. No exemplo gráfico, o ponto de inicialização é de (1500, 1000, 550), e o FKE só consegue adquirir uma estimativa perto da realidade no fim do voo do UAV.

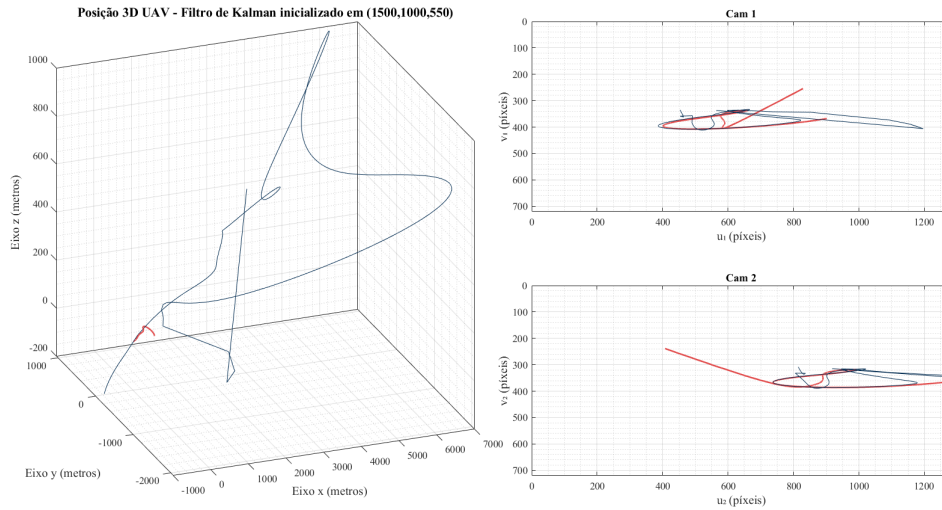


FIGURA 6.9: Inicialização do FKE num ponto distante.

A utilização da triangulação *stereo* como abordagem inicial para obter a posição inicial e adotando-a como ponto de partida para o FKE, resulta num seguimento mais preciso da trajetória real do UAV. No exemplo do gráfico seguinte, considerando o ponto de inicialização em (360, -200, 50) calculado através da triangulação *stereo*, obtém-se os resultados ilustrados na Figura 6.10.

## 6.1. Teste experimental dos dados sintéticos

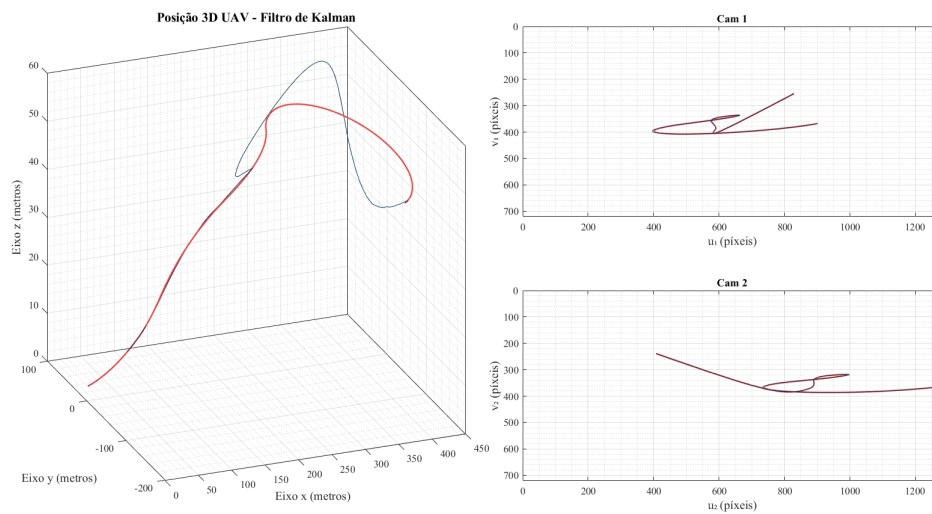


FIGURA 6.10: Inicialização do FKE num ponto calculado através da triangulação *stereo*.

Considerando o último exemplo como ponto de inicialização, o FKE varia considerando diferentes níveis de ruído. Caso se aumente o valor da variância do ruído do processo, o FKE irá permitir, de acordo com o gráfico ilustrado na Figura 6.11, que o estado estimado se desvie significativamente das estimativas anteriores. Em casos extremos de aumento da incerteza do processo, o modelo será praticamente igual ao método de triangulação *stereo* e perder-se-à a capacidade de previsão.

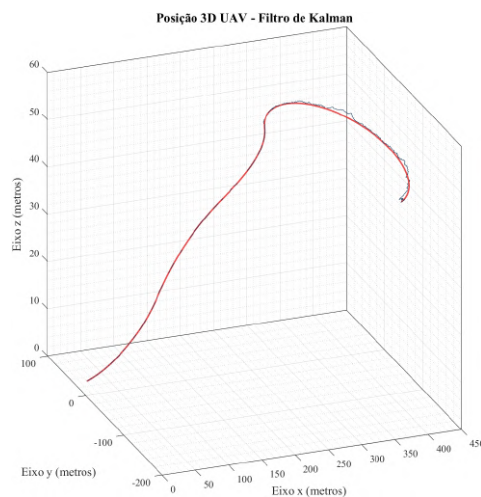


FIGURA 6.11: Aumento da variância do ruído do processo.

O método FKE também é sensível à presença de ruído nos dados de entrada. Adicionando os mesmos níveis de ruído utilizados no método anterior, obtém-se a

reação do filtro perante o ruído, de acordo com a Figura 6.12.

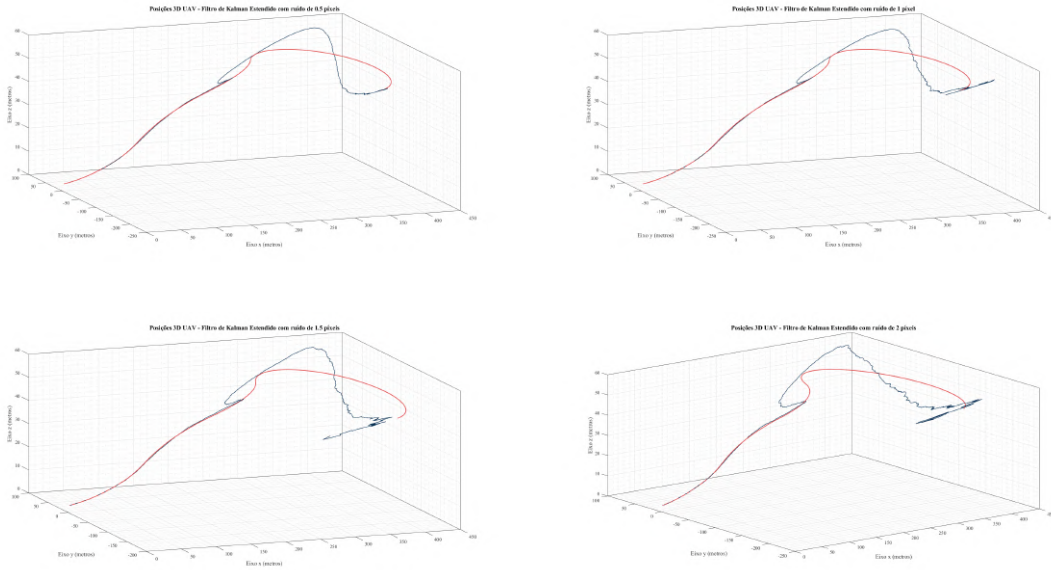


FIGURA 6.12: Evolução do ruído no filtro de Kalman estendido.

Caso se aumente a incerteza da variância do ruído das medições observadas pelas câmaras, o FKE irá priorizar a previsão do modelo. Em casos extremos, ao aumentar em demasia o peso atribuído às previsões de estimativa de posição, o sistema será praticamente de probabilidades. De acordo com a Figura 6.13, a atribuição do peso grande ao fator da previsão, resulta num afastamento maior entre a posição estimada e a real.

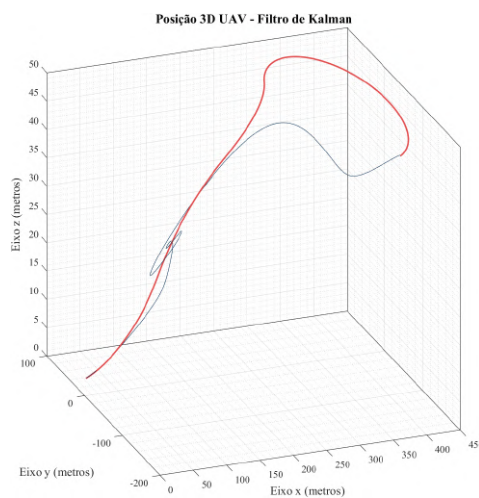


FIGURA 6.13: Aumento da variância do ruído da medição.

## 6.1. Teste experimental dos dados sintéticos

A velocidade inicial num FKE é estimada a partir das medições iniciais da posição do UAV. À medida que se adquirem mais dados de posição do UAV no filtro, a estimativa sofre ajustes de acordo com a informação que atualiza o filtro. Graficamente, as velocidades obtidas nos três eixos são visíveis de acordo com a Figura 6.14:

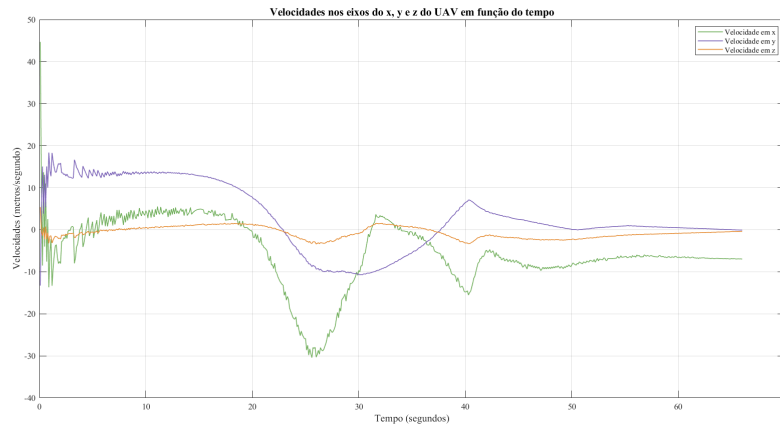


FIGURA 6.14: Velocidades do UAV em função do tempo dos dados sintéticos através do filtro de Kalman estendido.

Recorrendo aos dados do *ground truth* do UAV e da trajetória 3D calculada através do FKE, quantifica-se o erro entre as trajetórias, de acordo com o Anexo III, conforme ilustrado no gráfico da Figura 6.15.

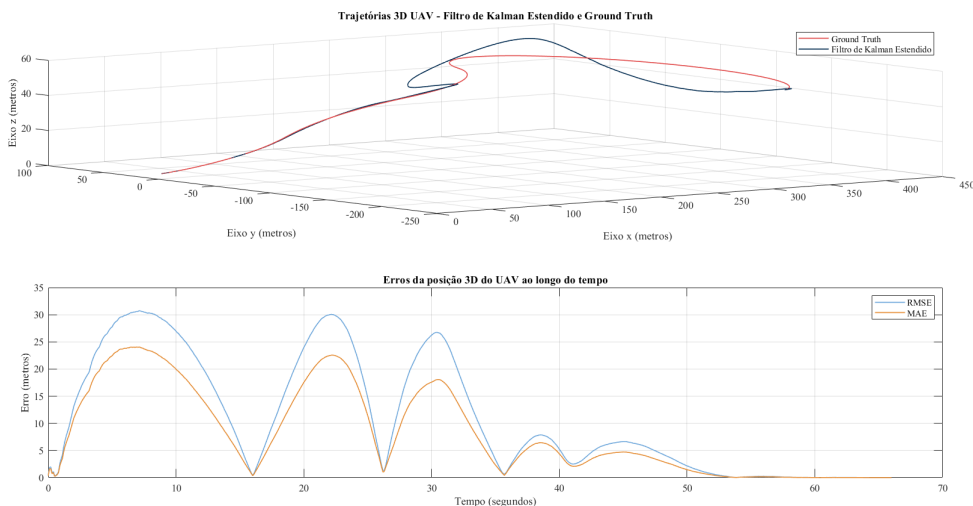


FIGURA 6.15: Erros da posição 3D do UAV ao longo do tempo no FKE.

Ao analisar a Figura 6.15 é possível aferir que, na fase inicial da trajetória o UAV, pequenas imprecisões na detecção acarretam erros significativos entre os 24 e 30 metros. É notável no gráfico, que à medida que o FKE cruza com a trajetória real o erro diminui.

## 6.2 Aplicação prática dos dados reais

No presente Secção é demonstrada a os dados reais são utilizados para obter estimativa de pose do UAV. A apresentação dos dados ilustram as aplicações que se consideraram mais adequadas por forma a contribuir para a aterragem do UAV e a sua utilidade no contexto real. No entanto, é fundamental realçar que os dados recolhidos tiveram limitações durante a aquisição, como a sincronização entre câmaras e os *timestamps* que têm um erro associado que aumenta a incerteza da estimativa. Em que a cada imagem foi captura. Ao invés, os *timestamps* refletem o momento em que as imagens foram guardadas no computador, gerando atrasos e assincronismo nos dados.

Não obstante dos problemas já mencionados, é possível visualizar ainda que de forma mesmo perceptível a aplicação prática dos dados. O objetivo é analisar os dados obtidos em dois ensaio distintos de dois cenários reais diferentes. Em que dos ensaios foi efetuado a bordo da fragata NRP Corte Real e o outro ensaio na Base Aérea nº1 da Força Aérea.

A análise desenvolvida permite retirar conclusões, apesar das limitações impostas pela discrepância nos dados, tendo em conta a identificação de padrões e tendências que o UAV tende a realizar para posteriores conclusões. Foram utilizados dois métodos distintos para a aplicação dos dados: o filtro de Kalman estendido (FKE) e a triangulação *stereo*.

A Secção divide-se em dois tópicos que abordam momentos de aquisição de dados distintos. Ambos recorrem à mesma estrutura de análise da aplicação prática, em que primeiro se descrevem as posições obtidas através das *bounding boxes* anotadas nas imagens capturadas por cada câmara. De seguida comparam-nas temporalmente e os erros associados que se constatarem. Aborda as velocidades calculadas em cada eixo, e simula a utilização dos dados pelos dois métodos propostos. Por fim, termina coma comparação das trajetórias obtidas com os dados GPS da posição do UAV obtidos através da telemetria.

### 6.2.1 Aplicação prática dos dados da aquisição de dados 2

Os dados da Secção 4.4 adquiridos a bordo da fragata NRP Corte Real, onde se assume que as câmaras 1 e 2 estão praticamente sincronizadas, ignorando a assincronia de 0,001 segundos. Desta forma, não é efetuado o sincronismo do *timestamp* entre câmaras.

Através das imagens capturadas pela câmara 1 e pela câmara 2, é possível obter a posição do UAV através do centro das *bounding boxes*. Sendo que as imagens captadas tem a resolução de 1280 por 720 píxeis, consegue-se verificar a posição do UAV tendo em consideração as coordenadas sobre o plano da imagem  $u$  e  $v$  de cada câmara, de acordo com a Figura 6.16.

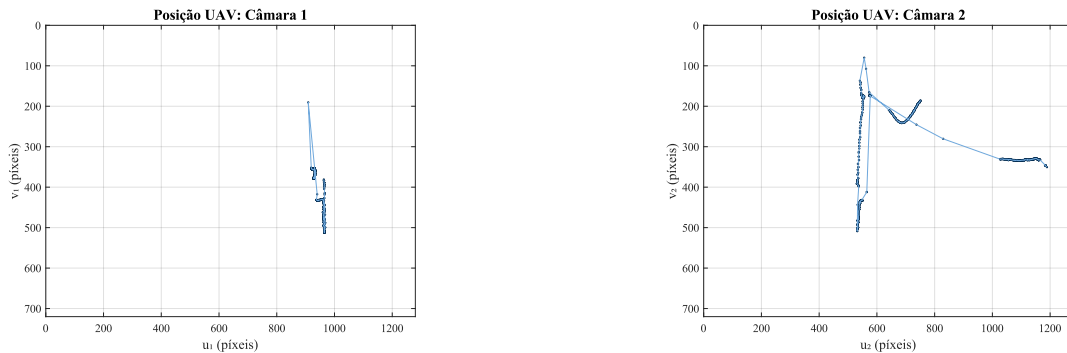


FIGURA 6.16: Posição do UAV das imagens capturadas pela câmara 1 (esquerda) e 2 (direita) na aquisição de dados 2.

Analisando a manobra do UAV, a aproximação foi efetuada pela alheta de bombordo do navio em direção ao convés. Através da análise do gráfico da câmara 1 é possível verificar que o UAV seguiu praticamente uma trajetória no azimute da câmara. Porém, ao analisar o gráfico proveniente da câmara 2 identifica-se que a trajetória de aproximação do UAV foi efetuada a partir de um ângulo oblíquo. Os resultados correspondem à trajetória real, no entanto, existiram falhas durante a aquisição de imagens, pois existem períodos da trajetória em que não foram obtidas posições do UAV.

As coordenadas na imagem  $u_1$  e  $v_1$  da sequência de imagens capturadas pela câmara 1 ao longo do ensaio, assim como as coordenadas na imagem  $u_2$  e  $v_2$ , ao longo do tempo podem ser representadas de acordo com a Figura 6.17.

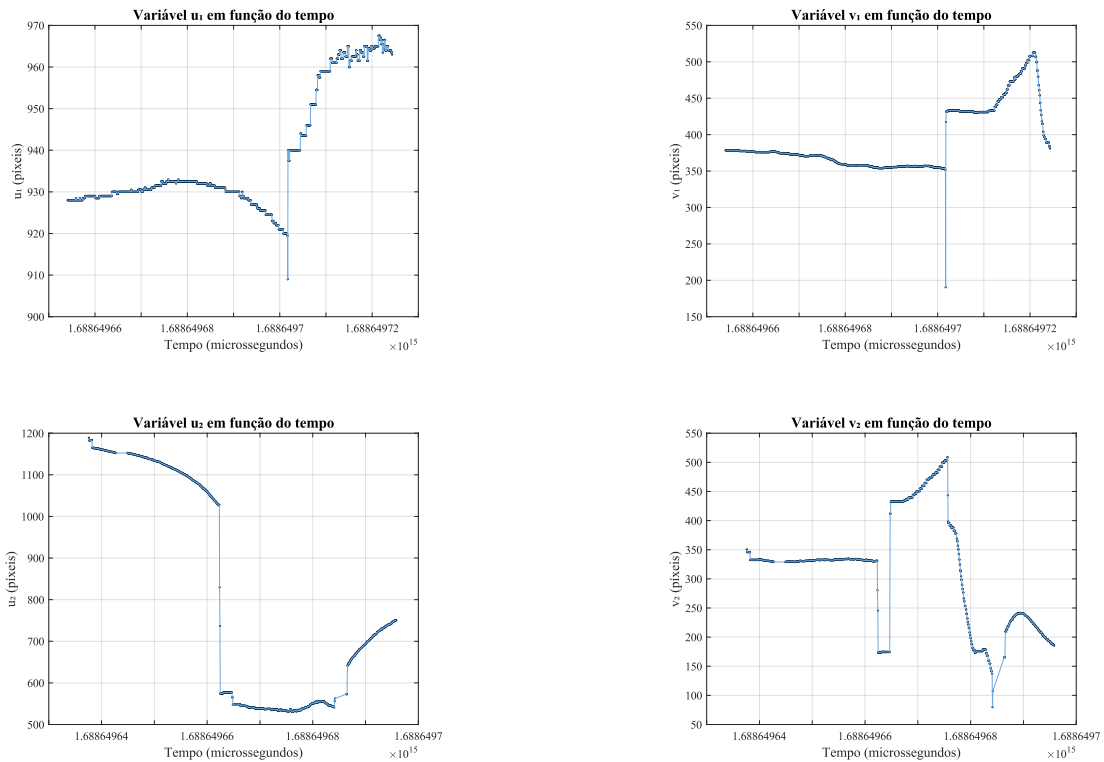


FIGURA 6.17: Coordenadas na imagem  $u_1$ ,  $v_1$ ,  $u_2$  e  $v_2$  em função do tempo da aquisição de dados 2

Tendo em consideração os gráficos descritos na Figura 6.17, é possível verificar que ao longo do tempo o UAV demonstra um comportamento irregular, pois não efetua uma trajetória contínua.

As características de aquisição de imagem das câmaras não permitiram uma recolha constante ao longo do tempo, o que se verifica nos saltos de posição repentinas do UAV, originando velocidade sem nexso do ponto de vista físico. É possível verificar as velocidades impraticáveis, considerando o UAV utilizado durante os ensaios, de acordo com o gráfico descrito na Figura 6.18:

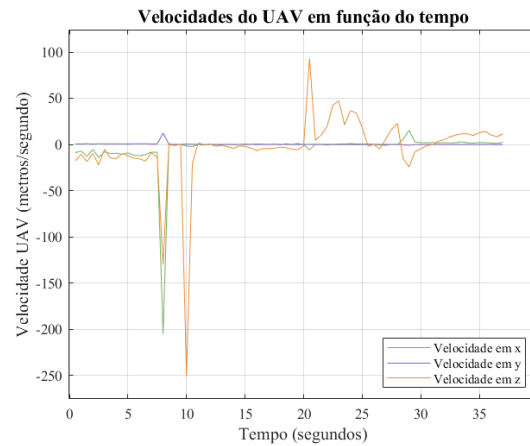


FIGURA 6.18: Velocidades do UAV em função do tempo da aquisição de dados 2.

Através da triangulação *stereo*, é possível obter a posição tridimensional do UAV à medida que este efetua a aproximação e aterragem. No gráfico da Figura 6.19 é possível verificar que o UAV aproxima-se praticamente de forma constante até aos 130 metros, tendo existido um salto até à próxima posição que encontra a 30 metros do marcador H. De seguida o UAV desce de forma mais lenta e volta a afastar-se. A trajetória de acordo com a triangulação *stereo* visualiza-se através da Figura 6.19:

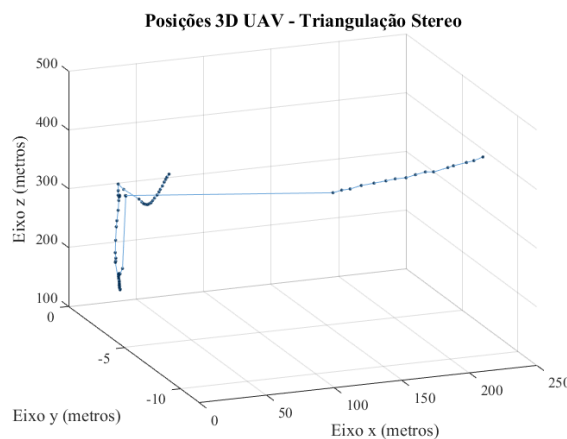


FIGURA 6.19: Posição 3D do UAV através da triangulação *stereo* da aquisição de dados 2.

De acordo com as velocidades inconstantes, neste caso o FKE tende a divergir. Apesar de se tratar de um método capaz de ser utilizado num sistema dinâmico não linear, o estado estimado é muito diferente do estado real, acrescentado o ruído e a incerteza. Desta forma, o FKE através da linearização e das aproximações não

consegue calcular com certeza a posição do UAV, e tende a divergir, como se verifica através da Figura 6.20.

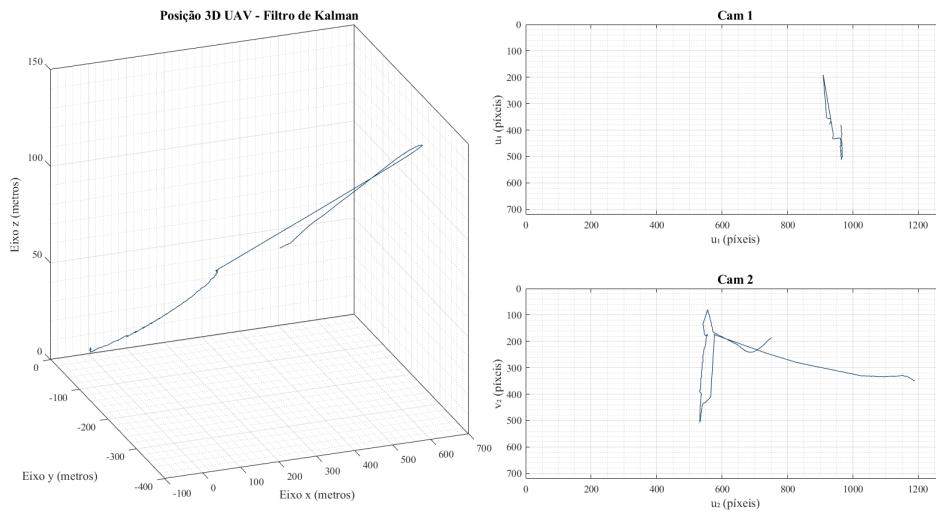


FIGURA 6.20: Posição 3D do UAV através do FKE da aquisição de dados 2.

Os dados obtidos pela telemetria registada pelo UAV durante o voo, estabeleceram uma relação de distância em relação ao marcador H durante a fase de aterragem, per, o que permite calcular a distância entre ambos. O processo foi realizado através de um código desenvolvido em *Matrix Laboratory* (MATLAB), que analisa a latitude e longitude e através da fórmula de Haversine, desenvolvida no Anexo II, calcula a distância entre o marcador H e o UAV. Graficamente a trajetória de acordo com a telemetria do UAV pode visualizar-se de acordo com a Figura 6.21.

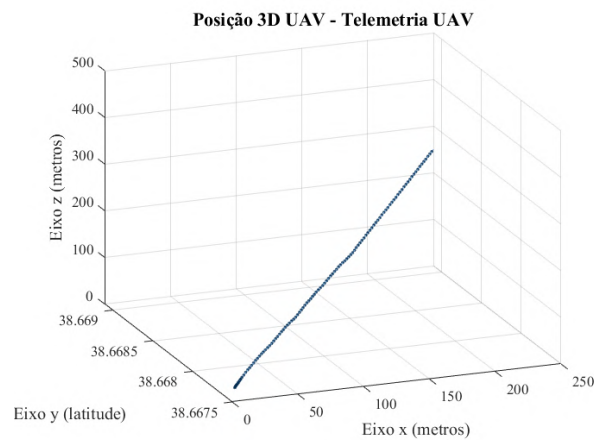


FIGURA 6.21: Posição 3D do UAV através da telemetria da aquisição de dados 2.

### 6.2.2 Aplicação prática dos dados da aquisição de dados 3

Na terceira aquisição, os dados foram obtidos na Base Aérea nº1 da Força Aérea, onde as câmaras 1 e 2 operaram em condições de sincronismo temporal diferentes, com uma discrepância de 3,653 segundos. Para a aplicação prática seguinte, os dados já estão alinhados temporalmente, tendo como referência a câmara 1. Relativamente à aquisição 2, acresce a análise prática de um *dataset* adquirido num cenário diferente e de um UAV com características distintas.

Através das imagens capturadas por ambas as câmaras, e posteriormente anotadas, é possível analisar as coordenadas na imagem  $u$  e  $v$ , considerando que a resolução de 1280 por 720 píxeis se manteve no presente cenário, de acordo com a Figura 6.22.

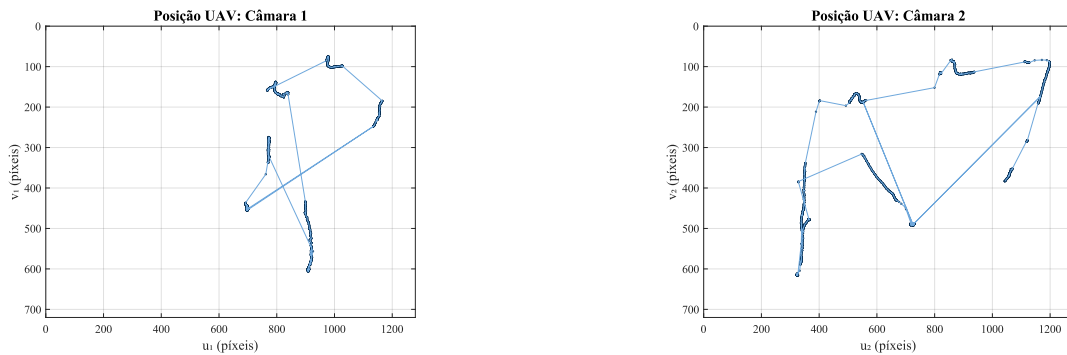


FIGURA 6.22: Posição do UAV das câmaras 1 (esquerda) e 2 (direita) na aquisição de dados 3.

Durante a aquisição de dados 3, é possível comparar os gráficos obtidos na Figura 6.22 e as imagens adquiridas. Em que o UAV se aproximou e efetuou a aterragem perto do marcador H, no pixel com as coordenadas  $[909.5, 607]$  na imagem da câmara 1 e com as coordenadas  $[323, 615.5]$  na imagem da câmara 2. De seguida o UAV descolou e afastou-se na direção onde iniciou a aproximação.

A representação gráfica da Figura 6.23 representa a evolução temporal das coordenadas na imagem  $u_1$  e  $v_1$  captadas pelas imagens da câmara 1 e as coordenadas na imagem  $u_2$  e  $v_2$  das imagens da câmara 2.

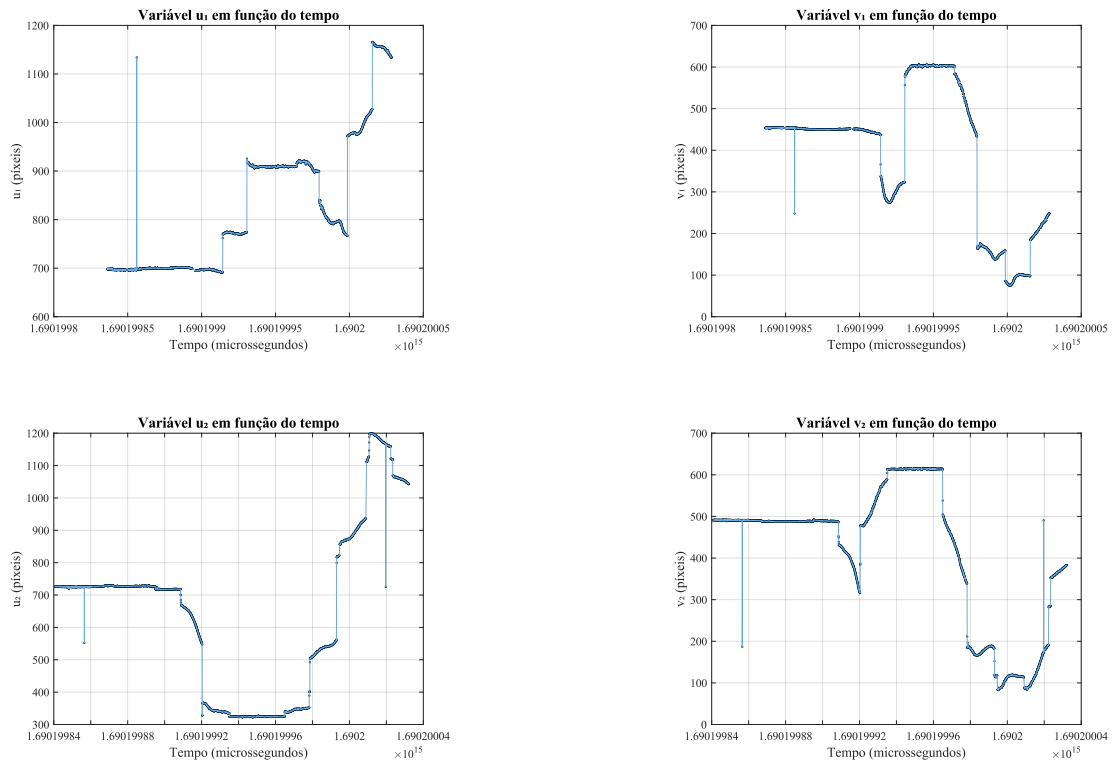


FIGURA 6.23: Coordenadas na imagem  $u_1$ ,  $v_1$ ,  $u_2$  e  $v_2$  em função do tempo da aquisição de dados 3

Após observação dos gráficos anteriores, é perceptível que o UAV efetua uma trajetória onde existem alguns períodos em que se verifica que as câmaras deixaram de adquirir imagens, devido às limitações anteriormente referidas. Na aquisição presente é notável também a existência de alguns *outliers* de posição devido à atribuição incorreta *timestamps* nas imagens, o que irá contribuir para o comportamento irregular dos métodos e velocidade do UAV.

A análise da trajetória delineada demonstra velocidades que, considerando o UAV utilizado nos testes, se revelam fisicamente impraticáveis. As velocidades resultam diretamente das lacunas ocorridas durante a aquisição de dados, inerentes às limitações das câmaras, como se observa na Figura 6.24:

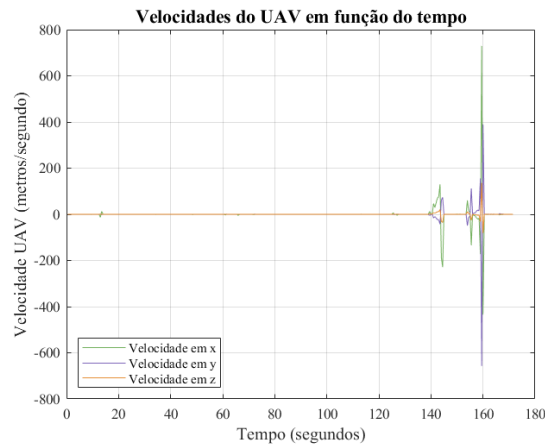


FIGURA 6.24: Velocidades do UAV em função do tempo da aquisição de dados 3.

Pelo método de triangulação *stereo*, o UAV inicia a aproximação a partir dos 225 metros e efetua a aterragem numa posição próxima do marcador H. De seguida, o UAV descola e afasta-se em direção ao ponto inicial onde se iniciou a trajetória. A representação visual da trajetória, não é tão perceptível como na aquisição de dados anterior, no entanto, de acordo com a triangulação *stereo* é possível visualizar na Figura 6.25.

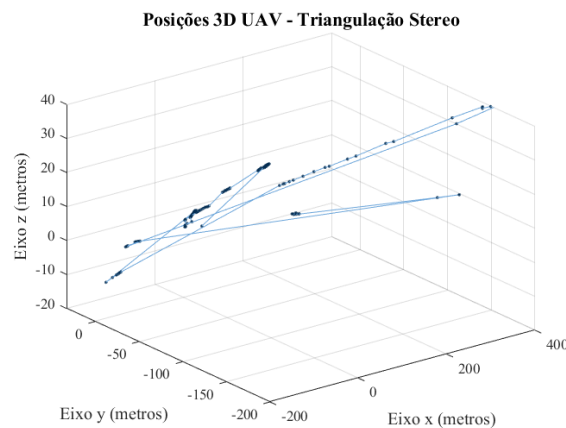


FIGURA 6.25: Posição 3D do UAV através da triangulação *stereo* da aquisição de dados 3.

Tendo em consideração as variações bruscas da velocidade, o FKE tende a divergir da trajetória efetuada pelo UAV. À semelhança da aquisição de dados anterior devido à divergência entre o estado estimado e o estado real em acumulação com o ruído e as incertezas inerentes, o modelo enfrenta dificuldades na determinação

correta da posição. De acordo com a Figura 6.26 denota-se as tendências a divergir a que o modelo está sujeito.

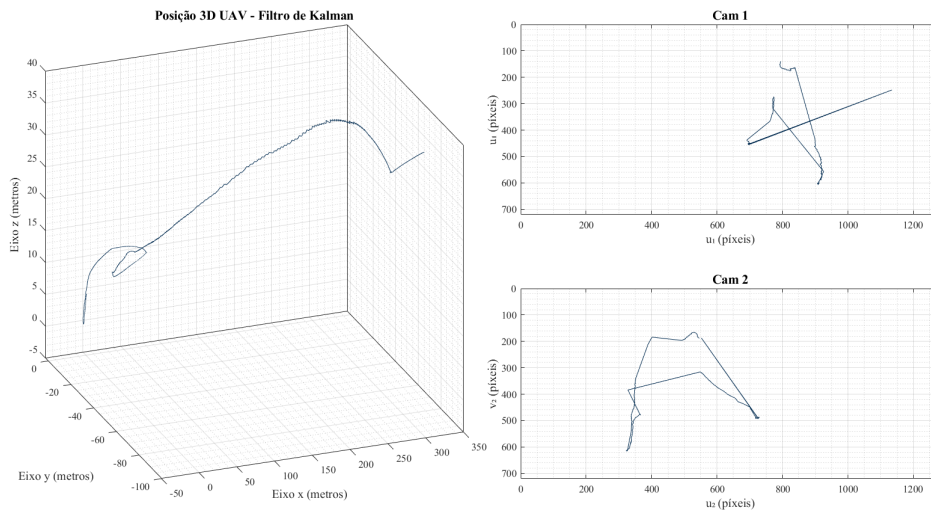


FIGURA 6.26: Posição 3D do UAV através do FKE da aquisição de dados 3.

A posição GPS é retirada dos dados de telemetria obtidos pelo UAV durante o voo. Através do código desenvolvido em MATLAB, calcula-se a distância ao marcador H através da equação de Haversine, descrita no Anexo II. Graficamente, a trajetória com as coordenadas GPS do UAV observa-se na Figura 6.27.

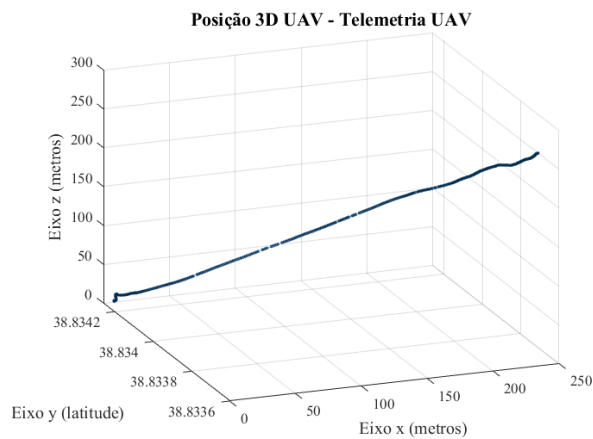


FIGURA 6.27: Posição 3D do UAV através da telemetria da aquisição de dados 3.

## 6.3 Simulação de outliers

*Outliers* são pontos que não se enquadram num conjunto de dados e podem induzir em erros e interferências nos cálculos dos modelos propostos (Hawkins 1980). Foram introduzidos alguns *outliers*, com valores distantes da trajetória real, nos dados sintéticos com o objetivo de simular falhas em cenários reais, e verificar como é que cada modelo reage a fatores imprevisíveis.

Os *outliers* impostos foram iguais para ambos os modelos, e visam a avaliar o desempenho de cada um perante a limitação. Deste modo, foram adicionados outliers ao longo de toda a trajetória. Os dados a que se recorreu para simular os *outliers* foram os dados sintéticos referidos na Secção 6.1.

Foi imposto o primeiro *outlier* logo no instante de  $t = 5,0$  segundos, o segundo foi adicionado meio percurso no instante  $t = 30,0$  segundos e por fim o último foi adicionado no fim da trajetória no instante de  $t = 66,0$  segundos.

A triangulação *stereo* demonstra ser sensível a conjuntos de dados com *outliers*, de acordo com a Figura 6.28. A presença de *outliers* causa degradação na precisão do método, pode introduzir distorção nas posições tridimensionais estimadas e na determinação de linhas epipolares.

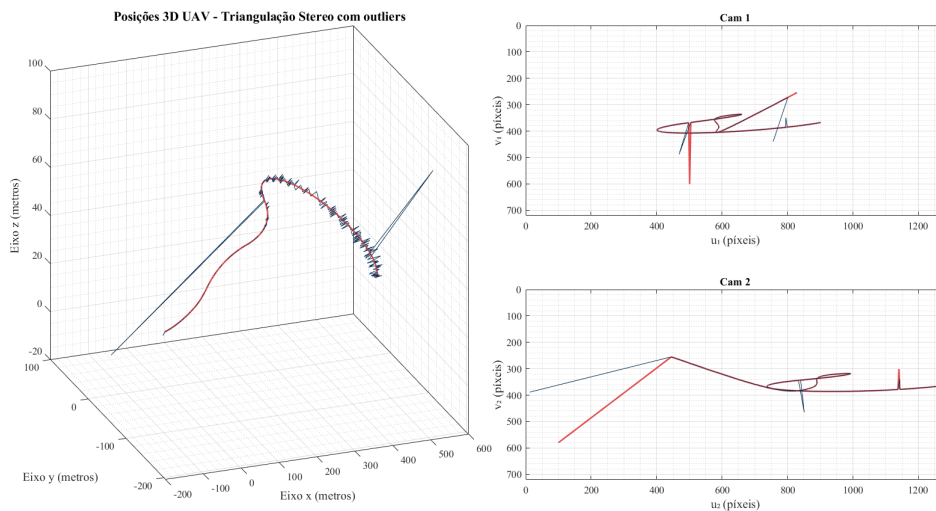


FIGURA 6.28: Triangulação *stereo* com *outliers*.

De acordo com o gráfico anterior, visualiza-se que a estimativa é que é influenciada pelos *outliers*, tanto no primeiro como no segundo, o método não assume que são pontos discrepantes e trata-os com a mesma importância dos restantes.

No entanto, é possível obter estimativas da posição mais precisas considerando a trajetória real a vermelho no gráfico, apesar da posição ser mais inconstante.

O filtro de Kalman estendido perante o primeiro *outlier* ainda tinha adquirido uma posição estimada perto da posição real. No entanto, o método após prever uma estimativa precisa de trajetória, não alterou significativamente a posição perante o segundo *outlier*, como é ilustrada na Figura 6.29.

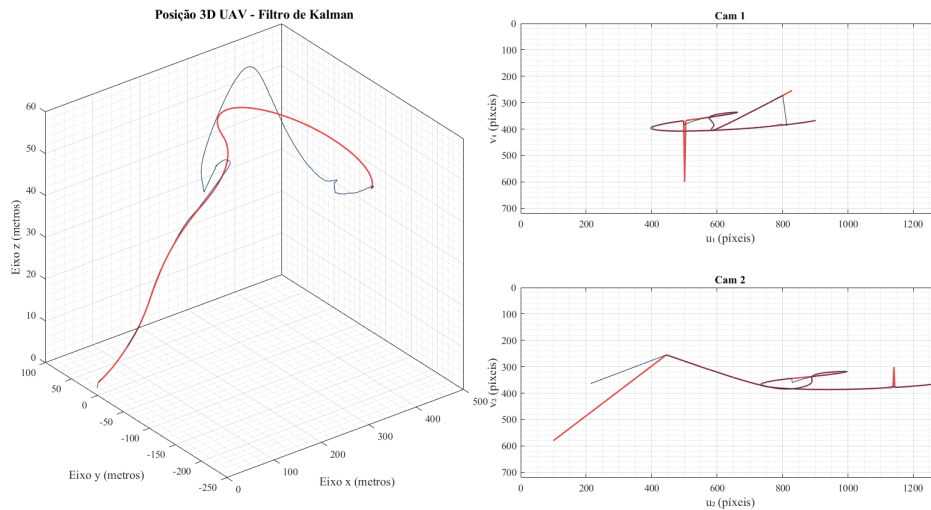


FIGURA 6.29: Filtro de Kalman estendido com *outliers*.

O comportamento do FKE perante *outliers* pode variar tendo em consideração o aumento das incertezas. Caso se aumente a incerteza do ruído das medições, implica que existe uma maior imprecisão nas observações, e consequentemente o filtro atribuirá um menor peso às medições individuais na atualização de pose do UAV, como se exemplifica de acordo com a Figura 6.30:

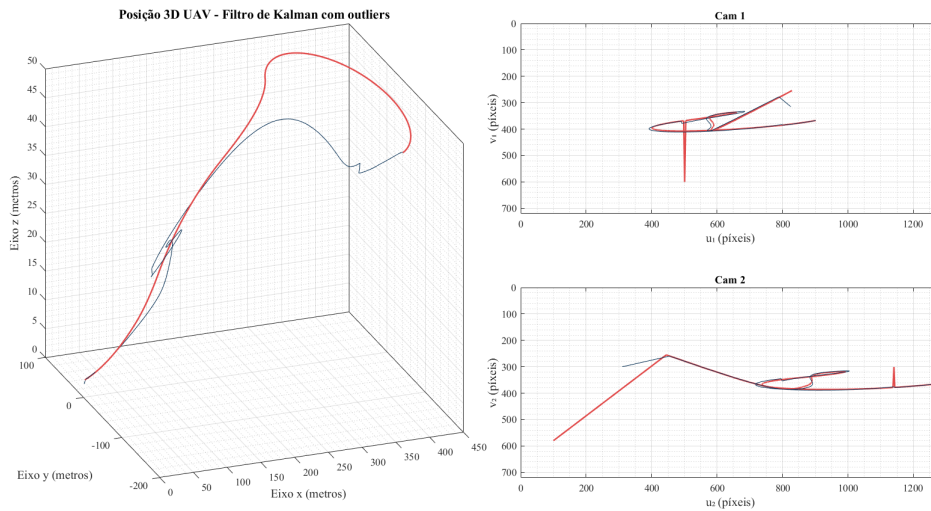


FIGURA 6.30: Filtro de Kalman estendido com *outliers* e maior incerteza.

## 6.4 Discussão dos resultados

As diferenças entre os métodos a que se recorreu, o filtro de Kalman e a triangulação *stereo*, são notáveis. A precisão da posição tridimensional calculada através da triangulação *stereo* é mais rigorosa, no entanto, efetua um maior número de correções do que o filtro de Kalman estendido. Para os mesmos valores de incerteza e ruído, é possível comparar ambas as trajetórias de acordo com a Figura 6.31:

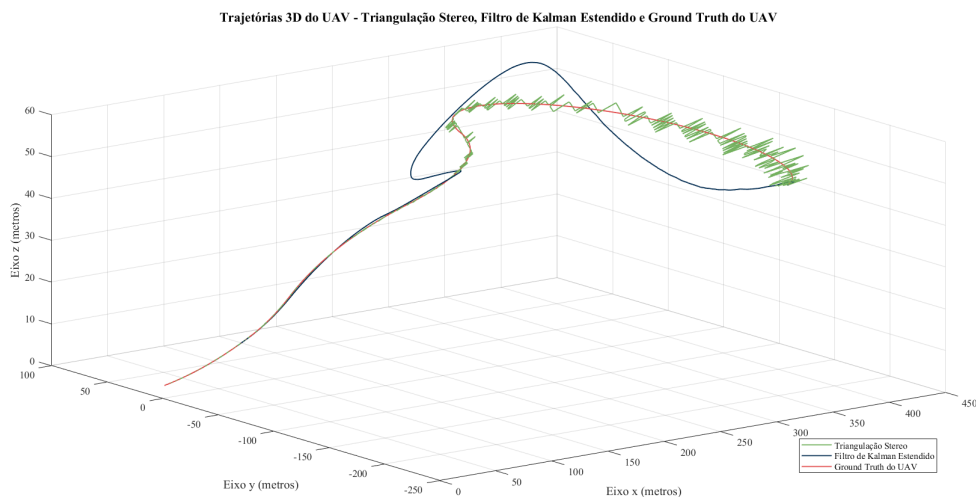


FIGURA 6.31: Trajetórias do UAV através da triangulação *stereo*, filtro de Kalman estendido e *ground truth* do UAV.

No que respeita aos níveis de ruído, visualizando o gráfico ilustrado na Figura 6.32, é possível verificar que ambos os métodos reagem de formas distintas. No primeiro gráfico foi adicionado ruído de 0,5 píxeis, e no segundo gráfico foi adicionado um ruído de 1,5 píxeis.

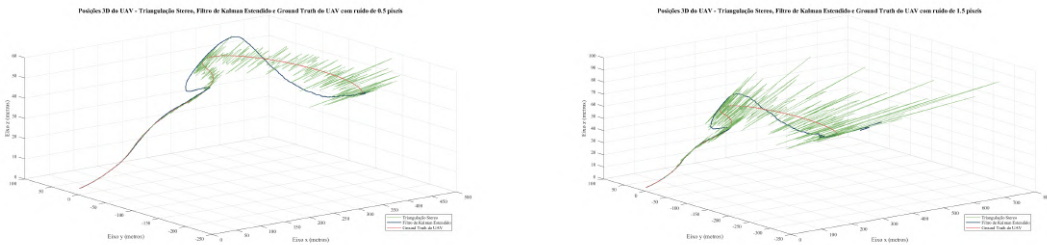


FIGURA 6.32: Triangulação *stereo* e filtro de Kalman estendido perante diferentes níveis de ruído.

A triangulação *stereo* exibe uma sensibilidade superior ao ruído. Conforme o ruído aumenta, a trajetória tende a dispersar e ficar mais instável. Enquanto que o FKE à medida que é introduzido ruído nos dados de entrada, a trajetória calculada tende a ajustar gradualmente devido às medições. Para níveis de ruído extremos o FKE demonstra uma performance mais robusta, devido à capacidade de atualizar a trajetória à medida que recebe dados.

O cálculo das velocidades vetoriais, considerando a velocidade nos eixos do  $x$ ,  $y$  e  $z$  em ambos os métodos, no espaço tridimensional, é possível comparar recorrendo à Figura 6.33:

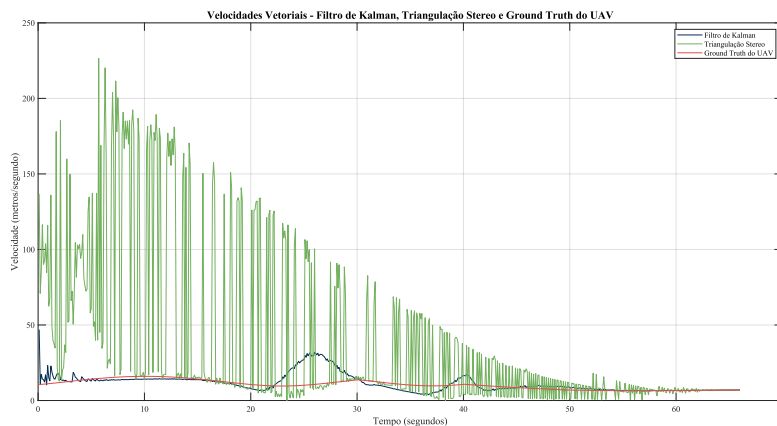


FIGURA 6.33: Trajetórias do UAV através da triangulação *stereo*, filtro de Kalman estendido e *ground truth* do UAV.

Devido às alterações de posição contínuas na triangulação *stereo*, as velocidades resultantes manifestam oscilações mais acentuadas. No entanto, no FKE as velocidades convergem de forma mais expedita em direção ao valor do *ground-truth* do UAV.

Através da análise do gráfico da Figura 6.34, compara-se o comportamento de ambos os métodos ao adicionar *outliers* à trajetória do UAV. Ambos os métodos se encontram sob as mesmas condições de incerteza e nível de ruído. Observa-se que o filtro de Kalman estendido suaviza medições ruidosas e por consequente as oscilações da velocidade também são suavizadas. Na triangulação *stereo* as velocidades estimadas ocorrem com maior variação devido às posições calculadas também oscilarem ao longo do tempo.

A partir do instante  $t = 45$  segundos, verifica-se que as velocidades estimadas pelo FKE demonstram um erro inferior comparando com as velocidades reais do UAV. No entanto, o mesmo é observado na triangulação *stereo* mas com um atraso de 10 segundos.

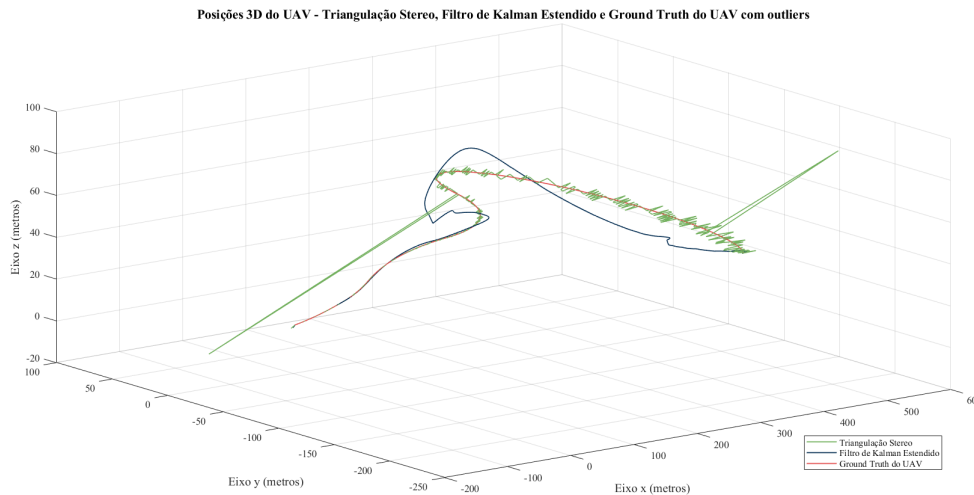


FIGURA 6.34: Posições 3D do UAV através da triangulação *stereo*, filtro de Kalman estendido e *ground truth* do UAV com outliers.

Ao analisar os erros RMSE e MAE, de acordo com a ilustração da Figura 6.35, conclui-se que na fase inicial da trajetória, a detecção do UAV apresenta maior precisão através da triangulação *stereo*. No entanto, a partir do instante  $t = 52$  segundos, no método FKE apesar de demorar mais a estabilizar, consegue estimar uma posição com maior precisão comparando com a triangulação *stereo*.

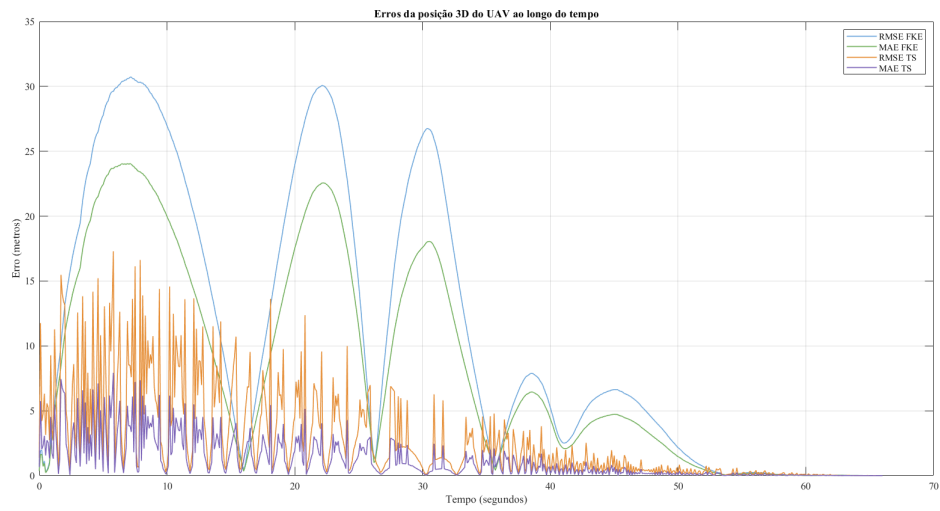


FIGURA 6.35: Erros da posição 3D do UAV ao longo do tempo no FKE e triangulação *stereo*.

# Capítulo 7

## Conclusões e Trabalho Futuro

O último capítulo da presente dissertação culmina com as conclusões e sugestões de trabalho futuro. A Secção 7.1 tem como objetivo apresentar as conclusões que englobam todo o trabalho apresentado. Na Secção 7.2 sugerem-se ideias para desenvolvimento em trabalho futuro.

### 7.1 Conclusões

A presente dissertação introduziu dois métodos com diferentes abordagens para o apoio à aterragem de UAVs em navios, o FKE e a triangulação *stereo*. No contexto, propôs-se a utilização de duas câmaras que captavam imagens ao longo da trajetória do UAV.

Para validação dos métodos adotados adquiriram-se três *datasets* em diferentes cenários reais. Num total foram captadas cerca de 106 832 imagens, das quais foram anotadas cerca de 19 125 imagens para o desenvolvimento e teste dos algoritmos propostos. Para a implementação dos dados nos métodos do filtro de Kalman estendido e na triangulação *stereo* foi necessário perceber a posição do UAV em relação às câmaras, através dos parâmetros extrínsecos.

Após se proceder ao sincronismo dos *timestamps* de ambas as câmaras e dos dados da telemetria do UAV, demonstra-se a aplicabilidade prática dos dados reais. Nos dados reais surgiram falhas de aquisição de imagens durante as aproximações do UAV e na atribuição precisa dos *timestamps*. Considerando as limitações inerentes anteriores, identificou-se a necessidade de exemplificar o funcionamento dos métodos através de dados sintéticos.

Na implementação prática dos dados reais, a trajetória efetuada pelo UAV é mais próxima da realidade através do método FKE, em ambas as aquisições de

dados. Destaca-se a capacidade do FKE na mitigação de ruído inerente aos dados reais, proporcionando uma representação com erro menor.

Na aplicação dos dados sintéticos, através de trajetórias encenadas, os dois métodos apresentam diferenças notáveis, no que corresponde ao modo de processamento dos dados. Concluiu-se que a precisão da posição tridimensional obtida pela triangulação *stereo* é mais rigorosa, porém, efetua ajustes mais frequentemente do que o FKE.

A triangulação *stereo* demonstrou-se mais sensível à presença de ruído, tendo tendência a dispersar e ficar instável com o seu aumento, do que o FKE que suaviza as medições ruidosas. Consequentemente, o FKE apresenta menor variação nas velocidades estimadas devido à menor oscilação das posições calculadas, do que a triangulação *stereo*. O FKE também demonstra ser mais robusto na presença de *outliers*, atenuando o peso à medida que atualiza com pontos ao longo da trajetória do UAV.

A combinação do FKE com a posição inicial previamente calculada pela triangulação *stereo* evidenciou resultados com maior robustez e precisão. A abordagem híbrida permitiu otimizar as estimativas das posições do UAV calculadas, oferecendo uma nova abordagem ao problema em questão.

Em síntese, o sistema proposto permite uma estimativa fiável da trajetória do UAV ao longo do tempo do UAV ao longo do tempo. Os resultados obtidos são promissores, podendo permitir assim contribuir para uma operação mais segura dos UAVs em ambiente marítimo

## 7.2 Trabalho Futuro

A dissertação elaborada ao longo do documento, onde se sugere uma abordagem diferente em relação a um sistema de apoio à aterragem de UAVs em navios, pode ter bastantes desenvolvimentos em trabalho futuros.

O sistema de câmaras proposto, em que se recorrem a duas câmaras IP, é considerado elementar, pois apenas é possível adquirir imagens em posições fixas das câmaras. Um dos campos a implementar em estudos posteriores será a utilização de câmaras com capacidade PTZ autónoma. Desta forma, mesmo que o UAV esteja fora do campo de visão, as câmaras procurem detetar o UAV movimentando-se. Ou até mesmo para a deteção de UAVs a uma longa distância em que seja necessário efetuar zoom.

Em relação à sincronização dos dados obtidos a partir das câmaras e telemetria do UAV, constatou-se desafios na realização da tarefa. Em trabalhos futuros sobre a temática, sugere-se a utilização de um meio, como por exemplo um GPS comum para as medições das câmaras, sendo apenas necessário sincronizar com os dados de telemetria do UAV. O método permitirá a sincronização dos dados provenientes de diferentes fontes, tornando o processamento mais ágil e preciso.

A implementação de métodos de aquisição da posição do UAV através de outros métodos diferentes, para além do filtro de Kalman e triangulação *stereo*, que sejam capazes de captar a posição do UAV e prever de forma mais precisa e expedita, como por exemplo, recorrer ao estado da arte atual como é o caso das redes neuronais profundas, *Deep Neural Networks* (DNN), poderá ser mais uma opção de desenvolvimento do tema.

Uma abordagem futura para validar e melhorar a estimativa do estado do UAV será efetuar os métodos em situações que decorram com voos reais, em tempo real. Será uma oportunidade de confrontar a posição e velocidade do UAV em cenários reais, acrescentando a complexidade de estimar a pose do UAV, no entanto, permitirá efetuar ajuste no momento exato.



# Bibliografia

- Abdujoub, S., J. McPhee e R. Irani (set. de 2020). «Methodologies for landing autonomous aerial vehicles on maritime vessels». Em: *Aerospace Science and Technology* 106, pp. 1–15. URL: <https://doi.org/10.1016/j.ast.2020.106169>.
- Ackerman, E. e E. Strickland (jan. de 2018). «Medical delivery drones take flight in east africa.» Em: *IEEE Spectrum* 55.1, pp. 34–35. URL: <https://doi.org/10.1109/MSPEC.2018.8241731>.
- Alam, M. e J. Oluoch (out. de 2021). «A survey of safe landing zone detection techniques for autonomous unmanned aerial vehicles (UAVs)». Em: *Expert Systems with Applications* 179, pp. 1–12. URL: <https://doi.org/10.1016/j.eswa.2021.115091>.
- Autelpilot, Autel - (2023). *Autel Robotics EVO II Dual 640T Drone*. <https://www.autelpilot.eu/collections/evo-ii-640t/products/autel-robotics-evo-ii-dual-640t-drone-with-enterprise-bundle?variant=41124599103545>. acessado em: 7 julho 2023.
- Azevedo, João (nov. de 2022). «UAVProtocol: Sistema de encaminhamento entre UAVs». Em: *Instituto Superior de Engenharia do Porto*, pp. 1–113. URL: <http://hdl.handle.net/10400.22/22144>.
- Barisic, A., M. Car e S. Bogdan (nov. de 2019). «Vision-based system for a real-time detection and following of UAV». Em: *2019 Workshop on Research, Education and Development of Unmanned Aerial Systems (RED UAS)*, pp. 156–159. URL: <https://doi.org/10.1109/REDUAS47371.2019.8999675>.
- Cho, G., J. Choi, G. Bae e H. Oh (nov. de 2022). «Autonomous ship deck landing of a quadrotor UAV using feed-forward image-based visual servoing». Em: *Aerospace Science and Technology* 130.107869, pp. 1–12. URL: <https://doi.org/10.1016/j.ast.2022.107869>.
- Christ, R. e R. Wernli (2014). «Navigational Sensors». Em: *The ROV Manual (Second Edition)* 17.2, pp. 453–475. URL: <https://doi.org/10.1016/B978-0-08-098288-5.00017-8>.

- DGRM (2018). *Zonas Marítimas sob Soberania e ou Jurisdição Portuguesa*. <https://www.dgrm.mm.gov.pt/am-ec-zonas-maritimas-sob-jurisdicao-ou-soberania-nacional>. acessado em: 26 fevereiro 2023.
- DJI (2023). *Matrice 600Pro*. <https://www.dji.com/pt/matrice600-pro>. acessado em: 25 julho 2023.
- Documentation, IBM (2010). *What is a data set?* <https://www.ibm.com/docs/en/zos-basic-skills?topic=more-what-is-data-set>. acessado em: 7 abril 2023.
- Dolgoplov, A., P. Kazantsev e N. Bezuhliy (dez. de 2016). «Ship Detection in Images Obtained from the Unmanned Aerial Vehicle (UAV)». Em: *Indian Journal of Science and Technology* 9.46, pp. 1–7. URL: <http://dx.doi.org/10.17485/ijst/2016/v9i46/107562>.
- Ferreira, Tiago (jul. de 2021). «6D UAV pose estimation for ship landing guidance». Em: *OCEANS 2021: San Diego – Porto*, pp. 1–10. URL: <https://doi.org/10.23919/OCEANS44145.2021.9705884>.
- Foix, S., G. Alenya e C. Torras (set. de 2014). «Lock-in Time-of-Flight (ToF) Cameras: A Survey». Em: *IEEE Sensors Journal* 11.9, pp. 1917–1926. URL: <https://doi.org/10.1109/JSEN.2010.2101060>.
- Garcia, G., J. Corrales, J. Pomares e F. Torres (dez. de 2009). «Survey of visual and force/tactile control of robots for physical interaction in Spain.» Em: *Sensors 2009* 9.12, pp. 9689–9733. URL: <https://doi.org/10.3390/s91209689>.
- Garmin (2023). *eTrex Vista HCx*. <https://www.garmin.com/en-US/p/8703>. acessado em: 12 abril 2023.
- Gonçalves, L. e B. Damas (jun. de 2022). «Automatic detection of rescue targets in maritime search and rescue missions using UAVs». Em: *International Conference on Unmanned Aircraft Systems (ICUAS 2022)*, pp. 1638–1643. URL: <https://doi.org/10.1109/ICUAS54217.2022.9836137>.
- Good, N., O. Aboutalib, B. Thai, N. Yamaoka, C. Kim, C. Wilkinson e D. Findlay (jan. de 2016). «Validation process of the physics-based modeling of navigation sensors for sea-based aviation automated landing.» Em: *AIAA Modeling and Simulation Technologies Conference*, pp. 1–12. URL: <https://doi.org/10.2514/6.2016-1919>.
- Guardian, The (2018). *'Uber for blood': how Rwandan delivery robots are saving lives*. <https://www.theguardian.com/global-development/2018/jan/02/rwanda-scheme-saving-blood-drone>. acessado em: 27 fevereiro 2023.

- Hariyama, M., N. Yokoyama e M. Kameyama (abr. de 2008). «Design of a trinocular-stereo-vision VLSI processor based on optimal scheduling.» Em: *IEICE Transactions on Electronics* 91.4, pp. 479–486. URL: <http://dx.doi.org/10.1093/ietele/e91-c.4.479>.
- Hassanalian, M. e A. Abdelkefi (mai. de 2017). «Classifications, applications, and design challenges of drones: A review.» Em: *Progress in Aerospace Sciences* 91, pp. 99–131. URL: <https://doi.org/10.1016/j.paerosci.2017.04.003>.
- Hawkins, Douglas (1980). «Identification of outliers.» Em: *Monographs on Statistics and Applied Probability* 11, pp. 1–188. URL: <https://doi.org/10.1007/978-94-015-3994-4>.
- Ikura, M., L. Miyashita e M. Ishikawa (jan. de 2020). «Real-time Landing Gear Control System Based on Adaptive 3D Sensing for Safe Landing of UAV.» Em: *2020 IEEE/SICE International Symposium on System Integration (SII)*, pp. 759–764. URL: <https://doi.org/10.1109/SII46433.2020.9026177>.
- Institute, U. S. Naval (2023). *Twilight manned flight*. <https://www.usni.org/magazines/proceeding/2016/april/twilight-manned-flight>. aceso em: 1 março 2023.
- Jung, S. e H. Kim (abr. de 2018). «Analysis of Amazon Prime Air UAV Delivery Service.» Em: *Journal of Knowledge Information Technology and Systems (JKITS)* 12.2, pp. 253–266. URL: <http://dx.doi.org/10.34163/jkits.2017.12.2.005>.
- Kalman, Rudolf (mar. de 1960). «A New Approach to Linear Filtering and Prediction Problems.» Em: *Journal of Basic Engineering* 82.1, pp. 1–11. URL: <https://doi.org/10.1115/1.3662552>.
- Konert, A. e T. Balcerzak (2021). «Military autonomous drones (UAVs)-from fantasy to reality. Legal and Ethical implications.» Em: *Transportation research procedia* 59, pp. 292–299. URL: <https://doi.org/10.1016/j.trpro.2021.11.121>.
- Li, W., Y. Ge e Z. Guan ang G. Ye (ago. de 2022). «Synchronized Motion-Based UAV–USV Cooperative Autonomous Landing.» Em: *Journal of Marine Science and Engineering* 10.9, pp. 1214–1214. URL: <https://doi.org/10.3390/jmse10091214>.
- Li, Y., H. Yuan, Y. Wang e B. Zhang (set. de 2022). «Maritime Vessel Detection and Tracking under UAV Vision.» Em: *2022 International Conference on Artificial Intelligence and Computer Information Technology (AICIT)*, pp. 1–4. URL: <https://doi.org/10.1109/AICIT55386.2022.9930166>.

- Tp-link (2023). *TL-SF1006P*. <https://www.tp-link.com/pt/business-networking/poe-switch/tl-sf1006p/>. acessado em: 12 abril 2023.
- Liu, F., J. Shan, B. Xiong e Z. Fang (mai. de 2022). «A Real-Time and Multi-Sensor-Based Landing Area Recognition System for UAVs». Em: *Drones* 6.5, pp. 1–17. URL: <https://doi.org/10.3390/drones6050118>.
- Ma, Y., Q. Li, L. Chu, Y. Zhou e C. Xu (jan. de 2021). «Real-time detection and spatial localization of insulators for UAV inspection based on binocular stereo vision.» Em: *Remote Sensing* 13.2, pp. 230–230. URL: <https://doi.org/10.3390/rs13020230>.
- Marcelino, Valentina (2021). *As Forças Armadas têm uma nova escola de drones e está no meio no Atlântico*. <https://www.dn.pt/sociedade/as-forcas-armadas-tem-uma-nova-escola-de-drones-e-esta-no-meio-no-atlantico-14015376.html>. acessado em: 28 fevereiro 2023.
- Meng, Y., W. Wang, H. Han e J. Ban (fev. de 2019). «A visual/inertial integrated landing guidance method for UAV landing on the ship». Em: *Aerospace Science and Technology* 85, pp. 474–480. URL: <https://doi.org/10.1016/j.ast.2018.12.030>.
- Miranda-Moya, A., H. Castañeda, J. Gordillo e H. Wang (mai. de 2017). «Classifications, applications, and design challenges of drones: A review». Em: *Mechatronics* 91, pp. 1–132. URL: <https://doi.org/10.1016/j.ast.2018.12.030>.
- Nex, F., C. Armenakis, M. Cramer, D. Cucci, M. Gerke, E. Honkavaara, A. Kukko, C. Persello e J. Skaloud (fev. de 2022). «UAV in the advent of the twenties: Where we stand and what is next». Em: *ISPRS Journal of Photogrammetry and Remote Sensing* 184, pp. 215–242. URL: <https://doi.org/10.1016/j.isprsjprs.2021.12.006>.
- Nguyen, P., K. Kim, Y. Lee e K. Park (set. de 2017). «Remote Marker-Based Tracking for UAV Landing Using Visible-Light Camera Sensor». Em: *Human Detection, Identification, and Recognition of Gesture and Behavior based on Thermal Camera, Sensors* 17.9, pp. 37–45. URL: <https://doi.org/10.3390/s17091987>.
- Nie, W., Z. Han, M. Zhou, L. Xie e Q. Jiang (jun. de 2021). «UAV Detection and Identification Based on WiFi Signal and RF Fingerprint». Em: *IEEE Sensors Journal* 21.12, pp. 13540–13550. URL: <https://doi.org/10.1109/JSEN.2021.3068444>.
- Oitzman, Mike (2022). *2022 Commercial UAV show recap*. <https://www.therobotreport.com/2022-commercial-uav-show-recap/>. acessado em: 28 fevereiro 2023.

- OpenCV (2023). *OpenCV: Open Source Computer Vision Library*. <https://github.com/opencv/opencv/blob/4.x/doc/pattern.png>. acessado em: 30 maio 2023.
- Pan, X., D. Ma, L. Jin e Z. Jiang (mai. de 2008). «Vision-Based Approach Angle and Height Estimation for UAV Landing». Em: *2008 Congress on Image and Signal Processing 3*, pp. 801–805. URL: <https://doi.org/10.1109/CISP.2008.78>.
- Pires, C., B. Damas e A. Bernardino (mar. de 2022). «An Efficient Cascaded Model for Ship Segmentation in Aerial Images». Em: *IEEE Access* 10, pp. 31942–31954. URL: <https://doi.org/10.1109/ACCESS.2022.3159667>.
- Portuguesa, Marinha (2023a). *NRP Sines*. [https://www.marinha.pt/pt/os\\_meios/patruilhasoceanicos/Paginas/nrp-sines.aspx](https://www.marinha.pt/pt/os_meios/patruilhasoceanicos/Paginas/nrp-sines.aspx). acessado em: 1 março 2023.
- (2023b). *NRP Vasco da Gama*. [https://www.marinha.pt/pt/os\\_meios/fragatas/Paginas/nrp-vasco-da-gama.aspx](https://www.marinha.pt/pt/os_meios/fragatas/Paginas/nrp-vasco-da-gama.aspx). acessado em: 1 março 2023.
- Rabelo, M., A. Brandão e M. Sarcinelli-Controller (mar. de 2021). «Landing a UAV on Static or Moving Platforms Using a Formation Controller». Em: *IEEE Systems Journal* 15.1, pp. 37–45. URL: <https://doi.org/10.1109/JSYST.2020.2975139>.
- Redmon, J., S. Divvala, R. Girshick e A. Farhadi (2017). «You Only Look Once: Unified, Real-Time Object Detection». Em: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788. URL: <http://pjreddie.com/yolo/>.
- República Portuguesa, Governo da (2020). *Drones vigiam florestas*. <https://www.portugal.gov.pt/pt/gc22/comunicacao/noticia?i=drones-vigiam-florestas->. acessado em: 28 fevereiro 2023.
- Ribeiro, M., B. Damas e A. Bernardino (out. de 2022). «Real-Time Ship Segmentation in Maritime Surveillance Videos Using Automatically Annotated Synthetic Datasets». Em: *Sensors* 2022 22.21, pp. 8090–8090. URL: <https://doi.org/10.3390/s22218090>.
- Riola, J., J. Diaz e J. Giron-Sierra (jun. de 2011). «The prediction of calm opportunities for landing on a ship: Aspects of the problem». Em: *OCEANS 2011 IEEE - Spain*, pp. 1–8. URL: <https://doi.org/10.1109/Oceans-Spain.2011.6003435>.
- Sadekar, Kaustubh (2020). *Understanding Lens Distortion*. <https://learnopencv.com/understanding-lens-distortion/>. acessado em: 9 junho 2023.

- Saj, V., B. Lee, D. Kalathil e M. Benedict (set. de 2022). «Robust Reinforcement Learning Algorithm for Vision-based Ship Landing of UAVs». Em: *Cornell University*, pp. 1–8. URL: <https://doi.org/10.48550/arXiv.2209.08381>.
- Santos, N., V. Lobo e A. Bernardino (out. de 2019). «AUTOLAND project: Fixed-wing UAV landing on a fast patrol boat using computer vision.» Em: *Oceans 2019 MTS/IEEE SEATTLE*, pp. 1–5. URL: <https://doi.org/10.23919/OCEANS40490.2019.8962869>.
- (fev. de 2020). «Directional Statistics for 3D Model-Based UAV Tracking». Em: *IEEE Access* 8, pp. 33884–33897. URL: <https://doi.org/10.1109/ACCESS.2020.2973970>.
- Santos, Nuno (2020a). «Fixed-wing UAV tracking in outdoor scenarios for autonomous landing». Em: *Universidade de Liaboa, Instituto Superior Técnico, Thesis PhD Degree*, pp. 1–167. URL: [https://www.academia.edu/49107749/Fixed\\_wing\\_UAV\\_tracking\\_in\\_outdoor\\_scenarios\\_for\\_autonomous\\_landing](https://www.academia.edu/49107749/Fixed_wing_UAV_tracking_in_outdoor_scenarios_for_autonomous_landing).
- (2020b). *Real UAV landing area tests*. [https://www.youtube.com/watch?v=qE1S-5QJBbQ&ab\\_channel=NunoPessanhaSantos](https://www.youtube.com/watch?v=qE1S-5QJBbQ&ab_channel=NunoPessanhaSantos). acessado em: xx fevereiro 2023.
- Silva, K., A. Morais, J. Morais e R. Gedraite (jul. de 2014). «Hardware para controle avançado de veículo aéreo não tripulado do tipo quadricóptero.» Em: *Horizonte Científico* 8.1, pp. 1–22. URL: <https://seer.ufu.br/index.php/horizontecientifico/article/view/24028>.
- Tarot (2011). *TAROT X6 Hex-Copter FPV Kit TL6X001*. <http://www.tarotrc.com/Product/Detail.aspx?Lang=en&Id=d85af58f-eb22-45ee-ad9e-3cbd1c6782bb>. acessado em: 3 junho 2023.
- Technology, Dahua (2023). *IPC-HDBW8232E-Z-SL*. <https://www.dahuasecurity.com/products/All-Products/Discontinued-Products/PTZ-Cameras/Anti-Corrosion-Cameras/IPC-HDBW8232E-Z-SL=S2>. acessado em: 11 abril 2023.
- Terven, J. e D. Cordova-Esparza (ago. de 2023). «A Comprehensive Review of YOLO: From YOLOv1 and Beyond». Em: *arXiv preprint arXiv:2304.00501*, pp. 1–34. URL: <https://doi.org/10.48550/arXiv.2304.00501>.
- Trong, T., Q. Hai, M. Van, B. Thai, T. Chi e T. Quang (dez. de 2021). «Autonomous Detection and Approach Tracking of Moving Ship on the Sea by VTOL UAV based on Deep Learning Technique through Simulated Real-time On-Air Image Acquisitions». Em: *2021 8th NAFOSTED Conference on Information*

- and *Computer Science (NICS)*, pp. 374–380. URL: <https://doi.org/10.1109/NICS54270.2021.9701509>.
- Williams, Kevin (dez. de 2004). «A Summary of Unmanned Aircraft Accident/Incident Data: Human Factors Implications». Em: *Civil Aerospace Medical Institute Federal Aviation Administration*, pp. 3–13. URL: <https://apps.dtic.mil/sti/citations/ADA460102>.
- Wojcik, A., P. Klapa, B. Mitka e I. Piech (nov. de 2019). «The use of TLS and UAV methods for measurement of the repose angle of granular materials in terrain conditions». Em: *Measurement* 146, pp. 780–791. URL: <https://doi.org/10.1016/j.measurement.2019.07.015>.
- Wu, J., Z. Jin, A. Liu, L. Yu e F. Yang (jan. de 2022). «Validation process of the physics-based modeling of navigation sensors for sea-based aviation automated landing.» Em: *Journal of the Franklin Institute* 359.1, pp. 556–577. URL: <https://doi.org/10.1016/j.jfranklin.2021.11.009>.
- Xia, B., I. Mantegh e W. Xie (out. de 2021). «Integrated Emergency Self-Landing Method for Autonomous UAS in Urban Aerial Mobility». Em: *2021 21st International Conference on Control, Automation and Systems (ICCAS)*, pp. 275–282. URL: <https://doi.org/10.23919/ICCAS52745.2021.9649955>.
- Xin, L., Z. Tang, W. Gai e H. Liu (out. de 2022). «Vision-Based Autonomous Landing for the UAV: A Review». Em: *Aerospace: Latest Advancements in Aeronautics and Astronautics: Celebrating the 70th Anniversary of Beihang University* 634.9, pp. 1–11. URL: <https://doi.org/10.3390/aerospace9110634>.
- Xu, G., Y. Zhang, S. Ji, Y. Cheng e Y. Tian (abr. de 2009). «Research on computer vision-based for UAV autonomous landing on a ship». Em: *Pattern Recognition Letters* 30.6, pp. 600–605. URL: <https://doi.org/10.1016/j.patrec.2008.12.011>.
- Yang, T., P. Li, H. Zhang, J. Li e Z. Li (mai. de 2018). «Monocular Vision SLAM-Based UAV Autonomous Landing in Emergencies and Unknown Environments». Em: *Autonomous Control of Unmanned Aerial Vehicles, Electronics* 7.73, pp. 1–18. URL: <https://doi.org/10.3390/electronics7050073>.
- Zhang, B., Z. Song e F. Zhao (jan. de 2022). «Overview of Propulsion Systems for Unmanned Aerial Vehicles». Em: *Energies* 15.2, pp. 1–25. URL: <https://doi.org/10.3390/en15020455>.
- Zhang, Yu-Jin (fev. de 2023). «Camera Calibration». Em: *Computer Vision: Principles, Algorithms and Applications*, pp. 37–65. URL: [https://doi.org/10.1007/978-981-19-7580-6\\_2](https://doi.org/10.1007/978-981-19-7580-6_2).

- Zhang, X., Z. He, Z. Ma, P. Jun e K. Yang (set. de 2021). «VIAE-Net: An end-to-end altitude estimation through monocular vision and inertial feature fusion neural networks for UAV autonomous landing.» Em: *Sensors* 21.18, pp. 6302–6302. URL: <https://doi.org/10.3390/rs13020230>.
- Zhang, X., Z. Shi, Z. Wu e J. Liu (jan. de 2020). «Sea surface ships detection method of UAV based on improved YOLOv3.» Em: *Eleventh International Conference on Graphics and Image Processing (ICGIP 2019)* 11373, pp. 213–225. URL: <https://doi.org/10.1117/12.2557479>.
- Zhang, Z., R. Deriche, O. Faugeras e Q. Luong (jan. de 1995). «A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry.» Em: *Artificial Intelligence* 78.1, pp. 87–119. URL: [https://doi.org/10.1016/0004-3702\(95\)00022-4](https://doi.org/10.1016/0004-3702(95)00022-4).
- Zhang, Zhengyou (set. de 1999). «Flexible camera calibration by viewing a plane from unknown orientations.» Em: *Proceedings of the Seventh IEEE International Conference on Computer Vision* 1, pp. 666–673. URL: <https://doi.org/10.1109/ICCV.1999.791289>.
- (nov. de 2000). «A flexible new technique for camera calibration.» Em: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22.11, pp. 1330–1334. URL: <https://doi.org/10.1109/34.888718>.
- Zhao, Z., P. Zheng, S. Xu e X. Wu (nov. de 2019). «Object Detection With Deep Learning: A Review.» Em: *IEEE Transactions on Neural Networks and Learning Systems* 30.11, pp. 3212–3232. URL: <https://doi.org/10.1109/TNNLS.2018.2876865>.
- Zhu, N., D. Bétaille e M. Berbineau (set. de 2018). «GNSS Position Integrity in Urban Environments: A Review of Literature.» Em: *Transactions on Intelligent Transportation Systems* 19.9, pp. 2762–2778. URL: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8248655&isnumber=8458340>.

## Anexo I - Desenvolvimento das derivadas parciais do Jacobiano

Considerando que o cálculo do Jacobiano será uma matriz com as dimensões  $2 \times 3$ , que é representada pela Equação I.1:

$$\left. \frac{\partial h()_{INT}}{\partial P_C} \right|_{x = h_{EXT}(\bar{X}_{k-1})} = \begin{bmatrix} \frac{du}{dx_C} & \frac{du}{dy_C} & \frac{du}{dz_C} \\ \frac{dv}{dx_C} & \frac{dv}{dy_C} & \frac{dv}{dz_C} \end{bmatrix} \quad (I.1)$$

Desenvolvendo as derivadas parciais de cada elemento da matriz anterior, esta é apresentada pela Equação I.2:

$$\left. \frac{\partial h()_{INT}}{\partial P_C} \right|_{x = h_{EXT}(\bar{X}_{k-1})} = \begin{bmatrix} A + 2\frac{x_C^2}{z_C^2} B & 2\frac{x_C y_C}{z_C^2} B & -2\frac{x_C(x_C^2 + y_C^2)}{z_C^3} B \\ 2\frac{x_C y_C}{z_C^2} B & A + 2\frac{y_C^2}{z_C^2} B & -2\frac{y_C(x_C^2 + y_C^2)}{z_C^3} B \end{bmatrix} \quad (I.2)$$

Em que se atribuem as variáveis A e B de forma a facilitar a leitura da matriz. Assim sendo, os valores atribuídos a cada variável são dados pela Equação I.3:

$$\begin{aligned} A &= \frac{f}{z_C} (1 + k_1 R + k_2 R^2) \\ B &= \frac{f}{z_C} (1 + k_1 R + k_2 R) \end{aligned} \quad (I.3)$$



## Anexo II - Desenvolvimento da fórmula de Haversine

A fórmula de Haversine é utilizada para determinar a distância mais curta entre dois pontos situados numa esfera. A fórmula recorre às latitudes dos dois pontos em questão  $\Phi_1$  e  $\Phi_2$ , e às longitudes  $\lambda_1$  e  $\lambda_2$  medidas na superfície para calcular a distância.

A fórmula de Haversine é dada pela Equação II.1:

$$\text{haversine}(\theta) = \sin^2\left(\frac{\theta}{2}\right) \quad (\text{II.1})$$

Desenvolvendo a Equação II.1, obtém o cálculo intermédio da fórmula, presente na Equação II.2:

$$\left(\frac{d}{r}\right) = \text{haversine}(\Phi_2 - \Phi_1) + \cos(\Phi_1)\cos(\Phi_2)\text{haversine}(\lambda_1 - \lambda_2) \quad (\text{II.2})$$

Em que se consegue deduzir a distância entre os dois pontos através da Equação II.3:

$$d = 2r \sin^{-1} \left( \sqrt{\sin^2\left(\frac{\Phi_2 - \Phi_1}{2}\right) + \cos(\Phi_1)\cos(\Phi_2)\sin^2\left(\frac{\lambda_2 - \lambda_1}{2}\right)} \right) \quad (\text{II.3})$$



## Anexo III - Definição de RMSE e MAE

A medida do erro *Root Mean Squared Error* (RMSE), em português, Erro Quadrático Médio (EQM), representa a medida utilizada para calcular a raiz quadrada da média dos quadrados dos desvios entre os valores observados (reais) e as previsões (hipóteses). O RMSE é dado pela Equação III.1:

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2} \quad (\text{III.1})$$

A medida do erro *Mean Absolut Error* (MAE), em português, Erro Médio Absoluto, calcula o valor médio das diferenças absolutas entre os valores observados (reais) e as previsões (hipóteses). O MAE é calculado de acordo com a Equação III.2:

$$MAE = \frac{1}{n} \sum_{j=i}^n |y_j - \hat{y}_j| \quad (\text{III.2})$$