



João Nuno Serrador da Silva

Licenciado em Ciências da Engenharia Eletrotécnica e de
Computadores

**Aterragem de Robôs Aéreos Multi-Rotor
em Plataformas Móveis: Uma Solução
Baseada em Comportamentos e em
Saliência Visual**

Dissertação para obtenção do Grau de Mestre em
Engenharia Eletrotécnica e de Computadores

Orientador: José António Barata de Oliveira,
Professor Auxiliar, FCT-UNL
Co-orientador: Pedro Figueiredo Santana,
Professor Auxiliar, ISCTE-IUL

Júri:

Presidente: Prof. Doutor Tiago Oliveira Machado de Figueiredo Cardoso
Arguente: Prof. Doutor João Paulo Branquinho Pimentão
Vogal: Prof. Doutor José António Barata de Oliveira

Copyright

Aterragem de Robôs Aéreos Multi-Rotor em Plataformas Móveis: Uma Solução Baseada em Comportamentos e em Saliência Visual

A Faculdade de Ciências e Tecnologia e a Universidade Nova de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

Dedico esta dissertação aos meus pais.

Agradecimentos

Gostaria de agradecer ao meu orientador Prof. José Barata pela oportunidade de realizar a dissertação na área de robótica participando num projeto europeu, pois tornou a dissertação um desafio ainda maior. Ao meu co-orientador Prof. Pedro Santana um agradecimento especial, em primeiro lugar pela disponibilidade e compreensão ao longo deste trabalho. Também pela motivação e conhecimento partilhado, pela exigência do rigor que me obrigou a fazer mais e melhor e por fim pela acessibilidade na comunicação. Considero um exemplo a seguir a nível profissional e pessoal, pela sua ambição em alcançar a excelência.

À Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa o meu agradecimento pelo prestígio no ensino a nível nacional e além fronteiras, pelo acolhimento e pelas condições oferecidas nos últimos anos. Também porque me permitiu conhecer pessoas que de alguma forma contribuíram positivamente para a minha evolução como profissional e como pessoa. Um agradecimento em particular aos colegas de laboratório: André Lourenço, André Parda, Eduardo Pinto, Francisco Marques, Paulo Rodrigues, Ricardo Mendonça e Ricardo Pombeiro. Também aos colegas de faculdade com quem passei grande parte do tempo durante estes anos de estudos: António Figueiredo, Bruno Galveia, Bruno Vilhena, Gonçalo Ferreira, José Baleia, Nuno Vilhena e Marcos Carvalho. Também um agradecimento a todos os outros que não mencionei mas que também tiveram o seu papel importante na minha obtenção deste título.

Às duas pessoas que permitiram o meu percurso académico e me acompanharam até ao final, os meus pais Fernanda e José Manuel. Sem eles, por vários motivos, não teria conseguido alcançar esta meta a que me propus. Nos bons e nos maus momentos sempre me aconselharam e ajudaram a seguir em frente e vencer esta etapa da minha vida. Ao meu irmão Mauro um obrigado pelas dicas e partilha de experiências académicas, pelo apoio e pelas palavras certas na altura certa.

Aos atuais e ex-elementos do Grupo Musical GeraçãoXXI do qual faço parte, em particular a Fernando Reis, João Paixim e Mauro Silva. Pela compreensão e entajuda prestada nos momentos mais difíceis e pelos momentos de companheirismo e diversão que me fizeram descontraír e ganhar força para ultrapassar as dificuldades e barreiras deste percurso.

Por fim um agradecimento muito especial à pessoa que me apoiou em todos os momentos e talvez a que mais aturou as alturas menos boas, a minha namorada Alexandra. Obrigado pela compreensão, paciência, apoio e amor nas alturas em que mais precisei.

Resumo

Esta dissertação apresenta um sistema para aterragem de veículos aéreos multi-rotor em plataformas móveis utilizando um sensor de visão monocular e GPS. O sistema aqui apresentado visa colmatar algumas das limitações dos sistemas atuais, tais como a dependência na existência de uma marca conhecida a identificar a plataforma de aterragem, cuja deteção é robusta apenas a baixas altitudes. Através da utilização de um modelo computacional de saliência visual aliado a um detetor de uma marca específica, esta dissertação apresenta uma solução robusta a um envelope mais alargado de altitudes. Por forma a permitir estender a pesquisa para além do campo visual instantâneo do veículo aéreo, o processo de aterragem é gerido por uma arquitetura comportamental capaz de controlar o posicionamento do veículo de acordo com estimativas da posição mais provável da plataforma, no caso da existência de comunicações diretas entre o veículo e a plataforma de aterragem. O caso de estudo utilizado para validar o sistema é composto por um veículo aéreo com quatro rotores e por uma plataforma de aterragem presente num veículo marítimo. Testes nos veículos físicos e simulados demonstram a capacidade do sistema em gerir a pesquisa e aterragem do veículo aéreo.

palavras-chave: saliência visual, robôs autónomos, comportamentos, veículos aéreos, aterragem.

Abstract

This dissertation presents a landing system to multi-rotor vertical take-off vehicles in a moving platform using a monocular vision and GPS sensors. The system introduced aims to overcome some of the current systems' limitations, in which the detection is robust only for low altitudes. Through the use of a computational model of visual saliency combined with a specific helipad detector, this dissertation presents a robust solution to a wide range of altitudes. In order to enable extending the search beyond the visual field of the aerial vehicle, the landing process is managed by a behaviour-based architecture able to handling the vehicle position according the expected position of the platform, in case of direct communication channel available between the vehicle and the platform. The case study used for system validating consisting of a four-rotor aerial vehicle and a landing platform on a marine vehicle. Tests on real and simulated vehicles show the system's capacity to manage the search and landing of the aerial vehicle.

keywords: visual saliency, autonomous robots, behavior-based, aerial vehicles, landing.

Conteúdo

Lista de Figuras	xv
Lista de Tabelas	xvii
1 Introdução	1
1.1 MR-VTOLs: Tipos e Aplicações	1
1.2 O Problema da Aterragem	2
1.3 Estrutura da Dissertação	3
2 Trabalho Relacionado	5
2.1 Aterragem de MR-VTOLs	5
2.2 Modelos Baseados em Comportamentos	9
2.3 Modelos de Saliência Visual na Robótica	10
3 Modelo Proposto	13
3.1 Visão Global do Sistema	13
3.1.1 O Sistema Robótico	13
3.1.2 A Tarefa	15
3.1.3 A Solução	16
3.2 Pesquisa Visual	18
3.2.1 Detetor Dedicado	18
3.2.2 Modelo de Saliência Visual Base	20
3.2.3 Inclusão da Aprendizagem em Função da Altitude	22
3.2.4 Pesquisa no Mapa de Saliência em Função da Altitude	23
3.2.5 Seguimento no Mapa de Saliência	24
3.3 A Arquitetura Baseada em Comportamentos	25

3.3.1	O Ciclo Geral	26
3.3.2	O Levantamento da Plataforma	26
3.3.3	A Aterragem na Plataforma	28
4	Resultados Experimentais	35
4.1	Instalação Experimental	35
4.1.1	Simulação	35
4.1.2	Real	36
4.2	Parametrização do Sistema	37
4.2.1	Simulação	37
4.2.2	Real	39
4.3	Experiências	40
4.3.1	Simulação	40
4.3.2	Real	47
5	Conclusões e Trabalho Futuro	53
5.1	Conclusões	53
5.2	Trabalho Futuro	55
	Bibliografia	57
	Appendices	
A	Referenciais	61
B	Momentos Geométricos	63
C	Modelo de Saliência Base	65

Lista de Figuras

2.1	Exemplo de um sistema de controlo baseado em comportamentos	9
3.1	Referenciais	14
3.2	Diagrama da relação de projeção baseado no modelo <i>pinhole</i>	15
3.3	Descrição das 4 tarefas do MR-VTOL	16
3.4	Marcas específicas: simulador e real	18
3.5	Rotina percecional do detetor dedicado.	20
3.6	Comparação entre aprendizagem da saliência por altitudes	23
3.7	Comparação entre aprendizagem da saliência por altitudes	24
3.8	Pesquisa de segmentos isolados de saliência	24
3.9	Exemplo de seguimento da plataforma através de <i>CamShift</i>	25
3.10	Diagrama de aprendizagem na subida do MR-VTOL.	27
3.11	Combinação linear pesada dos mapas de utilidade	29
3.12	Gráfico representativo do esquecimento nos mapas georreferenciados	30
3.13	Esquecimento do mapa de utilidade M_p representativo da posição da plataforma	31
3.14	Exemplo de dinâmica dos mapas	31
3.15	Duração estimada da pesquisa da plataforma	32
3.16	Diagrama da deteção da plataforma	33
4.1	Imagens do simulador Kelpie	36
4.2	Fotografia do MR-VTOL utilizado nos testes	37
4.3	Diagrama representativo das comunicações entre MR-VTOL e ROS	37
4.4	Posições relativas do plano testado	41
4.5	Início da aprendizagem durante o levantamento	41
4.6	Estrutura do ficheiro YAML	42

4.7	Mapas georreferenciados no início do plano de execução em simulação	42
4.8	Primeiro objeto de interesse durante o plano de execução em simulação	43
4.9	Mapas de saliência e zona de seguimento na primeira descida	43
4.10	Segundo objeto de interesse durante o plano de execução em simulação	43
4.11	Mapas de saliência e zona de seguimento na segunda descida	43
4.12	Mapa georreferenciados durante o plano de execução em simulação	44
4.13	Terceiro objeto de interesse durante o plano de execução em simulação	44
4.14	Mapas de saliência e zona de seguimento na terceira descida	44
4.15	Mapas georreferenciados no final do plano de execução em simulação	45
4.16	Quarto objeto de interesse durante o plano de execução em simulação	45
4.17	Mapas de saliência e zona de seguimento na quarta descida	46
4.18	Resultado do detetor dedicado durante uma descida	46
4.19	Detetor dedicado nos testes em ambiente aquático	47
4.20	Saliência visual nos testes em ambiente aquático	48
4.21	Pesquisa visual de zonas de interesse num teste com o robô	49
4.22	Aprendizagem da aparência da plataforma num levantamento com o robô	50
4.23	Aterragem do robô na plataforma	51

Lista de Tabelas

3.1	Descrição dos campos do modelo baseado em contrastes, W	22
3.2	Descrição dos mapas de utilidade.	29
4.1	Resultados dos testes em simulação do plano de execução.	46
4.2	Resultados da aterragem realizada pelo robô em ambiente terrestre	49

Lista de acrónimos e símbolos

α	Ponderação na classificação do detetor dedicado
α_p	Limiar para a classificação da deteção da plataforma
α_s	Ponderação na classificação de segmentos isolados de saliência
\bar{d}	Média das dissemelhanças das deteções dedicadas consideradas positivas
Δt	Duração da pesquisa no mapa de utilidade
Δt_e	Duração da pesquisa num elemento do mapa de utilidade
ΔX	Distância no plano do solo no referencial mundo
Δx	Distância no plano da imagem
\dot{p}	Derivada da intensidade de uma dada posição num dado mapa de utilidade
γ_r	Raio de posições próximas à posição atual do MR-VTOL
\mathbf{H}_c	Vetor dos momentos geométricos da circunferência c
\mathbf{H}_r	Vetor dos momentos geométricos de referência
p	Posição no mapa de utilidade
\mathbf{W}	Vetor de pesos para cálculo da saliência
\mathbf{W}^h	Vetor de pesos para cálculo da saliência para altitude h
\mathcal{K}	Referencial do MR-VTOL
\mathcal{V}	Referencial da plataforma
\mathcal{W}	Referencial da área de trabalho do MR-VTOL
BY	Mapa de dupla-oposição de cores azul-amarelo
C_{bu}^C	Mapa de conspicuidade de cor
C_{bu}^I	Mapa de conspicuidade de intensidade
F	Campo neuronal
GPS	Sistema de Posicionamento Global
IMU	Unidade de Medidas Inerciais

I	Mapa de intensidade
MR – VTOL	<i>Multirotor Vertical Take Off and Landing</i>
P^C	Mapa de feromona de cor
P^I	Mapa de feromona de intensidade
RG	Mapa de dupla-oposição de cores vermelho-verde
ROI	<i>Region of Interest</i>
S	Mapa de saliência
UTM	Sistema de coordenadas cartesianas bidimensional para localizações na Terra
VTOL	<i>Vertical Take Off and Landing</i>
YAML	<i>Ain't Markup Language</i>
σ	Varição máxima de tamanho dos segmentos no mapa de saliência em percentagem
σ_r	Varição do raio esperado para a plataforma no detetor dedicado
τ	Constante temporal associada ao esquecimentos nos mapas de utilidade
ξ_m	Valor empírico de referência para a dissemelhança máxima
a_e	Área no solo correspondente a um elemento no mapa de utilidade
a_w	Área de pesquisa no solo
C	Conjunto de circunferências
c	Circunferência
c_p	Comprimento do veículo que transporta a plataforma em metros
D	Conjunto de dissemelhança das deteções dedicadas consideradas positivas
d_c	Distância euclidiana dos momentos geométricos da circunferência c à referência
e_1	Margem de segurança
e_g	Erro associado ao GPS
f	Comprimento focal da câmara
h	Altitude do MR-VTOL ao solo
h_d	Altitude que inicia/termina a aplicação do detetor dedicado
h_{max}	Altitude máxima do MR-VTOL durante a aprendizagem
h_{min}	Altitude mínima do MR-VTOL durante a aprendizagem
i	Média de intensidade de um segmento isolado de saliência
l	Altura da imagem de entrada em pixels
l_p	Largura da plataforma em metros

M_a	Mapa de utilidade representativo da componente aleatória
M_f	Mapa de utilidade combinação linear dos restantes mapas
M_l	Mapa de utilidade representativo dos locais visitados
M_p	Mapa de utilidade representativo da posição da plataforma
M_v	Mapa de utilidade representativo da posição atual do MR-VTOL
n_d	Índice da imagem quando $h = h_d$
n_{min}	Índice da imagem quando $h = h_{min}$
p	Possibilidade da plataforma conhecida ter sido encontrada
q	Região de Interesse correspondente à plataforma
r	Raio da circunferência esperado
r_{ref}	Rácio entre tamanho do segmento isolado de saliência e o de referência
w	Peso atribuído ao mapa de utilidade

Capítulo 1

Introdução

1.1 MR-VTOLs: Tipos e Aplicações

Os veículos aéreos não tripulados multi-rotor, MR-VTOLs¹, podem ser utilizados em vigilância, salvamento, percepção sensorial remota, investigação científica e em operações que colocariam em risco vidas humanas. Devido às suas características multi-rotor, os MR-VTOLs diferenciam-se dos helicópteros pela capacidade de levantamento e aterragem verticais, pela elevada estabilidade, agilidade e possibilidade de utilização em espaços confinados. O desenvolvimento tecnológico destes veículos tende a diminuir a dimensão, melhorar a percepção sensorial e aumentar a autonomia energética, assim como o poder de processamento computacional.

Atualmente os MR-VTOLs são utilizados em aplicações de construção, em que o veículo transporta um determinado material e coloca no local pretendido, e.g., transporte e colocação de vigas e pilares, (Lindsey et al., 2011). No transporte de carga, em que se coloca o problema de um robô pequeno dispor de pouca capacidade de carga, então para cargas superiores, utilizam-se vários robôs para transportar a mesma peça (e.g., (Michael et al., 2011)). Contudo, nesta situação as pequenas dimensões dos robôs resultam numa diminuição da agilidade do conjunto. Em busca e salvamento, e.g., enviando um veículo para um edifício em ruínas para aceder aos estragos provocados por desastres naturais ou para monitorizar níveis de radiação, poluição (e.g., (Murphy and Cycon, 1999)). Neste caso pode encontrar-se uma pessoa ferida entre os escombros, que de outra forma não seria encontrada pois a presença de um humano no local para procurar outros, seria arriscada. Como se pode verificar, este tipo de veículos é utilizado em várias aplicações não estando terminado o seu estudo e evolução. O número de aplicações possíveis aumenta quando se trata de um sistema multi-robôs em coordenação e cooperação (e.g., (Mellinger et al., 2013)). Este aumento deve-se ao facto de um sistema apresentar maior potencial para realizar tarefas impossíveis de alcançar com um único robô, mesmo que de outro tipo, ou mesmo por um humano.

Um caso particular da aplicação em monitorização ambiental é a experiência RIVERWATCH² (Pinto et al., 2014) do Projeto Europeu FP7 de integração de larga escala ECHORD (*European*

¹Do inglês, *Multirotor Vertical Take-off and Landing*.

²Mais informações em: <http://riverwatchws.cloudapp.net>

Clearing House for Open Robotics Development), no qual se enquadrou esta dissertação. A experiência consiste num sistema autónomo multi-robôs para monitorização ambiental fluvial, composto por um veículo autónomo de superfície (ASV³) e um MR-VTOL transportado pelo de superfície. O veículo aéreo colmata as limitações de perceção do ambiente ao nível da superfície. O veículo de superfície pode ser equipado com um painel fotovoltaico cuja energia produzida serviria para se sustentar e armazenar para sustentar o veículo aéreo. Na interação dos robôs as perceções individuais do ambiente são partilhadas entre ambos, de modo a obter resultados difíceis de obter por um só robô. Deste modo, através da perceção visual do veículo aéreo é possível traçar um plano de navegação seguro para o veículo de superfície. Durante a operação de monitorização, no caso de ser detetado uma ameaça ou problema ao sistema, o veículo aéreo levanta para inspecionar a zona onde o veículo de superfície navega. A visão aérea pode detetar obstáculos ou uma alteração da margem, ao longo da rota que o veículo de superfície segue. Esta situação permite a alteração da rota atempadamente, evitando acidentes ou poupando tempo de navegação. Após realizar a tarefa, o MR-VTOL regressa ao veículo de superfície efetuando pesquisa e aterragem no mesmo, sendo este, o problema que esta dissertação pretende resolver.

1.2 O Problema da Aterragem

Apesar da elevada atividade de desenvolvimento, os MR-VTOL ainda carecem de certas capacidades que os impedem de serem utilizados em larga escala. Nomeadamente, o problema da aterragem devido a falhas provenientes de diversos fatores, e.g., a escolha e o processamento da perceção sensorial que tomam a posição principal na qualidade e fiabilidade da aterragem. Num veículo tripulado, apesar dos valores dos sensores disponíveis, o piloto tem ainda a perceção visual humana. O mesmo não acontece num veículo não tripulado em que a aproximação à visão humana nos robôs é realizada com o processamento de vídeo através de uma câmara. Assim, o problema atual passa por existir uma rotina de processamento capaz de detetar e decidir se o veículo está perante o local onde é suposto realizar a aterragem.

Existem alguns métodos candidatos à resolução do problema da aterragem automática, que generalizando se dividem em dois tipos. O primeiro tipo recorre à utilização de um local conhecido para aterrar, tipicamente uma marca específica (e.g., (Saripalli et al., 2003)). Desta forma pode-se aterrar num local pré-definido e conhecido. No entanto este tipo de aterragem apresenta a limitação de ser viável exclusivamente para baixas altitudes, pois à medida que a altitude aumenta, as características da marca deixam de ser perceptíveis na imagem. O segundo tipo tem como objetivo escolher um local seguro para aterragem, sem recorrer a marcas (e.g., (Garcia-Pardo et al., 2002)). Uma limitação deste tipo de aterragem é o facto de o veículo aterrar num local diferente do que se pretendia, i.e., no caso de existir um local mais seguro do que aquele onde se quer que o veículo aterre. No entanto serve para uma aterragem em segurança, em situações de emergência.

Com o objetivo de superar estas limitações, esta dissertação propõe um modelo que resolve o seguimento e deteção da plataforma para altitudes elevadas, mantendo a fiabilidade da deteção de uma determinada marca específica visualmente distinta para baixas altitudes. Para isso, o processo de pesquisa da plataforma divide-se em global e local. A pesquisa global consiste numa arquite-

³Do inglês, *Autonomous Surface Vehicle*.

tura comportamental multi-valor baseada em votação para uma pesquisa numa área abrangente. Em cada posição selecionada pela pesquisa global, aplica-se a pesquisa local que consiste numa pesquisa visual baseada em saliência visual, utilizando informação adquirida aquando do levantamento do MR-VTROL. Como opção possui um detetor dedicado para baixas altitudes capaz de detetar uma marca específica visualmente distinta, com a finalidade de ter um ajuste fino na decisão. Desta forma o sistema pode ser aplicado a qualquer plataforma em movimento, independentemente do que a está a movimentar, e.g., uma outra plataforma a ser movimentada por um robô terrestre. No caso particular desta tese, a plataforma apresenta uma marca específica devido ao facto de estar num barco e assim evitar falsos positivos levando a que o veículo aterre na água. Um considerável conjunto de testes em simulação e num robô real demonstram que em ambos os casos foram obtidos os resultados esperados.

1.3 Estrutura da Dissertação

Esta dissertação tem a seguinte estrutura:

Capítulo 2 analisa a literatura sobre aterragem em plataformas em movimento.

Capítulo 3 descreve o sistema proposto para deteção baseada em comportamentos e em saliência visual.

Capítulo 4 apresenta o aparato experimental e os resultados obtidos.

Capítulo 5 apresenta um conjunto de conclusões, principais contribuições e oportunidades de investigação na área.

Capítulo 2

Trabalho Relacionado

Ao longo da investigação de veículos aéreos autónomos têm surgido várias abordagens para o tema da aterragem automática, visto a importância que representa na utilização deste tipo de veículos, independentemente da sua função. Estas abordagens assentam essencialmente em dois pilares de suporte, GPS¹ e visão por computador. Neste capítulo são apresentadas algumas destas abordagens, trabalhos sobre modelos baseados em comportamentos e por fim trabalhos relacionados com saliência visual aplicada na robótica.

2.1 Aterragem de MR-VTOLs

Esta secção apresenta algumas soluções atuais para este problema, bem como as suas limitações. Na aterragem de veículos aéreos tripulados, apesar de existir informação captada por diversos sensores, sistemas de controlo que atuam baseados nos sensores, sistemas auxiliares de aterragem, entre outros, possui a vantagem de ter tripulação. Esta mais valia, comparativamente aos não tripulados, acrescenta toda a perceção sensorial e raciocínio por parte dos humanos a bordo. Assim, na tentativa de aproximação aos humanos, a aterragem automática de veículos aéreos continua a ser um problema desafiante e necessário para, e.g., a sua aplicação em veículos autónomos de pequenas dimensões. Como exemplo, Edwards et al. (2007) implementou um subsistema de visão a bordo do veículo aéreo para auxiliar a aterragem automática. Para tal, foi utilizada uma FPGA² como fonte de processamento e aquisição de dados. Para a deteção do local de aterragem através da visão é utilizada uma marca de uma só cor destacada do fundo, proporcionando uma deteção com um erro associado reduzido. Após a deteção é calculado o centro de massa da marca e assim definida a posição em que o veículo aéreo aterrará. A partir deste ponto, a tarefa de aterragem não depende da visão. Constata-se assim a aproximação à visão humana em que a deteção de uma mancha cuja cor se destaca do fundo é rápida e positiva. No entanto este sistema não inclui seguimento da marca, efetuando a aterragem na posição selecionada através de GPS, restringindo o sistema a locais com sinal GPS.

¹Do inglês, *Global Positioning System*

²Do inglês. *Field Programmable Gate Array*.

Uma solução atual para a aterragem de MR-VTOLs passa por todo o controlo ser realizado baseado na informação proveniente de um GPS equipado no veículo (Wendel et al., 2006). Entre outros sensores, esta solução utiliza uma IMU³, um dispositivo eletrónico com tamanho na ordem dos centímetros, constituído por um conjunto de três acelerómetros, três giroscópios e opcionalmente três magnetómetros, aplicados aos três eixos ortogonais do objeto em que está aplicada. Este dispositivo permite obter a velocidade, posição e orientação. Esta solução utiliza ainda um altímetro e um magnetómetro, necessitando de pouco poder computacional e capacidade de carga, o que leva a um baixo consumo de energia. São implementados dois modos de funcionamento para os casos onde há, ou não, sinal GPS. Esta abordagem apresenta as limitações do ruído, erro e ausências do sinal GPS e o facto de necessitar de comunicação com a plataforma terrestre para obter a posição onde vai aterrar.

Outra solução é uma abordagem bioinspirada nos insetos utilizando fluxo ótico (Horn and Schunck, 1981). (Herisse et al., 2012) utiliza uma textura plana móvel no solo para o controlo da posição vertical do veículo através do fluxo ótico para efetuar a aterragem automática. Esta abordagem utiliza uma IMU, uma câmara e quatro controladores, não existindo processamento a bordo. Um dos controladores faz a comunicação com a estação no solo onde é processada a informação sensorial e envia comandos para o MR-VTOL, impondo assim uma latência entre perceção e ação. Outra limitação aqui identificada é relativa à altitude máxima de 3 m, devido ao facto do fluxo ótico ser significativo apenas para baixas altitudes.

Outra solução baseia-se num sistema de visão a bordo do MR-VTOL que pesquisa pontos de referência naturais no ambiente, com base num mapa 3D de referência, para localizar o MR-VTOL no espaço (Wendel et al., 2011; Bosch et al., 2006). Esta pesquisa restringe-se a edifícios, árvores e outros elementos do ambiente que não alterem a sua geometria com o passar do tempo. Assim, a tarefa de aterragem é guiada pelos pontos de referência próximos ao local pré-definido para aterrar. O mapa 3D de referência é construído previamente a partir de imagens e vídeos, ambos capturados por câmaras de alta qualidade, num computador comum pois necessita de poder de processamento gráfico. Esta abordagem tem as limitações de necessitar de um estudo prévio do local onde o veículo aéreo opera e da existência de pontos de referência que cumpram a regra. Algumas abordagens no âmbito da pesquisa visual por pontos de referência utilizam técnicas idênticas ao SLAM⁴ com a finalidade de localizar o MR-VTOL. Também utilizando pontos de referência, (Cesetti et al., 2010) recorre ao algoritmo SIFT⁵ para que as características dos pontos de referência na imagem se mantenham para diferentes altitudes. O local de aterragem é selecionado em imagens de alta definição aéreas ou de satélite, a partir do qual são pesquisados e seguidos os pontos de referência e por fim, mais próximo do solo o piso é analisado comparando com uma base de conhecimento de texturas através de fluxo ótico. Esta abordagem utiliza os sensores GPS, câmara e sistema de referência de atitude e rumo AHRS⁶. Identificam-se as limitações de todo o processamento ser realizado por um computador em terra, necessitando de comunicação e da necessidade de existirem pontos de referência ideais, próximo ao local escolhido para a aterragem. No caso particular desta dissertação, o local de aterragem está rodeado de água onde não existem pontos de referência. No entanto a utilização do fluxo ótico poderia complementar o detetor dedicado proposto.

³Do inglês, *Inertial Measurement Unit*. Unidade de Medidas Inercias.

⁴Do inglês *Simultaneous Localization And Mapping*. Técnica para construir um mapa de um ambiente, ao mesmo tempo que se localiza.

⁵Do inglês, *Scale Invariant Feature Transform*

⁶Do inglês, *Attitude Heading Reference System*

Algumas abordagens foram estudadas com o objetivo de encontrar um local seguro para aterrizar. Desta forma, o local de aterragem não é previamente determinado pelo utilizador mas sim escolhido pelo sistema. Este tipo de abordagem pode ter aplicação para situações de emergência para uma aterragem segura. (Johnson et al., 2005) cria um mapa de relevo para pesquisar um local seguro. É utilizado um sistema de visão a bordo constituído por uma câmara e um altímetro. Esta abordagem tem como limitação o solo apresentar características relevantes que possibilitem o seguimento dos mesmos na imagem. Garcia-Pardo et al. (2002) apresenta um sistema com uma arquitetura baseada em comportamentos e é constituído por uma câmara, uma IMU, um GPS e um sensor conta rpm para motor. O processamento de imagem é realizado em terra através de uma ligação *wireless*, existindo uma latência nos resultados. Esta abordagem contempla ambientes desconhecidos e foi testada em diferentes condições de voo, no entanto a identificação do local seguro depende, entre outras, da assunção de que os obstáculos nas imagem estão bem segmentados de modo a serem identificados com pouco processamento. A abordagem de Templeton et al. (2007) também pesquisa locais seguros para aterragem mas em um helicóptero com um comprimento de 2,2 m, com as respetivas desvantagens associadas, comparativamente aos MR-VTOLs. O sistema é composto por 2 controladores, um GPS, uma IMU, uma câmara e uma comunicação *wireless* com uma estação terrestre para supervisão. A escolha do local é realizada a partir de uma reconstrução 3D criada pelo sistema de visão do veículo. Relativamente ao caso particular desta dissertação, uma abordagem de pesquisa de um local seguro poderia ser complementar ao modelo proposto, aplicável na situação em que o MR-VTOL está no ar, tem os níveis de energia baixos e não encontra a plataforma para aterrizar. De notar que este tipo de abordagem não substitui o modelo proposto.

Outro tipo de abordagem passa por procurar a plataforma desejada através da sua temperatura utilizando uma câmara térmica FLIR⁷. Yilmaz et al. (2001) apresentou um sistema capaz de seguir um alvo, definido pelo utilizador, através de um algoritmo estimador, *Mean Shift* (Comaniciu and Meer, 2002). Para os casos em que este algoritmo falha devido a movimentações bruscas do veículo, é aplicada uma compensação de movimento, assegurando a continuidade do seguimento do alvo. Esta abordagem não foi testada num veículo aéreo mas sim num computador utilizando uma sequência de imagens aéreas. Assim, tem-se a limitação do elevado poder computacional exigido para uma solução a bordo do veículo em tempo real. (Paravati et al., 2009) efetua o seguimento de objetos através da aplicação de um algoritmo genético. Este algoritmo visa otimizar a pesquisa, neste caso do alvo, através de uma sequência de imagens FLIR. A limitação deste sistema deve-se à necessidade de um utilizador para selecionar o alvo na imagem, no início do programa. Outra solução com infra-vermelhos (IR) é através de uma câmara da consola Nintendo Wii. Em (Wenzel et al., 2011) é apresentado um sistema constituído por um veículo aéreo e outro terrestre sem necessidade de comunicação para efetuar levantamento, seguimento e aterragem na plataforma. A posição relativa é determinada a partir da forma do padrão de LEDs IR, com a limitação de que o veículo terrestre não pode sair do campo de visão do aéreo. O MR-VTOL possui os sensores: 2 controladores, acelerómetro, magnetómetro, GPS, sensor de pressão, comunicação rádio e uma câmara Wii. A plataforma de aterragem contém 4 LEDs distribuídos de forma estratégica no plano horizontal e noutro perpendicular a este. Esta abordagem é limitada pela a sua utilização exclusiva em locais sem radiação solar pois interfere com a captação da câmara e pela altitude máxima do veículo aéreo de 2.5 m em relação à plataforma.

As abordagens referidas possuem, entre outras, a limitação e desvantagem de não existir uma marca única que permita aterrizar com maior probabilidade no local desejado. Para garantir a ater-

⁷Do inglês, *Forward Looking Infrared*

ragem num local único, surge a necessidade de utilizar marcas específicas como plataformas de aterragem para que possam ser detetadas por um sistema de visão. O trabalho de De Wagter and Mulder (2005) considera um xadrez como plataforma de aterragem, detetado por reconhecimento de padrões e fluxo ótico. Esta plataforma distingue-se pelo alto contraste entre preto e branco, pela dificuldade em encontrar semelhante no meio ambiente introduzindo falsos positivos e pelo baixo processamento de imagem exigido. Esta abordagem utiliza uma câmara com inclinação ativa, GPS e IMU, efetua reconstrução 3D do terreno durante a navegação. Contudo, utiliza um helicóptero cujo controlo é mais difícil e a estabilidade para veículos de pequena dimensão é menor, comparativamente com um MR-VTOL. Nas abordagens (Sharp et al., 2001) e (Yuan et al., 2005) aplicam-se plataformas constituídas por quadrados brancos dentro de um quadrado preto em fundo branco e o MR-VTOL é equipado por câmara, GPS e IMU. Apesar da plataforma e sistema serem idênticos, a segunda abordagem diferencia-se por não necessitar de calibração prévia da câmara pois esta é realizada no início da navegação. Uma marca muito comum no que respeita a qualquer tipo de VTOLs consiste na letra H. Existem diversas plataformas com esta marca com diferenças como tipo de letra, tamanho, contrastes de cores, contornos adicionais, etc. (Saripalli et al., 2002, 2003) é exemplo de utilização de uma plataforma com H em que utiliza a visão para deteção e reconhecimento do alvo e a visão juntamente com o GPS para a navegação. A deteção da plataforma é realizada com uma série de operações em cada imagem aérea: binarização, filtro de mediana, segmentação por objetos através da intensidade do pixel, momentos geométricos invariantes de cada objeto permitindo o reconhecimento do objeto procurado. O seguimento do alvo é efetuado apenas com o reconhecimento de objetos, sendo que esta informação é passada para o sistema de controlo cuja arquitetura é baseada em comportamentos. Saripalli and Sukhatme (2006, 2007) vem melhorar a abordagem anterior com a implementação de um filtro de Kalman para o seguimento da plataforma e um planeamento da trajetória através das velocidades do veículo aéreo e da plataforma. A deteção e seguimento da plataforma são realizados desde os 12 m de altitude até se efetuar a aterragem. Ambas abordagens utilizam um VTOL pelo que acarretam as limitações de estabilidade para veículos de pequenas dimensões. Feng et al. (2007) segue uma abordagem idêntica às de Saripalli et al. (2003) com a diferença de utilizar um classificador Bayesiano para os objetos presentes na imagem. Outra abordagem com a utilização de marcas utiliza quatro câmaras, (Meier et al., 2011), com a hipótese de poderem operar como dois pares estéreo e sem recorrer ao GPS. Este sistema possui uma qualidade de visão superior aos sistemas com visão monocular conseguindo melhores características da marca. A marca utilizada é um quadrado preto com quadrados brancos dentro.

Estas abordagens possuem a limitação da altitude, pois à medida que a altitude aumenta, mais difícil se torna a visualização das características da marca na imagem. A abordagem de Lange et al. (2008, 2009), visa solucionar a questão da altitude utilizando uma marca com vários níveis de grandeza, sendo facilmente detetável a diferentes altitudes. É utilizado um MR-VTOL com 53 cm de diâmetro equipado com GPS, IMU, câmara, sonar e sensor de fluxo ótico. O controlo é realizado por posição de GPS, sendo que na ausência de sinal é utilizado o sensor de fluxo ótico para controlar a velocidade e posição do veículo. A plataforma é constituída por 4 anéis brancos em fundo preto, tendo o anel maior 45 cm de diâmetro. Os rácios entre raio interior e exterior dos anéis são 85%, 75%, 65% e 50%, do maior para o menor. Apesar desta abordagem detetar a plataforma com erros inferiores a 5 cm, a conjugação dos parâmetros do processamento de imagem como a binarização, podem levar a conclusões erradas pois a diferença entre os rácios dos círculos é mínima. Assim, esta abordagem não resolve o problema da deteção para elevadas altitudes pois para uma marca com 45 cm de diâmetro, a altitude máxima testada são 2 m.

Todas as abordagens aqui descritas não solucionam o problema que esta dissertação pretende resolver. Apesar da diversidade de métodos e equipamentos, não existe uma abordagem que cumpra todos os requisitos propostos. Assim, esta dissertação visa a reunião e alteração de métodos já investigados para solucionar o problema da aterragem automática de MR-VTOLs.

2.2 Modelos Baseados em Comportamentos

Os robôs necessitam de uma forte componente perceptual para que possam agir ou reagir em função do ambiente que os rodeia. Para alcançar determinado objetivo, o robô utiliza sensores para, e.g., preparação e verificação de movimentos, presença de objeto, análise de vídeo para uma finalidade específica. Também se surgir uma situação de perigo, e.g., uma rajada de vento mais forte num robô aéreo, este deverá concluir o sucedido através da análise sensorial e assim voltar ao ponto onde estava. Desta forma, de um modo generalizado a percepção sensorial está na base da robótica.

No comportamento de um humano, parte das suas ações são resultado da análise de um conjunto de percepções, eventualmente para atingir um objetivo ou simplesmente para regularizar o funcionamento conhecido como normal. A título de exemplo, uma pessoa que está a trabalhar numa sala onde a temperatura está acima de 30 °C, vai procurar ligar a ventilação por forma a diminuir a temperatura. Do mesmo modo, na robótica, um comportamento processa informação proveniente de sensores e executa uma ação. Um sistema baseado em comportamentos está dividido em comportamentos que são independentes, i.e., funcionam isoladamente, no entanto um sistema complexo contém comportamentos compostos que utiliza um ou mais comportamentos primitivos. A descrição de um sistema deste tipo torna mais simples a leitura e compreensão do seu modelo, principalmente na fase de modelação, pois cada elemento liga a percepção à ação. Desta forma, os comportamentos funcionam em paralelo no sistema, sendo possível um comportamento ativar e desativar outros comportamentos. Como exemplo, um veículo aéreo durante a descida pretende estar centrado num determinado ponto X do solo. Neste caso tem-se o comportamento descer e o comportamento centrar no ponto X . A Fig. 2.1 apresenta o sistema de exemplo em que o nó de coordenação definirá a ação do veículo. Para o tipo competitivo o veículo ou desce ou se centra, enquanto que no tipo cooperativo o veículo desce e simultaneamente centra-se no ponto X .

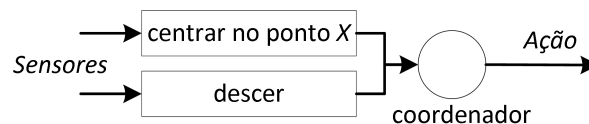


Figura 2.1: Exemplo de um sistema de controlo baseado em comportamentos

As arquiteturas baseadas em comportamentos não têm uma estrutura única específica, tendo alguns fatores que se adaptam de acordo com o sistema pretendido. Assim, diferenciam-se três dimensões (Arkin, 1998):

Estratégias de Coordenação - do tipo competitivo, um comportamento é independente e para determinados atuadores é o único ativo, tornando assim o sistema robusto. No entanto dois comportamentos podem estar em conflito num mesmo atuador, dando ordens controversas

quando estão ativos. Do tipo cooperativo, os comportamentos precisam de ajustes na sua contribuição no comportamento global, o comportamento global pode ser contrário ao esperado por alguns dos comportamentos que o constituem, no entanto o comportamento global apresenta uma solução suave e mais direta ao objetivo.

Complexidade de Comportamento - relacionado com a quantidade de tipos primitivos e compostos, i.e., comportamentos que têm na sua constituição um ou mais comportamentos primitivos. É preferível um sistema de baixa complexidade em que existe maioritariamente comportamentos primitivos e compostos pouco complexos. Desta forma, o sistema pode ser descrito até chegar a um comportamento final que engloba todos os outros.

Codificação da Resposta Comportamental - interface dos comportamentos dividem-se em dois tipos. Codificação discreta em que há uma quantidade finita de ações possíveis para comportamentos. Aplicável a ambos tipos de coordenação. Codificação contínua em que a quantidade de ações é infinita, e.g., potência fornecida a um motor para atingir dada aceleração. Geralmente aplicada à coordenação do tipo cooperativo.

Para uma leitura mais detalhada sobre sistemas baseados em comportamentos: Arkin (1998); Pfeifer et al. (2001); Pirjanian (1999).

Particularmente, um sistema de controlo baseado em comportamentos para um veículo aéreo foi implementado por Saripalli et al. (2002) e Montgomery (1999) utilizando um helicóptero. O sistema hierárquico divide o problema de controlo em vários comportamentos de fraca complexidade, cada um responsável por uma tarefa. Os comportamentos de baixo nível funcionam como reflexo e mantêm a estabilidade aérea do veículo. O Controlo de Direção atua no rotor da cauda em função do sensor IMU, a fim de manter o rumo desejado. O Controlo da Altitude utiliza o sonar para regular o Coletivo e o regulador de potência para os rotores. O Controlo de Inclinação e Rolamento mantém os ângulos recebidos pelo Controlo Lateral. Estes comportamentos são implementados com controladores proporcionais. Por fim, no topo do controlo tem-se o Controlo de Navegação que passa informação para o Controlo de Direção, Controlo de Altitude e Controlo Lateral. Desta forma tem-se um comportamento global complexo, constituído por vários comportamentos simples, que por sua vez são compostos por comportamentos primitivos.

2.3 Modelos de Saliência Visual na Robótica

De um modo simplificado, a saliência produz um mapa que identifica as zonas que se destacam do fundo. São utilizados dois métodos diferentes: com utilização de conhecimento à priori do objeto de interesse, método de cima para baixo⁸; sem utilização de conhecimento à priori do objeto de interesse, método de baixo para cima⁹. Durante processo de obtenção do mapa de saliência, são efetuadas diversas operações, e.g., relacionados com cor e intensidade, utilizando exclusivamente a imagem captada pela câmara (Mendonça, 2012), apresentados na Secção 3.2.2.

A saliência visual tem vindo a ser investigada e aplicada na robótica. Em (Ruesch et al., 2008),

⁸Conhecido como *Top-down* na literatura anglo-saxónica.

⁹Conhecido como *Bottom-up* na literatura anglo-saxónica.

a saliência visual é combinada com a saliência sonora de modo a guiar a atenção visual do robô humanoide, resultando na movimentação de olhos e pescoço. Nesta abordagem são criados vários mapas de saliência com diferentes informações com a finalidade de obter um mapa final. Este mapa é combinação dos restantes, sendo que cada mapa tem o seu peso de importância relativamente ao objetivo a alcançar. Entre os vários mapas tem-se um mapa de habituação que regista a forma como o sistema se acostuma aos estímulos mais repetitivos relativamente às novas zonas de saliência que vão aparecendo. Outro mapa é o de inibição que regista as últimas zonas de atenção visual, que tem o objetivo de não dar atenção a zonas que a tiveram num passado recente. Este mapa tem esquecimento associado a uma constante temporal.

Santana and Correia (2011) apresenta um sistema de navegação para um robô todo-o-terreno baseado em visão estéreo. Esta abordagem utiliza um conjunto de agentes virtuais que habitam a imagem de entrada e pesquisam por obstáculos de forma coletiva e ordenada. A atividade destes agentes é modificada em função da ação que o robô pretende realizar. Assim, a cognição de enxame permite detetar obstáculos salientes na imagem, reduzindo o processamento de análise de imagem e consequentemente o poder computacional necessário. Outra abordagem por Santana et al. (2011) apresenta uma abordagem híbrida para deteção de obstáculos em ambiente todo-o-terreno através de visão estéreo. A aplicação da saliência visual em uma das imagens de entrada permite isolar os obstáculos de maiores dimensões, enquanto que uma nuvem de pontos 3D visa detetar obstáculos mais pequenos. Assim, a saliência visual está em paralelo com um método complementar, melhorando os resultados do conjunto.

Santana et al. (2012) apresenta um sistema de visão para a deteção de trilhos para veículos terrestres em ambientes complexos. Utiliza uma abordagem baseada em redes neuronais e inteligência de enxame para influenciar o processo de saliência visual na aprendizagem da aparência do trilho. A camada de rede neuronal extrai características de baixo nível enquanto que a camada de inteligência de enxame extrai características de alto nível. A base do modelo de saliência desta abordagem está na base do modelo aplicado nesta dissertação.

Em Meger et al. (2008) tem-se a saliência visual aplicada num robô cujo objetivo é fazer reconhecimento de objetos num ambiente próximo do real, rotulando-os. Neste caso, a saliência visual está a ser aplicada em imagens capturadas por uma câmara estéreo que, considerando as distâncias calculadas aos obstáculos, resulta em zonas da imagem que representam possíveis objetos. Posteriormente outra camada do sistema analisa esta zona até rotular o objeto.

Todt and Torras (2000) utiliza a saliência visual para reconhecimento de marcas naturais em ambientes ao ar livre. Esta abordagem introduz o tema da atenção visual para deteção de marcas naturais sem conhecimento à priori das mesmas. Como resultado final têm-se zonas da imagem possíveis de conter uma marca natural.

Conclui-se que o potencial da saliência visual na robótica aumenta quando em cooperação com um método complementar específico para atingir o objetivo proposto. Esta dissertação segue este tipo de abordagem ao nível da pesquisa visual, em que a saliência visual cobre uma parcela na janela de altitudes que o sistema suporta, deixando a restante parcela para um método complementar. No entanto, esta dissertação propõe uma pesquisa global que vai além do campo visual. Esta pesquisa coordena a pesquisa visual de modo a evitar deslocações desnecessárias do MR-VTOL, poupando no consumo energético e diminuindo o tempo de pesquisa da plataforma de aterragem.

Capítulo 3

Modelo Proposto

Este capítulo descreve toda a arquitetura do sistema de visão proposto e desenvolvido para o problema da aterragem automática de MR-VTOLs. A Secção 3.1 faz uma introdução à visão geral do sistema proposto. A Secção 3.2 aborda todo o processamento de imagem com referência à saliência visual. A Secção 3.3 detalha a arquitetura baseada em comportamentos proposta, que inclui o levantamento e aterragem do MR-VTOL.

3.1 Visão Global do Sistema

Nesta secção são apresentadas algumas informações sobre os robôs utilizados, os requisitos que o sistema pretende cumprir e a solução proposta para tal, onde consta uma visão global de todo o sistema.

O modelo proposto nesta dissertação consiste num sistema de aterragem que aplica a saliência visual ao problema da detecção e seguimento da plataforma, complementando com um detetor dedicado para baixas altitudes. Este sistema possuiu uma camada de comportamentos relacionados com a pesquisa da plataforma, em que utiliza informações como os locais aéreos visitados e posição atual do MR-VTOL para prever a localização atual da plataforma. Desta forma o veículo aéreo terá comportamentos baseados nas informações passadas e atuais do veículo aéreo e plataforma, reduzindo assim o número de falsos positivos inerentes à pesquisa da plataforma através da visão.

3.1.1 O Sistema Robótico

Esta secção apresenta a relação de referenciais das entidades envolvidas e o modelo e calibração da câmara do MR-VTOL.

A existência de várias entidades em movimento num ambiente, sugere a especificação dos referenciais utilizados. A Fig. 3.1 mostra os 3 referenciais considerados. O referencial \mathcal{K} para o

MR-VTOL, o \mathcal{V} para a plataforma e o \mathcal{W} para a área de trabalho do MR-VTOL.

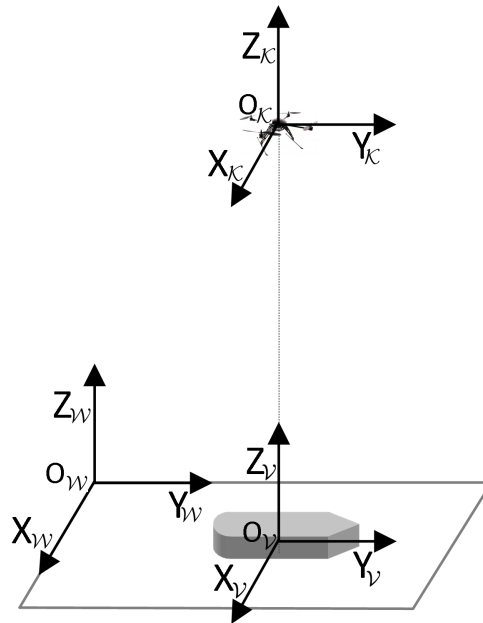


Figura 3.1: Representação dos 3 referenciais do sistema. Referencial \mathcal{K} : MR-VTOL; \mathcal{V} : plataforma; \mathcal{W} : área de trabalho do MR-VTOL, coincidente com o referencial mundo.

O MR-VTOL e a plataforma estão equipados com sensores de posicionamento global, GPS. No entanto, todo o sistema funciona com o sistema cartesiano, e por isso existe uma conversão bidirecional de coordenadas. Cada posição GPS obtida por um sensor é convertida para o sistema de coordenadas UTM¹. Esta unidade resulta da divisão longitudinal da Terra em 60 zonas, sendo que cada uma das zonas tem coordenadas relativas, Este e Norte. As fórmulas de conversão estão no Apêndice A. Assim, o referencial \mathcal{W} da área de trabalho do MR-VTOL fica associado com o referencial mundo no sistema de coordenadas UTM. Os referenciais \mathcal{K} e \mathcal{V} são obtidos através de altímetros equipados no MR-VTOL e na plataforma.

A câmara que equipa o MR-VTOL é monocular e necessita ser calibrada antes da sua fixação ao veículo. A calibração resulta numa matriz de valores referentes às características internas da câmara. O processo é realizado através de um método que analisa os movimentos realizados com um xadrez impresso num placar rígido plano. Este método deteta as arestas dos quadrados do xadrez, que cruzando com a informação das medidas reais do xadrez no sistema métrico, permite efetuar os cálculos necessários para obter as características da câmara. Neste caso o interesse da calibração remete para a aquisição do comprimento focal, f . Através do conhecimento deste valor e da altitude do MR-VTOL ao solo, h , é possível criar uma relação entre distâncias em pixels na imagem, Δx e a distância em metros no solo, ΔX , segundo a relação de projeção baseado no modelo *pinhole*:

$$\Delta x = \frac{f \cdot \Delta X}{h} \quad (3.1)$$

Estas relações são válidas pois assume-se a perpendicularidade do MR-VTOL ao plano do chão, sendo que a câmara está aplicada de forma que o eixo ótico está alinhado com $Z_{\mathcal{W}}$, como mostra a

¹Do inglês, *Universal Transverse Mercator*. UTM é um sistema de coordenadas cartesianas bidimensional para localizações na Terra.

Fig. 3.2 em que o plano oxy é o plano da imagem.

Neste trabalho a altitude do MR-VTOL ao solo pode ser estimada através de dois métodos: (1) pelo valor de saída do altímetro ou (2) através da Eq. 3.1, se estiver presente na imagem um objeto bem definido e cujas dimensões sejam conhecidas. O segundo método é utilizado quando a informação do altímetro está inacessível. Esta situação pode surgir por falha de comunicação entre o sistema de processamento e o sensor, falha do sensor aquando da leitura ou valor lido inválido, i.e., fora dos limites dos valores esperados.

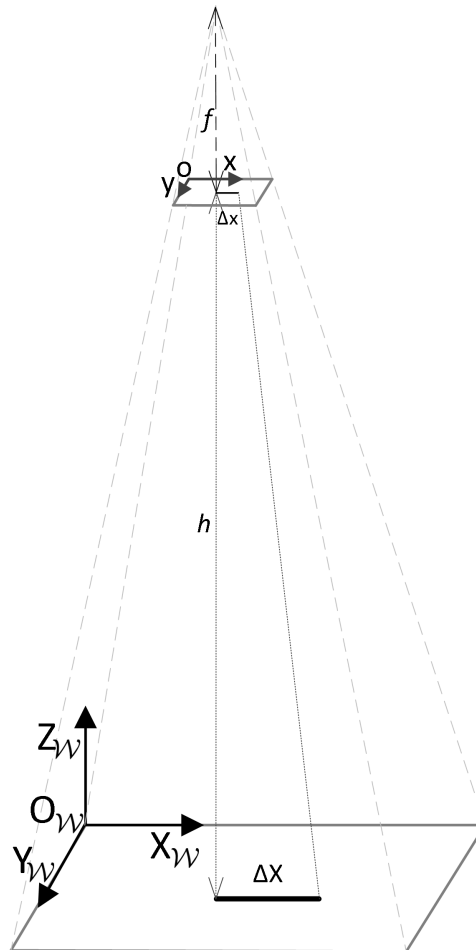


Figura 3.2: Diagrama da relação de projeção baseado no modelo *pinhole*. Esta relação é possível devido à assunção da verticalidade do MR-VTOL. oxy é o plano da câmara.

3.1.2 A Tarefa

Uma vantagem associada à cooperação entre o MR-VTOL e a plataforma consiste na reunião de informação perceptual de ambos, tendo em vista uma decisão mais robusta. A informação recolhida por um, complementa a recolhida pelo outro, podendo assim definir a ação de ambos. Contudo, esta cooperação pode não existir por diversos motivos, pelo que nesse caso, o sistema está preparado para continuar o seu fluxo de trabalho. De seguida são apresentados os cenários possíveis de levantamento e aterragem do MR-VTOL e a respetiva ligação à cor do trajeto na Fig. 3.3.

O MR-VTOL levanta a partir da plataforma e segue o seu plano de tarefas, enquanto o veículo que

transporta a plataforma pode também seguir o seu plano de tarefas. Quando o MR-VTOL necessita aterrar, começa por pesquisar a plataforma no seu campo de visão atual, deteta-o e aterriza; situação a verde na figura. Quando a plataforma não é encontrada, o MR-VTOL terá que deslocar-se para a posição que recebe da plataforma através do canal comum de comunicação; situação a vermelho na figura. Ao chegar à posição, o sistema de visão guia o MR-VTOL até à plataforma. Pode no entanto surgir a situação de falhar a comunicação entre os veículos. Neste caso o MR-VTOL terá de efetuar uma busca através da visão em forma de espiral, a partir da sua posição atual até encontrar a plataforma; situação a castanho na figura. De notar que a espiral consiste no percurso do MR-VTOL pelas localizações onde efetua a pesquisa visual, sendo que em cada localização pode efetuar descida para verificar a presença da plataforma. Se durante esta busca, a comunicação voltar a existir, o MR-VTOL navega para a posição que recebe da plataforma e aterriza pelo sistema de visão; situação a azul na figura.

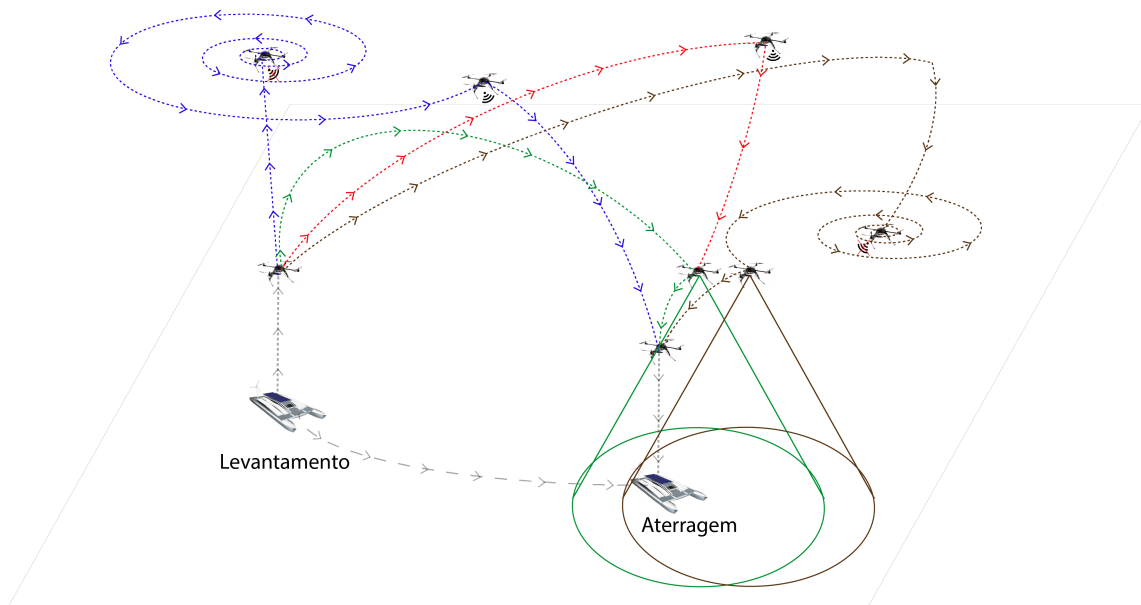


Figura 3.3: Descrição das 4 tarefas do MR-VTOL em que as linhas tracejadas correspondem aos seus trajetos. A cinzento: o levantamento e a aterragem são comuns a todas as tarefas. A verde: o MR-VTOL levanta a plataforma e segue o seu plano de tarefas, enquanto a plataforma de aterragem segue um outro percurso. Quando o MR-VTOL necessita de aterrar, pesquisa a plataforma no seu campo de visão atual, deteta-o e aterriza. A vermelho: situação em que a plataforma não é encontrada, o MR-VTOL desloca-se para a posição que recebe da plataforma através do canal comum de comunicação e aterriza. A castanho: situação de falha na comunicação entre os veículos. O MR-VTOL efetua uma busca através da visão em forma de espiral, a partir da sua posição atual até encontrar a plataforma e aterriza. A azul: durante a busca em espiral, a comunicação volta a existir, o MR-VTOL navega para a posição que recebe da plataforma e aterriza.

3.1.3 A Solução

O sistema proposto consiste no processamento do vídeo capturado pela câmara que equipa o MR-VTOL, de modo a que este possa efetuar levantamento, pesquisa e aterragem numa plataforma em movimento. No que respeita à pesquisa da plataforma divide-se em global e local. A pesquisa global é realizada através de um conjunto de mapas de utilidade cuja escolha e a aplicação dos mesmos, permite obter a posição mais provável da plataforma e proporciona uma dinâmica de pesquisa em espiral. Para cada posição provável da plataforma é aplicada a pesquisa local, realizada durante a descida. No caso da plataforma não ser detetada, recorre-se novamente à pesquisa global e assim

sucessivamente até à deteção da mesma.

Durante a subida é efetuada uma aprendizagem da aparência da plataforma, baseada em saliência visual. Esta aprendizagem será utilizada na pesquisa e seguimento da plataforma aquando da descida. Entre outros sensores, cada veículo possui um sistema de posicionamento global, GPS. Assume-se a existência de uma comunicação direta entre os dois veículos para que a posição atual de ambos possa ser partilhada, estando prevista a ausência de comunicações. O MR-VTOL guarda informação das posições das pesquisas anteriores e deste modo prevê qual a posição atual mais provável da plataforma. Após a subida, o MR-VTOL terá um plano de ações com diversas tarefas fora do âmbito deste trabalho.

Na descida, efetua-se a pesquisa global (ver Secção 3.3.3) e para cada posição provável da plataforma aplica-se a pesquisa local (ver Secção 3.3.3), em que são utilizados dois detetores cuja seleção depende da altitude atual do MR-VTOL. Esta pesquisa, ao contrário da global, só utiliza informação sensorial. Um detetor baseado em saliência visual é aplicado para detetar regiões de interesse, prioriza-las e fazer o seguimento da região durante a descida. A partir de determinada altitude passa a ser aplicado um detetor dedicado (ver Secção 3.2.1) que explora o facto da plataforma ter uma marca específica. Este detetor apresenta resultados com uma maior fiabilidade pois são pesquisadas e comparadas características da marca. No entanto não é aplicado para elevadas altitudes devido à fraca definição das características da marca na imagem. No final da descida, atingida a altitude mínima de segurança, a região de interesse é classificada a fim de saber se corresponde à plataforma. Esta classificação baseia-se nos resultados obtidos pelo detetor dedicado. Caso a região de interesse não contenha a plataforma, passa-se à região seguinte da prioritização. Se a plataforma não for encontrada em nenhuma das regiões de interesse, parte-se para o próximo elemento mais provável dos mapas de utilidade, correspondendo a uma pesquisa local noutra porção do solo.

Entrando em mais detalhe na descida, no caso de serem analisadas todas as regiões de interesse presentes na imagem e a plataforma não foi encontrada, por-se-á em causa a aprendizagem realizada na subida. Assim, é aplicada uma saliência sem aprendizagem de modo a surgirem novas regiões de interesse. No caso de falha, o MR-VTOL recorre à posição seguinte mais provável nos mapas de utilidade. Com estes mapas, uma pesquisa em espiral sucede quando a informação que o MR-VTOL possui sobre a posição da plataforma está incorreta ou não existe. Neste caso a pesquisa é realizada a partir da posição atual do MR-VTOL, seguindo para locais próximos do mesmo, onde ainda não foi realizada uma pesquisa. Nesta pesquisa em espiral, entre cada elemento do mapa aplica-se a pesquisa local e se a plataforma não for detetada, volta à pesquisa global.

Como a plataforma é conhecida à priori, o detetor dedicado tem valores esperados para as medidas da mesma. Para conhecer a relação entre distâncias em pixels na imagem e distâncias em metros no solo é necessária a calibração da câmara, que permite a aplicação da relação de projeção baseada no modelo de *pinhole*. A calibração é efetuada uma só vez para determinada câmara, requerendo uma nova calibração se a câmara for substituída.

A partir do sistema apresentado, identificam-se comportamentos como descer, subir e mover, cuja complexidade é menor comparativamente ao sistema como um todo. Os comportamentos propostos são apresentados na Secção 3.3.

3.2 Pesquisa Visual

A utilização de um modelo computacional de pesquisa visual tem como principal objetivo a pesquisa e seguimento da plataforma no campo visual do MR-VTOL. Para uma maior robustez o sistema necessita de aprender as características visuais que tornam a plataforma saliente no seu ambiente, i.e., um modelo da plataforma. O facto de se saber que o MR-VTOL levanta da plataforma é explorado para guiar o processo de aprendizagem. Concretamente o sistema assume a presença de uma marca conhecida bem definida a baixas altitudes para determinar a posição relativa da plataforma. Dado que a aparência da plataforma varia com a altitude a que esta é observada, a aprendizagem do modelo da plataforma é realizada em função da altitude. Com este processo de aprendizagem, o sistema fica menos dependente da capacidade de visualizar todos os detalhes da marca presente na plataforma. Nesta pesquisa serão apresentados 3 patamares de altitudes de referência h_{min} , h_d e h_{max} .

Na Secção 3.2.1 é apresentado o detetor da marca que representa a plataforma. A Secção 3.2.2 faz uma breve explicação sobre modelo de saliência visual utilizado nesta dissertação. A Secção 3.2.3 explica o processo de aprendizagem da aparência da marca e do barco aquando da subida do MR-VTOL. A Secção 3.2.4 referencia a aplicação da aprendizagem para pesquisar a plataforma através da saliência visual e a Secção 3.2.5 apresenta o algoritmo de seguimento do resultado desta pesquisa.

3.2.1 Detetor Dedicado

O detetor dedicado é utilizado na pesquisa visual a baixas altitudes e consiste numa rotina de processamento de imagem para a deteção de uma marca conhecida à priori pelo sistema.

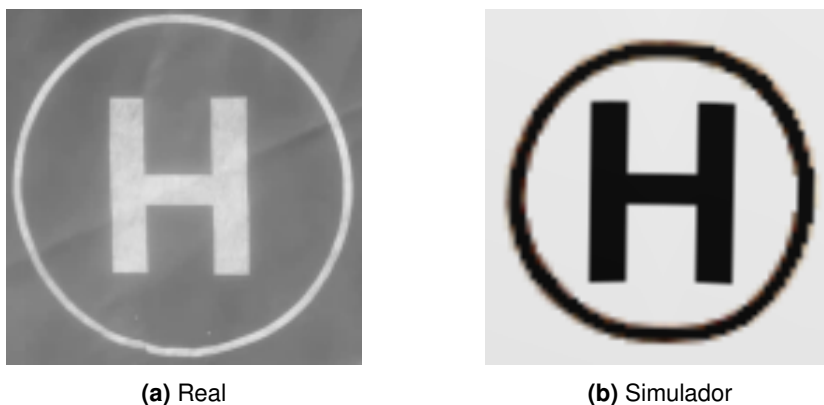


Figura 3.4: Marcas específicas para deteção da plataforma. Estas marcas são utilizadas no simulador e nos testes reais e por isso servem de referência para o detetor dedicado classificar a região de interesse. A comparação dos momentos geométricos da marca de referência com os da região provável, decide a presença da plataforma procurada.

A marca escolhida foi a letra 'H' circunscrita por uma circunferência em branco e preto para um melhor contraste, mostrada na Fig. 3.4. Para a marca de referência é calculado e guardado o vetor dos seus momentos geométricos invariantes, \mathbf{H}_r (Hu, 1962). Os momentos geométricos invariantes são descritores que permitem descrever uma imagem em escala de cinzentos de forma invariante

com a escala, rotação e translação. Um vetor de momentos geométricos invariante contém 7 valores cujo o processo de cálculo pode ser consultado no Apêndice B.

O módulo detetor começa por pesquisar circunferências de raio Δx na imagem, sendo este o valor aproximado correspondente ao raio real da circunferência, r , a uma altitude h_{max} , através da Eq 3.2, obtida da Eq. 3.1.

$$\Delta x = \frac{f \cdot r}{h_{max}} \quad (3.2)$$

De modo a introduzir flexibilidade no sistema, a pesquisa é realizada com uma janela de valores de raio com uma variação de σ_r relativamente ao esperado. Assim, esta pesquisa tem como resultado um conjunto de circunferências, C . Para cada circunferência, $c \in C$ é criada uma região de interesse retangular, justa ao objeto, e é calculado o vetor dos seus momentos geométricos, \mathbf{H}_c . Os momentos \mathbf{H}_c são comparados com os momentos de referência, \mathbf{H}_r , utilizando a distância euclidiana dos respetivos campos dos momentos, d_c , de forma a obter a sua dissemelhança:

$$d_c = \sqrt{\sum_{n=1}^7 (H_r[n] - H_c[n])^2} \quad (3.3)$$

Após o cálculo da dissemelhança pela Eq. 3.3, que utiliza todos os valores dos momentos geométricos invariantes, será considerada a marca de levantamento, a região de interesse cuja dissemelhança é menor e inferior a um dado valor empírico de referência ξ_m :

$$q = \arg \min_{c \in C} (d_c : d_c < \xi_m) \quad (3.4)$$

onde q é a região de interesse que corresponde à marca de levantamento. Se a condição da Eq. 3.4 não se verificar, significa que nenhuma das circunferências encontradas na imagem é a marca de levantamento. O valor da dissemelhança é posteriormente escalado por ξ_m , por forma a obter valores entre $[0, 1]$, para o cálculo da possibilidade de detetar a plataforma depois de uma descida. A Fig. 3.5 resume o processo de deteção utilizando o detetor dedicado.

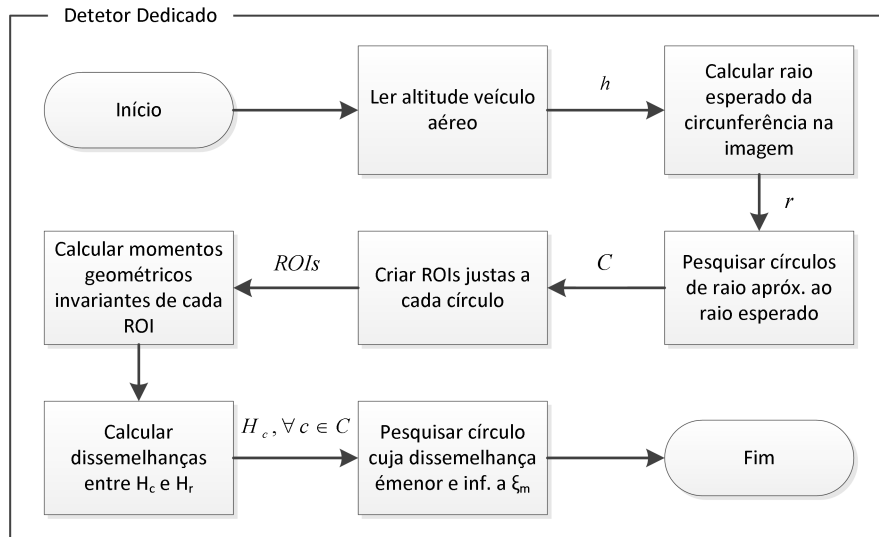


Figura 3.5: Rotina perceptual do detetor dedicado. Este detetor é aplicado a baixas altitudes, i.e., inferiores a h_d , e tem como objetivo a classificação da presença da marca de referência.

3.2.2 Modelo de Saliência Visual Base

Esta secção descreve de forma generalizada o processo de saliência visual base e o modelo de saliência utilizado nesta dissertação.

O processamento computacional que resulta na saliência visual, pretende replicar os fenómenos que acontecem ao nível do córtex visual. Pretende-se que no final do método sejam identificadas zonas e objetos salientes, i.e., que se destacam, na imagem. No final tem-se um sistema de visão por computador que objetiva a aproximação à visão humana, em que numa fotografia identifica o que se destaca através de cor, textura, tamanho, etc.

De seguida, é descrito o fluxo de operações da saliência visual aplicada na robótica utilizada por Itti et al. (1998) e Frinrop (2006), que serviu de base para todo o trabalho até alcançar o algoritmo utilizado nesta dissertação. O modelo base começa por calcular 3 mapas de características a partir da imagem de entrada. Estes mapas são: (1) de intensidade, I ; (2) de dupla-oposição de cores vermelho-verde, RG ; (3) de dupla-oposição de cores azul-amarelo, BY (ver fórmulas no Apêndice C). Cada um dos mapas passa por um mecanismo de centro-periferia, criando mapas de conspicuidade em que as zonas que se destacam da periferia apresentam maior luminosidade. O mapa de intensidade resulta no mapa de conspicuidade de intensidade C_{bu}^I . Uma média ponderada dos dois mapas de conspicuidade de oposição de cores geram um único mapa de conspicuidade de cor C_{bu}^C . Na obtenção dos mapas de conspicuidade são realizadas normalizações que fazem os mapas com menor número de elementos salientes ficarem mais fortes. Desta forma, a normalização atribui mais significado aos mapas que apresentam menos ruído na saliência. O mapa de saliência visual final, S resulta da média aritmética dos dois mapas de conspicuidade, C_{bu}^I e C_{bu}^C .

Este modelo computacional de atenção visual baseado em saliência visual é suficiente para a deteção de regiões de interesse, tendo em conta a sua arquitetura de baixo para cima. No entanto a sua aplicação pode originar um elevado número de falsos positivos, i.e., identificar zonas como

salientes mas na realidade não correspondem ao que se pretende.

Baseado neste modelo, Santana et al. (2010) propôs um modelo computacional baseado em inteligência de enxame, para detecção de trilhos como objeto de interesse. O sistema utiliza o mesmo fluxo de operações do modelo base explicado anteriormente até obter os mapas de conspicuidade C_{bu}^C e C_{bu}^I . Estes mapas são depois filtrados utilizando o conhecimento à priori das formas típicas do objeto de interesse. Um enxame de formigas virtuais são colocadas em cada um dos mapas, resultando em dois mapas de feromona, P^C e P^I . A média aritmética destes dois mapas resulta no mapa de saliência final S .

Para que o resultado de um mapa de feromona influencie os restantes, é utilizado um campo neuronal F alimentado pelo mapa S . Este campo neuronal integra feromona entre imagens para depois influenciar a criação de cada mapa de feromona, existindo assim uma componente de histórico. O princípio básico de F consta de um neurónio por pixel que excita os vizinhos e inibe os restantes, promovendo o agrupamento perceptual e reduzindo ambiguidade no foco de atenção. Por fim, de salientar que o modelo contém um mecanismo de compensação do campo neuronal em função da movimentação da câmara entre imagens.

Este modelo tem a capacidade de manter o objeto de interesse com alto valor provável mesmo na presença de distratores, graças à interação entre o campo neuronal e a capacidade sensor-motor das formigas virtuais. No entanto se a conspicuidade do objeto de interesse for baixa durante algumas imagens, por diferentes causas, o campo neuronal não é suficiente para que o enxame não saia do objeto de interesse, propagando-se pela sua periferia. Assim, estas limitações foram superadas com o modelo proposto por Santana et al. (2012), explicado de seguida.

O modelo de saliência utilizado nesta dissertação resulta de modificações ao modelo explicado anteriormente, de modo a incorporar um mecanismo de aprendizagem em tempo real do conhecimento de cima para baixo do objeto de interesse a ser procurado. Assim, a aparência do objeto de interesse que se pretende detetar, é atualizado ao longo da aplicação da saliência visual. Concretamente, os mapas de conspicuidade C_{bu}^C e C_{bu}^I são gerados a partir da imagem capturada I , pelo processo explicado anteriormente. Após a primeira aprendizagem, o cálculo destes mapas é influenciado por um modelo baseado em contrastes, W , que consiste num conjunto de pesos aprendidos em tempo real, representando a importância de uma dada característica visual para a detecção do objeto de interesse. Cada um dos pesos corresponde à ponderação de cada um dos mapas apresentados, no cálculo do mapa de saliência final. Desta forma, a ponderação levará a uma maior importância aos mapas em que o objeto de interesse está saliência e menos importância aos mapas em que a zona mais saliente não corresponde ao objeto de interesse. A Tabela 3.1 indica a ordem dos pesos dos mapas em W . Assim, são obtidos os mapas de conspicuidade C_{td}^C através dos 4 mapas de cores e C_{td}^I através dos 2 mapas de intensidade. Depois, cada um deles é normalizado e realizada a sua média aritmética resultando no mapa de saliência final S .

Este modelo, (Santana et al., 2012), tem uma componente de mapas de feromona que não é utilizada nesta dissertação e por isso não está aqui detalhada. No entanto, esta componente utiliza o conhecimento da aparência do objeto de interesse para definir os níveis de feromona depositados pelo enxame, enquanto que no modelo anterior, a quantidade de feromona depositado era constante. Assim, os mapas apresentam maior quantidade de feromona nas regiões da imagem em que a aparência é idêntica à de um objeto de interesse.

Índice	Descrição
0	Intensidade On-Off
1	Intensidade Off-On
2	Vermelho-Verde On-Off
3	Vermelho-Verde Off-On
4	Azul-Amarelo On-Off
5	Azul-Amarelo Off-On

Tabela 3.1: Descrição dos campos do modelo baseado em contrastes, W .

3.2.3 Inclusão da Aprendizagem em Função da Altitude

A saliência visual aplicada a uma imagem resulta num mapa de saliência em que o objeto de interesse fica saliente, neste caso, a plataforma. Como a aprendizagem do modelo da plataforma é realizada durante o levantamento do MR-VTOL, a aparência da plataforma altera com a posição da câmara. Deste modo, o mapa de saliência calculado para uma determinada altitude, pode não ser o mais indicado para uma outra altitude. Surge a necessidade de ajustar os parâmetros do cálculo dos mapas de saliência em função da altitude do MR-VTOL.

Durante o levantamento assume-se que a plataforma é detetada com sucesso na imagem através da aplicação do detetor dedicado (ver Secção 3.2.1). A partir da posição da plataforma na imagem são calculados os pesos de cada um dos 6 mapas referidos na Tabela 3.1, com o objetivo de criar um mapa de saliência em que o objeto de interesse mais se destaca. No entanto, estes pesos aplicados na saliência de uma outra imagem aérea a uma altitude superior de mais 10 m, podem remeter a falsos positivos. Assim, ao ajustar estes pesos para esta nova altitude consegue-se destacar o objeto de interesse, visto que a aparência do mesmo está diferente na imagem. A Fig. 3.6 e a Fig. 3.7 apresentam comparativos em que a primeira linha utiliza os mesmo pesos independentemente da altitude e a linha de baixo utiliza os pesos aprendidos para a altitude atual. Outro parâmetro da aprendizagem em função da altitude é um retângulo na imagem, justo ao objeto de interesse. Este retângulo é utilizado na descida como tamanho esperado do objeto de interesse na imagem, para a altitude atual. É obtido a partir do mapa de saliência calculado com os pesos correspondentes à altitude atual. Não se utiliza o diâmetro da circunferência como referência pois à medida que a altitude aumenta, a zona saliente passa a ser retangular pois o objeto de interesse passa a ser o veículo que contém a plataforma. Esta alteração do objeto de interesse não é imposta e resulta das pequenas dimensões da plataforma na imagem para elevadas altitudes. Desta forma, o processo da saliência vai ajustando o objeto de interesse entre plataforma e veículo que a transporta, em função da altitude.

No final do primeiro levantamento do MR-VTOL, o sistema aprendeu o modelo da plataforma que resultou num conjunto de pesos e num retângulo em função da altitude. Este processo está detalhado na Secção 3.3.2.

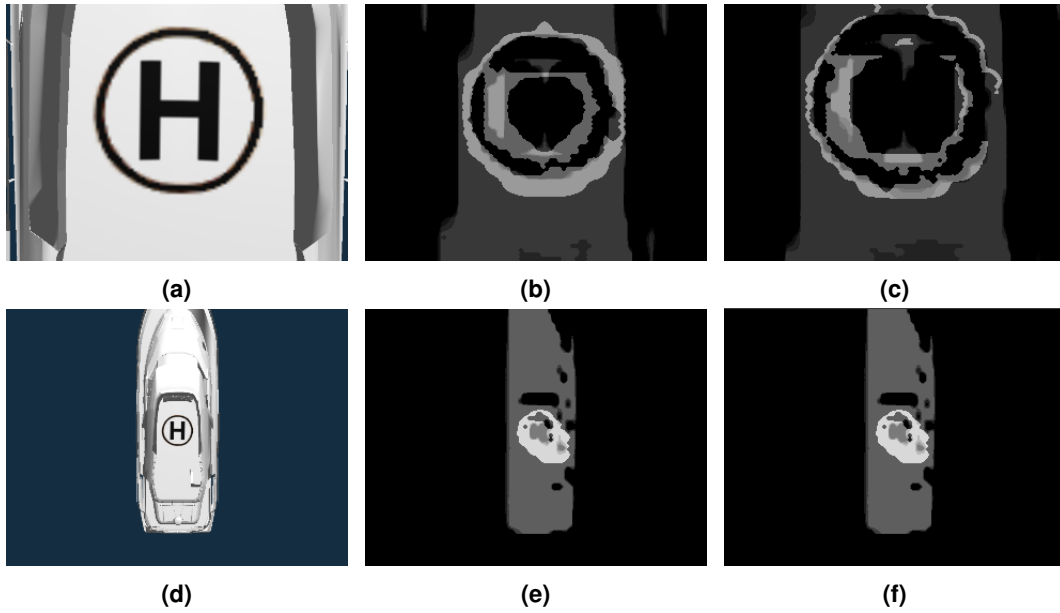


Figura 3.6: Comparação entre aprendizagem contínua da saliência em (b) e (e) e aprendizagem por altitudes em (c) a (f). As imagens (a) e (d) correspondem às imagens de entrada.

3.2.4 Pesquisa no Mapa de Saliência em Função da Altitude

A aprendizagem do modelo da plataforma realizada durante o primeiro levantamento permite posteriormente calcular o mapa de saliência de uma imagem capturada a determinada altitude. Ao fazer este cálculo, a plataforma deverá estar saliente em relação ao fundo, permitindo identificar a sua localização.

A pesquisa visual que utiliza o mapa de saliência é realizada na pesquisa local (ver Secção 3.3.3). Para tal é necessário o conhecimento da altitude atual do MR-VTOL proveniente do altímetro. Nesta fase, o veículo deverá estar próximo da altitude máxima, h_{max} , onde são identificadas as posições prováveis de conter a plataforma. Começa-se por calcular o mapa de saliência utilizando os pesos correspondentes à altitude atual. No mapa de saliência são identificados segmentos isolados de saliência e filtrados pelo tamanho. Como tamanho de referência utiliza-se o comprimento, c_p , do veículo que transporta a plataforma, convertido para pixels, em função da altitude atual do MR-VTOL. Este filtro exclui os segmentos cujo lado maior do retângulo justo ao mesmo, tem uma variação superior a σ relativamente ao tamanho de referência. Os segmentos restantes são depois classificados pela intensidade da saliência e proximidade do seu tamanho ao de referência. A Fig. 3.8 apresenta algumas fases deste processo, num caso em que os dois segmentos isolados de saliência passaram no filtro.

De seguida, para cada segmento é calculada a média da intensidade da saliência, i e o rácio, r_{ref} , do tamanho do segmento com o de referência, ambos normalizados através de valores empíricos de referência. O valor ponderado, u , obtido pela Eq. 3.5, resulta da ponderação dos dois critérios através do fator α_s . A classificação resulta numa lista destes valores, em que o segmento mais provável de conter a plataforma apresenta o valor mais elevado.

$$u = \alpha_s \cdot i + (1 - \alpha_s) \cdot r_{ref} \quad (3.5)$$

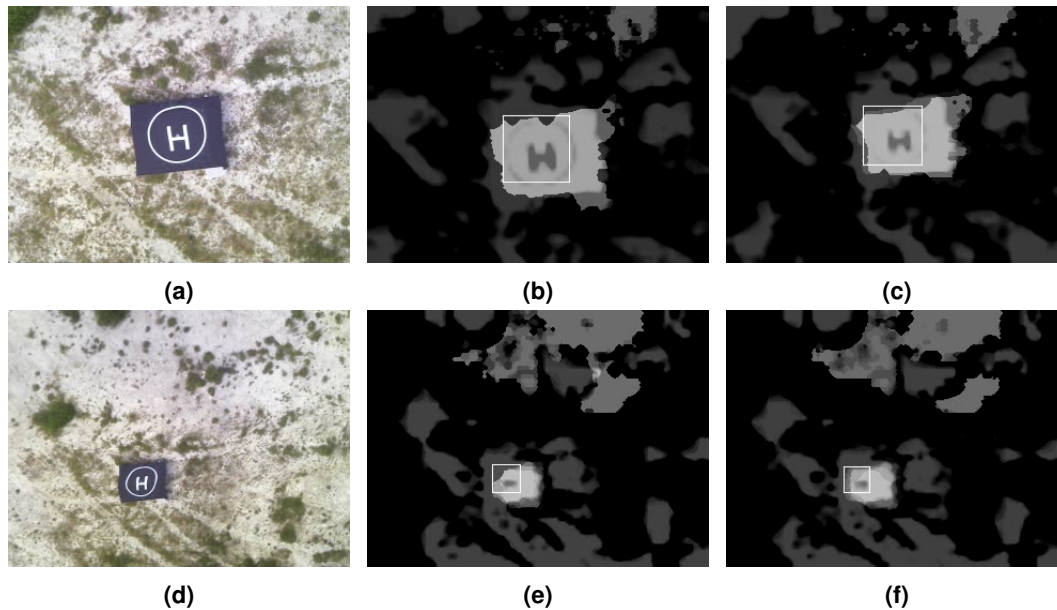


Figura 3.7: Comparação entre aprendizagem contínua da saliência em (b) e (e) e aprendizagem por altitudes em (c) a (f). As imagens (a) e (d) correspondem às imagens de entrada.

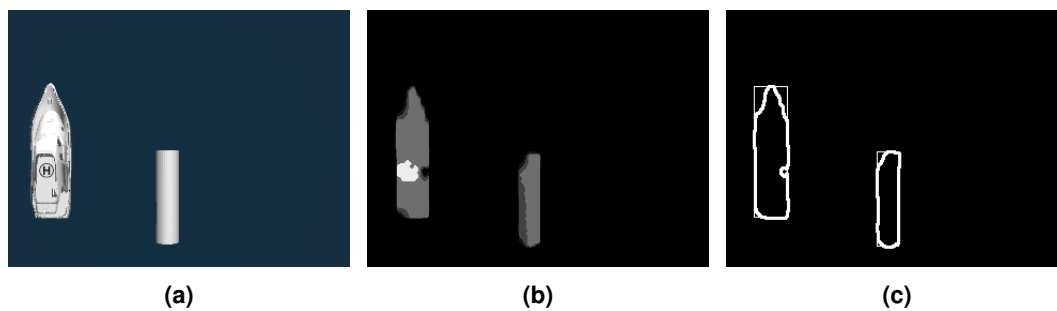


Figura 3.8: Pesquisa de segmentos isolados de saliência. (a) é a imagem de entrada; (b) é o mapa de saliência correspondente ; (c) são os contornos dos elementos encontrados após aplicação do filtro.

Após a classificação é calculada a posição absoluta dos objetos salientes no campo de visão. O veículo efetua a descida às posições absolutas conforme a prioridade definida, até encontrar a plataforma. Após visitar todas as posições absolutas encontradas no mapa de saliência e a plataforma não foi encontrada, coloca-se em causa a aprendizagem realizada aquando do levantamento. Assim, efetua-se uma pesquisa idêntica à descrita, utilizando a saliência sem conhecimento à priori da plataforma. Se no final a plataforma não foi encontrada conclui-se que a plataforma não está presente na porção de terreno em análise.

3.2.5 Seguimento no Mapa de Saliência

No levantamento e na aterragem do MR-VTOL o sistema de visão tem um papel fundamental para uma aterragem bem sucedida. Durante o levantamento é recolhida informação sobre a aparência da plataforma, pelo que se a informação não estiver correta, a aterragem não terá sucesso. Assim, surge a necessidade de garantir que o MR-VTOL se eleva de modo a que o seu campo de visão esteja centrado com a plataforma.

A detecção visual da plataforma produz um mapa de saliência em que mostra a plataforma saliente do fundo, portanto através de um algoritmo que siga esta zona saliente é possível centrar o MR-VTOL na plataforma. Este algoritmo segue uma dada região de imagem para imagem, adaptando-se às suas alterações geométricas na imagem. Para tal, aplica-se um estimador, o algoritmo *CamShift*² (Allen et al., 2004) em cada mapa de saliência. Este algoritmo estimador necessita de inicialização que se resume a indicar qual a zona da imagem que se pretende seguir. Para tal é utilizado o detetor dedicado, visto ser o detetor cuja detecção é mais pormenorizada. Assim, no início do levantamento aplica-se o detetor dedicado para inicializar o processo de seguimento. A partir deste momento o seguimento é realizado através da aplicação do estimador ao mapa de saliência.

O seguimento da plataforma é também realizado aquando da descida, para que o MR-VTOL desça centrado na plataforma. De outra forma, numa situação de vento, o MR-VTOL podia começar a desviar do seu alvo até que este saísse do seu campo de visão. Na Fig. 3.9 está um exemplo prático de seguimento no mapa de saliência utilizando o *CamShift*. Na figura, à esquerda tem-se o resultado do detetor dedicado em que as circunferências a castanho representam as zonas candidatas a plataforma e a vermelho a plataforma detetada. Note-se que o resultado desta aplicação do detetor dedicado não foi utilizada para o seguimento, apenas para comparação da localização real da plataforma e da localização estimada pelo *CamShift*. Na figura, à direita tem-se o mapa de saliência em que o nível de cinzentos indicam as zonas salientes, tendo a plataforma como objeto de interesse. O contorno quadrado a branco representa o resultado do estimador.

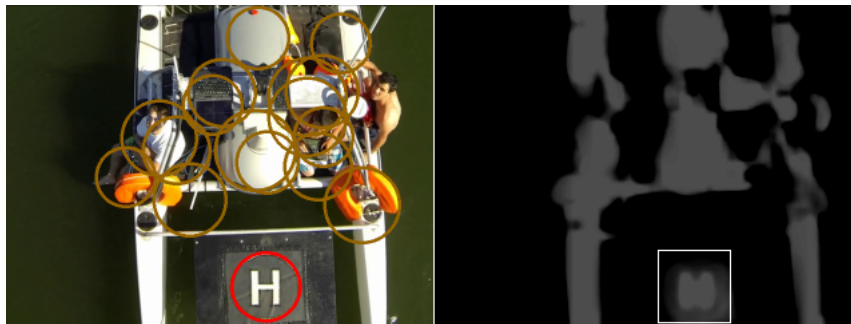


Figura 3.9: Exemplo de seguimento da plataforma através de *CamShift*. À esquerda a imagem de entrada com o resultado do detetor dedicado, em que as circunferências a castanho são as zonas candidatas a plataforma e a vermelho a plataforma detetada. Este resultado não é utilizado no seguimento. À direita está o respetivo mapa de saliência, em que o nível de cinzentos corresponde a zonas salientes na imagem. Marcada com um contorno retângulo branco está o resultado do *CamShift*.

3.3 A Arquitetura Baseada em Comportamentos

Uma arquitetura complexa pode dividir-se em módulos de menor complexidade até que cada módulo presente na arquitetura apresente um funcionamento simples. Utilizando este princípio, a arquitetura proposta é baseada em comportamentos em que o comportamento global é dividido em comportamentos menos complexos. Na Secção 3.3.1 é apresentada uma visão global do ciclo normal de funcionamento do sistema, entrando de seguida em detalhe sobre o levantamento na Secção 3.3.2 e sobre a aterragem na Secção 3.3.3.

²Do inglês: *Continuously Adaptive Mean Shift Algorithm*.

3.3.1 O Ciclo Geral

As ações do MR-VTOL resumem-se ao levantamento, execução de um plano de tarefas e aterragem. Tendo em conta a complexidade deste comportamento global, divide-se em 3 comportamentos, considerado a máquina de estados do sistema. Ao efetuar uma subida aprende a aparência da plataforma por forma a poder realizar uma pesquisa visual quando for necessário aterrar. Após a aprendizagem o veículo navega de acordo com um plano de tarefas, deixando em pausa todo o sistema de visão abordado nesta dissertação, disponibilizando a câmara para missões do plano. Terminadas as tarefas, o veículo regressa à plataforma efetuando as pesquisas necessárias até que a encontra e aterra.

3.3.2 O Levantamento da Plataforma

Para que todo o sistema funcione é necessário o MR-VTOL efetuar um levantamento a partir da plataforma para proceder à aprendizagem das formas e texturas da mesma. Para tal utiliza-se a saliência visual com o método de baixo para cima, i.e., com a utilização do conhecimento à priori da plataforma, por forma a que a plataforma se destaque mapa de saliência. Esta aprendizagem servirá para a pesquisa da plataforma a elevadas altitudes, aquando da descida. A altitude em que se inicia a aprendizagem, h_{min} , é calculada de modo a garantir que a plataforma está presente na imagem com l pixels de altura:

$$h_{min} = \frac{f \cdot (l_p + e_g + e_1)}{l} \quad (3.6)$$

onde l_p é a largura da plataforma, e_g é o erro associado ao GPS e e_1 é uma margem. Assim, a Eq. 3.6 obtida a partir da Eq. 3.1, contempla uma margem para eventuais erros de aquisição de dados ou incertezas associadas.

A aprendizagem inicia-se com a aplicação da saliência que utiliza o conhecimento à priori da plataforma. Da saliência resulta um conjunto de 6 pesos, $\mathbf{W} = (W_0, \dots, W_5)$, a partir dos quais é possível calcular o mapa da saliência. Assim, sabendo os pesos, em qualquer momento se calcula o mapa de saliência. Como a saliência visual da plataforma altera com a altitude, a aprendizagem final consiste num conjunto de pesos para cada altitude unitária, \mathbf{W}^h :

$$\mathbf{W}^h = (W_0^h, \dots, W_5^h) \quad (3.7)$$

O comportamento Subida é descrito pelos comportamentos β_H , β_{Sal} e β_{Subir} presentes na Fig. 3.10.

Os comportamentos β_H e β_{Sal} têm como objetivo centrar o MR-VTOL na plataforma. Para tal, o comportamento β_H , aplicado se o MR-VTOL está a uma altitude entre h_{min} e h_d , utiliza o Detetor Dedicado (ver Secção 3.2.1) na pesquisa visual. O comportamento β_{Sal} , aplicado se o MR-VTOL está a uma altitude entre h_d e h_{max} , utiliza a pesquisa visual no mapa de saliência (ver Secção 3.2.4).

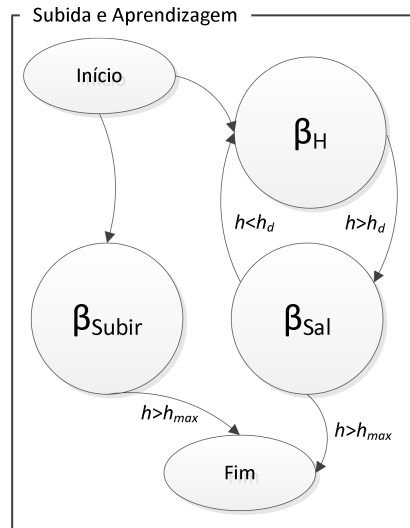


Figura 3.10: Diagrama de aprendizagem na subida do MR-VTOL. Os comportamentos β_{Subir} e β_{H} atuam em atuadores desacoplados.

Nesta fase são utilizados os pesos \mathbf{W} aprendidos no ciclo de aprendizagem anterior.

Existe um processo que resulta nas deslocções que o MR-VTOL necessita efetuar para ficar centrado na plataforma, enquanto realiza a aprendizagem da plataforma. Este processo começa por obter a região de interesse proveniente do Detetor Dedicado ou do estimador *CamShift*, dependendo da altitude atual do MR-VTOL, para criar uma máscara. Esta máscara resume-se a uma imagem a preto, com a área quadrada da região de interesse a branco. É utilizada no método da saliência visual em que a zona a branco corresponde à zona da imagem que se pretende destacar no mapa de saliência. Desta forma são obtidos os pesos que correspondem à aprendizagem da aparência da plataforma para a altitude atual do MR-VTOL. Após a aprendizagem é calculada a deslocção que o MR-VTOL necessita realizar para ficar centrado na plataforma. No caso da plataforma sair do campo visual do MR-VTOL, não é possível centrar através da visão. Assim, é utilizado um mecanismo baseado em GPS que resulta nas deslocções necessárias para que o MR-VTOL volte para a posição de onde iniciou o levantamento. Desta forma, o MR-VTOL desloca-se para a posição de levantamento mantendo a altitude atual, onde a plataforma volta a estar no campo de visão. Este método é necessário em condições de ventos fortes, para que o MR-VTOL volte à posição onde estava quando a plataforma saiu do campo visão, possibilitando a continuação da aprendizagem.

O comportamento β_{Subir} tem como objetivo aumentar a altitude do MR-VTOL quando os comportamentos β_{H} e β_{Sal} não estão ativos, o que indica que o MR-VTOL se encontra centrado com a plataforma. Após aumentar a altitude passa a estar ativo o comportamento β_{H} ou β_{Sal} , dependendo da altitude atual do MR-VTOL.

Quando o MR-VTOL atinge h_{max} , termina o processo de aprendizagem, com toda a informação necessária para que seja possível obter a saliência a qualquer altitude, i.e., entre a mínima e a máxima. A partir deste ponto, a aprendizagem na subida não é obrigatória, no entanto poderá ser feita, sobrepondo-se aos resultados anteriores. A Fig. 3.10 resume o processo de aprendizagem apresentado.

Pode acontecer que na subida, não foi efetuada aprendizagem para um intervalo de altitudes,

e.g., entre 10 m e 11 m, logo, na descida utilizam-se os parâmetros da altitude mais próxima à atual. O conjunto de parâmetros necessitam ser guardados e restaurados a partir da estrutura implementada no *software*. Deste modo, o resultado da aprendizagem é guardado num ficheiro YAML que contém todos os parâmetros aprendidos.

3.3.3 A Aterragem na Plataforma

Esta secção diferencia e explica os dois tipos de pesquisa da plataforma propostas nesta dissertação. A pesquisa global e a pesquisa local complementam-se pois a primeira baseia-se em registo de dados e comunicações para definir uma localização absoluta provável da plataforma e a segunda aplica rotinas de visão por computador para a deteção e seguimento da plataforma.

Pesquisa Global

O controlo de um robô aumenta de complexidade com o número de sensores em consideração. Cada sensor tem o seu propósito e efetua a sua leitura, pelo que se o controlo de um robô passar por um único controlador, a ação depende exclusivamente dos valores obtidos pelo sensor. Torna-se um caso mais complexo quando uma ação depende da leitura de diversos sensores, alguns até concorrentes. Esta situação acontece se dois ou mais sensores resultam numa ação para o mesmo atuador, suscetível a existirem conflitos. Com vista a resolver esta questão, Brooks (1986) propôs uma arquitetura comportamental em que cada ação resulta da cooperação dos vários comportamentos em função da sua informação sensorial. Um caso particular desta arquitetura é proposta por Rosenblatt (1997), designada por arquitetura comportamental multivalor baseada em votação. Neste caso cada comportamento deposita votos nas várias ações possíveis, em função da sua informação sensorial. No final é executada a ação com maior número de votos. Esta secção explica a arquitetura comportamental multivalor baseada em votação proposta nesta dissertação. De seguida são apresentados os mapas que permitem ao MR-VTOL obter a posição com maior utilidade de visitar, quando pretende aterrar.

De um modo geral, um mapa georreferenciado é uma fotografia de satélite representativa do solo e associada a uma coordenada global. No caso particular desta dissertação, assume-se que um mapa georreferenciado é uma grelha 2D representativa da área de trabalho do MR-VTOL cujos elementos representam a utilidade, de acordo com um dado critério, de visitar uma dada porção do terreno de área a_e . Para tal, é utilizada uma imagem de 8 bits em que o valor 0 corresponde a uma porção sem utilidade de visitar e o valor 255 a uma porção com a maior utilidade de visitar.

Resumindo, para que o MR-VTOL obtenha uma estimativa da sua posição relativamente à plataforma são necessários cinco mapas. No mapa M_v as posições num raio γ_r da posição atual do MR-VTOL são utilizadas para determinar que locais próximos ao MR-VTOL são mais prováveis de encontrar a plataforma. No mapa M_a as posições têm uma utilidade aleatória com o objetivo de resolver situações de bloqueio³ no mapa final. No mapa M_p as posições são utilizadas para determinar que locais são mais prováveis de encontrar a plataforma tendo em conta o seu trajeto; numa

³Conhecido como *deadlock* na literatura anglo-saxónica.

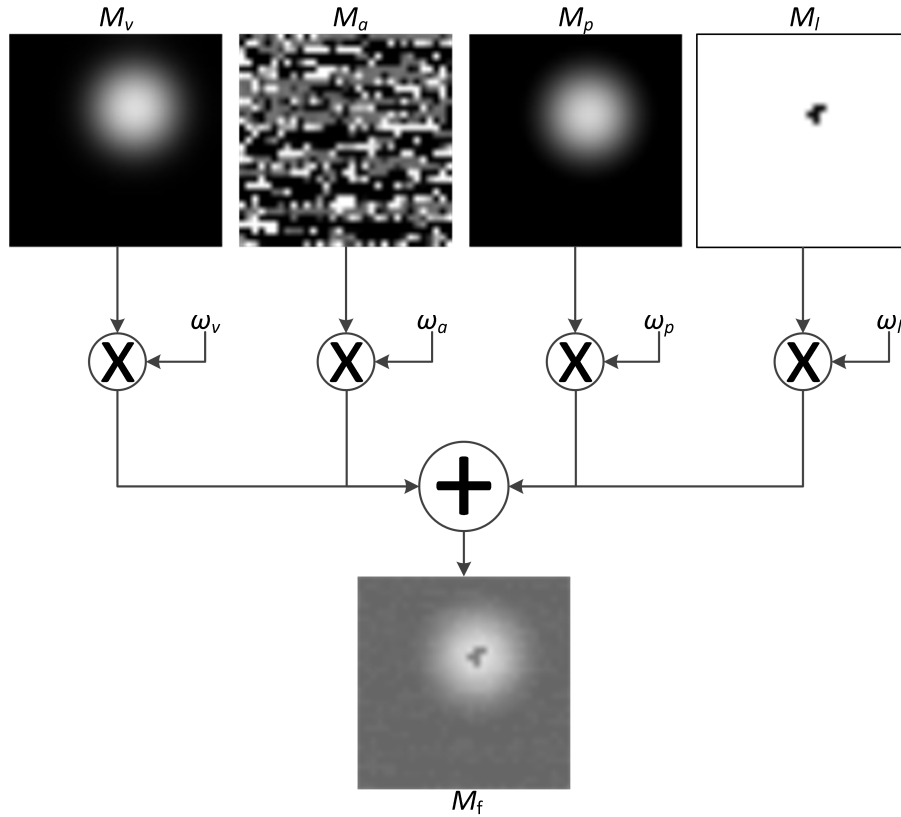


Figura 3.11: Exemplo de uma combinação linear pesada dos mapas de utilidade. Os símbolos \oplus e \otimes representam uma soma e uma multiplicação, respetivamente, ambos posição a posição. O mapa M_v representa a posição atual do MR-VTOL. O mapa M_a representa a componente aleatória. O mapa M_p representa a posição da plataforma. O mapa M_l representa os locais visitados. O mapa M_f é a combinação linear pesada dos restantes

situação em que o MR-VTOL e a plataforma não perderam contacto, este mapa seria suficiente pois o MR-VTOL teria sempre conhecimento da posição exata da plataforma. O mapa M_l contém a representação dos locais visitados pelo MR-VTOL em que não foi encontrada a plataforma. O mapa M_f é o mapa que resulta da combinação linear pesada dos anteriores:

$$M_f = \sum_i (w_i \cdot M_i), i \in \{v, a, p, l\} \quad (3.8)$$

onde, M_i são os mapas utilizados e w_i os respetivos pesos no cálculo do mapa M_f . A Fig. 3.11 representa um exemplo de combinação linear dos mapas, em que os símbolos \oplus e \otimes representam uma soma e uma multiplicação, respetivamente, ambos posição a posição.

Mapa	Ponderação	Descrição
M_v	w_v	Locais num raio próximo ao MR-VTOL
M_a	w_a	Ruído (componente aleatória)
M_p	w_p	Localização provável da plataforma
M_l	w_l	Locais visitados sem presença da plataforma

Tabela 3.2: Descrição dos mapas de utilidade.

O mapa M_f é utilizado para obter a localização mais provável da plataforma, em que a posição

$\mathbf{p} = (x_p, y_p)$ do mapa com maior utilidade corresponderá à próxima posição a visitar. Os mapas M_p e M_l têm esquecimento associado a uma constante temporal τ :

$$\tau \cdot \dot{\mathbf{p}} = -\mathbf{p} \quad (3.9)$$

em que \mathbf{p} é uma posição (x_p, y_p) de uma imagem e $\dot{\mathbf{p}}$ a sua derivada em ordem ao tempo. A Fig. 3.12 apresenta o nível de decaimento na intensidade de cada elemento do mapa em função do tempo, para diferentes valores de τ .

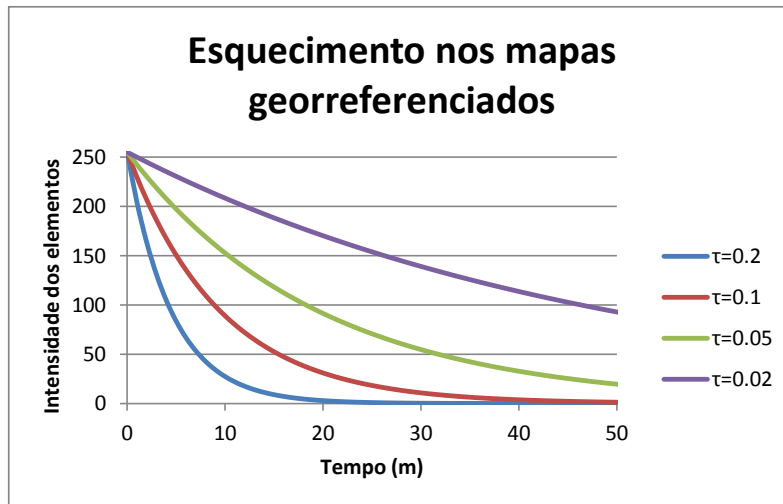


Figura 3.12: Gráfico representativo do esquecimento nos mapas georreferenciados, para diferentes valores de constante temporal, segundo a Eq. 3.9.

Desta forma, se o MR-VTOL perder a comunicação com a plataforma, este poderá visitar um local já visitado onde não encontrou a plataforma. Na Fig. 3.13, a comunicação perdeu-se na segunda iteração, que corresponde à imagem da primeira linha e segunda coluna. Como o MR-VTOL deixou de receber informação da posição da plataforma, a última informação que tem vai diminuindo de intensidade, que significa que a informação perde utilidade.

No caso do MR-VTOL não ter a localização atual da plataforma porque, por exemplo, perderam a comunicação entre si, a dinâmica dos mapas resulta numa pesquisa em espiral a partir da posição atual do MR-VTOL. Esta dinâmica deve-se à utilização do mapa M_v e do ajuste de sensibilidade da ponderação de cada mapa na combinação linear que gera o mapa M_f . De notar que o facto da pesquisa ser em espiral não é pré-definido, i.e., a ponderação dos mapas no cálculo do mapa final, M_f , faz com que a sequência das posições com maior utilidade descrevem uma espiral. Desta forma, se o MR-VTOL receber informação sobre a posição da plataforma, o respetivo mapa M_p é atualizado e conseqüentemente será a posição no mapa final com maior utilidade de ser visitado. Na Fig. 3.14 consta a dinâmica dos mapas, mostrando a evolução do mapa final M_f durante uma pesquisa da plataforma, na situação em que o MR-VTOL não tem conhecimento da localização da plataforma. Verifica-se visualmente que o círculo correspondente à última localização conhecida da plataforma vai diminuindo de intensidade, o que significa uma diminuição de utilidade em visitar a área. Verifica-se também que os locais visitados onde não foi encontrada a plataforma não têm utilidade de ser visitados novamente. Contudo, à medida que o tempo passa, se o MR-VTOL continuar na pesquisa pela plataforma, estas zonas voltam a ter utilidade no mapa final. Por último, a

componente da posição atual do MR-VTOL guia a pesquisa em torno dos locais já visitados, apesar da não estar perceptível à visão humana na imagem.

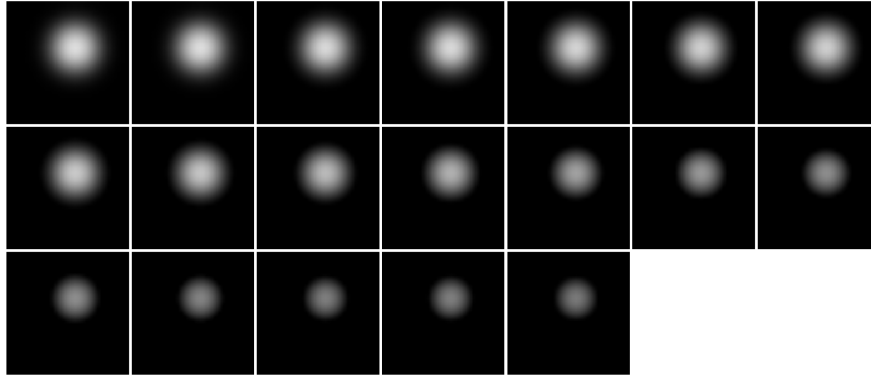


Figura 3.13: Esquecimento do mapa de utilidade M_p representativo da posição da plataforma. Apresenta uma evolução de cima para baixo e da esquerda para a direita.

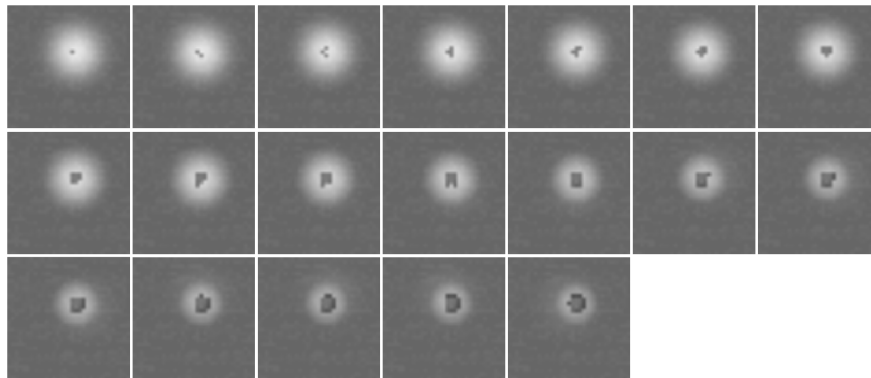


Figura 3.14: Exemplo de dinâmica dos mapas numa situação em que o MR-VTOL não tem informação da localização da plataforma. O mapa final apresenta uma evolução de cima para baixo e da esquerda para a direita.

A pesquisa em espiral da plataforma tem uma duração difícil de estimar, por parte do MR-VTOL, devido à imprevisão do número de candidatos a conter a plataforma em cada posição do mapa georreferenciado. No entanto, poderia ser implementado um plano de emergência permitindo ao MR-VTOL gerir os seus recursos energéticos. Para tal, o tempo de pesquisa, Δt , em função da área, a_w , pode ser estimado por:

$$\Delta t(a_w) = \Delta t_e \cdot \frac{a_w}{a_e} \quad (3.10)$$

onde Δt_e é o tempo médio da pesquisa por elemento e a_e é a área no solo, em metros, correspondente a um elemento do mapa. Utilizando este método, o MR-VTOL poderia fazer a gestão das baterias prevenindo o término de energia durante o voo e fazer uma aterragem segura de emergência. O valor da duração da pesquisa num elemento do mapa de utilidade, Δt_e , poderia ser atualizado automaticamente pelo MR-VTOL baseado na estatísticas de experiências anteriores. A Fig. 3.15

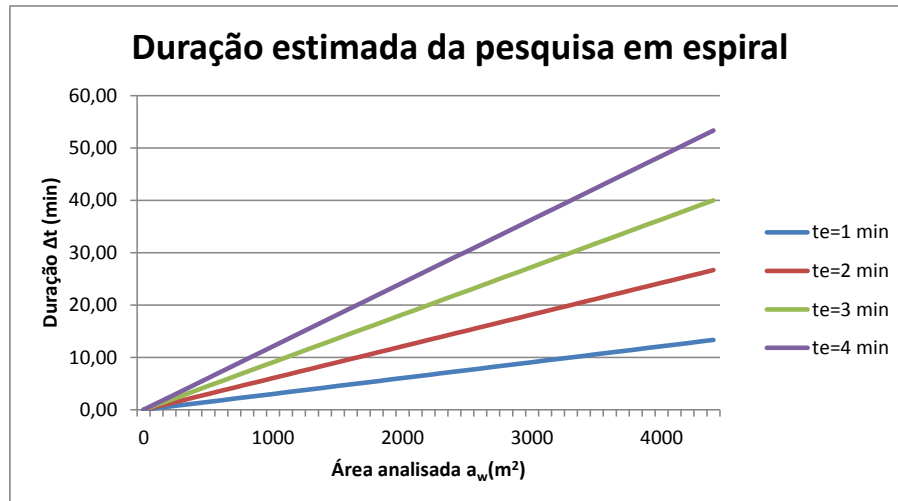


Figura 3.15: Duração estimada da pesquisa da plataforma, segundo a Eq. 3.10. Este método tem aplicação no caso em que o MR-VTOL pretende aterrar e perdeu a comunicação com a plataforma.

Pesquisa Local

A pesquisa local já foi abordada neste documento mencionando algumas das suas características. No entanto esta secção apresenta formalmente e explica o processo de pesquisa local proposto para deteção e seguimento visual da plataforma.

Aquando da descida, para detetar a plataforma são utilizados dois processos dependendo da altitude atual do MR-VTOL: (1) cálculo da saliência e (2) aplicação do detetor dedicado. Estes processos são complementares pois a saliência visual atua em altitudes em que a plataforma não está definida na imagem aérea e o detetor dedicado em altitudes em que a plataforma se destaca na imagem, tornando possível a extração de características da mesma. Esta complementaridade resulta num detetor aplicável a uma janela de altitudes mais ampla relativamente às abordagens anteriores.

Aquando da descida, se o MR-VTOL estiver a uma altitude compreendida entre h_{max} e h_d , a pesquisa visual é realizada com recurso à saliência visual tendo em conta o modelo da plataforma aprendido durante subida (ver Secção 3.3.2). O resultado desta pesquisa alimenta o mecanismo de seguimento da plataforma (ver Secção 3.2.5) para que o MR-VTOL se centre com a plataforma detetada. Após estar centrado, o MR-VTOL diminui a sua altitude. Este comportamento denomina-se β_{Sal} .

O comportamento β_H é aplicado para baixas altitudes, i.e., entre h_d e h_{min} , em que a pesquisa visual é realizada com recurso ao detetor dedicado (ver Secção 3.2.1). Durante este comportamento é recolhida informação para decidir se o objeto seguido pelo MR-VTOL corresponde à plataforma de onde levantou. Esta informação consiste no conjunto de dissemelhanças entre os momentos geométricos invariantes de referência e o da marca detetada. Ao aplicar o detetor dedicado, se encontrar a plataforma, o MR-VTOL centra-se na marca, diminui de altitude e o valor da respetiva dissemelhança, d_c (ver Eq. 3.3), é adicionado ao conjunto de dissemelhanças D :

$$D \leftarrow D \cup \{d_c\}. \quad (3.11)$$

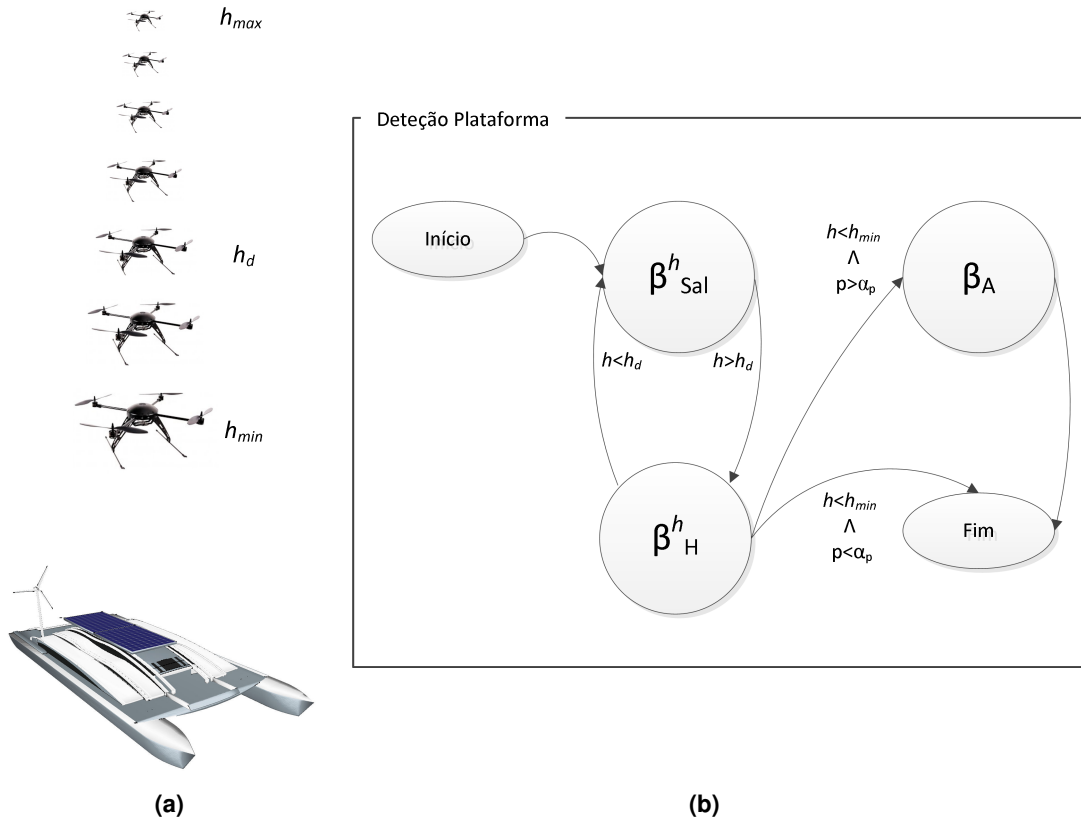


Figura 3.16: Aplicação do método de detecção da plataforma em função da altitude atual do MR-VTOL, por cada imagem. (a) Limiares de altitude. (b) Comportamento composto Detecção Plataforma, β_{Plat} , composto pelos comportamentos primitivos β_{Sal} , β_{H} e β_{A} .

À altitude h_{min} termina o comportamento β_{H} , sendo realizada a tomada de decisão em relação à plataforma conhecida. Esta decisão tem como base a análise de todas as imagens processadas pelo comportamento β_{H} , nomeadamente na média do conjunto de dissimilaridades D , \bar{d} :

$$\bar{d} = \frac{1}{|D|} \sum_{d \in D} d \quad (3.12)$$

Finalmente, a possibilidade do objeto detetado ser a plataforma, p , é:

$$p = \alpha \bar{d} + (1 - \alpha) \frac{|D|}{n_{\text{min}} - n_d}, \quad (3.13)$$

onde n_{min} o índice da imagem adquirida à altitude h_{min} , n_d é o índice da imagem adquirida à altitude h_d e α é o fator de ponderação de dois critérios.

Na Eq. 3.13, a primeira parcela corresponde ao critério de qualidade das deteções enquanto que a segunda parcela corresponde ao critério de quantidade de deteções realizadas. Desta forma, uma única deteção com uma dissimilaridade muito reduzida, não indica uma deteção da plataforma. É então necessário algumas deteções com um valor médio de dissimilaridade reduzido. A análise de várias aplicações do detetor ajuda a minimizar o contributo de falsos positivos para a tomada de

decisão. Se o valor de p for superior a α_p , então a plataforma foi detetada e o MR-VTOL efetua o comportamento de aterragem, β_A .

Capítulo 4

Resultados Experimentais

4.1 Instalação Experimental

Nesta secção são apresentados *hardwares* e *softwares* utilizados. Todo o modelo proposto foi completamente implementado na linguagem de programação C++ e compatível com *Robotics Operation System* (ROS)¹ (Quigley et al., 2009). O ROS é uma plataforma de desenvolvimento de *software* para robôs, com licenciamento gratuito e utiliza bibliotecas de fonte aberta. É um sistema distribuído e modular permitindo a sua customização em função do que se pretende, existindo mais de 3000 pacotes de software complementar. Esta plataforma permite implementar um sistema distribuído, i.e., funcionando em computadores diferentes, conectados a uma rede de trabalho comum. Este sistema distribuído comunica através de nós ROS onde podem ser publicadas e subscritas mensagens em determinado tópico. Desta forma, um nó ROS ao subscrever um determinado tópico ROS, irá receber informação publicada por outro nó ROS. De notar que os nós ROS terão que estar registados no mesmo universo ROS. Assim, esta plataforma permite a comunicação entre o MR-VTOL e o veículo que contém a plataforma. Outra vantagem na utilização do ROS nesta dissertação refere-se ao facto de permitir a implementação de um simulador com os requisitos necessários. Desta forma, foi possível a realização de testes em simulação antes de utilizar os robôs reais. Permitiu ainda um estudo à sensibilidade da parametrização utilizada, sem recorrer aos robôs, evitando possíveis acidentes com os mesmos.

4.1.1 Simulação

Para a simulação foi utilizado um Pentium(R) Core2Duo CPU T5600 1,83 GHz com 3 GB de memória RAM, processamento gráfico Nvidia 512 MB de memória, correndo Linux de 32bits, com a distribuição Ubuntu 12.04 (*Precise Pangolin*) e uma biblioteca de visão computacional OpenCV 2.4.5 (Bradski and Kaehler, 2008). A imagem recebida como visão aérea tem resolução de 640×480 .

A simulação do modelo proposto nesta dissertação foi realizada no Kelpie (Mendonca et al.,

¹ROS: www.ros.org

2013), também desenvolvido no âmbito da experiência RIVERWATCH do projecto ECHORD. Este simulador consiste numa visão aérea controlável sobre uma textura de água com 3 elementos a flutuar. Dois deles são barcos, um com uma plataforma de aterragem e o terceiro elemento é um barril com a dimensão e cor aproximados do barco. Na Fig. 4.1 estão presentes o barco com a plataforma de aterragem e o barril.



Figura 4.1: Imagens obtidas do simulador Kelpie em vistas aéreas com diferentes perspetivas

4.1.2 Real

Para os testes reais foi utilizado um MR-VTOL com 4 motores MK3638 sem escovas, com controlador BL Ctrl V2.0, e hélices em fibra de carbono 12×4.5 , distribuição de energia Okto XL V8, bateria de 4 células LiPo 11000 mAh 4S 40C (MaxAmps) 14.8 V. Contém também um módulo externo de GPS, U-blox LEA-6H, um modem de telemetria, Xbee Pro v2 e um recetor RC 2.4 GHz, Hott 16 (Graupner) para permitir comunicação com telecomando MC-32 Hott (Graupner), em casos de emergência. O controlador deste sistema é VRBrain² em ARM Cortex-M4 a 192 MHz. A computação de vídeo em ROS é realizado por um ODROID-XU³ - Exynos5 Octa Cortex-A15 4×1.6 GHz e Cortex-A7 4×1.2 GHz, com 2 GB LPDDR3 de memória RAM. A câmara USB que equipa o veículo é uma Logitech C525. As imagens têm uma resolução de 1280×720 e uma frequência de 10 imagens por segundo. A comunicação de testes para uma estação terrestre é realizada através de um amplificador WI-FI de 1 W, Pico Station 2 (Ubiquit Networks, Inc.). A Fig. 4.2 mostra uma fotografia do robô descrito.

Como já referido, os componentes do sistema utilizam o ROS para comunicar, o que sugere a existência de uma ligação do controlador geral do MR-VTOL, VRBrain ao ROS. Assim, é criada uma ligação física entre o VRBrain e o computador ODROID através de um módulo RS-232-USB, utilizando o protocolo MAVLink. Através deste nó ROS, é possível publicar e/ou subscrever tópicos para enviar e receber informação. Este nó publica tópicos com informações como: altitude, posição GPS, imagem capturada pela câmara; e recebe tópicos com informações como deslocamentos a realizar, alterações de modos de voo. Um esquema destas comunicações é apresentado na Fig. 4.3.

²VRBrain: <http://vrbrain.wordpress.com>

³ODROID: <http://www.hardkernel.com>



Figura 4.2: Fotografia do MR-VTOL utilizado nos testes

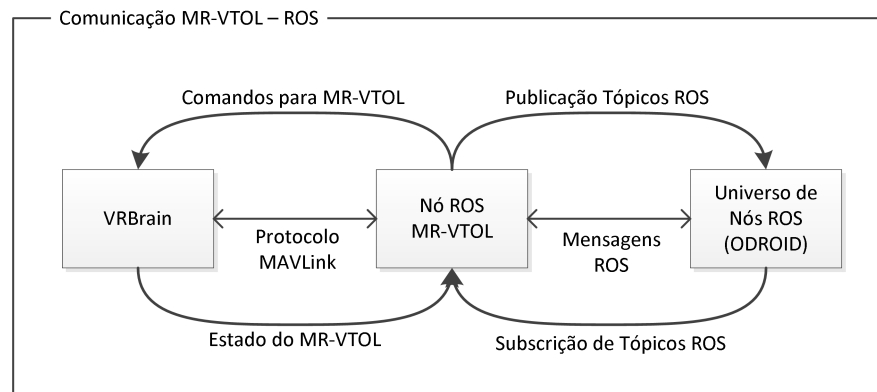


Figura 4.3: Diagrama representativo das comunicações entre o controlador do MR-VTOL, VRBrain e o ROS do ODROID.

4.2 Parametrização do Sistema

Nesta secção são apresentados os valores atribuídos às variáveis do sistema referidas ao longo da Secção 3. Entre estes valores alguns são medidos, outros estimados e outros são valores empíricos de referência.

4.2.1 Simulação

O tamanho da marca foi deduzido pelo modelo do barco do simulador, resultando num raio de $r = 0.75$ (ver Secção 3.2.1). Já o comprimento do barco que transporta a plataforma é $c_p = 11.91$ (ver Secção 3.2.4). Através deste valor, foi efetuado o cálculo inverso da Eq 3.1 de modo a obter um comprimento focal, f , de 790.56 (ver Secção 3.1.1). Estes valores não se aproximam a valores reais, pois trata-se de um simulador virtual. A altura da imagem $l = 480$ (ver Secção 3.3.2) pixels provém do simulador utilizado. Este valor, apesar de reduzido tendo em conta as resoluções atuais comuns das câmaras, é aceitável para o processamento de vídeo em tempo real pois quanto maior a resolução, maior o tempo de resposta.

As altitudes de referência utilizadas na Pesquisa Local (ver Secção 3.3.3) foram obtidas através de experiências, à exceção da altitude mínima de $h_{min} = 4$ (ver Secção 3.3.2). Este valor resulta da Eq 3.6, em que a largura da plataforma corresponde a duas vezes o valor do raio r , $l_p = 1.5$. Ao erro associado ao GPS, e_g , foi atribuído o valor unitário devido ao facto de não ser diretamente aplicado no simulador pois as posições são absolutas. Também a margem de segurança e_1 tem o valor unitário, pois considera-se suficiente tendo em conta o seu peso na equação. Por diversos testes consecutivos verificou-se que o detetor dedicado a partir de $h_d = 15$ resulta num aumento de falsos positivos. Por fim, a altitude máxima de $h_{max} = 51$ foi assim definida por aproximação a um caso real pois o MR-VTOL não necessitará de navegar em altitudes superiores (ver Secção 3.2).

Para o mapa georreferenciado foi definida uma área quadrada com 1000 m de lado. Esta área de 100 Ha é dividida em porções de área quadrada $a_e = 900 \text{ m}^2$ correspondendo assim a um elemento no mapa de utilidade (ver Secção 3.3.3). O mapa georreferenciado final obtém-se através de uma média ponderada dos 4 mapas existentes. Esta ponderação foi obtida através de diversas experiências no simulador de modo a resultar no comportamento esperado da pesquisa em forma de espiral. Assim, chegaram-se aos valores de $w_v = 10\%$ para o mapa que regista os locais num raio próximo ao MR-VTOL, $w_a = 3\%$ para a componente aleatória de ruído, $w_p = 47\%$ para o mapa que contém a localização provável da plataforma e de $w_l = 40\%$ para o mapa que regista os locais já visitados onde a plataforma não foi detetada. Os últimos dois mapas referidos apresentam esquecimento associado a uma constante temporal $\tau_e = 5 \text{ s}$. Este valor foi obtido tendo em conta a velocidade média da plataforma de 1 m/s e a área de cada elemento do mapa georreferenciado. O mapa M_v tem um raio de pesquisa $\gamma_r = 5$ que corresponde a aproximadamente 250 m no solo. De modo a prever a duração da pesquisa em determinada área de terreno, pela Eq 3.10, é assumido uma duração da pesquisa num elemento do mapa de utilidade $t_e = 60 \text{ s}$. Este valor foi obtido através das várias experiências realizadas e contempla a descida do MR-VTOL até um ponto provável de conter a plataforma. A duração obtida pela Eq 3.10, poderá ser superior à duração real pois para cada elemento do mapa georreferenciado, o MR-VTOL não detetará possíveis zonas de conter a plataforma.

Quando o MR-VTOL está a uma altitude h_{max} e pretende aterrar, efetua uma pesquisa de segmentos isolados de saliência no mapa de saliência. O resultado desta pesquisa é filtrado pelo tamanho com uma variação de $\sigma = 25\%$, relativamente ao tamanho de referência (ver Secção 3.2.4). Este valor foi determinado segundo a experiência. Os segmentos que passam no filtro são depois prioritizados em função da intensidade da saliência e do rácio do tamanho relativamente à referência, com uma ponderação $\alpha_s = 0.7$.

Na deteção da plataforma, o detetor dedicado pesquisa círculos com uma variação de raio de $\sigma_r = 25\%$ relativamente ao esperado (ver Secção 3.2.1). Este detetor classifica depois as diversas zonas potenciais de conter a marca através da sua dissemelhança à marca de referência. Para tal, só são consideradas dissemelhanças inferiores a $\xi_m = 7$, valor este obtido através de experiências realizadas. A classificação da deteção da plataforma utiliza um fator de ponderação de $\alpha = 0.7$ para os dois critérios, qualidade e quantidade de deteções (ver Secção 3.3.3). No final, o limiar para esta classificação é de $\alpha_p = 70\%$, pois pelas experiências realizadas, os valores abaixo deste, apresentam alguns falsos positivos.

4.2.2 Real

O tamanho da marca utilizada na plataforma, tem um raio de $r = 0.305$ m (ver Secção 3.2.1). Já o comprimento do barco que transporta a plataforma é $c_p = 2.5$ m (ver Secção 3.2.4). Através da calibração da câmara obteve-se um comprimento focal, f , de 396.70 (ver Secção 3.1.1). Apesar da câmara capturar imagens com uma resolução de 1280×720 , é realizada uma conversão para 640×480 , ficando a altura da imagem $l = 480$ pixels (ver Secção 3.3.2). Este valor, apesar de reduzido tendo em conta as resoluções atuais comuns das câmaras, é aceitável para o processamento de vídeo em tempo real pois quanto maior a resolução, maior o tempo de resposta.

As altitudes de referência utilizadas na Pesquisa Local (ver Secção 3.3.3) foram obtidas através de experiências, à exceção da altitude mínima de $h_{min} = 4$ m (ver Secção 3.3.2). Este valor resulta da Eq 3.6, em que ao erro associado ao GPS, e_g , foi atribuído o valor unitário, próximo ao erro associado ao dispositivo GPS. Também a margem de segurança e_1 tem o valor unitário, pois considera-se suficiente tendo em conta o seu peso na equação. Por diversos testes consecutivos verificou-se que o detetor dedicado a partir de $h_d = 15$ m resultada num aumento de falsos positivos. Por fim, a altitude máxima de $h_{max} = 30$ m foi assim definida pois o MR-VTOL não necessitará de navegar em altitudes superiores (ver Secção 3.2).

Para o mapa georreferenciado foi definida uma área quadrada com 1000 m de lado. Esta área de 100 Ha é dividida em porções de área quadrada $a_e = 330$ m² correspondendo assim a um elemento no mapa de utilidade (ver Secção 3.3.3). O mapa georreferenciado final obtém-se através de uma média ponderada dos 4 mapas existentes. Esta ponderação foi obtida através de diversas experiências no simulador de modo a resultar no comportamento esperado da pesquisa em forma de espiral. Assim, chegaram-se aos valores de $w_v = 10\%$ para o mapa que regista os locais num raio próximo ao MR-VTOL, $w_a = 3\%$ para a componente aleatória de ruído, $w_p = 47\%$ para o mapa que contém a localização provável da plataforma e de $w_l = 40\%$ para o mapa que regista os locais já visitados onde a plataforma não foi detetada. Os últimos dois mapas referidos apresentam esquecimento associado a uma constante temporal $\tau_e = 5$ s. Este valor foi obtido tendo em conta a velocidade média da plataforma de 1 m/s e a área de cada elemento do mapa georreferenciado. O mapa M_v tem um raio de pesquisa $\gamma_r = 5$ que corresponde a aproximadamente 250 m no solo. De modo a prever a duração da pesquisa em determinada área de terreno, pela Eq 3.10, é assumido uma duração da pesquisa num elemento do mapa de utilidade $t_e = 60$ s. Este valor foi obtido através das várias experiências realizadas e contempla a descida do MR-VTOL até um ponto provável de conter a plataforma. A duração obtida pela Eq 3.10, poderá ser superior à duração real pois para cada elemento do mapa georreferenciado, o MR-VTOL não detetará possíveis zonas de conter a plataforma.

Quando o MR-VTOL está a uma altitude h_{max} e pretende aterrar, efetua uma pesquisa de segmentos isolados de saliência no mapa de saliência. O resultado desta pesquisa é filtrado pelo tamanho com uma variação de $\sigma = 25\%$, relativamente ao tamanho de referência (ver Secção 3.2.4). Este valor foi determinado segundo a experiência. Os segmentos que passam no filtro são depois prioritizados em função da intensidade da saliência e do rácio do tamanho relativamente à referência, com uma ponderação $\alpha_s = 0.7$.

Na deteção da plataforma, o detetor dedicado pesquisa círculos com uma variação de raio de

$\sigma_r = 25\%$ relativamente ao esperado (ver Secção 3.2.1). Este detetor classifica depois as diversas zonas potenciais de conter a marca através da sua dissemelhança à marca de referência. Para tal, só são consideradas dissemelhanças inferiores a $\xi_m = 7$, valor este obtido através de experiências realizadas. A classificação da deteção da plataforma utiliza um fator de ponderação de $\alpha = 0.7$ para os dois critérios, qualidade e quantidade de deteções (ver Secção 3.3.3). No final, o limiar para esta classificação é de $\alpha_p = 70\%$, pois pelas experiências realizadas, os valores abaixo deste, apresentam alguns falsos positivos.

4.3 Experiências

Nesta secção são apresentadas experiências realizadas no simulador e nos robôs reais, utilizando a parametrização descrita na Secção 4.2.

4.3.1 Simulação

Esta secção apresenta um plano de execução que testa todas as funcionalidades do sistema no simulador apresentado, Kelpie. Para tal foi simulado o pior cenário do sistema, ou seja, após o levantamento bem sucedido, quando o MR-VTOL pretende aterrar, não tem a plataforma no seu campo de visão e a última localização conhecida da plataforma foi a de levantamento. Este cenário corresponde à situação apresentada a azul na Fig. 3.3. Desta forma o sistema recorre aos mapas georreferenciados até encontrar a plataforma. Utilizando os valores (x, y) do simulador, o MR-VTOL está à altitude h_{max} na posição $(25, 118)$ e a última informação da plataforma é $(70, 120)$. Entretanto como a plataforma se moveu, está agora na posição $(5, 98)$, correspondendo a uma distância em linha reta de 68 m. Resumindo, o MR-VTOL está a sensivelmente 28 m da plataforma mas como não tem essa informação recorre aos mapas georreferenciados até encontrar a plataforma. De modo a introduzir possíveis falsos positivos, na última posição conhecida da plataforma, encontra-se um barco igual ao de levantamento mas não contém a plataforma. A Fig. 4.4 mostra as posições relativas mencionadas, sendo que o mapa de cima corresponde a um excerto do mapa georreferenciado e o de baixo ao sistema de coordenadas do simulador.

O início do levantamento do MR-VTOL, mostrado na Fig. 4.5, começa pela aplicação do detetor dedicado, em que o seu resultado é utilizado para inicializar a saliência visual e o estimador que faz o seguimento da plataforma. A imagem correspondente ao resultado do detetor dedicado, (a) e (d), em que as circunferências a castanho são as zonas candidatas à plataforma e a vermelho a plataforma detetada. A partir deste resultado é criada a máscara, (b) e (e), para dar início à aprendizagem da saliência visual que por sua vez resulta no mapa (c) e (f). O contorno retângulo a branco corresponde ao estimador inicializado a partir do resultado do detetor dedicado.

Após o levantamento é guardada a aprendizagem da plataforma num ficheiro YAML com a estrutura mostrada na Fig. 4.6. Este ficheiro começa por apresentar o número de altitudes em que foi guardada informação e as dimensões do mapa de saliência utilizado. De seguida, passa às estruturas por altitude em que é guardada a altitude, retângulo resultante do estimador *CamShift* e os 6 pesos para o cálculo do mapa de saliência.

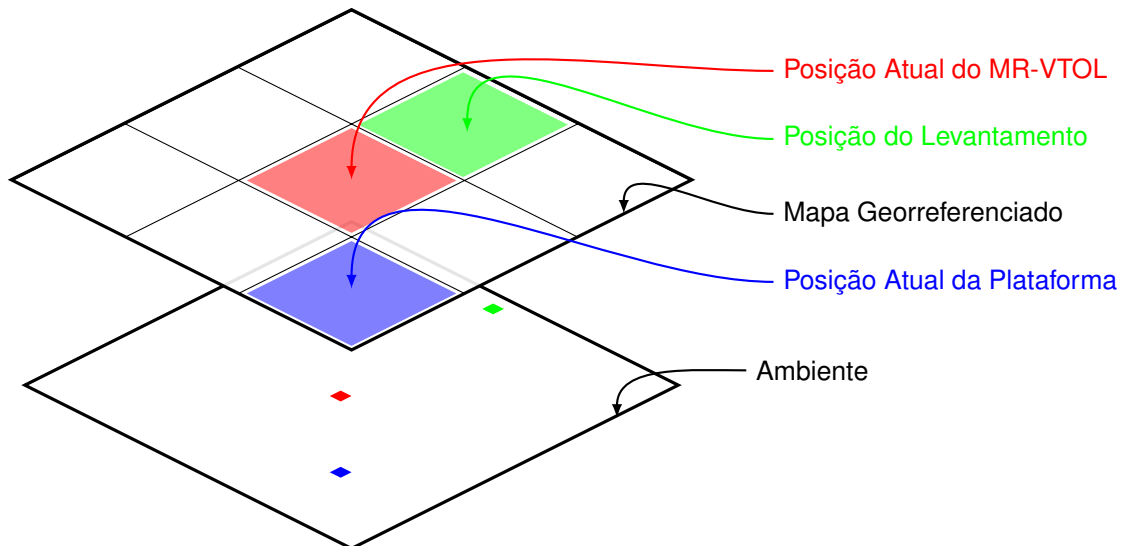


Figura 4.4: Posições relativas do plano testado. A figura de baixo corresponde ao ambiente, referencial \mathcal{W} , e a figura de cima corresponde ao mapa georreferenciado. A verde está a porção de solo onde se realizou o levantamento do MR-VTOL da plataforma; a vermelho a porção de solo correspondente à posição atual do MR-VTOL; a azul a porção de solo correspondente à posição atual da plataforma.

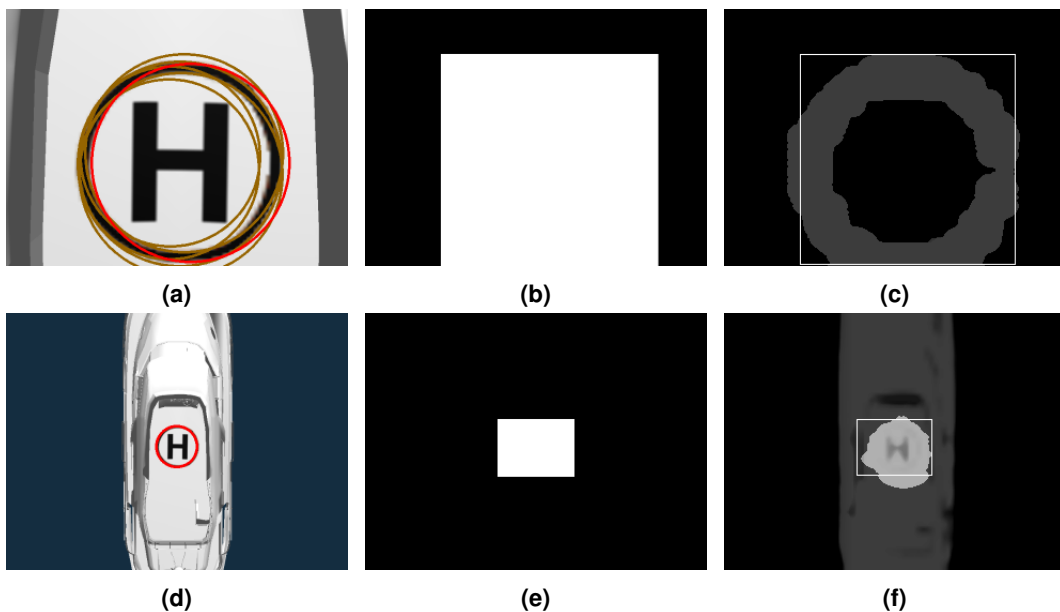


Figura 4.5: Início da aprendizagem durante o levantamento a uma altitude de 4 m. Em (a) está o resultado do detetor dedicado, em que as circunferências a castanho são as zonas candidatas à plataforma e a vermelho a plataforma detetada. Em (b) está a máscara criada a partir do resultado do detetor dedicado. Em (c) está o mapa de saliência, em que o contorno retângulo a branco corresponde ao estimador inicializado a partir do resultado do detetor dedicado. As figuras (d), (e) e (f) correspondem aos passos descritos, para uma altitude de 15 m.

Neste ponto, o MR-VTOL está disponível para executar um plano de atividades e terminado o plano, aterriza na plataforma. De seguida é descrito o processo de aterragem tendo para isso sido criado o pior cenário, como já referido. Chegado o momento do MR-VTOL aterrar, começa-se por pesquisar a plataforma no campo visual atual. Como a aplicação dos dois tipos de saliência visual na imagem aérea do MR-VTOL não resulta em objetos de interesse, o sistema recorre aos mapas georreferenciados para obter a próxima posição mais provável de conter a plataforma. A Fig. 4.7

apresenta o estado dos 4 mapas georreferenciados e o mapa final, após duas posições do mapa não apresentarem objetos de interesse no campo de visão. Em primeiro lugar está o mapa M_v que representa a posição atual do MR-VTOL, seguido do mapa M_a que representa a componente aleatória, em terceiro o mapa M_p que representa a posição da plataforma, em quarto o mapa M_l que representa os locais visitados e por fim M_f que é a combinação linear pesada dos restantes.

```
%YAML:1.0
structsize:
  size: 26
maps_size:
  width: 320
  height: 240
struct0:
  pos: 3
  rec: [ 71, 42, 202, 197 ]
  mat: !!opencv-matrix
    rows: 1
    cols: 8
    dt: f
    data: [ 9.87260699e-01, 8.26159954e-01, 2.44317665e+01, 0., 0.,
      2.82010136e+01, 9.07436550e-01, 2.60920010e+01 ]
```

Figura 4.6: Estrutura do ficheiro YAML. Em *size* tem-se o número de estruturas guardadas; *maps_size* são as dimensões do mapa de saliência; *struct0* é o primeiro elemento aprendido onde consta a altitude do MR-VTOL em *pos*, o retângulo do estimador em *rec* e os 6 pesos para o cálculo do mapa de saliência em *mat*.

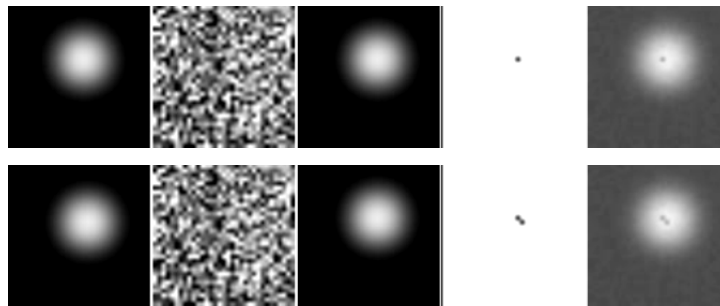


Figura 4.7: Mapas georreferenciados no início do plano de execução em simulação. Da esquerda para a direita tem-se: o mapa M_v representa a posição atual do MR-VTOL; o mapa M_a representa a componente aleatória; o mapa M_p representa a posição da plataforma; o mapa M_l representa os locais visitados; o mapa M_f é a combinação linear pesada dos restantes.

Na terceira utilização dos mapas georreferenciados, surge um objeto de interesse no campo de visão. Este objeto resultou da aplicação da saliência visual, utilizando os pesos da aprendizagem e é moestrado na Fig. 4.8. Neste caso surge um objeto de interesse que corresponde ao barco que não contem a plataforma. De notar que a localização deste barco é a conhecida pelo MR-VTOL como a última localização da plataforma.

O MR-VTOL efetua uma descida ao objeto de interesse a fim de analisar e decidir se está perante a plataforma para que se possa efetuar a aterragem. A Fig 4.9 apresenta o mapa de saliência para diferentes altitudes, em que o retângulo branco corresponde à zona na qual o MR-VTOL se está a centrar enquanto desce.

Durante esta descida, a aplicação do detetor dedicado não encontrou a plataforma em nenhuma altitude. Assim, o MR-VTOL volta para a altitude h_{max} para aplicar a saliência sem conhecimento à priori do objeto de interesse, cujo resultado está na Fig. 4.10. O resultado da aplicação desta saliência visual resultou também num objeto de interesse, o objeto já analisado. É realizada a

descida ao objeto de interesse e os mapas de saliência para diversas altitudes são mostrados na Fig. 4.11.



Figura 4.8: Primeiro objeto de interesse do plano de execução em simulação. (a) imagem capturada pela câmera do MR-VTOL; (b) objeto de interesse resultante da aplicação da saliência visual, utilizando os pesos aprendidos durante o levantamento.



Figura 4.9: Mapas de saliência e zona de seguimento na primeira descida. Os retângulos a branco correspondem à zona que o MR-VTOL utiliza para se manter centrado no objeto de interesse durante a descida.



Figura 4.10: Segundo objeto de interesse do plano de execução em simulação. (a) imagem capturada pela câmera do MR-VTOL; (b) objeto de interesse resultante da aplicação da saliência visual que não utiliza os pesos aprendidos durante o levantamento.

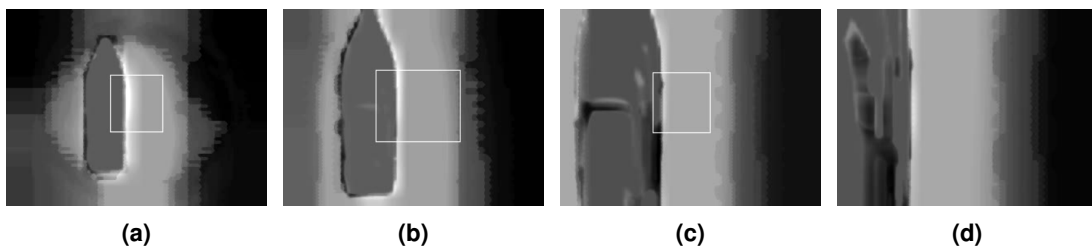


Figura 4.11: Mapas de saliência e zona de seguimento na segunda descida. Os retângulos a branco correspondem à zona que o MR-VTOL utiliza para se manter centrado no objeto de interesse durante a descida.

O detetor dedicado, mais uma vez, não detetou a plataforma durante a descida e por isso o

sistema recorreu novamente aos mapas georreferenciados para obter a próxima localização a visitar. O estado dos mapas após esta operação é mostrado na Fig. 4.12. Nesta localização foi encontrado um objeto de interesse, mostrado na Fig. 4.13. O MR-VTOL centra-se no objeto e realiza uma descida. Alguns mapas de saliência desta descida são mostrados na Fig. 4.14. Novamente o detetor dedicado não encontrou nenhuma vez a plataforma. De notar que no início da descida, o barco que contém a plataforma está presente no campo de visão do MR-VTOL, no entanto, à altitude h_d onde começa a ser aplicado o detetor dedicado, o barco já não está no campo de visão.

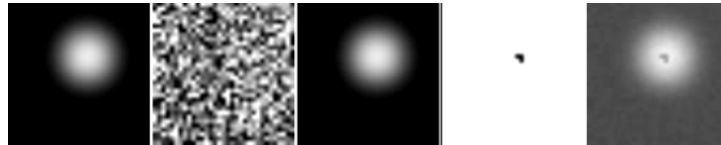


Figura 4.12: Mapas georreferenciados durante o plano de execução em simulação. Da esquerda para a direita tem-se: o mapa M_v representa a posição atual do MR-VTOL; o mapa M_a representa a componente aleatória; o mapa M_p representa a posição da plataforma; o mapa M_l representa os locais visitados; mapa M_f é a combinação linear pesada dos restantes.



Figura 4.13: Terceiro objeto de interesse do plano de execução em simulação. (a) imagem capturada pela câmara do MR-VTOL; (b) objeto de interesse resultante da aplicação da saliência visual, utilizando os pesos aprendidos durante o levantamento.

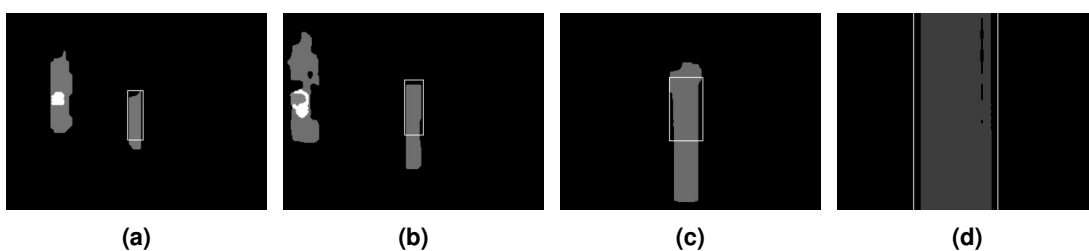


Figura 4.14: Mapas de saliência e zona de seguimento na terceira descida. Os retângulos a branco correspondem à zona que o MR-VTOL utiliza para se manter centrado no objeto de interesse durante a descida.

Como o objeto de interesse não contém a plataforma e não existiam mais objetos no campo de visão quando foi encontrado o objeto acabado de analisar, o MR-VTOL subiu e aplicou a saliência que não utiliza os pesos aprendidos. O resultado desta saliência visual foi nulo, i.e., não foram identificados objetos de interesse. Assim, recorreu-se novamente aos mapas georreferenciados. A Fig. 4.15 apresenta o estado dos mapas após cada utilização dos mesmos. Nas utilizações dos mapas presentes na imagem, não foram identificados objetos de interesse no campo de visão do MR-VTOL, independentemente do tipo de saliência visual aplicada.

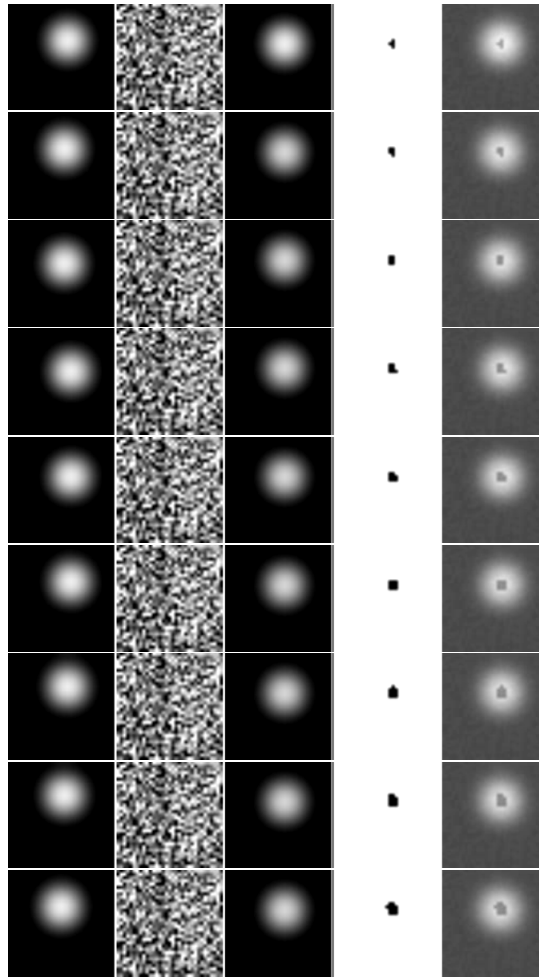


Figura 4.15: Mapas georreferenciados no final do plano de execução em simulação. Da esquerda para a direita tem-se: o mapa M_v representa a posição atual do MR-VTOL; o mapa M_a representa a componente aleatória; o mapa M_p representa a posição da plataforma; o mapa M_l representa os locais visitados; mapa M_f é a combinação linear pesada dos restantes.

Após visitar 9 localizações sem encontrar objetos de interesse no campo de visão, o MR-VTOL visitou uma localização com dois objetos de interesse, resultantes da aplicação da saliência visual aprendida durante o levantamento. Os objetos mostrados na Fig. 4.16 foram ordenados por prioridade sendo que o primeiro a ser visitado é o objeto cuja média da intensidade no mapa é superior. Os mapas de saliência respondentes à descida do primeiro objeto são mostrados na Fig. 4.17.



Figura 4.16: Quarto objeto de interesse do plano de execução em simulação. (a) corresponde à imagem capturada pela câmara do MR-VTOL; (b) mostra o objeto de interesse resultante da aplicação da saliência visual, utilizando os pesos aprendidos durante o levantamento.

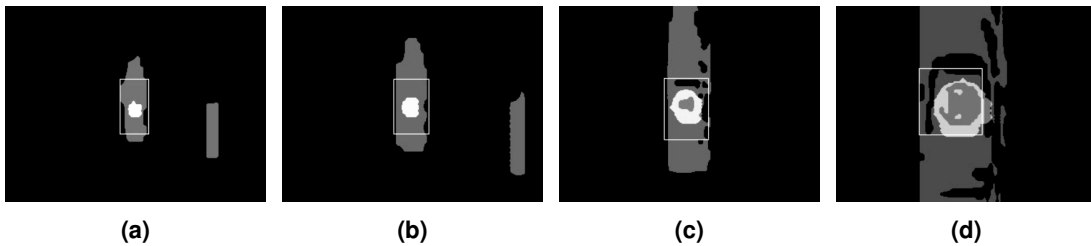


Figura 4.17: Mapas de saliência e zona de seguimento na quarta descida. Os retângulos a branco correspondem a zona que o MR-VTOL utiliza para se manter centrado no objeto de interesse durante a descida.

Durante a descida, à altitude h_d o detetor dedicado começou a apresentar resultados positivos na detecção da plataforma. Os resultados da aplicação deste detetor ao longo da descida são mostrados na Fig. 4.18. Ao alcançar a altitude mínima, h_{min} , as detecções da plataforma foram avaliadas e desta forma terminou o processo de pesquisa da plataforma com 76% de possibilidade do MR-VTOL estar perante a plataforma.

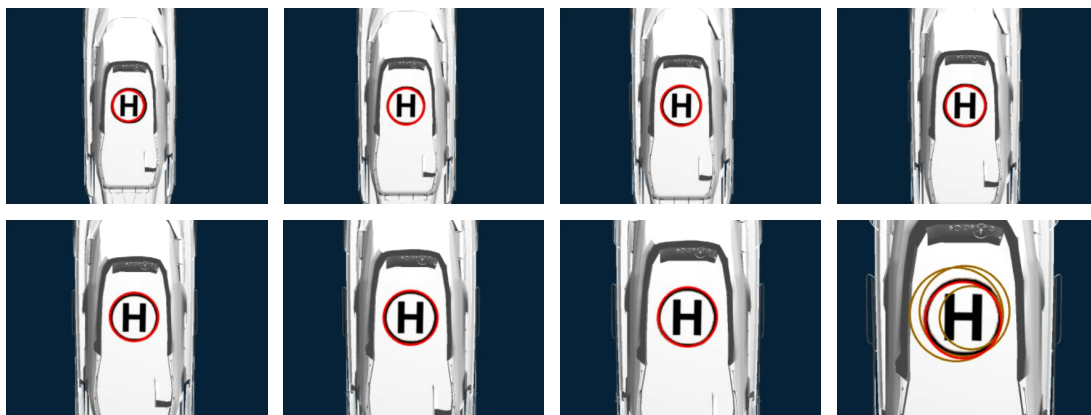


Figura 4.18: Resultado do detetor dedicado durante uma descida. As imagens apresentam evolução de cima para baixo, da esquerda para a direita.

Foram executados 10 planos consecutivos utilizando a aprendizagem da saliência realizada durante o primeiro levantamento. Os resultados estão resumidos na Tabela 4.1.

N	# GeoMaps Pos	% Detetor H	Área (m ²)	Duração (mm:ss)
1	9	86,77	8100	05:11
2	13	89,03	11700	11:58
3	13	89,15	11700	09:51
4	14	79,71	12600	21:36
5	19	90,77	17100	29:08
6	11	89,47	9900	10:05

Tabela 4.1: Resultados dos testes em simulação do plano de execução.

4.3.2 Real

Alguns testes foram realizados com a plataforma transportada pelo barco em ambiente aquático. Estes testes serviram principalmente para avaliar os contrastes e reflexos da água na visão do MR-VTOL. A Fig. 4.19 apresenta os resultados do detetor dedicado e da saliência, onde a zona indicada pelo detetor dedicado como incluindo a plataforma se encontra marcada com uma circunferência vermelha. Da mesma forma, a Fig. 4.20 apresenta os mapas de saliência durante uma subida em que a zona mais saliente corresponde à plataforma para baixas altitudes e ao barco para elevadas altitudes.

O detetor dedicado apresentou detecções da plataforma no local correto. No entanto, o número de círculos detetados (a castanho na imagem) podem ser reduzido através de ajuste nos parâmetros de afinação do sistema, no que respeita à pesquisa de círculos na imagem. Verificou-se ainda que apesar da alteração da aparência da água na imagem com a altitude, a saliência visual continua a indicar o barco que contem a plataforma com a zona mais saliente. Assim, também o iterador *CamShift* efetua corretamente o seguimento da zona mais saliente do mapa de saliência para que o MR-VTOL se centre nesta zona.

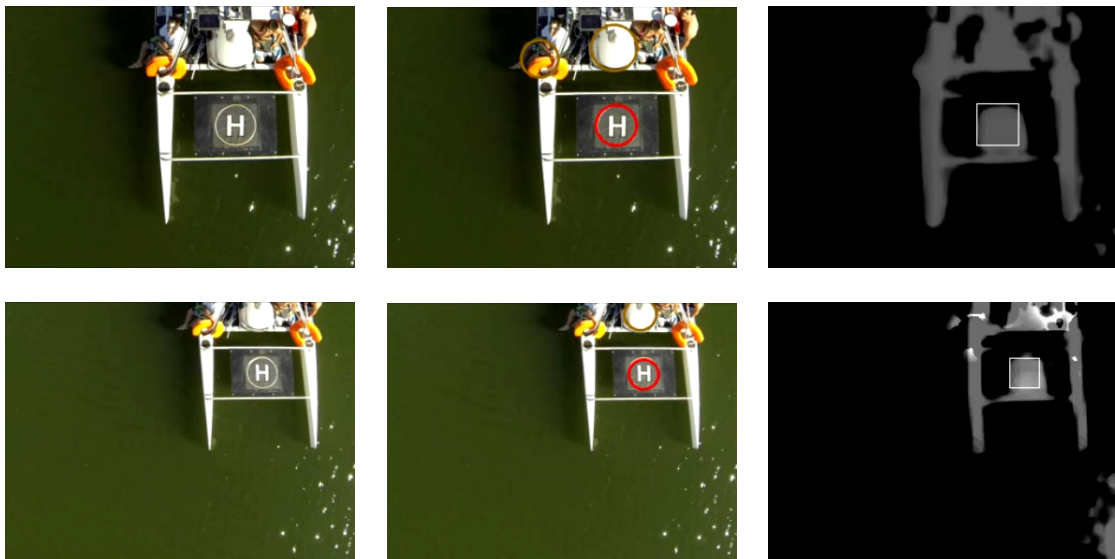


Figura 4.19: Detetor dedicado nos testes em ambiente aquático durante um levantamento. As imagens de baixo correspondem a 21 iterações a seguir às imagens de cima. À esquerda está a imagem de entrada, no centro a zona indicada pelo detetor dedicado como incluindo a plataforma encontra-se marcada com uma circunferência vermelha e à direita o respetivo mapa de saliência com um retângulo a branco que corresponde à zona indicada pelo estimador *CamShift* como incluindo a plataforma.

Os testes às funcionalidades do sistema, aprendizagem durante o levantamento e pesquisa local da plataforma, foram realizados num terreno descampado com a plataforma colocada no chão. A Fig. 4.22 apresenta as zonas indicadas pelo detetor dedicado e pelo iterador *CamShift* como incluindo a plataforma, em algumas iterações da aprendizagem durante um levantamento. A figura divide-se em 3 colunas de duas imagens que correspondem a uma posição onde foram aprendidos ou atualizados os pesos para a altitude atual do MR-VTOL. De notar que devido ao vento, o MR-VTOL pode efetuar mais do que uma aprendizagem para o mesmo intervalo de altitudes porque ao deixar de estar centrado na plataforma, o MR-VTOL centra-se na plataforma sem ter em conta a altitude onde realizou a última aprendizagem. Este procedimento faz com que não seja desperdiçado

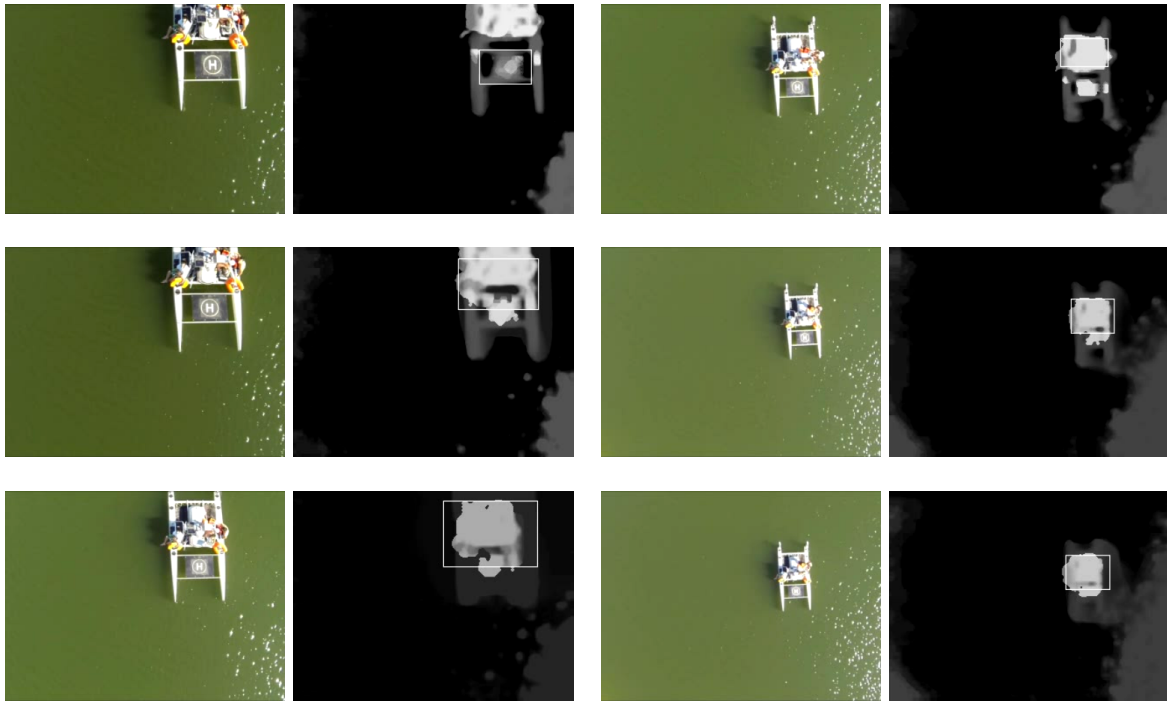


Figura 4.20: Saliência visual nos testes em ambiente aquático durante um levantamento. As imagens estão agrupadas em pares (imagem de entrada e respectivo mapa de saliência) e divididas em 2 colunas. Entre cada imagem existem os respectivos intervalos de iterações: 4, 27, 36, 63, 23. No mapa de saliência, o retângulo a branco corresponde à zona indicada pelo estimador *CamShift* como incluindo a plataforma.

tempo durante um levantamento com muito vento, pois o MR-VTOL estaria constantemente a voltar à altitude onde parou a aprendizagem. Neste caso, os pesos calculados anteriormente são atualizados. A primeira aprendizagem foi efetuada aos 3 m de altitude e terminou aos 18 m.

Após o levantamento, o MR-VTOL foi desviado para uma zona em que a plataforma não está no seu campo de visão. De modo a testar o sistema em situações em que aparência do veículo que transporta a plataforma sofreu alterações, foram removidas as estruturas metálicas que continham a plataforma. Deste modo, a plataforma foi colocada diretamente no terreno, alterando assim a visão da plataforma por parte do MR-VTOL.

Para a aterragem, aos 9,71 m de altitude, o MR-VTOL foi colocado sobre a plataforma para que esta estivesse no seu campo de visão e foi iniciado o processo de pesquisa local da plataforma. A Fig. 4.21 apresenta o resultado desta pesquisa em (b) onde o retângulo a branco corresponde à zona de interesse possível de conter a plataforma. Na Fig. 4.23 são apresentados pares de imagens distribuídas em 3 colunas em que a figura da esquerda corresponde à imagem de entrada e a da direita ao respectivo mapa de saliência em que o quadrado a branco corresponde ao resultado do estimador. O detetor dedicado é aplicado em todas as iterações pois a altitude do MR-VTOL é inferior à altitude h_d . Quando este detetor deteta a plataforma, a imagem de entrada na figura contém a zona detetada representada por uma circunferência a vermelho.

Quando o MR-VTOL atinge a altitude 4.45 m, é efetuada a decisão sobre a deteção da plataforma resultando em 72 % de possibilidade do MR-VTOL estar perante a plataforma. Para esta decisão são consideradas todas as iterações desde o início da aterragem pois o detetor dedicado foi aplicado em todas elas. Durante a descida o detetor dedicado foi aplicado em 62 iterações, sendo que em 46

a plataforma foi detetada com sucesso. Apesar do elevado número de iterações durante a descida, existe uma percentagem de possibilidade do MR-VTOL estar perante a plataforma acima do valor definido como seguro, 70%. No entanto esta percentagem é baixa devido ao facto do elevado número de iterações para os 5.26 m de altitude percorridos. A Tabela 4.2 resume os valores dos resultados obtidos na aterragem.

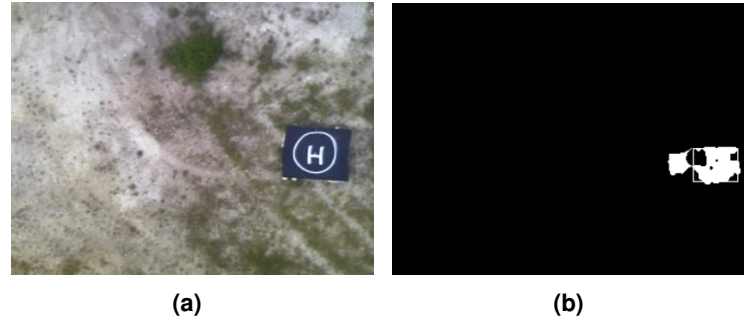


Figura 4.21: Pesquisa visual de zonas de interesse num teste com o robô. Em (a) está a imagem de entrada do sistema e em (b) o objeto presente no mapa de saliência que corresponde a uma zona de interesse, possível de conter a plataforma. Este mapa de saliência foi obtido através dos pesos aprendidos aquando do levantamento.

Descrição	Valor
Possibilidade de deteção da plataforma	72 %
Altitude inicial	9,71 m
Altitude final	4.45 m
Diferença de altitudes	5.26 m
# aplicações Detetor Dedicado	62
# deteções do Detetor Dedicado	46

Tabela 4.2: Resultados da aterragem realizada pelo robô em ambiente terrestre



Figura 4.22: Aprendizagem da aparência da plataforma num levantamento com o robô. A sequência de imagens tem evolução de cima para baixo e por colunas de duas imagens da esquerda para a direita. Para cada posição de aprendizagem surgem duas imagens: à esquerda a imagem de entrada com a plataforma detetada representada com uma circunferência a vermelho; à direita o respetivo mapa de saliência com os pesos atualizados e com um retângulo a branco que representa o resultado do estimador.



Figura 4.23: Aterragem do robô na plataforma. A sequência de imagens tem evolução de cima para baixo e por colunas de duas imagens da esquerda para a direita. Para cada altitude processada surgem duas imagens: à esquerda a imagem de entrada com a plataforma detetada representada com uma circunferência a vermelho; à direita o respetivo mapa de saliência com os pesos para a altitude atual e com um retângulo a branco que representa o resultado do estimador.

Capítulo 5

Conclusões e Trabalho Futuro

Este capítulo apresenta as conclusões aos resultados experimentais obtidos e alguns tópicos de evolução do sistema em trabalhos futuros.

5.1 Conclusões

Esta dissertação propõe um sistema de aterragem para veículos aéreos multi-rotor baseados em comportamentos e em saliência visual, colocando metodologias já estudadas em cooperação para alcançar o seu objetivo. O sistema é capaz de detectar a plataforma para altas altitudes e também de detectar com ajuste mais fino para baixas altitudes. Este objetivo conseguiu-se sem recorrer a uma marca demasiado grande para as altas altitudes e sem a necessidade de uma marca mais pequena dentro da principal para as baixas altitudes. Assim, com uma única marca simples cumpriu-se a questão da altitude. Relativamente ao problema da comunicação entre os veículos, foi proposto um método que contorna a situação de falha. Assim, o MR-VTOL é dotado de um sistema de pesquisa visual pela plataforma, sem ter conhecimento da sua localização. No entanto, esta situação poderá levar a um desfecho indesejável, caso os recursos energéticos do MR-VTOL não sejam suficientes para suportar a duração desta pesquisa.

Confirmou-se a versatilidade da saliência visual na robótica, desta vez aplicado a deteção e seguimento de uma plataforma móvel. Ao longo do desenvolvimento do sistema surgiram algumas dúvidas quanto aos resultados da saliência visual para o objetivo proposto. Foram efetuadas algumas alterações ao método, de modo a diminuir o tempo de resposta, pois influencia diretamente o tempo de resposta do robô.

O estimador utilizado *CamShift* mostrou algumas limitações em situações em que o objeto de interesse desaparece da imagem e depois volta a aparecer e mesmo quando o objeto altera significativamente de aparência, entre imagens. Foram realizados alguns ajustes e condições de atuação do seguidor, eliminando os casos não favoráveis à sua utilização. Desta forma foi possível manter um estimador cujos requisitos de processamento são diminutos, comparativamente a outros estimadores.

Relativamente aos resultados do plano de execução em simulação, obtiveram-se os resultados esperados. O pior cenário foi reproduzido e o MR-VTOL demorou 4 minutos e 41 segundos a encontrar a plataforma de onde levantara. Esta duração poderia ser maior caso fossem encontrados mais objetos de interesse, existissem mais distratores no ambiente, ou simplesmente pela ordem das localizações a visitar, resultantes da aplicação dos mapas georreferenciados. Fora do plano de execução foram simuladas situações de ventos fortes em que o MR-VTOL é levado pelo vento, perdendo os objetos de interesse que estava a seguir. Em todas as situações o sistema conseguiu voltar à localização onde estava e retomar o comportamento em curso.

Relativamente aos resultados dos testes com o robô em ambiente aquático, verificou-se que a saliência visual destaca a plataforma ainda que a sua aparência altere. Também os reflexos do sol na água não provocaram alterações nos resultados expectáveis da saliência visual. A textura e cor da água alteram na imagem com a altitude como se verifica nas imagens apresentadas nesta dissertação. Ao nível dos testes com o robô em ambiente terrestre, os resultados corresponderam ao expectável, durante o levantamento, pesquisa, seguimento e aterragem na plataforma. Verificou-se uma robustez na deteção, na situação em que ocorre uma alteração no veículo que transporta a plataforma entre o levantamento e aterragem, nomeadamente na saliência visual que continua a identificar a plataforma como a zona mais saliente. Nos testes, a aterragem propriamente dita, foi realizada por um sistema externo que controla o MR-VTOL para altitudes inferiores à altitude mínima deste sistema.

Concretamente, a saliência visual elimina os distratores presentes no campo de visão, neste caso consiste na vegetação distribuída de forma aleatória ao longo do terreno descampado. O estimador utilizado apresentou bons resultados acompanhando sempre o objeto de interesse na imagem, mesmo nas situações de alterações significativas entre iterações. Estas situações são, por exemplo, quando o MR-VTOL está centrado num objeto e na iteração seguinte o objeto está quase a sair do campo de visão ou sai mesmo do campo de visão. Quando na iteração seguinte, o objeto volta a aparecer na imagem, o estimador continua o seguimento que estava a efetuar. O detetor dedicado mostrou-se também robusto ao nível do ruído existente na imagem. Foram realizadas deteções positivas da plataforma em imagens cuja plataforma tinha a circunferência distorcida. Este ruído foi provocado pelas vibrações que a estrutura do MR-VTOL transporta para a câmara. Esta situação foi provocada para servir de teste do sistema ao ruído, uma vez que o MR-VTOL tem um mecanismo de amortecimento de vibrações entre a sua estrutura e a câmara.

O sistema de aterragem desenvolvido nesta dissertação não só resolve o problema da aterragem do MR-VTOL como permite também a evolução na utilização dos mapas georreferenciados. Assim, a pesquisa global utilizada neste trabalho pode ser alterada para outra finalidade, por exemplo, navegação em função de outras variáveis representadas por um novo mapa georreferenciado. O facto de existir um detetor dedicado à plataforma utilizada, permite a aplicação deste sistema a uma outra qualquer plataforma, necessitando apenas da substituição do módulo detetor dedicado.

5.2 Trabalho Futuro

Nesta secção são apresentadas novas funcionalidades, melhorias e complementos que poderiam ser implementados em futuros trabalhos com a perspectiva de evoluir o sistema, dotando-o de mais capacidades.

Uma melhoria ao sistema de aterragem passaria por alterar o modelo de saliência visual utilizado de modo a incluir o factor de escala na aprendizagem da plataforma. Desta forma, o nível de intensidade da plataforma no mapa de saliência seria superior ao conseguido nesta dissertação.

Outro aspeto melhorável seria a existência de cópias de segurança de ficheiros YAML de aprendizagens para diferentes horas do dia. As cores e texturas da plataforma e veículo que a transportam, alteram na imagem com as alterações de intensidade de luz solar. Assumindo que o MR-VTOL realiza a aprendizagem da plataforma durante o levantamento e sabendo que a autonomia de voo do MR-VTOL não é suficiente para que a diferença de luz solar seja significativa, será aplicada a última aprendizagem. No entanto, a existência das cópias de segurança das aprendizagens salvaguarda a situação de falha da aprendizagem durante a subida, sendo aplicada a aprendizagem que melhor se adapta à hora atual.

Outra melhoria consiste na adição de um método de aterragem de emergência através da deteção de um local seguro. Esta aterragem seria utilizada quando o MR-VTOL não tem informação da localização da plataforma, a plataforma não é encontrada pela pesquisa em espiral e os recursos energéticos do veículo estão na reserva. Assim, o sistema utiliza as reservas de energia para aterrar o veículo, evitando um acidente por falta de energia.

Por fim, uma melhoria ao detetor dedicado passaria pelo reconhecimento de informação normalizada sobre uma nova plataforma, i.e., o detetor receber características normalizadas sobre uma nova plataforma. Desta forma o detetor dedicado seria dinâmico, capacitando o sistema de aprender mais do que uma plataforma.

O trabalho desenvolvido nesta dissertação pode ser aplicado a qualquer MR-VTOL autónomo em que seja necessário o veículo levantar e aterrar numa mesma plataforma. No entanto, através de algumas alterações seria possível aplicar num sistema de entregas de encomendas em que o MR-VTOL levanta de uma plataforma no armazém e desloca-se à casa do comprador do produto. Esta aterragem teria que ser realizada de outra forma pois não existe uma plataforma de aterragem na casa de cada cliente. Após a entrega, o MR-VTOL regressaria ao armazém onde a aterragem seria realizada com auxílio da plataforma. Outra aplicação possível seria na pesquisa por sobreviventes num local de naufrágio em que exista uma elevada agitação marítima. Podia também ser utilizado para localização de pessoas numa multidão em que não é possível encontrar a pessoa que se procura, através de um detetor dedicado à pesquisa de faces, aplicado à imagem aérea. Em armazéns cuja altura a que estão as caixas com os artigos seja demasiado elevada, para fazer por exemplo, inventário. Neste caso, seria evitado o deslocamento de uma pessoa através do mecanismo elevador para verificar a existência de determinado número de caixas, pois esta tarefa estaria a cargo da visão do MR-VTOL.

Bibliografia

- Allen, J. G., Xu, R. Y., and Jin, J. S. (2004). Object tracking using camshift algorithm and multiple quantized feature spaces. In *Proceedings of the Pan-Sydney Area Workshop on Visual Information Processing*, pages 3–7. Australian Computer Society, Inc.
- Arkin, R. C. (1998). *Behavior-based robotics [electronic resource]*. MIT press.
- Bosch, S., Lacroix, S., and Caballero, F. (2006). Autonomous detection of safe landing areas for an uav from monocular images. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5522–5527. IEEE.
- Bradski, G. and Kaehler, A. (2008). *Learning OpenCV: Computer vision with the OpenCV library*. O'reilly.
- Brooks, R. A. (1986). A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, 2(1):14–23.
- Cesetti, A., Frontoni, E., Mancini, A., and Zingaretti, P. (2010). Autonomous safe landing of a vision guided helicopter. In *Proceedings of the IEEE/ASME International Conference on Mechatronics and Embedded Systems and Applications (MESA)*, pages 125–130. IEEE.
- Comaniciu, D. and Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):603–619.
- De Wagter, C. and Mulder, J. (2005). Towards vision-based uav situation awareness. In *Proceedings of the AIAA Guidance, Navigation and Control Conference*, pages 15–18.
- Edwards, B., Archibald, J., Fife, W., and Lee, D.-J. (2007). A vision system for precision mav targeted landing. In *Proceedings of the International Symposium on Computational Intelligence in Robotics and Automation (CIRA)*, pages 125–130. IEEE.
- Feng, L., Chen Ben, M., and Yew, L. K. (2007). Integration and implementation of a low-cost and vision-based uav tracking system. In *Proceedings of the Chinese Control Conference (CCC)*, pages 731–736. IEEE.
- Frintrop, S. (2006). *VOCUS: A visual attention system for object detection and goal-directed search*, volume 3899. Springer.
- Garcia-Pardo, P. J., Sukhatme, G. S., and Montgomery, J. F. (2002). Towards vision-based safe landing for an autonomous helicopter. *Robotics and Autonomous Systems*, 38(1):19–29.
- Herisse, B., Hamel, T., Mahony, R., and Russotto, F.-X. (2012). Landing a vtol unmanned aerial vehicle on a moving platform using optical flow. *IEEE Transactions on Robotics*, 28(1):77–89.

- Horn, B. K. and Schunck, B. G. (1981). Determining optical flow. *Artificial Intelligence*, 17(1):185–203.
- Hu, M.-K. (1962). Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, 8(2):179–187.
- Itti, L., Koch, C., and Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11):1254–1259.
- Johnson, A., Montgomery, J., and Matthies, L. (2005). Vision guided landing of an autonomous helicopter in hazardous terrain. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 3966–3971. IEEE.
- Lange, S., Sünderhauf, N., and Protzel, P. (2008). Autonomous landing for a multirotor uav using vision. In *Proceedings of the International Conference on Simulation, Modeling, and Programming for Autonomous Robots (SIMPAN)*, pages 482–491.
- Lange, S., Sunderhauf, N., and Protzel, P. (2009). A vision based onboard approach for landing and position control of an autonomous multirotor uav in gps-denied environments. In *Proceedings of the International Conference on Advanced Robotics (ICAR)*, pages 1–6. IEEE.
- Lindsey, Q., Mellinger, D., and Kumar, V. (2011). Construction of cubic structures with quadrotor teams. In *Proceedings of Robotics: Science and Systems*, Los Angeles, CA, USA.
- Meger, D., Forssén, P.-E., Lai, K., Helmer, S., McCann, S., Southey, T., Baumann, M., Little, J. J., and Lowe, D. G. (2008). Curious george: An attentive semantic robot. *Robotics and Autonomous Systems*, 56(6):503–511.
- Meier, L., Tanskanen, P., Fraundorfer, F., and Pollefeys, M. (2011). Pixhawk: A system for autonomous flight using onboard computer vision. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 2992–2997. IEEE.
- Mellinger, D., Shomin, M., Michael, N., and Kumar, V. (2013). Cooperative grasping and transport using multiple quadrotors. In *Distributed autonomous robotic systems*, pages 545–558. Springer.
- Mendonça, R., Santana, P., Marques, F., Lourenço, A., Silva, J., and Barata, J. (2013). Kelpie: A ros-based multi-robot simulator for water surface and aerial vehicles. In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 3645–3650. IEEE.
- Mendonça, R. A. M. (2012). A learning approach to swarm-based path detection and tracking. Master's thesis, Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa.
- Michael, N., Fink, J., and Kumar, V. (2011). Cooperative manipulation and transportation with aerial robots. *Autonomous Robots*, 30(1):73–86.
- Montgomery, J. F. (1999). *Learning helicopter control through teaching by showing*. PhD thesis, University of Southern California.
- Murphy, D. W. and Cycon, J. (1999). Applications for mini vtol uav for law enforcement. In *Proceedings of the Enabling Technologies for Law Enforcement and Security*, pages 35–43. International Society for Optics and Photonics.

- Paravati, G., Sanna, A., Pralio, B., and Lamberti, F. (2009). A genetic algorithm for target tracking in flir video sequences using intensity variation function. *IEEE Transactions on Instrumentation and Measurement*, 58(10):3457–3467.
- Pfeifer, R. et al. (2001). *Understanding intelligence*. The MIT Press.
- Pinto, E., Santana, P., Marques, F., Mendonça, R., Lourenço, A., and Barata, J. (2014). On the design of a robotic system composed of an unmanned surface vehicle and a piggybacked vtol. In *To appear in Proceedings of DoCEIS*. Springer.
- Pirjanian, P. (1999). Behavior coordination mechanisms-state-of-the-art. Technical report, Citeseer.
- Quigley, M., Conley, K., Gerkey, B., Faust, J., Foote, T., Leibs, J., Wheeler, R., and Ng, A. Y. (2009). Ros: an open-source robot operating system. In *Proceedings of the ICRA Workshop on Open Source Software*, volume 3.
- Rosenblatt, J. K. (1997). Damn: A distributed architecture for mobile navigation. *Journal of Experimental & Theoretical Artificial Intelligence*, 9(2-3):339–360.
- Ruesch, J., Lopes, M., Bernardino, A., Hornstein, J., Santos-Victor, J., and Pfeifer, R. (2008). Multi-modal saliency-based bottom-up attention a framework for the humanoid robot icub. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 962–967. IEEE.
- Santana, P., Alves, N., Correia, L., and Barata, J. (2010). Swarm-based visual saliency for trail detection. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 759–765. IEEE.
- Santana, P. and Correia, L. (2011). Swarm cognition on off-road autonomous robots. *Swarm Intelligence*, 5(1):45–72.
- Santana, P., Guedes, M., Correia, L., and Barata, J. (2011). Stereo-based all-terrain obstacle detection using visual saliency. *Journal of Field Robotics*, 28(2):241–263.
- Santana, P., Mendonça, R., Correia, L., and Barata, J. (2012). Neural-swarm visual saliency for path following. *Applied Soft Computing*.
- Saripalli, S., Montgomery, J. F., and Sukhatme, G. S. (2002). Vision-based autonomous landing of an unmanned aerial vehicle. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, volume 3, pages 2799–2804. IEEE.
- Saripalli, S., Montgomery, J. F., and Sukhatme, G. S. (2003). Visually guided landing of an unmanned aerial vehicle. *IEEE Transactions on Robotics and Automation*, 19(3):371–380.
- Saripalli, S. and Sukhatme, G. S. (2006). Landing on a moving target using an autonomous helicopter. In *Proceedings of the Field and Service Robotics*, pages 277–286. Springer.
- Saripalli, S. and Sukhatme, G. S. (2007). Landing a helicopter on a moving target. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 2030–2035. IEEE.
- Sharp, C. S., Shakernia, O., and Sastry, S. S. (2001). A vision system for landing an unmanned aerial vehicle. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, volume 2, pages 1720–1727. IEEE.

- Templeton, T., Shim, D. H., Geyer, C., and Sastry, S. S. (2007). Autonomous vision-based landing and terrain mapping using an mpc-controlled unmanned rotorcraft. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 1349–1356. IEEE.
- Todt, E. and Torras, C. (2000). Detection of natural landmarks through multiscale opponent features. In *Proceedings of the 15th International Conference on Pattern Recognition*, volume 3, pages 976–979. IEEE.
- Wendel, A., Irschara, A., and Bischof, H. (2011). Natural landmark-based monocular localization for mavs. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 5792–5799. IEEE.
- Wendel, J., Meister, O., Schlaile, C., and Trommer, G. F. (2006). An integrated gps/mems-imu navigation system for an autonomous helicopter. *Aerospace Science and Technology*, 10(6):527–533.
- Wenzel, K. E., Masselli, A., and Zell, A. (2011). Automatic take off, tracking and landing of a miniature uav on a moving carrier vehicle. *Journal of Intelligent & Robotic Systems*, 61(1-4):221–238.
- Yilmaz, A., Shafique, K., Lobo, N., Li, X., Olson, T., and Shah, M. (2001). Target-tracking in flir imagery using mean-shift and global motion compensation. In *Proceedings of the IEEE Workshop on Computer Vision Beyond Visible Spectrum*, pages 54–58.
- Yuan, Z., Gong, Z., Wu, J., Chen, J., and Rao, J. (2005). A real-time vision-based guided method for autonomous landing of a rotor-craft unmanned aerial vehicle. In *Proceedings of the IEEE International Conference Mechatronics and Automation*, volume 4, pages 2212–2215. IEEE.

Apêndice A

Referenciais

Conversão de coordenadas GPS (latitude, longitude), respetivamente (φ, λ) , para coordenadas UTM (Este, Norte), respetivamente (E, N) .

$$t = \sinh \left(\tanh^{-1} \sinh \varphi - \frac{2\sqrt{n}}{1+n} \tanh^{-1} \left(\frac{2\sqrt{n}}{1+n} \sin \varphi \right) \right) \quad (\text{A.1})$$

$$\xi' = \tan^{-1} \left(\frac{t}{\cos(\lambda - \lambda_0)} \right), \eta' = \tanh^{-1} \left(\frac{\sin(\lambda - \lambda_0)}{\sqrt{1+t^2}} \right) \quad (\text{A.2})$$

$$\sigma = 1 + \sum_{j=1}^3 2j\alpha_j \cos(2j\xi') \cosh(2j\eta'), \tau = \sum_{j=1}^3 2j\alpha_j \sin(2j\xi') \sinh(2j\eta') \quad (\text{A.3})$$

$$(\text{A.4})$$

$$E = E_0 + k_0 A \left(\eta' + \sum_{j=1}^3 \alpha_j \cos(2j\xi') \sinh(2j\eta') \right) \quad (\text{A.5})$$

$$N = N_0 + k_0 A \left(\xi' + \sum_{j=1}^3 \alpha_j \sin(2j\xi') \cosh(2j\eta') \right) \quad (\text{A.6})$$

$$k = \frac{k_0}{a} \sqrt{\left\{ 1 + \left(\frac{1-n}{1+n} \tan \varphi \right)^2 \right\} \frac{\sigma^2 + \tau^2}{t^2 + \cos^2(\lambda - \lambda_0)}} \quad (\text{A.7})$$

$$\gamma = \tan^{-1} \left(\frac{\tau\sqrt{1+t^2} + \sigma t \tan(\lambda - \lambda_0)}{\sigma\sqrt{1+t^2} - \tau t \tan(\lambda - \lambda_0)} \right) \quad (\text{A.8})$$

Conversão de coordenadas UTM (Este, Norte, Zona, Hemisfério), respetivamente $(E, N, Z, Hemi)$, para coordenadas GPS (latitude, longitude), respetivamente (φ, λ) .

$$\xi = \frac{N - N_0}{k_0 A}, \eta = \frac{E - E_0}{k_0 A} \quad (\text{A.9})$$

$$\xi' = \xi - \sum_{j=1}^3 \beta_j \sin(2j\xi) \cosh(2j\eta), \eta' = \eta - \sum_{j=1}^3 \beta_j \cos(2j\xi) \sinh(2j\eta) \quad (\text{A.10})$$

$$\sigma' = 1 - \sum_{j=1}^3 2\beta_j \cos(2j\xi) \cosh(2j\eta), \tau' = \sum_{j=1}^3 2\beta_j \sin(2j\xi) \sinh(2j\eta) \quad (\text{A.11})$$

$$\chi = \sin^{-1} \left(\frac{\sin \xi'}{\cosh \eta'} \right) \quad (\text{A.12})$$

$$\varphi = \chi + \sum_{j=1}^3 \delta_j \sin(2j\chi) \quad (\text{A.13})$$

$$\lambda_0 = \text{Zone} \cdot 6^\circ - 183^\circ \quad (\text{A.14})$$

$$\lambda = \lambda_0 + \tan^{-1} \left(\frac{\sinh \eta'}{\cos \xi'} \right) \quad (\text{A.15})$$

$$k = \frac{k_0}{a} \sqrt{\left\{ 1 + \left(\frac{1-n}{1+n} \tan \varphi \right)^2 \right\} \frac{\cos^2 \xi' + \sinh^2 \eta'}{\sigma'^2 + \tau'^2}} \quad (\text{A.16})$$

$$\gamma = \text{Hemi} \cdot \tan^{-1} \left(\frac{\tau' + \sigma' \tan \xi' \tanh \eta'}{\sigma' + \tau' \tan \xi' \tanh \eta'} \right) \quad (\text{A.17})$$

Apêndice B

Momentos Geométricos

Este apêndice apresenta o processo de cálculo dos momentos geométricos invariantes de uma imagem I , de coordenadas (x, y) , sendo $j, i = 0, 1, 2, \dots$ e $j + i$ a ordem dos momentos.

Momentos Espaciais

$$m_{ji} = \sum_{x,y} (I(x, y) \cdot x^j \cdot y^i) \quad (\text{B.1})$$

Momentos Centrais

$$\mu_{ji} = \sum_{x,y} (I(x, y) \cdot (x - \bar{x})^j \cdot (y - \bar{y})^i) \quad (\text{B.2})$$

Centros de Massa

$$\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}} \quad (\text{B.3})$$

Momentos Centrais Normalizados

$$\eta_{ji} = \frac{\mu_{ji}}{\mu_{00}^{(i+j)/2+1}} \quad (\text{B.4})$$

$$\begin{aligned}
 M1 &= \eta_{20} + \eta_{02} \\
 M2 &= (\eta_{20} + \eta_{02})^2 + 4\eta_{11}^2 \\
 M3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
 M4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
 M5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\
 &\quad (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
 M6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
 M7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - \\
 &\quad (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]
 \end{aligned}$$

Em que $\eta_{ji} = nu_{ji}$.

Apêndice C

Modelo de Saliência Base

O modelo base começa por calcular 3 mapas de características a partir da imagem de entrada. Estes mapas são: (1) de intensidade, I ; (2) de dupla-oposição de cores vermelho-verde, RG ; (3) de dupla-oposição de cores azul-amarelo, BY . Seguem-se as equações que permitem esta separação, em que r , g e b são respetivamente os canais vermelho, verde e azul.

$$I = \frac{r + g + b}{3} \quad (\text{C.1})$$

Os canais vermelho, verde, azul e amarelo normalizados, respetivamente:

$$r_n = \frac{r - (g + b)}{2} \quad (\text{C.2})$$

$$g_n = \frac{g - (r + b)}{2} \quad (\text{C.3})$$

$$b_n = \frac{b - (r + g)}{2} \quad (\text{C.4})$$

$$y_n = \frac{(r + g)}{2} - \frac{|r - g|}{2} - b \quad (\text{C.5})$$

Para separar intensidade de saturação, cada um destes canais é normalizado por I , resultando respetivamente em R , G , B e Y . Para obter os dois mapas de dupla-oposição de cores:

$$BY = B - Y \quad (\text{C.6})$$

$$RG = R - G \quad (\text{C.7})$$

