



MINISTÉRIO DA  
CIÊNCIA, TECNOLOGIA  
E INOVAÇÕES



sid.inpe.br/mtc-m21c/2021/02.05.01.02-TDI

**CARACTERIZAÇÃO E DESENVOLVIMENTO DE UM  
SISTEMA AUTOADAPTATIVO PARA ESTIMAÇÃO DA  
POSIÇÃO DE VEÍCULOS AÉREOS NÃO TRIPULADOS  
BASEADO EM IMAGENS**

Gabriel Fornari

Tese de Doutorado do Curso de Pós-Graduação em Computação Aplicada, orientada pelos Drs. Valdivino Alexandre de Santiago Júnior e Elcio Hideiti Shiguemori, aprovada em 22 de dezembro de 2020.

URL do documento original:

<<http://urlib.net/8JMKD3MGP3W34R/44582LP>>

INPE  
São José dos Campos  
2020

**PUBLICADO POR:**

Instituto Nacional de Pesquisas Espaciais - INPE  
Gabinete do Diretor (GBDIR)  
Serviço de Informação e Documentação (SESID)  
CEP 12.227-010  
São José dos Campos - SP - Brasil  
Tel.:(012) 3208-6923/7348  
E-mail: pubtc@inpe.br

**CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO DA PRODUÇÃO INTELLECTUAL DO INPE - CEPPII (PORTARIA Nº 176/2018/SEI-INPE):****Presidente:**

Dra. Marley Cavalcante de Lima Moscati - Centro de Previsão de Tempo e Estudos Climáticos (CGCPT)

**Membros:**

Dra. Carina Barros Mello - Coordenação de Laboratórios Associados (COCTE)

Dr. Alisson Dal Lago - Coordenação-Geral de Ciências Espaciais e Atmosféricas (CGCEA)

Dr. Evandro Albiach Branco - Centro de Ciência do Sistema Terrestre (COCST)

Dr. Evandro Marconi Rocco - Coordenação-Geral de Engenharia e Tecnologia Espacial (CGETE)

Dr. Hermann Johann Heinrich Kux - Coordenação-Geral de Observação da Terra (CGOBT)

Dra. Ieda Del Arco Sanches - Conselho de Pós-Graduação - (CPG)

Silvia Castro Marcelino - Serviço de Informação e Documentação (SESID)

**BIBLIOTECA DIGITAL:**

Dr. Gerald Jean Francis Banon

Clayton Martins Pereira - Serviço de Informação e Documentação (SESID)

**REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:**

Simone Angélica Del Ducca Barbedo - Serviço de Informação e Documentação (SESID)

André Luis Dias Fernandes - Serviço de Informação e Documentação (SESID)

**EDITORAÇÃO ELETRÔNICA:**

Ivone Martins - Serviço de Informação e Documentação (SESID)

Cauê Silva Fróes - Serviço de Informação e Documentação (SESID)



MINISTÉRIO DA  
CIÊNCIA, TECNOLOGIA  
E INOVAÇÕES



sid.inpe.br/mtc-m21c/2021/02.05.01.02-TDI

## CARACTERIZAÇÃO E DESENVOLVIMENTO DE UM SISTEMA AUTOADAPTATIVO PARA ESTIMAÇÃO DA POSIÇÃO DE VEÍCULOS AÉREOS NÃO TRIPULADOS BASEADO EM IMAGENS

Gabriel Fornari

Tese de Doutorado do Curso de Pós-Graduação em Computação Aplicada, orientada pelos Drs. Valdivino Alexandre de Santiago Júnior e Elcio Hideiti Shiguemori, aprovada em 22 de dezembro de 2020.

URL do documento original:

<<http://urlib.net/8JMKD3MGP3W34R/44582LP>>

INPE  
São José dos Campos  
2020

Dados Internacionais de Catalogação na Publicação (CIP)

---

Fornari, Gabriel.

F767c Caracterização e desenvolvimento de um sistema autoadaptativo para estimação da posição de veículos aéreos não tripulados baseado em imagens / Gabriel Fornari. – São José dos Campos : INPE, 2020.

xxiv + 157 p. ; (sid.inpe.br/mtc-m21c/2021/02.05.01.02-TDI)

Tese (Doutorado em Computação Aplicada) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2020.

Orientadores : Drs. Valdivino Alexandre de Santiago Júnior e Elcio Hideiti Shiguemori.

1. Sistemas autoadaptativos. 2. Veículos Aéreos Não Tripulados (VANTs). 3. Visão computacional. 4. Estimação da posição por imagens. 5. Aprendizado de máquina. I. Título.

CDU 004.02:629.014.8

---



Esta obra foi licenciada sob uma Licença [Creative Commons Atribuição-NãoComercial 3.0 Não Adaptada](https://creativecommons.org/licenses/by-nc/3.0/).

This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](https://creativecommons.org/licenses/by-nc/3.0/).



MINISTÉRIO DA  
CIÊNCIA, TECNOLOGIA  
E INOVAÇÕES



## INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

### DEFESA FINAL DE TESE DE GABRIEL FORNARI

No dia 22 de dezembro de 2020, às 10h, por videoconferência, o(a) aluno(a) mencionado(a) acima defendeu seu trabalho final (apresentação oral seguida de arguição) perante uma Banca Examinadora, cujos membros estão listados abaixo. O(A) aluno(a) foi APROVADO(A) pela Banca Examinadora por unanimidade, em cumprimento ao requisito exigido para obtenção do Título de Doutor em Computação Aplicada. O trabalho precisa da incorporação das correções sugeridas pela Banca Examinadora e revisão final pelo(s) orientador(es).

Título: "Caracterização e Desenvolvimento de um sistema Autoadaptativo para Estimção da Posição de Veículos Aéreos Não Tripulados baseado em Imagens"

Eu, Rafael Duarte Coelho dos Santos, como Coordenador e Presidente do Conselho do Curso de Computação Aplicada, em conformidade com o Presidente da Banca Examinadora Dr. Lamartine Nogueira Frutuoso Guimarães, assino esta ATA em nome de todos os membros.

#### Membros da Banca

Dr. Lamartine Nogueira Frutuoso Guimarães, Presidente IEAv.  
Dr. Valdivino Alexandre de Santiago Junior, Orientador(a) INPE  
Dr. Elcio Hideiti Shiguemori, Orientador(a) IEAv  
Dr. Haroldo Fraga de Campos Velho, Membro da Banca, INPE  
Dr. Thales Sehn Korting, Membro da Banca, INPE  
Dr. Jurandy Gomes de Almeida Junior, Convidado(a) UNIFESP  
Dr. Marcos Ricardo Omena de Albuquerque Maximo, Convidado(a) ITA



Documento assinado eletronicamente por **Rafael Duarte Coelho dos Santos, Tecnologista**, em 02/02/2021, às 13:03 (horário oficial de Brasília), com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).

A autenticidade deste documento pode ser conferida no site <http://sei.mctic.gov.br/verifica.html>, informando o código verificador **6454686** e o código CRC **1AA6E1BD**.



---

**Referência:** Processo nº 01340.000786/2021-18

SEI nº 6454686

*“E já se encontra a lua sob nossos pés;  
Pouco é o tempo que nos é concedido  
E há mais coisas para ver do que estas que vês”*

DANTE ALIGHIERI  
em “A Divina Comédia”





## AGRADECIMENTOS

Agradeço, antes de mais nada, aos meus orientadores, Drs. Elcio e Valdivino, pela paciência que tiveram comigo e pelas inúmeras discussões que foram feitas ao longo desta jornada e que culminaram na elaboração desta tese. Agradeço também aos integrantes do PITER, que me acolheram e revelaram suas experiências em uma área até então completamente nova para mim.

Sou muito grato ao professor da Universidad Autónoma de Nuevo León, Dr. Manuel Jiménez, e também a sua equipe pelas contribuições que deram a este trabalho. Da mesma forma, deixo aqui meu especial agradecimento ao professor da École Centrale de Nantes, Dr. Vincent Fremont, por sua dedicação e ajuda que contribuíram de maneira fundamental na minha formação. Também agradeço ao Dr. José Renato Garcia Braga por sua contribuição significativa de algoritmos e conhecimento, que serviram de alicerces para esta tese. Da mesma forma, agradeço ao aluno de iniciação científica Hércules Carlos, pelo seu apoio na obtenção dos resultados deste trabalho.

Agradeço o apoio financeiro promovido pelo CNPq, permitindo que eu me dedicasse integralmente à pesquisa. Agradeço também aos colegas e funcionários do INPE por seu respeito e companheirismo nas horas que passamos juntos.

Serei eternamente grato aos meus pais e irmão, que me confortaram ao longo desta jornada e me deram forças nas horas difíceis. Por último, agradeço a todos aqueles que, embora não citados diretamente aqui, desempenharam papel de igual importância na minha formação.



## RESUMO

Ambientes operacionais dinâmicos requerem que sejam desenvolvidos sistemas que possam se adaptar a mudanças não determinísticas das condições operacionais. Na literatura, áreas com diferentes nomenclaturas estudam o processo de adaptação em sistemas de engenharia, tais como sistemas autoadaptativos e sistemas reconfiguráveis. Recentemente, os Veículos Aéreos Não Tripulados (VANTs) passaram a ser utilizados em diferentes aplicações. Tais veículos necessitam de um sistema de estimação da posição para serem capazes de realizar o controle e navegação da aeronave. Apesar de ser por meio do Sistema Global de Navegação por Satélite (GNSS) que a estimativa de posição de VANTs em tempo real é mais usualmente identificada, devido a uma série de problemas associados ao uso de satélites no cálculo da posição, metodologias alternativas estão sendo propostas para o mesmo fim, como o uso de sensores imageadores e visão computacional. Nesse sentido, diferentes algoritmos de visão computacional podem ser utilizados para compor o sistema de navegação do VANT, que, por sua vez, se encontra imerso em um ambiente operacional cujas condições podem variar dinamicamente. A presente tese de doutorado objetiva contribuir para o processo de navegação autônoma de VANTs por meio da visão computacional identificando o mais adequado, entre vários algoritmos de visão computacional, para a estimação da posição de VANTs por imagens obtidas sob diferentes condições ambientais. Para esse processo, é proposto o uso de diferentes técnicas de Aprendizado de Máquina, que incluem Máquinas de Vetores de Suporte, Árvores de Classificação e Redes Neurais Artificiais. Os resultados revelam que a abordagem autoadaptativa melhora o desempenho, para estimar a posição de VANTs por meio de imagens, se comparada a uma abordagem não adaptativa estática, demonstrando a viabilidade dessa pesquisa.

Palavras-chave: Sistemas Autoadaptativos. Veículos Aéreos Não Tripulados (VANTs). Visão Computacional. Estimação da Posição por Imagens. Aprendizado de Máquina.



# CHARACTERIZATION AND DEVELOPMENT OF A SELF-ADAPTIVE SYSTEM FOR ESTIMATING THE POSITION OF UNMANNED AERIAL VEHICLES BASED ON IMAGES

## ABSTRACT

Dynamic operating environments demand the development of systems able to self-adapt to non-deterministic changes in operating conditions. In the scientific literature, different nomenclatures are used to name the study of adaptation process in engineering systems, such as self-adaptive systems and reconfigurable systems. Recently, Unmanned Aerial Vehicles (UAV) have been used in different applications. Such vehicles need a pose estimation system and, consequently, a system able to control and navigate the aircraft. Despite the fact that it is by means of Global Navigation Satellite System (GNSS) that the estimation of the position of real-time UAVs is usually estimated, there is a series of problems associated to the use of satellites in the computation of the position, and new alternative methodologies are being proposed for the same goal, as the use of imaging sensors and computer vision. In this sense, different algorithms of computer vision can be used to compose the navigation system of the UAV, which, in turn, is immersed in an operating environment whose conditions can change dynamically. This PhD thesis aims to contribute to the process of autonomous navigation of UAVs using computer vision identifying the most appropriate, among several computer vision algorithms, for the pose estimation system of UAVs based on images obtained under different environmental conditions. For this strategy, several Machine Learning methods are used, including Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Classification Trees. The results revealed that the adaptive methodology can increase performance, to estimate the position of UAVs using images, if compared to a nonadaptive static methodology, demonstrating the viability of this research.

Keywords: Self-Adaptive Systems. Unmanned Aerial Vehicles (UAVs). Computer Vision. Pose Estimation based on Images. Machine Learning.



## LISTA DE FIGURAS

	<u>Pág.</u>
1.1 Processo metodológico envolvendo as principais atividades desenvolvidas neste trabalho. . . . .	9
2.1 Resultados da classificação dos artigos segundo a propriedade “área do conhecimento”. . . . .	15
2.2 Classificação proposta para sistemas reconfiguráveis. . . . .	16
2.3 Resultados da classificação dos artigos segundo a propriedade “dispositivo utilizado”. . . . .	17
2.4 Resultados da classificação dos artigos segundo a propriedade “caso de uso”. . . . .	18
2.5 Resultados da classificação dos artigos segundo a propriedade “país dos autores”. . . . .	19
2.6 Ilustração do funcionamento do SLAM com imagens aéreas. . . . .	29
3.1 Fluxograma da técnica de estimação da posição baseado em reconhecimento de regiões. . . . .	37
3.2 Imagem capturada por VANT antes e depois de sofrer ajuste de escala. . . . .	41
3.3 Imagem capturada por VANT antes e depois de sofrer ajuste de rotação. . . . .	42
3.4 Convolução entre uma imagem e uma janela de tamanho $3 \times 3$ . . . . .	43
3.5 Exemplo do efeito do filtro da média em uma imagem aérea. . . . .	45
3.6 Exemplo do efeito do filtro de Gauss em uma imagem aérea. . . . .	46
3.7 Exemplo do efeito do filtro da mediana em uma imagem aérea com ruído. . . . .	47
3.8 Exemplo do efeito do operador de Sobel-Feldman em uma imagem aérea. . . . .	49
3.9 Exemplo do efeito do operador de Canny em uma imagem aérea. . . . .	50
3.10 Exemplo do efeito da detecção de bordas realizado por uma RNA. . . . .	52
3.11 Exemplo da técnica de congruência de fase na detecção de bordas. . . . .	53
3.12 Exemplo do uso de máscara para demarcar a região válida de uma imagem. . . . .	57
3.13 Exemplo de árvore de classificação binária. . . . .	63
3.14 Exemplo de hiperplano em um conjunto de dados linearmente separável. . . . .	65
3.15 Modelo não linear de um neurônio artificial. . . . .	67
3.16 Estrutura de um perceptron multicamadas. . . . .	69
4.1 Visão geral da abordagem autoadaptativa. . . . .	72
4.2 Modelo abstrato para seleção de algoritmos proposto por John Rice. . . . .	73
4.3 Posição e orientação de VANTs no espaço e efeito em imagens. . . . .	77

4.4	Combinações de algoritmos possíveis para estimação da posição por imagens. . . . .	82
4.5	Esquema em que classificadores são utilizados na seleção de algoritmos. . . . .	84
5.1	Trajectoria percorrida por um VANT durante o experimento em São José dos Campos, Brasil. . . . .	88
5.2	Trajectoria percorrida por um VANT durante a realização do experimento em Revinge, Suécia. . . . .	90
5.3	Hexacóptero utilizado na realização dos experimentos em Monterrey, México. . . . .	91
5.4	Trajectoria percorrida por um VANT durante o primeiro experimento em Monterrey, México. . . . .	92
5.5	Trajectoria percorrida por um VANT durante o segundo experimento em Monterrey, México. . . . .	93
5.6	Trajectoria percorrida por um VANT durante o terceiro experimento em Monterrey, México. . . . .	94
6.1	Análise comparativa entre as posições estimadas pelo GNSS e pela técnica de reconhecimento por imagens. . . . .	96
6.2	Erro da estimação da posição utilizando a abordagem não adaptativa aplicada ao conjunto de imagens do experimento realizado no Brasil. . . . .	100
6.3	Erro da estimação da posição utilizando a abordagem adaptativa ideal aplicada ao conjunto de imagens do experimento realizado no Brasil. . . . .	101
6.4	Erro da estimação da posição utilizando a abordagem adaptativa por regiões aplicada ao conjunto de imagens do experimento realizado no Brasil. . . . .	102
6.5	Exemplos de imagens dos subconjuntos referente ao experimento realizado no Brasil. . . . .	103
6.6	Erro da estimação da posição utilizando a abordagem não adaptativa aplicada ao conjunto de imagens do experimento realizado na Suécia. . . . .	106
6.7	Erro da estimação da posição utilizando a abordagem adaptativa ideal aplicada ao conjunto de imagens do experimento realizado na Suécia. . . . .	107
6.8	Erro da estimação da posição utilizando a abordagem adaptativa por regiões aplicada ao conjunto de imagens do experimento realizado na Suécia. . . . .	108
6.9	Exemplos de imagens dos subconjuntos referente ao experimento realizado na Suécia. . . . .	109



6.10	Erro da estimação da posição utilizando a abordagem não adaptativa aplicada ao conjunto de imagens do primeiro experimento realizado no México. . . . .	112
6.11	Erro da estimação da posição utilizando a abordagem adaptativa ideal aplicada ao conjunto de imagens do primeiro experimento realizado no México. . . . .	113
6.12	Erro da estimação da posição utilizando a abordagem adaptativa por regiões aplicada ao conjunto de imagens do primeiro experimento realizado no México. . . . .	114
6.13	Exemplos de imagens dos subconjuntos referente ao primeiro experimento realizado no México. . . . .	115
6.14	Erro da estimação da posição utilizando a abordagem não adaptativa aplicada ao conjunto de imagens do segundo experimento realizado no México. . . . .	117
6.15	Erro da estimação da posição utilizando a abordagem adaptativa ideal aplicada ao conjunto de imagens do segundo experimento realizado no México. . . . .	118
6.16	Erro da estimação da posição utilizando a abordagem adaptativa por regiões aplicada ao conjunto de imagens do segundo experimento realizado no México. . . . .	119
6.17	Exemplos de imagens dos subconjuntos referente ao segundo experimento realizado no México. . . . .	120
6.18	Erro médio das três abordagens de cada um dos conjuntos de dados analisados. . . . .	124
6.19	Taxa de boa correspondência das três abordagens de cada um dos conjuntos de dados analisados. . . . .	124
6.20	Comportamento do desempenho e acurácia em relação ao aumento do número de subconjunto utilizados. . . . .	125
B.1	Estrutura de uma imagem RGB em suas bandas, vermelho, verde e azul.	144
B.2	Modelo simplificado de uma câmera digital em duas dimensões. . . . .	146
D.1	Vizinhos mais próximos para uma distância de 1 elemento gráfico. . . . .	151
D.2	Exemplo do cálculo da matriz de cocorrência para direções de 0° e 90°. . . . .	152



## LISTA DE TABELAS

	<u>Pág.</u>
2.1	Questões de pesquisa da RSL e seus respectivos objetivos. . . . . 14
2.2	Total de artigos selecionados para cada base de dados. . . . . 15
2.3	Trabalhos de estimação da posição de VANTs por meio de imagens a partir do reconhecimento de regiões. . . . . 30
3.1	Parâmetros de entrada necessário para estimação da posição por imagens. 35
4.1	Texturas de Haralick invariantes a tons de cinza e equações correspondentes. . . . . 80
5.1	Número de experimentos realizados e suas respectivas localizações. . . . . 87
6.1	Configurações de entrada utilizados no treinamento das técnicas de aprendizado de máquina. . . . . 98
6.2	Acurácia das técnicas de aprendizado de máquina por região referente ao experimento realizado no Brasil. . . . . 104
6.3	Resultados de EM e TBC ao considerar a taxa de acerto dos classificadores do conjunto de imagens do experimento realizado no Brasil. . . . . 104
6.4	Acurácia das técnicas de aprendizado de máquina por região referente ao experimento realizado na Suécia. . . . . 110
6.5	Resultados de EM e TBC ao considerar a taxa de acerto dos classificadores do conjunto de imagens do experimento realizado na Suécia. . . . . 110
6.6	Acurácia das técnicas de aprendizado de máquina por região referente ao primeiro experimento realizado no México. . . . . 116
6.7	Resultados de EM e TBC ao considerar a taxa de acerto dos classificadores do conjunto de imagens do primeiro experimento realizado no México. . . . . 116
6.8	Acurácia das técnicas de aprendizado de máquina por região referente ao segundo experimento realizado no México. . . . . 121
6.9	Resultados de EM e TBC ao considerar a taxa de acerto dos classificadores do conjunto de imagens do segundo experimento realizado no México. . . . . 121
6.10	Acurácia de classificação das técnicas de aprendizado de máquina considerando todos os conjuntos de imagens. . . . . 123
E.1	Exemplo de planilha contendo as características das imagens. . . . . 153

F.1	Matriz de confusão em porcentagem da árvore de classificação. . . . .	155
F.2	Matriz de confusão em porcentagem da SVM. . . . .	156
F.3	Matriz de confusão em porcentagem da MLP. . . . .	156
F.4	Composição das combinações de algoritmos de cada subconjunto de imagens. . . . .	157

## LISTA DE ABREVIATURAS E SIGLAS

ANAC	–	Agência Nacional de Aviação Civil
ANATEL	–	Agência Nacional de Telecomunicações
ASIC	–	<i>Application Specific Integrated Circuits</i>
CC	–	Correlação Cruzada
CCD	–	<i>Charged Coupled Devices</i>
CCN	–	Correlação Cruzada Normalizada
CCNZ	–	Correlação Cruzada Normalizada de Média Zero
DECEA	–	Departamento de Controle do Espaço Aéreo
EM	–	Erro Médio
FOV	–	<i>Field of View</i>
FPGA	–	<i>Field Programmable Gate Array</i>
GLCM	–	<i>Gray-Level Co-occurrence Matrix</i>
GNSS	–	<i>Global Navigation Satellite System</i>
GPS	–	<i>Global Positioning System</i>
GSD	–	<i>Ground Sample Distance</i>
IFOV	–	<i>Instantaneous Field of View</i>
IMU	–	<i>Inertial Measurement Unit</i>
INS	–	<i>Inertial Navigation System</i>
KLT	–	Kanade-Lucas-Tomasi
MLP	–	<i>MultiLayer Perceptron</i>
PITER	–	Processamento de Imagens em Tempo Real
PMD	–	Peso Máximo de Decolagem
RAM	–	<i>Random Access Memory</i>
RBAC-E	–	Regulamento Brasileiro da Aviação Civil Especial
RGB	–	<i>Red, Green, Blue</i>
RNA	–	Redes Neural Artificial
RPA	–	Aeronave Remotamente Pilotada
RPAS	–	Sistema de Aeronave Remotamente Pilotada
RPS	–	Estação de Pilotagem Remota
RSL	–	Revisão Sistemática da Literatura
SIFT	–	<i>Scale-Invariant Feature Transform</i>
SLAM	–	<i>Simultaneous Localization and Mapping</i>
SRD	–	Sistema Reconfigurável Dinamicamente
SURF	–	<i>Speeded-Up Robust Features</i>
SVM	–	<i>Support Vector Machines</i>
TBC	–	Taxa de Boa Correspondência
VANT	–	Veículo Aéreo Não Tripulado



## SUMÁRIO

	<u>Pág.</u>
<b>1 INTRODUÇÃO</b> . . . . .	<b>1</b>
1.1 Motivação . . . . .	4
1.2 Objetivos e hipótese . . . . .	7
1.3 Processo metodológico . . . . .	8
1.4 Contribuições e limitações . . . . .	10
1.5 Organização do texto . . . . .	11
<b>2 SISTEMAS RECONFIGURÁVEIS E VEÍCULOS AÉREOS NÃO TRIPULADOS</b> . . . . .	<b>13</b>
2.1 RSL sobre SRDs . . . . .	13
2.1.1 Protocolo de Pesquisa . . . . .	13
2.1.2 Resultados e análise . . . . .	15
2.1.3 Descobertas e Agenda de Pesquisa . . . . .	20
2.2 Veículos Aéreos Não Tripulados . . . . .	22
2.2.1 Regulamentação no Brasil . . . . .	23
2.2.2 Nomenclatura e definição . . . . .	24
2.2.3 Navegação e estimação da posição . . . . .	26
2.2.3.1 Trabalhos relacionados . . . . .	30
2.3 Considerações finais sobre este capítulo . . . . .	34
<b>3 REGISTRO DE IMAGENS E APRENDIZADO DE MÁQUINA</b> <b>35</b>	
3.1 Conversão para escala de cinza . . . . .	38
3.2 Transformações geométricas . . . . .	39
3.2.1 Ajuste de escala . . . . .	40
3.2.2 Ajuste de rotação . . . . .	41
3.3 Filtros no Domínio do Espaço . . . . .	42
3.3.1 Filtro da Média . . . . .	44
3.3.2 Filtro de Gauss . . . . .	45
3.3.3 Filtro da Mediana . . . . .	46
3.4 Detecção de Bordas . . . . .	47
3.4.1 Operador de Sobel . . . . .	48
3.4.2 Operador de Canny . . . . .	49
3.4.3 Redes Neurais Artificiais . . . . .	51

3.4.4	Congruência de Fase . . . . .	51
3.5	Registro de imagens . . . . .	53
3.5.1	Correlação Cruzada . . . . .	54
3.5.2	Correlação Cruzada Normalizada . . . . .	56
3.5.3	O uso de máscaras . . . . .	57
3.5.4	Correlação Cruzada com Máscara . . . . .	58
3.6	Cálculo da posição . . . . .	60
3.7	Aprendizado de máquina . . . . .	61
3.7.1	Árvore de classificação . . . . .	62
3.7.2	Máquinas de Vetor de Suporte . . . . .	63
3.7.3	Redes Neurais Artificiais . . . . .	67
3.7.3.1	Perceptron de Múltiplas Camadas . . . . .	69
3.8	Considerações finais sobre este capítulo . . . . .	70
<b>4</b>	<b>ABORDAGEM AUTOADAPTATIVA PARA ESTIMAÇÃO DA POSIÇÃO POR IMAGENS . . . . .</b>	<b>71</b>
4.1	O problema da seleção de algoritmos . . . . .	72
4.2	Estratégia autoadaptativa para estimação da posição por imagens . . . . .	75
4.2.1	Espaço de problemas . . . . .	76
4.2.2	Espaço de características . . . . .	78
4.2.3	Espaço de algoritmos . . . . .	79
4.2.4	Espaço de desempenhos . . . . .	81
4.2.5	Aprendizado de Máquina como mapa de seleção . . . . .	83
4.3	Considerações finais sobre este capítulo . . . . .	84
<b>5</b>	<b>AVALIAÇÕES EXPERIMENTAIS . . . . .</b>	<b>87</b>
5.1	Experimento realizado no Brasil . . . . .	87
5.2	Experimento realizado na Suécia . . . . .	89
5.3	Experimentos realizados no México . . . . .	90
5.4	Considerações finais sobre este capítulo . . . . .	94
<b>6</b>	<b>RESULTADOS . . . . .</b>	<b>95</b>
6.1	Resultados do experimento realizado no Brasil . . . . .	99
6.2	Resultados do experimento realizado na Suécia . . . . .	105
6.3	Resultados dos experimentos realizados no México . . . . .	111
6.3.1	Primeiro experimento . . . . .	112
6.3.2	Segundo experimento . . . . .	117
6.4	Análise dos resultados . . . . .	122



6.5	Considerações finais sobre este capítulo . . . . .	126
<b>7</b>	<b>CONCLUSÕES . . . . .</b>	<b>127</b>
7.1	Trabalho futuros . . . . .	128
	<b>REFERÊNCIAS BIBLIOGRÁFICAS . . . . .</b>	<b>131</b>
	<b>APÊNDICE A - <i>STRING</i> DE BUSCA DA RSL . . . . .</b>	<b>141</b>
	<b>APÊNDICE B - ASPECTOS BÁSICOS DA IMAGEM DIGITAL. . . . .</b>	<b>143</b>
	<b>APÊNDICE C - CÁLCULO DA CORRELAÇÃO CRUZADA NOR-</b>	
	<b>MALIZADA COM MÁSCARA . . . . .</b>	<b>147</b>
	<b>APÊNDICE D - CÁLCULO DA MATRIZ DE COCORRÊNCIA . . . . .</b>	<b>151</b>
	<b>APÊNDICE E - ESTRUTURA DAS PLANILHAS DE TREINA-</b>	
	<b>MENTO . . . . .</b>	<b>153</b>
	<b>APÊNDICE F - MATRIZ DE CONFUSÃO DAS TÉCNICAS DE</b>	
	<b>APRENDIZADO DE MÁQUINA . . . . .</b>	<b>155</b>



# 1 INTRODUÇÃO

Ambientes operacionais dinâmicos apresentam características que variam constantemente e, por esse motivo, se tornam difíceis de serem controlados. Organismos imersos nesses ambientes podem ter a adaptação como uma característica imprescindível para sua sobrevivência, i.e., a capacidade de um organismo alterar seu próprio comportamento pode se revelar crucial para que o mesmo modifique sua forma de agir ante adversidades. Em outras palavras, toda vez que um ambiente muda, o organismo nele presente deve ser capaz de se adaptar a tais mudanças.

A necessidade de adaptação é uma realidade cada vez mais necessária no projeto de novos sistemas (KRUPITZER et al., 2015). É desejável que o próprio sistema seja capaz de se ajustar por meio do gerenciamento de seus recursos ou, ao menos, de possuir a capacidade de ser ajustado. A adaptação é um fator que exige a capacidade de mudança nas configurações de um sistema, i.e., a adaptação possui como requisito a reconfiguração. Para que um sistema seja adaptativo e possa alcançar o controle do seu gerenciamento, ele precisa antes ser reconfigurável.

A reconfigurabilidade de um sistema está associada ao quanto as características de um sistema podem mudar, i.e., o quão flexível suas configurações são. Qualquer sistema que possua habilidade para mudar de estado de maneira reversível é considerado um sistema reconfigurável (LYKE et al., 2015). Nesse sentido, sistemas digitais (e.g., computadores pessoais), podem ser considerados sistemas reconfiguráveis, pois apresentam a capacidade de retornar ao estado original. Da mesma forma, um cartão de Memória de Acesso Aleatório (*Random Access Memory* - RAM), cujos *bits* mudam de estado para armazenar diferentes tipos de informação, também se caracteriza como um sistema reconfigurável. Sistemas desse tipo podem apresentar diferentes níveis de complexidade. Um sistema reconfigurável pode ser simples como um relógio digital ou complexo como uma rede composta por centenas de computadores, em que a interconexão multiplica a severidade do sistema como um todo. Além disso, tais sistemas podem ter associados conhecimentos que envolvem diferentes disciplinas e técnicas, tais como matéria reconfigurável e sistemas eletrônicos digitais e analógicos (LYKE et al., 2015).

Sistemas não reconfiguráveis, por sua vez, não são necessariamente inferiores em todos os aspectos quando comparados a sistemas reconfiguráveis. Estes últimos podem apresentar perda de desempenho por conta de sobrecarga no controle de subsistemas, aumento da complexidade do projeto e redução da confiabilidade (LYKE et al., 2015). Além disso, o simples fato de um sistema ser capaz de modificar-se não o

torna mais apto a executar uma tarefa, nem tampouco é garantia de obter maior eficiência. Nesse contexto, os usos que um sistema pode fazer por meio de suas possíveis configurações está diretamente associado à sua capacidade de adaptação.

Dessa forma, a adaptabilidade pode ser entendida como a habilidade que um sistema possui de explorar e gerenciar seus domínios de configurações e, assim, usá-los com um determinado objetivo (SALEHIE; TAHVILDARI, 2009). Embora reconfigurabilidade e adaptabilidade sejam termos com significados similares, existe uma diferença entre eles. Enquanto a reconfigurabilidade está relacionada à flexibilidade de um determinado sistema, a adaptabilidade está relacionada aos usos que podem ser feitos dessa flexibilidade.

Quando um sistema possui a capacidade de modificar-se em resposta a um ambiente operacional, ele é denominado autoadaptativo (KRUPITZER et al., 2015). Sendo, aqui, ambiente operacional definido como qualquer objeto observável pelo sistema (SALEHIE; TAHVILDARI, 2009). Sistemas autoadaptativos estão relacionados a características de autogerenciamento, tais como autoconfiguração, autocura, auto-otimização e autoproteção, i.e., a área de sistemas adaptativos estuda características que um sistema pode ou não apresentar. Por exemplo, um roteador pode otimizar a transferência de pacotes modificando suas configurações internas (auto-otimização), ou uma técnica de *backup* presente em um disco rígido de um computador pode recuperar arquivos após algum tipo de falha (autocura).

Esses sistemas podem apresentar vantagens no que se refere ao conforto do usuário e à sua eficiência operacional. Sistemas autoadaptativos apresentam características cada vez mais desejadas, uma vez que são capazes de ajustarem suas configurações de maneira automatizada, i.e., são capazes não apenas de executar uma tarefa, mas ainda de modificar seus parâmetros internos enquanto a executam (KRUPITZER et al., 2015). Um computador capaz de entrar em modo de economia de energia quando se encontra ocioso é mais adequado que um computador que não possui essa característica. Assim como uma aeronave híbrida capaz de modificar a estrutura de suas asas e rotores para executar voos verticais ou voos horizontais onde for mais conveniente (HEREDIA et al., 2012; ABNOUS et al., 2017).

Sistemas autoadaptativos podem ser comparados a sistemas autônomos, considerando autoadaptação como uma forma de prover autonomia ao sistema. Porém, enquanto o primeiro está relacionado às características que um sistema dispõe, o segundo está associado ao nível de dependência de operadores humanos para funcionar corretamente (KRUPITZER et al., 2015; SAE International, 2018). A autonomia pode

então ser entendida como uma escala, em que o nível mais baixo é composto por sistemas completamente dependentes de seres humanos e o nível mais alto composto por sistemas completamente independentes.

Uma classificação de sistemas dependentes/independentes pode exprimir na realidade as tomadas de decisões impostas em ambientes específicos. Por exemplo, o padrão J3016, publicado pela *SAE International*, propõe uma escala com 6 níveis de autonomia para carros *on-road*, sendo estes ([SAE International, 2018](#)):

- Nível 0: o motorista é completamente responsável por gerenciar todos os subsistemas do veículo, como o acelerador, a direção e o freio. O sistema pode apenas auxiliar o motorista com avisos e assistências momentâneas.
- Nível 1: o sistema auxilia o motorista com o gerenciamento de um dos subsistemas do carro, como o acionamento do freio ou acelerador, sempre que necessário.
- Nível 2: o sistema de veículo é capaz de auxiliar o motorista com mais de um subsistema ao mesmo tempo, como acionar os freios e guiar a direção do carro para que este não saia da pista.
- Nível 3: o sistema, sob condições específicas, é capaz de gerenciar todos os seus subsistemas sem precisar de um condutor humano. Porém, sempre que as condições não são satisfeitas, o controle é transferido para o motorista, o que o torna condição necessária para operar o veículo.
- Nível 4: o sistema é capaz de operar em determinados ambientes sem depender de um condutor, i.e., o controle do veículo não é mais transferido ao motorista. Neste caso, todos os instrumentos controlados pelo motorista, como o volante e os pedais, se tornam desnecessários.
- Nível 5: não há mais restrições de ambientes para a operação do sistema. Neste nível, assim como no nível 4, o veículo só é guiado pelo motorista caso este tome a decisão de fazê-lo.

Sistemas autônomos podem apresentar inúmeras vantagens em relação à confiabilidade da execução das tarefas ou mesmo em tomadas de decisão em tempo hábil. Sistemas dessa natureza podem integrar a interface homem/máquina na busca por melhores condições de segurança, economia e bem estar. Um sistema com alto nível de autonomia é desejável não apenas para carros, mas para uma ampla gama

de aplicações (CORKE, 2011; LYKE et al., 2015). Porém, o desenvolvimento de um sistema autônomo está longe de ser uma tarefa simples devido aos fatores que estão relacionados ao próprio sistema. Quando este apresenta dinamismo, as variáveis envolvidas se somam e aumentam a complexidade do controle, exigindo uma maior capacidade do sistema como um todo.

De uma maneira geral, sistemas autônomos móveis, sejam eles terrestres, aquáticos ou aéreos, devem ser capazes de percorrer uma trajetória para executar sua missão (CORKE, 2011). Assim, eles devem ser auxiliados por mecanismos periféricos de forma a receber informações específicas para a devida tomada de decisão. Para isso, tais sistemas devem ser capazes de, por meio de sensores projetados para estes fins, navegar em segurança por ambientes desconhecidos.

Os Veículos Aéreos Não Tripulados (VANTs) têm sido amplamente estudados nos últimos anos, considerando a difusão de suas aplicações em diferentes áreas. Também conhecidos como *drones*, os VANTs se beneficiam da ausência de um piloto a bordo, o que permite que sejam utilizados em ambientes de difícil acesso. Além disso, dispensam uma cabine de controle e a aparelhagem nela contida, tornando-se mais compactos e leves. Apesar de não serem uma invenção recente (NONAMI, 2007), foi apenas nas últimas décadas que os VANTs se tornaram populares e passaram a chamar atenção por seu potencial de uso em diversos tipos de aplicações (OTTO et al., 2018).

Em 2018, por exemplo, um VANT foi utilizado pela primeira vez para resgatar dois indivíduos em situação de afogamento na costa da Austrália (BRADY, 2018). E, em 2019, também pela primeira vez na história, um VANT foi utilizado no transporte de órgãos entre dois hospitais (WAKEFIELD, 2019). As aplicações ainda se estendem a outras áreas, como monitoramento de infraestruturas físicas, como torres de eletricidade, estradas, oleodutos e gasodutos, monitoramento do uso da terra e transporte de mercadorias (OTTO et al., 2018).

## 1.1 Motivação

Embora VANTs possam ser operados remotamente por rádio controle, é desejável que sejam autônomos, ou que possuam um alto grau de autonomia. Para isso, seu sistema interno deve ser capaz de realizar uma série de tarefas em prol de um determinado objetivo, i.e., um VANT com alto grau de autonomia depende de uma série de subsistemas que o auxiliam a atingir um objetivo específico. O VANT deve ser capaz, por exemplo, de alçar voo, pousar, avaliar seu nível de combustível, desviar

de obstáculos, realizar manobras e planejar rotas (KANELLAKIS; NIKOLAKOPOULOS, 2017).

Além de planejar rotas de maneira integrada com os elementos do sistema, estruturando o ambiente operacional, um VANT deve ser capaz de percorrê-las de forma estável e segura. A segurança do trajeto é então dependente do planejamento inicial e das tomadas de decisão que se fazem necessárias durante o percurso. Uma condição necessária para percorrer a rota é conhecer a posição do veículo enquanto navega pelo ambiente. Em outras palavras, para que um VANT se mova em direção a um alvo, é fundamental que este seja capaz de identificar sua atual posição. Dessa forma, a estimação da posição pode ser considerada um dos subsistemas de navegação do VANT, sendo responsável por manter sua posição atualizada durante o voo (GROVES, 2013). Nesse sentido, o cálculo da posição se apresenta como um problema fundamental a ser resolvido, de forma a possibilitar as tomadas de decisão e atuação no controle do sistema para que este possa prover autonomia ao veículo.

O reposicionamento do veículo envolve uma complexidade de cálculos de todo o sistema integrado, buscando solucionar, em curto espaço de tempo, a sua orientação no espaço. Em ambientes fechados, o controle pode se tornar relativamente mais simples, uma vez que pontos fixos de posicionamento podem auxiliar no controle do veículo. Entretanto, com o aumento das distâncias entre o veículo e o ponto de referência, a tendência é que este se encontre fora dos limites possíveis de controle, o que caracteriza um problema no deslocamento do veículo (JIANG et al., 2018).

Uma possível solução é a utilização de um *Inertial Navigation System* (INS). O INS é baseado em sensores inerciais, também chamados de *Inertial Measurement Unit* (IMU), responsáveis por calcular a variação de movimento de um corpo. Os dados providos pelos sensores são então processados e integrados para se obter a posição, velocidade e atitude do corpo. Como o erro proveniente dos sensores se acumula a cada integração, a qualidade destes está diretamente associada à precisão da posição estimada. Sistemas inerciais de alto custo, no valor de 100 mil dólares, possuem erro horizontal inferior a 1500 m na primeira hora. Sistemas inerciais de baixo custo possuem erro que pode chegar a ordens de magnitude maior e não são qualificados para atender a requisitos de diversas aplicações comerciais (GROVES, 2013). O sistema inercial, por si só, não consegue estimar a posição absoluta de um veículo, porém, é capaz de estimar a mudança na sua posição e, dessa forma, inferir sua posição a partir de uma referência conhecida (GROVES, 2013; TANG et al., 2018).

Pelo motivo exposto acima, VANTs de baixo custo não navegam utilizando unica-

mente o INS, mas uma combinação deste com um sistema baseado em satélites. A utilização de dados de Sistemas de Navegação Global por Satélite (do inglês, *Global Navigation Satellite System* - GNSS), tal como o *Global Positioning System* (GPS), combinado com o INS, é assim uma outra possibilidade. Os sistemas GNSS fazem uso de um conjunto de sinais provenientes de satélites artificiais para calcular a posição de um objeto na superfície terrestre em relação a um sistema cartográfico (GAO et al., 2018; GROVES, 2013). Porém, sistemas de navegação baseados em sinais de rádio estão condicionados à comunicação entre fonte e receptor de sinais; sempre que a comunicação falha, o sistema também apresentará falha. Por esse motivo, sistemas de estimação da posição por satélites estão sujeitos à interferência na comunicação, que pode ser causada por fenômenos naturais, como eventos ionosféricos, ou artificiais, cujo objetivo é deliberadamente bloquear ou manipular o sinal (GROVES, 2013).

Uma terceira solução possível é a utilização de características físicas do terreno ou do ambiente para estimar o movimento e a posição de um veículo. As condições físicas do espaço podem então ser utilizadas para apoiar e estabelecer condições favoráveis ao seu posicionamento e redirecionamento. Diversas características podem ser utilizadas para se inferir a posição de um corpo, tais como a morfologia do terreno, presença de obstáculos, pressão do ar, além dos campos geomagnético e gravitacional. Quando essas características são mensuradas, processadas e comparadas a informações conhecidas ou pré-estabelecidas, elas podem ser úteis no auxílio do balizamento do veículo. Dessa forma, quando ocorre uma condição de orientação a partir de referências conhecidas, pode-se obter a posição absoluta de um corpo. Por outro lado, quando esse conjunto de características é comparado entre si pode-se obter sua movimentação, i.e., sua posição relativa (GROVES, 2013; ŠAŠAK et al., 2019).

As características presentes no ambiente podem ser detectadas por meio de diferentes tipos de sensores. Dependendo da característica das informações, cada tipo de sensor pode ser mais ou menos adequado ao propósito ao qual está destinado. Recentemente, o uso de sensores imageadores ópticos, aliado a técnicas de visão computacional, trouxe avanços relevantes na área de navegação de VANTs (AL-KAFF et al., 2018; CADENA et al., 2016; KANELLAKIS; NIKOLAKOPOULOS, 2017; SILVA, 2016; BRAGA, 2018). Nesse tipo de abordagem, o sistema deve ser capaz de processar as informações provenientes do mundo real traduzidas pelos sensores a fim de calcular os movimentos realizados pelo veículo.

Em contrapartida, o fato do veículo se movimentar imerso em um ambiente dinâ-



mico, com situações que ocorrem espontânea e aleatoriamente, aumenta a dificuldade de se projetar um sistema baseado em imagens. Por exemplo, um sensor ao detectar o movimento aparente da superfície da água pode levar ao cálculo de uma velocidade superior à velocidade real do veículo. Ou a intensidade da luminosidade natural, que varia conforme a hora do dia, poderia levar à incompatibilidade das imagens capturadas em diferentes horários. Além disso, os movimentos do VANT influenciam as características detectadas pelos sensores, pois estes podem observar o mesmo objeto de diferentes distâncias e ângulos, transmitindo ao sistema uma condição incompleta da realidade. Além disso, podem ocorrer mudanças na composição das imagens devido às características das regiões de sobrevoo, que podem envolver, por exemplo, um conjunto urbano ou um espaço florestal.

Uma primeira motivação para este trabalho é, portanto, o fato de que é importante realizar uma caracterização para as diferentes situações ou cenários que um VANT pode enfrentar, como descrito anteriormente. Como este trabalho se propõe a usar imagens para estimar a posição do VANT, é interessante que diversos ambientes (e.g., diferentes coberturas de solo) sejam analisados de forma a se utilizar técnicas de processamento digital de imagens adequadas em cada contexto. Apesar da quantidade significativa de trabalhos na área de navegação autônoma de VANTs baseada em imagens (ABDI et al., 2016; BRAGA, 2018; CONTE; DOHERTY, 2009; SILVA, 2016; STEPANOV; TISHCHENKO, 2017), percebeu-se a necessidade de lidar com imagens que representem uma maior variabilidade de cenários que um VANT pode encontrar.

Uma vez que se tem uma caracterização adequada, pode-se pensar na possibilidade da exploração da autoadaptação para estimar a posição de VANTs por imagens (FORNARI; SANTIAGO JÚNIOR, 2019). Em outras palavras, o sistema teria a inteligência adequada para selecionar, de um conjunto de combinações de técnicas de processamento digital de imagens, a mais apropriada para um determinado cenário. A adaptação pode-se provar uma abordagem útil para sistemas baseados em imagens uma vez que a mudança do ambiente durante o voo é uma característica inerente ao problema.

## 1.2 Objetivos e hipótese

O objetivo geral do presente trabalho é contribuir para a melhoria do processo de estimação da posição de VANTs por imagens pela apresentação de um método autoadaptativo. Tal método pode selecionar, de forma inteligente, os melhores algoritmos de processamento digital de imagens para esse propósito. Esse objetivo geral pode

ser subdivido nos seguintes objetivos específicos:

- Caracterizar cenários. Esse objetivo específico se relaciona à caracterização do desempenho de conjuntos de técnicas de processamento digital de imagens de acordo com diferentes cenários e situações de voo. Desse modo, é possível determinar quais técnicas são mais adequadas para uma certa condição de voo;
- Prover autoadaptação para escolha dos melhores algoritmos. Nesse caso, um sistema inteligente deve ser desenvolvido de forma que, baseando-se na caracterização dos cenários, possa selecionar, de forma automática, os melhores algoritmos para propiciar um melhor desempenho do VANT no que tange à estimação da sua posição.

A hipótese que se deseja validar nessa pesquisa pode ser formulada da seguinte forma.

Uma abordagem autoadaptativa para seleção automática de técnicas de processamento digital de imagens é melhor, em termos da estimativa da posição do VANT, do que uma abordagem não adaptativa.

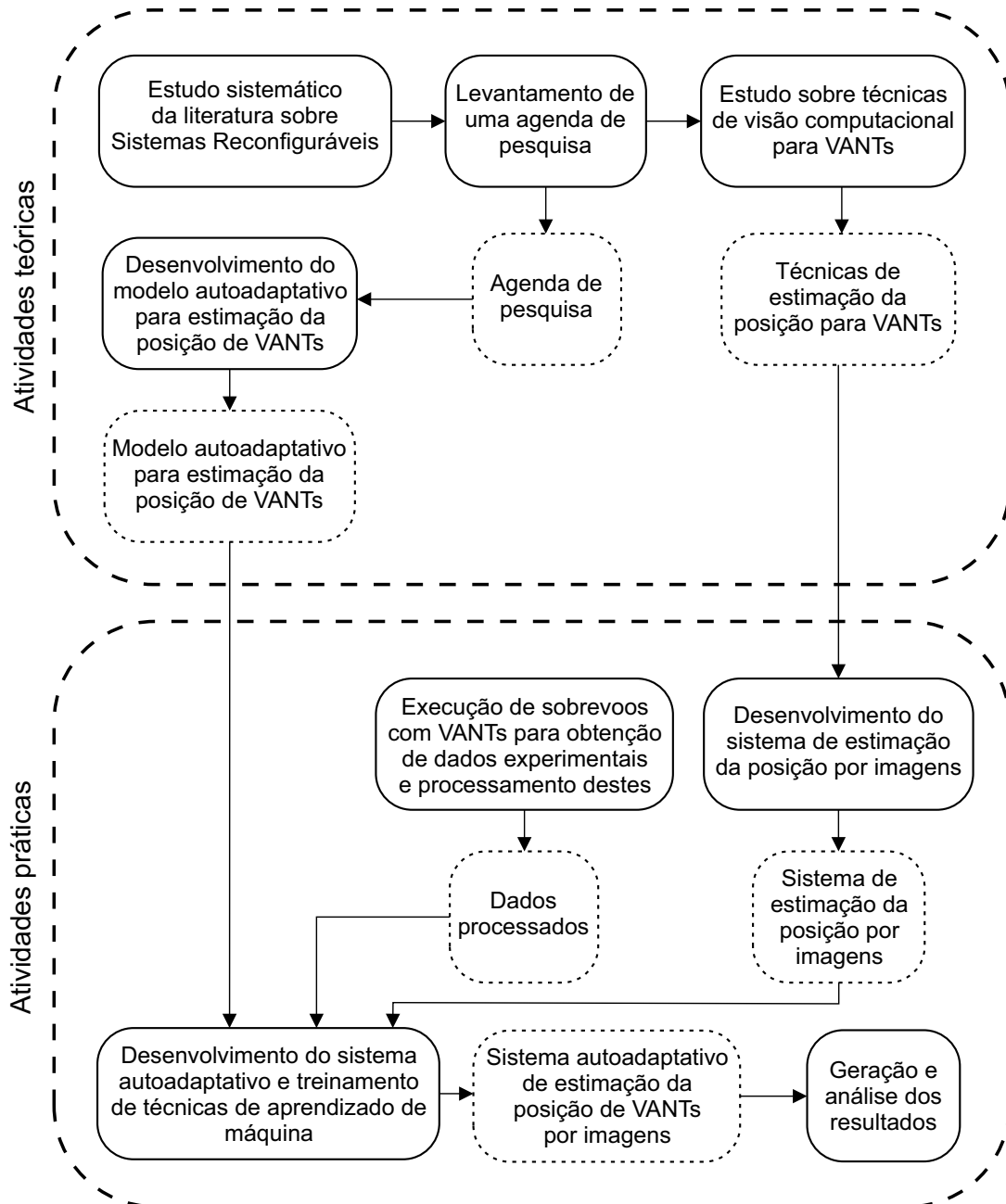
Ao final deste trabalho, a hipótese acima poderá ser aceita, o que justificaria o desenvolvimento da abordagem autoadaptativa, ou rejeitada.

### **1.3 Processo metodológico**

Para alcançar os objetivos propostos neste trabalho, uma série de tarefas foram desenvolvidas. Na Figura 1.1 podem ser vistas as principais atividades aqui produzidas, divididas entre teóricas e práticas. Dentre as tarefas teóricas, destacam-se a produção de uma agenda de pesquisa resultante de um estudo sistemático da literatura, e um estudo acerca das metodologias de estimação da posição de VANTs por imagens presentes na literatura científica. O primeiro estudo serviu de base para o desenvolvimento de um modelo teórico autoadaptativo para estimação da posição de VANTs. O segundo, por sua vez, resultou no desenvolvimento e implementação do sistema de estimação da posição por imagens. Este, somado ao modelo teórico autoadaptativo e também aos dados obtidos por meio de experimentos realizados com VANTs, resultaram no desenvolvimento do sistema autoadaptativo, do qual

fazem parte as técnicas de aprendizado de máquina. Tal sistema foi utilizado para gerar os resultados necessários para validar a hipótese proposta.

Figura 1.1 - Processo metodológico envolvendo as principais atividades desenvolvidas neste trabalho.



As tarefas desenvolvidas ao longo deste trabalho de doutorado podem ser divididas entre teóricas e práticas. Além disso, elas podem ser classificadas entre resultados (linha tracejada) e atividades (linha contínua).

Fonte: Produção do autor.

## 1.4 Contribuições e limitações

A contribuição que este trabalho propõe está apoiada em dois pontos principais: a necessidade de desenvolver técnicas de estimação da posição para VANTs alternativas às existentes; e, a possibilidade da exploração da autoadaptação para escolher as melhores técnicas de processamento digital de imagens, visando uma melhoria da estimação da posição de VANTs. O primeiro item está relacionado ao uso de sistemas de navegação/posicionamento já consolidados, como o INS e o GPS, enquanto o segundo está associado à característica que as imagens proporcionam de registrar em tempo real as mudanças ocorridas no ambiente.

Nesse contexto, a primeira contribuição deste trabalho se apresenta de forma mais relevante com a atuação em buscar soluções paralelas ou inovadoras, e se baseia no desenvolvimento de abordagens alternativas de sistemas de estimação da posição. Uma dessas possibilidades se encontra no uso das informações provenientes de sensores imageadores, em que a captura de imagens e o seu respectivo processamento apresentam a vantagem de revelar e transmitir as informações do ambiente em tempo real. Nesse tipo de abordagem, algoritmos de processamento de imagens são capazes de utilizar informações do ambiente providas por câmeras de vídeo para calcular a posição e movimento de um corpo. Dessa forma, é possível estimar a posição sem apresentar os mesmos problemas que o INS e o GPS apresentam.

Porém, apesar de apresentar vantagens, como baixo custo, baixo peso e baixo consumo de energia em alguns modelos, essa abordagem enfrenta outros problemas, que precisam buscar alternativas para o desenvolvimento de soluções adequadas. Um desses desafios está relacionado com o aumento da carga computacional e mudanças repentinas das características presentes no ambiente operacional (AL-KAFF *et al.*, 2018). Neste último, surge a segunda contribuição para o presente trabalho, que é aproveitar e explorar a mudança das características do ambiente para decidir a forma como o processo de estimação da posição será feito. Tal sistema deve possuir inteligência suficiente para modificar seu comportamento de forma a se adaptar da melhor forma possível ao ambiente.

Além disso, relacionado a essa segunda contribuição, está um processo de caracterização que levou em consideração um conjunto bastante significativo de combinações de técnicas de processamento digital de imagens e diferentes cenários, para identificar as melhores abordagens de acordo com diferentes situações de voo.

Os ambientes escolhidos para análise neste trabalho incluem regiões externas, em

que o VANT percorre uma trajetória entre 30 e 100 metros acima do solo. Ao todo, o conjunto de dados inclui voos com 3 *drones* diferentes e 5 trajetórias em localidades distintas, que apresentam características variadas, tais como a presença de ruas, edificações e outras coberturas do solo. Porém, uma limitação desse trabalho é o fato de que os sensores precisam de luz para operar e, desse modo, a abordagem aqui proposta não é adequada para voos noturnos.

Adicionalmente, há diferentes técnicas de visão computacional que podem ser utilizadas para extrair informações de imagens a fim de estimar a posição do VANT (SILVA, 2016; JAIMES, 2016; GALINDO, 2016; SILVA FILHO, 2016; KANELLAKIS; NIKOLAKOPOULOS, 2017; AL-KAFF et al., 2018). A técnica explorada neste trabalho é baseada no trabalho de Braga (2018), que utiliza a correlação para registro de imagens aliada a etapas de pré-processamento das mesmas. Dessa forma, o presente trabalho se utiliza da mesma técnica, mas busca ampliar o espaço de algoritmos utilizados em cada etapa de processamento. Porém, outras técnicas, tal como Odometria Visual, poderiam ter sido exploradas no trabalho.

Por último, pode-se dizer que os resultados obtidos no presente trabalho foram gerados por um processo de simulação, em que conjuntos de dados registrados por VANTs foram utilizados após a realização dos experimentos. Em outras palavras, não houve uma implementação da abordagem autoadaptativa de forma que a mesma pudesse ser avaliada, em tempo real, embarcada em VANTs. Essa outra limitação do trabalho.

## 1.5 Organização do texto

Esta tese de doutorado encontra-se organizada em 7 capítulos, incluindo este. O seguinte, Capítulo 2, é dedicado a apresentar um estudo sistemático da literatura sobre sistemas reconfiguráveis, assim como fundamentos sobre VANTs. Em seguida, no Capítulo 3, é apresentada a técnica de estimação da posição baseada em reconhecimento de regiões e registro de imagens bem como os algoritmos a serem empregados na técnica. No Capítulo 4, é apresentada a ideia na qual se fundamenta a proposta desta tese, que é a investigação de uma abordagem autoadaptativa na estimação da posição de VANTs por meio de imagens. No Capítulo 5 são apresentados e descritos os experimentos realizados com *drones* e os dados obtidos a partir destes. Os resultados, baseados nos dados obtidos por meio dos experimentos, são analisados e discutidos no Capítulo 6. Por último, no Capítulo 7 são apresentadas as considerações finais e as propostas para investigações futuras nesta linha de pesquisa.



## 2 SISTEMAS RECONFIGURÁVEIS E VEÍCULOS AÉREOS NÃO TRIPULADOS

O presente capítulo apresenta uma revisão de conceitos relacionados a VANTs. No entanto, primeiramente será apresentado um resumo de uma Revisão Sistemática da Literatura (RSL) desenvolvida no contexto deste trabalho, a qual objetivou obter uma visão do estado da arte de Sistemas Reconfiguráveis Dinamicamente (SRDs) (FORNARI; SANTIAGO JÚNIOR, 2019). Essa revisão foi importante para compreender os esforços científicos realizados na área de SRDs e, assim, orientar a presente pesquisa. É importante perceber que o tipo de sistema de estimação da posição de VANTs a que esta tese de doutorado se refere pode também ser considerado um SRD.

O presente capítulo encontra-se organizado da seguinte forma. Na primeira Seção, 2.1, é apresentada a RSL sobre SRDs. Em seguida, na Seção 2.2, são discutidos aspectos gerais dos VANTs, como nomenclatura e algumas das definições dadas a esse tipo especial de aeronaves (Subseção 2.2.2) e aspectos envolvendo a navegação e estimação da posição operada por tais veículos, com ênfase nas técnicas de visão computacional (Subseção 2.2.3). Por último, a Seção 2.3 apresenta as considerações finais sobre o capítulo.

### 2.1 RSL sobre SRDs

Uma RSL foi desenvolvida com o objetivo de investigar a área de SRDs, abordando 4 questões de pesquisa, como pode ser visto na Tabela 2.1.

#### 2.1.1 Protocolo de Pesquisa

Com base nas questões de pesquisa, um protocolo de pesquisa foi definido e diversas palavras-chave sobre o tema foram criadas para formar a base da *string* de busca. Via inspeção visual e dos limites mínimo e máximo dos resultados, a *string* de busca foi refinada até se chegar em sua versão final, que pode ser vista no Apêndice A. Apenas os metadados dos trabalhos, i.e., o título, o resumo e as palavras-chave foram considerados na busca. Cinco bases de dados foram consideradas: *Scopus*, *ACM*, *Web of Science*, *Science Direct* e *IEEE*.

Como parte do protocolo de revisão, os critérios de inclusão e exclusão foram definidos. Os critérios de inclusão envolvem artigos publicado em jornais, revistas, conferências, livros e *workshops*, enquanto notas editoriais e *white papers* foram excluídos. Além disso, artigos que não trouxessem nenhuma discussão sobre SRDs ou

Tabela 2.1 - Questões de pesquisa da RSL e seus respectivos objetivos.

Id	Questão de Pesquisa	Objetivo
QP1	Qual é o conceito e a classificação dada aos SRDs na literatura científica?	Entender qual é a definição de SRDs e quais as características que eles podem apresentar e verificar se existe alguma taxonomia de classificação desses sistemas.
QP2	Quais metodologias, técnicas, infraestruturas de sistemas computacionais e estratégias são usadas no contexto de SRDs?	Compreender como os SRDs podem ser desenvolvidos na prática, i.e., quais os meios que existem para se criar um sistema reconfigurável dinamicamente.
QP3	Quais os domínios de aplicação que os SRDs são usados?	Compreender em que situações os SRDs são usados, na prática, e com que finalidade.
QP4	Quais países lideram o número de publicações em SRDs?	Delinear o número de artigos publicados por país associando os seus autores às suas respectivas instituições.

não apresentassem um caso de uso para SRDs não seriam aceitos. Apenas artigos escritos na língua inglesa e publicados entre 1995 e 2016 foram considerados.

A quantidade de artigos selecionados seguindo o protocolo apresentado, sucintamente, acima pode ser vista na Tabela 2.2. Na primeira fase, um total de 616 artigos foram retornados das 5 bases de dados utilizando as *strings* de busca. A partir desse conjunto primário de artigos, os critérios de inclusão e exclusão foram aplicados. Cerca de 18,3% (113) dos artigos eram duplicados e 35,7% (220) eram editoriais ou não estavam relacionados a SRDs. Um total de 72 artigos foi escolhido por atender diretamente aos critérios de inclusão, do qual 4 apresentaram conteúdo e autores semelhantes e apenas a versão mais completa foi mantida, e 3 apresentaram conteúdo fora do escopo desse trabalho e também foram excluídos. Dessa maneira, 72 artigos foram lidos, mas apenas 65 foram selecionados para extrair informações.

Além disso, outros 5 trabalhos que não foram obtidos pela *string* de busca foram incluídos na leitura. Desses 5 trabalhos, apenas 1 foi removido por não atender aos critérios de inclusão. No total, 65 artigos foram obtidos via a *string* de busca e outros 4 foram completamente lidos, totalizando 69.



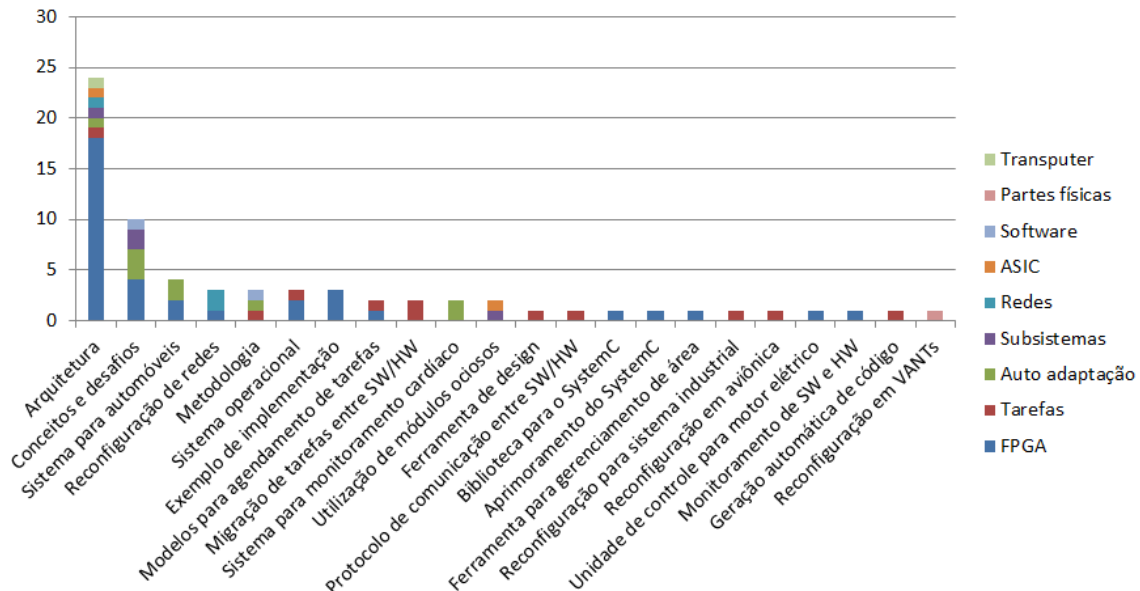
Tabela 2.2 - Total de artigos selecionados para cada base de dados.

Fonte	1º Passo	2º Passo	3º Passo
Scopus	273	24	24
ACM	96	15	11
Web of Science	38	8	6
Science Direct	147	17	17
IEEE	62	8	7
<b>Total</b>	<b>616</b>	<b>72</b>	<b>65</b>

### 2.1.2 Resultados e análise

**QP1.** A primeira questão de pesquisa foi definida como “Qual é o conceito e a classificação dada aos SRDs na literatura científica?”. O principal objetivo dessa questão é entender o que a literatura identifica como SRDs. Para isso, o tipo de reconfiguração e a área do conhecimento foram extraídos dos artigos, como pode ser visto na Figura 2.1.

Figura 2.1 - Resultados da classificação dos artigos segundo a propriedade “área do conhecimento”.



Fonte: Produção do autor.

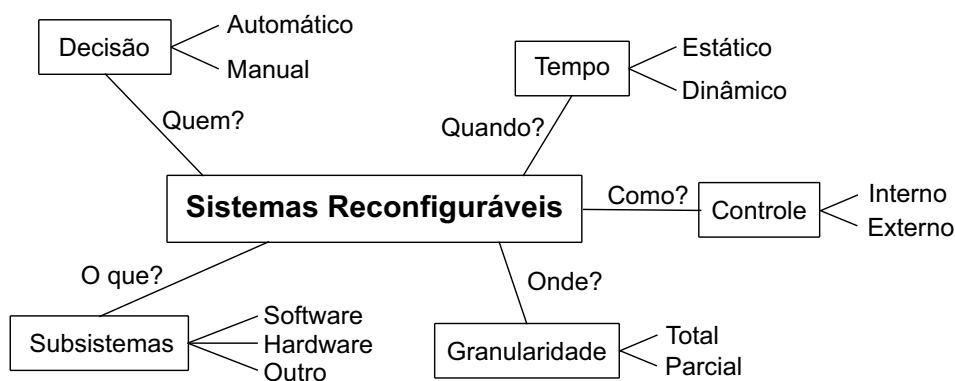
Como pode ser visto, mais da metade dos trabalhos (36) apresenta sistemas reconfiguráveis como FPGAs (*Field Programmable Gate Array*), i.e., a discussão presente

nesses artigos associa a reconfiguração de um sistema à reconfiguração das unidades lógicas de uma FPGA. Já a classe “autoadaptação” está em terceiro lugar, com 9 artigos, incluindo trabalhos teóricos que definem conceitos sobre sistemas autoadaptativos e também trabalhos práticos que buscam aumentar a flexibilidade de sistemas. Por exemplo, [Krupitzer et al. \(2015\)](#) apresentam uma taxonomia sobre sistemas autoadaptativos e seus conceitos. Já os trabalhos com aplicações práticas incluem adaptação para automóveis e monitoramento cardíaco ([WAGNER et al., 2014](#); [ZONG et al., 2015](#)).

Além disso, [Chowdhury et al. \(2014\)](#) apresentam um VANT desmontável que pode ser adaptado para se tornar um VANT de asa fixa ou com rotores. Nesse caso, a reconfiguração necessita de uma pessoa para montar ou desmontar suas peças, i.e., a decisão e a forma de reconfiguração são, necessariamente, manuais. Apesar de um único trabalho ter apresentado tais características, ele é importante para mostrar que os outros trabalhos apresentam sistemas automáticos, o que significa que existe uma unidade de controle responsável não apenas pela reconfiguração, mas também pela tomada de decisão da mesma.

Baseado nas informações extraídas dos trabalhos, é proposta uma nova classificação para sistemas reconfiguráveis, apresentada na Figura 2.2. A classificação leva em conta as características discutidas anteriormente e a taxonomia apresentada em [Krupitzer et al. \(2015\)](#).

Figura 2.2 - Classificação proposta para sistemas reconfiguráveis.



Fonte: Produção do autor.

Além disso, é proposta uma nova definição para SRDs, que é bem próxima da

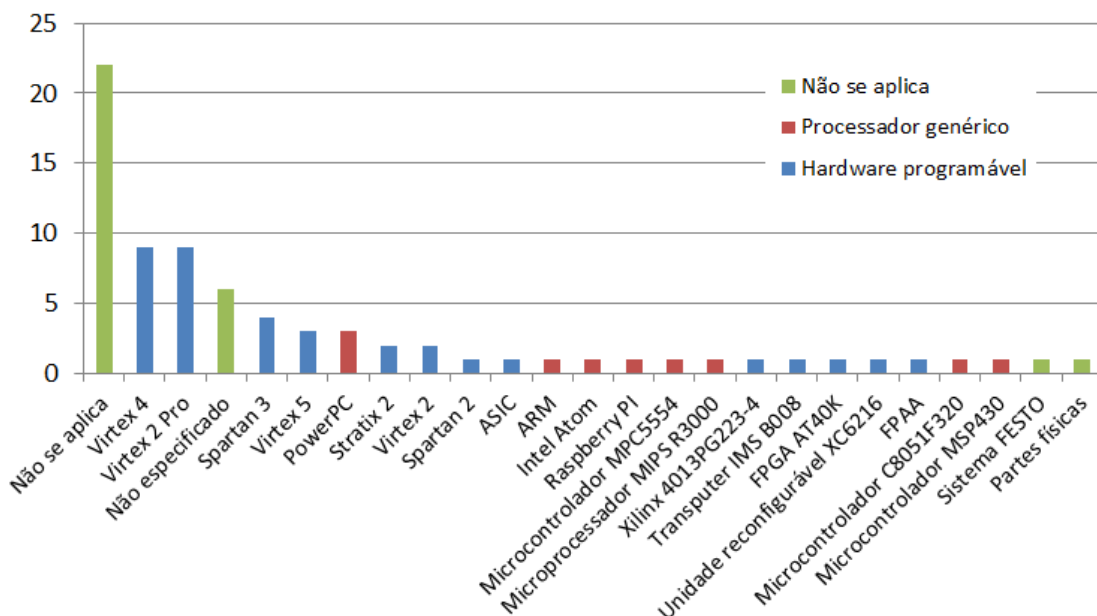
apresentada por Józwiak e Nedjah (2009):

**Definição 1.** Um SRD é um sistema cujos subsistemas podem ser modificados ou ter suas configurações alteradas durante seu funcionamento para atingir um objetivo específico.

É importante salientar que essa definição é geral e não especifica do quê um sistema precisa ser constituído, nem como sua reconfiguração é dada, nem quantas vezes ele pode ser reconfigurado.

**QP2.** A segunda questão de pesquisa é “Quais metodologias, técnicas, infraestruturas de sistemas computacionais e estratégias são usadas no contexto de SRDs?”. O foco dessa questão é entender quais os dispositivos utilizados em SRDs e como esses são reconfigurados. Para isso, foi extraído dos artigos, com exceção dos trabalhos teóricos, o dispositivo engajado nos testes e a estratégia de reconfiguração. Os resultados podem ser vistos na Figura 2.3.

Figura 2.3 - Resultados da classificação dos artigos segundo a propriedade “dispositivo utilizado”.

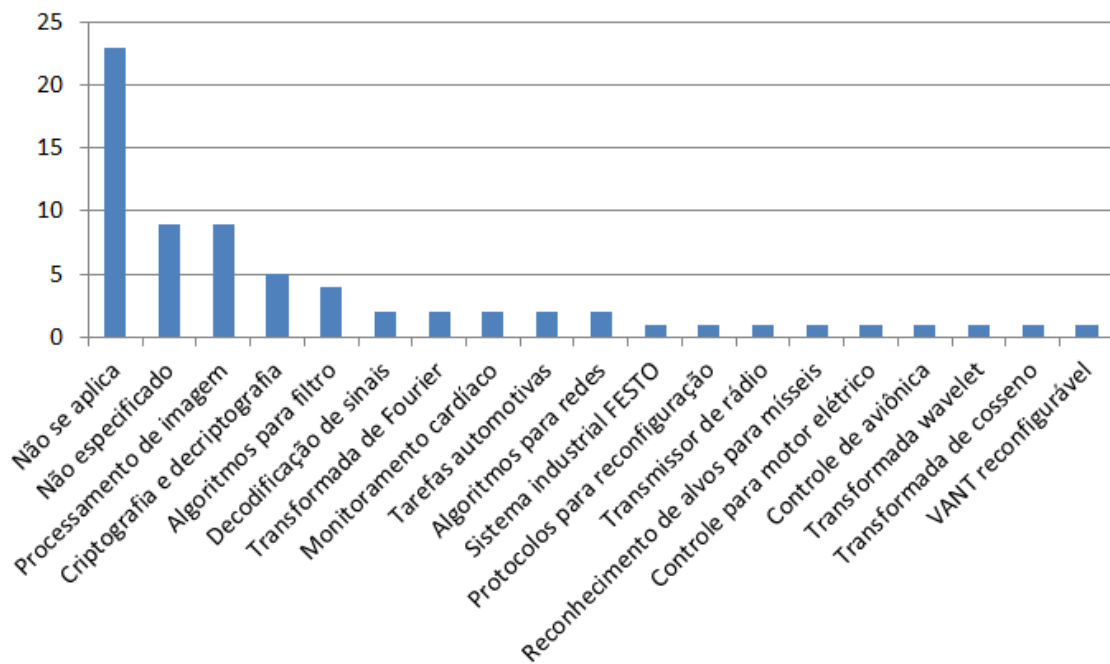


Fonte: Produção do autor.

Quanto ao dispositivo utilizado, as FPGAs somam a maioria, como foi discutido anteriormente. Quanto à estratégia de reconfiguração, pode-se afirmar que todo SRD é baseado em uma unidade de controle, mesmo que esse seja uma pessoa, responsável por tomar a decisão de reconfigurar o sistema. Além disso, o conceito de reconfiguração em um sistema é um conceito genérico, mas a forma como ele vai ser implementado é único para cada um deles, podendo ou não envolver conceitos de adaptação.

**QP3.** A terceira questão de pesquisa é “quais os domínios de aplicação em que os SRDs são usados?”. Com o intuito de respondê-la, os casos de uso e exemplos de aplicações apresentados nos artigos foram extraídos, como mostra a Figura 2.4.

Figura 2.4 - Resultados da classificação dos artigos segundo a propriedade “caso de uso”.



Fonte: Produção do autor.

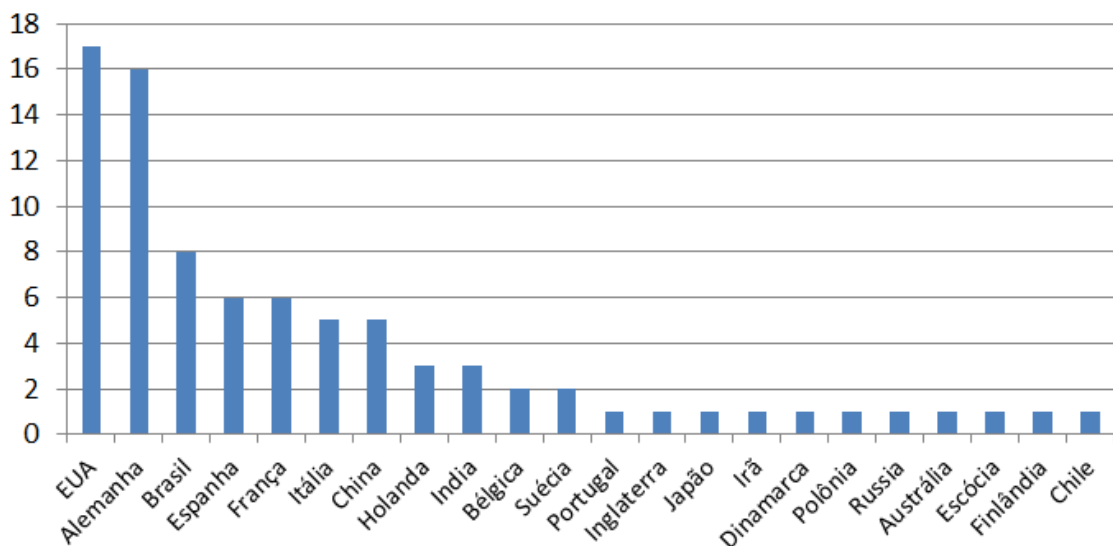
De maneira geral, o grande número de trabalhos relacionados a FPGAs pode ter criado uma tendência nos resultados. Conseqüentemente, os resultados não apresentaram áreas de aplicação vinculadas à domínios de aplicação específicos, mas sim a áreas genéricas, como processamento de imagens e de sinais. Além disso, a maioria dos artigos apresenta esforços para obter ganho em processamento e melhor adminis-

tração de recursos dos dispositivos, como memória e unidade lógicas, e demonstram pouca preocupação com a adaptabilidade no sentido de se flexibilizar ao ambiente.

**QP4.** A quarta e última questão de pesquisa é “Quais países lideram o número de publicações em SRDs?”. Para responder essa pergunta, o país onde os autores estão vinculados foi extraído para cada um dos artigos.

A Figura 2.5 mostra os resultados obtidos para cada país. Como pode ser visto, os dois países que lideram o número de publicações são Estados Unidos e Alemanha. Em um segundo grupo, pode-se ver Brasil, Espanha, França, China e Itália. Embora seja difícil explicar a distribuição apresentada na Figura 2.5, pode-se dizer que os países que lideram, Estados Unidos e Alemanha, apresentam algumas tendências. Os dispositivos mais citados nesta revisão são produzidos por duas empresas, Altera e Xilinx, ambas situadas no vale do silício, nos Estados Unidos. Talvez o aprimoramento desse tipo de dispositivos seja uma forte área de pesquisa. Por outro lado, todos os artigos desta revisão cujo assunto envolvia automóveis incluía, ao menos, um autor alemão, o que revela uma preocupação em estudar reconfiguração nessa área (ANTHONY et al., 2008; CLAUS; STECHELE, 2010; WAGNER et al., 2014; ZELLER et al., 2013).

Figura 2.5 - Resultados da classificação dos artigos segundo a propriedade “país dos autores”.



Fonte: Produção do autor.

### 2.1.3 Descobertas e Agenda de Pesquisa

As principais descobertas dessa RSL sobre SRDs são:

**Conceito de SRDs.** Baseado nos conceitos de reconfiguração obtidos durante a leitura dos trabalhos, foi proposta a definição de que um SRD é um sistema cujos subsistemas podem ser modificados ou ter suas configurações alteradas durante seu funcionamento para atingir um determinado objetivo, que é semelhante à definição proposta por outros trabalhos (LYKE et al., 2015; JÓZWIAK; NEDJAH, 2009). Essa definição também engloba o conceito de computação reconfigurável, termo proposto em 1960 para designar sistemas eletrônicos capazes de modificar sua estrutura interna a fim de aumentar a eficiência no processamento de tarefas específicas (ESTRIN, 2002). Também, como resultado dessa pesquisa, foi proposta uma nova taxonomia para SRDs.

**SRDs como sistemas digitais.** Embora Lyke et al. (2015) proponha uma taxonomia para sistemas reconfiguráveis que compreende desde matéria reconfigurável até meios analógicos e digitais, nesta revisão foram detectados apenas sistemas digitais, sem considerar o trabalho sobre o VANT reconfigurável que não apresenta reconfiguração dinâmica. Isso significa que as estratégias para reconfiguração dinâmica são baseadas em software ou hardware, sendo que a maioria dos trabalhos apresentou reconfiguração em hardware.

**Pesquisa em FPGAs.** Mais da metade dos trabalhos (52,17%) apresentou estudos relacionados a FPGA. Acredita-se que, devido a uma maior evolução e comercialização desses dispositivos, a área de pesquisa envolvendo tais dispositivos ganhou espaço. O propósito dessa área de pesquisa não está vinculado a nenhum projeto maior, mas sim à exploração dos recursos do próprio dispositivo. Por exemplo, alguns trabalhos buscam estratégias para diminuir o tempo de reconfiguração do FPGA, e propõe diferentes arquiteturas para isso (KESSAL et al., 2003; MÖLLER et al., 2006; WANG et al., 2009). Embora diversos trabalhos tenham demonstrado o grande poder dos FPGAs, projetar sistemas para esses dispositivos é uma tarefa muito árdua e ainda é necessário projetistas com bastante experiência engajados nos projetos.

**Baixo número de casos de uso industriais.** Os aqui chamados casos de uso industriais são aqueles trabalhos que descrevem soluções para problemas reais, como o caso do VANT reconfigurável (CHOWDHURY et al., 2014) ou o do assistente de direção para automóveis (WAGNER et al., 2014). Como a adaptação de sistemas é uma necessidade cada vez maior, esperava-se encontrar uma quantidade maior de

trabalhos relacionados a isso.

Uma agenda de pesquisa para a área de SRDs é a seguinte:

1. Tratar a área de pesquisa sobre computação reconfigurável, que inclui estudos com FPGAs, como uma área independente da área sobre SRDs. A computação reconfigurável é uma área que propõe arquiteturas para processamento de dados capaz de se modificar a fim de criar unidades de processamento específicas (ESTRIN, 2002). Comparada com outras arquiteturas, a computação reconfigurável está entre a arquitetura de Von Neumann e Circuitos Integrados de Aplicação Específica (do inglês, *Application Specific Integrated Circuits* - ASIC), sendo que a primeira possui alta flexibilidade, mas baixo desempenho, e a segunda baixa flexibilidade e alto desempenho (BOBDA, 2007). Ao contrário dos processadores de propósito geral, que são estáticos, o objetivo da computação reconfigurável é ser dinâmica, capaz de se adaptar às necessidades das aplicações. Além disso, a computação reconfigurável pode ser vista como parte do domínio dos SRDs, o que permite classificar a computação reconfigurável como uma subárea independente de SRDs. Assim, tratá-los individualmente pode ajudar a melhor definir seus conceitos e objetivos. Enquanto na computação reconfigurável as arquiteturas são discutidas no nível de *bits* e o objetivo é economizar memória ou melhorar o desempenho, em sistemas adaptativos o nível da abstração é maior e o objetivo é fornecer a um sistema complexo completo um certo tipo de inteligência. Dessa forma, acredita-se que tratá-las de maneira individual ajude a melhor definir os seus conceitos. Além disso, essa divisão pode ajudar a definir os objetivos de futuras revisões da literatura.
2. Explorar as metodologias usadas em SRDs e em FPGAs, i.e., definir as métricas usadas para a tomada de decisão, a forma como a tomada de decisão ocorre e em que ela é baseada, para cada uma das áreas propostas no item anterior;
3. Estudar as dificuldades encontradas na implementação de SRDs, as soluções existentes e os desafios que ainda existem;
4. Estimular o desenvolvimento de metodologias e técnicas para SRDs, e a aplicação delas a uma quantidade significativamente maior a estudos de caso industriais ou criados em institutos de pesquisa e desenvolvimento.

Particularmente, a investigação e desenvolvimento de SRDs no contexto aeroespacial, focando em aspectos como autoadaptação, dado que sistemas aeroespaciais estão sujeitos a ambientes de operação cujas características podem variar, muitas vezes, de forma imprevisível;

5. Investigar o uso de outras plataformas de hardware além de FPGA. Como foi mencionado nesta RSL, pode-se dizer que FPGA domina o contexto de SRDs. Seria interessante investigar mais a fundo outras plataformas de hardware, como microcontroladores e processadores de uso genérico, para o desenvolvimento de SRDs.

Em particular, a presente tese de doutorado se relaciona com o item 4 da agenda de pesquisa acima, e, nesse sentido, concentra esforços no desenvolvimento de técnicas de autoadaptação aplicadas ao contexto aeroespacial, em específico aos VANTs.

## 2.2 Veículos Aéreos Não Tripulados

Veículos não tripulados, sejam eles terrestres (ANDERSON et al., 2014), aquáticos (YAN et al., 2010), subaquáticos (YUH, 2000) ou aéreos (KANELLAKIS; NIKOLAKOPOULOS, 2017) diferem de veículos tradicionais por não haver necessidade de carregarem uma tripulação a bordo responsável por sua operação. Por esse motivo, tais veículos apresentam vantagens, tais como a facilitação de operação em áreas de difícil acesso, custo operacional reduzido e a criação de modelos com formas e tamanhos variados sem a necessidade da estrutura necessária para acomodar uma tripulação. Diferentes aplicações se beneficiaram dessas características ao longo do tempo, como, em específico na área aeroespacial, o *Mars Exploration Rover* na exploração de Marte (MAURETTE, 2003), e a sonda *Rosetta* na investigação do interior do cometa 67P (KOFMAN et al., 2007).

Com o advento dos microcomponentes, o desenvolvimento de novos materiais e a melhoria na tecnologia de sensores, surge uma tendência de modernização nos modelos de veículos não tripulados, onde estes passam a ganhar protagonismo em diferentes tipos de aplicações (OTTO et al., 2018). Nesse contexto, os VANTs se popularizaram e passaram a ser comercializados mais amplamente atingindo também as aplicações de caráter civil, gerando preocupação quanto aos riscos que poderiam causar mediante sua má utilização, além de carecerem de uma nomenclatura própria, tornando-se popularmente conhecidos como *drones*. Nesse sentido, a presente seção busca esclarecer algumas das questões acerca dessas aeronaves, apresentando a regulamentação de acordo com o órgão responsável brasileiro, a Agência Nacional de Aviação Civil



(ANAC), e discutindo aspectos como nomenclatura, definição e classificações.

### 2.2.1 Regulamentação no Brasil

No Brasil, o órgão responsável pela regulamentação de VANTs é a ANAC e que o fez por meio do Regulamento Brasileiro da Aviação Civil Especial número 94 (RBAC-E 94), no qual descreve os requisitos gerais para aeronaves não tripuladas de uso civil (AGÊNCIA NACIONAL DE AVIAÇÃO CIVIL, 2017). O documento se aplica a todas as aeronaves não tripuladas de uso civil, também chamadas de apenas aeronaves não tripuladas, que operam no território brasileiro ou possuem cadastro na ANAC.

O regulamento proíbe o transporte de pessoas e animais nesse tipo especial de aeronaves, além de não permitir seu uso de maneira autônoma, i.e., livre de um operador. Por esse motivo não são reconhecidas, em território brasileiro, aeronaves completamente independente de pilotos, o que leva a ANAC a chamar esse tipo particular de aeronaves como Aeronave Remotamente Pilotada, utilizando o acrônimo RPA (do inglês, *Remotely Piloted Aircraft*). Além disso, a ANAC adota o termo aeromodelo para aeronaves não tripuladas remotamente pilotadas cuja finalidade seja exclusivamente recreação e lazer. Como o regulamento exige um operador, mas ao mesmo tempo proíbe que esteja a bordo da aeronave, a operação se dá a partir de uma Estação de Pilotagem Remota (RPS - *Remote Pilot Station*), que contém os aparelhos necessários para o controle do veículo. O conjunto da aeronave, RPA, e da estação, RPS, forma o Sistema de Aeronave Remotamente Pilotada (RPAS - *Remotely-Piloted Aircraft System*), que é o termo oficial utilizado para denominar as aeronaves não tripuladas no Brasil.

As RPAS são classificadas de acordo com o Peso Máximo de Decolagem (PMD) da RPA, i.e., o peso máximo permitido para que uma aeronave tente decolar, o que inclui não apenas a massa da própria aeronave, mas também outras cargas envolvidas na missão. A classificação é dada da seguinte forma:

- Classe 1: RPA com PMD maior que 150kg;
- Classe 2: RPA com PMD maior que 25kg e menor ou igual a 150kg; e
- Classe 3: RPA com PMD menor ou igual a 25kg.

O regulamento exige que todos os pilotos remotos que operarem uma RPA sejam maiores de 18 anos de idade. Um piloto pode operar apenas um RPAS por vez, sendo a substituição de pilotos durante o voo permitida. Toda RPA, independente

da classe, precisa estar cadastrada junto à ANAC com a devida identificação na aeronave, com exceção de RPA com PMD menor ou igual a 250 gramas. Da mesma forma, a operação envolvendo RPA com PMD acima de 250 gramas deve possuir seguro com cobertura de danos a terceiro.

Além da subdivisão de aeronaves com PMD maior e menor que 250 gramas, o regulamento ainda divide RPA da classe 3 com PMD maior que 250 gramas que operarem no espaço abaixo de 400 pés acima do nível do solo e aquelas que operarem acima de 400 pés. O tipo de operação também influencia a documentação requerida, sendo operações em que é possível manter contato visual com a aeronave sem o auxílio de equipamentos menos exigente. Por último, ainda é necessário estar atento a outros possíveis documentos exigidos por parte do Departamento de Controle do Espaço Aéreo (DECEA), da Agência Nacional de Telecomunicações (ANATEL), e de outros órgãos competentes.

### 2.2.2 Nomenclatura e definição

Veículos aéreos livres de tripulação a bordo são denominados de diferentes modos. No Brasil, o termo utilizado pela ANAC é RPA, quando se refere apenas à aeronave, e RPAS, quando se refere ao sistema completo no qual a RPA é parte integrante. Atualmente, os órgãos responsáveis pelo espaço aéreo dos Estados Unidos, o Departamento de Defesa e a Administração Federal de Aviação, bem como a Agência Europeia para Segurança da Aviação, utilizam o termo Aeronave Não Tripulada (do inglês, *Unmanned Aircraft*) para a aeronave em si, e o termo Sistema de Aeronave Não Tripulada (do inglês, *Unmanned Aircraft System*) para designar todos os componentes, incluindo a aeronave, necessários para a realização do voo (DALMAGKIDIS et al., 2012). Neste trabalho, por se tratar de um estudo que explora um sistema que independe do uso de uma estação remota para operar aeronaves e também por ser comumente usado na literatura científica, o termo utilizado para designar aeronaves não tripuladas será Veículo Aéreo Não Tripulado (VANT), que deriva do inglês *Unmanned Aerial Vehicle*, juntamente com o termo *drone*, como são popularmente conhecidos.

A grande variedade de *drones* desenvolvida ao longo do tempo leva a múltiplas possíveis classificações baseadas nas características presentes nas aeronaves, como altura máxima de voo, tempo máximo de voo, tamanho, massa, alcance e PMD (AL-KAFF et al., 2018). Além dessas, pode-se considerar outra classificação baseada na forma como o veículo se sustenta no ar. Nesse sentido, de acordo com Kanellakis e Nikolopoulos (2017), os VANTs podem ser classificados em 4 categorias: veículos com

apenas um rotor (helicópteros), veículos com múltiplos rotores, na qual se enquadram quadricópteros, hexacópteros, octocópteros, e assim por diante a depender do número de rotores utilizados, veículos de asa fixa, e veículos híbridos. As principais características de cada um dos grupos está listada a seguir:

- **Único rotor.** Os helicópteros são caracterizados por apresentarem um rotor principal, responsável por manter a sustentação da aeronave e um rotor secundário localizado na cauda, responsável pelo controle e estabilização da aeronave. Veículos deste tipo são capazes de decolar e pousar verticalmente, sem a necessidade de uma pista de pouso ou decolagem. Além disso, podem carregar uma considerável carga útil e possuem grande tempo de voo.
- **Múltiplos rotores.** Essa é a classe mais popular na comunidade de VANTs, sendo os quadricópteros e hexacópteros os mais comuns. Se caracterizam por apresentarem diversos rotores de mesmo tamanho que servem para dar sustentação ao veículo. Assim como os helicópteros, podem decolar e pousar verticalmente. São ágeis e podem operar em ambientes fechados, porém apresentam limitada carga útil e tempo de voo.
- **Asa fixa.** Se caracterizam por apresentarem uma asa e um propulsor horizontal, que serve para dar velocidade ao veículo e, conseqüentemente, permitir que o fluxo de ar nas asas promova a sustentação da aeronave. Sua principal vantagem é a velocidade e eficiência. Também possuem maior capacidade para carga útil do que a classe anterior. As desvantagens incluem a necessidade de uma área para decolagem e pouso e a necessidade de uma velocidade horizontal mínima para se manterem em voo.
- **Híbridos.** Esta classe inclui um novo tipo de veículo, que pode ser entendido como uma mistura de veículos de asa fixa e veículos com rotores. Dessa forma, o objetivo dos veículos híbridos seria de manter as melhores características dos dois grupos, sendo capazes de decolar e pousar verticalmente, e eficiência e velocidade de voo dos veículos de asa fixa. A principal desvantagem se dá pela complexidade de desenvolvimento de uma aeronave com tais características.

Apesar de não estarem incluídos na classificação acima, os dirigíveis, ou zepelins, também podem ser considerados uma categoria própria, conforme proposto por Al-Kaff et al. (2018). Uma classe mais geral a ser considerada é a de aeróstatos, i.e.,

aeronaves mais leves que o ar, que incluem não apenas dirigíveis, mas também balões. Porém, segundo a Administração Federal de Aviação dos Estados Unidos, balões não podem ser considerados aeronaves não tripuladas porque não possuem total controle sobre a direção nos 3 eixos do espaço. A definição completa é dada a seguir (FEDERAL AVIATION ADMINISTRATION, 2008):

“Um dispositivo utilizado ou destinado a ser utilizado para voar no ar que não possui piloto a bordo. Isso inclui todas as classes de aviões, helicópteros, dirigíveis, e aeronaves de elevação translacional que não possuam piloto a bordo. Aeronaves não tripuladas são entendidas por incluir apenas aeronaves controláveis nos três eixos e, por isso, são excluídos os balões tradicionais.”

Algumas definições não deixam claro se balões são ou não são considerados VANTs, como a apresentada por Kanellakis e Nikolakopoulos (2017):

“Aeronave sem um piloto humano a bordo. O controle é realizado por um computador de bordo, controle remoto ou combinação de ambos.”

Além dessas, outras definições foram propostas para esse tipo de aeronaves, como, por exemplo, a que exige que os veículos sejam reutilizáveis, excluindo da classe mísseis, projéteis e torpedos (DALAMAGKIDIS et al., 2012). Aqui, o que se entende por aeronave não tripulada, denominada doravante de VANT ou *drone*, compreende todo veículo capaz de se sustentar e se guiar no ar sem que um humano possa, de dentro dele, operá-lo. O controle pode ser realizado por um operador humano desde que este se encontre em uma estação remota capaz de se comunicar com a aeronave. Porém, vale ressaltar que a forma como o controle é realizado não é relevante para o escopo deste trabalho, uma vez que o mesmo visa explorar técnicas de estimação da posição por meio do registro automático de imagens. Dessa forma, todo corpo que se movimenta no ar é um possível alvo para aplicação do sistema aqui explorado.

### 2.2.3 Navegação e estimação da posição

Uma das características que tornam os VANTs úteis em diversas aplicações é o amplo uso de sensores. Diferentes tipos de sensores podem ser utilizados com diferentes finalidades. Em ambientes fechados pode haver a necessidade de realizar o mapeamento 3D do ambiente, a detecção de objetos e o desvio de obstáculos. Para isso,

sensores de ultrassom, laser e câmeras estéreo e monoculares podem ser empregadas (AL-KAFF et al., 2018).

A área de pesquisa que estuda a utilização de sensores imageadores, como as câmeras fotográficas, para realizar a varredura e percepção de ambientes recebe o nome de visão computacional. Nessa área, a imagem é o dado de entrada por meio do qual são extraídas as características do mundo real (SZELISKI, 2011). Em VANTs, suas aplicações incluem não apenas estimação da posição, mas também modelagem 3D, vigilância e reconhecimento, inspeção de estruturas, desvio de obstáculos, entre outras (KANELLAKIS; NIKOLAKOPOULOS, 2017; AL-KAFF et al., 2018). Dessa forma, pode-se dizer que o uso de visão computacional em VANTs traz contribuições significativas para o aumento de suas capacidades. Por exemplo, utilizando o conceito de controle servo visual, foram propostos trabalhos de decolagem e pouso verticais sobre um alvo, navegação de linhas de alta tensão, estabilização em órbita circular e pouso autônomo sobre uma plataforma móvel (KANELLAKIS; NIKOLAKOPOULOS, 2017).

Contudo, independente da aplicação, é imprescindível que *drones* sejam capazes de navegar de forma precisa pelo ambiente em que estão imersos. Assim, uma peça que desempenha papel fundamental é o sistema de navegação embarcado na aeronave. Na área da robótica, navegação e posição são termos que possuem significados distintos. O primeiro engloba não apenas a ciência acerca da posição de um corpo, mas também o planejamento e gerenciamento da rota percorrida por ele evitando possíveis obstáculos e colisões. Já o segundo explora as estimações da posição e da velocidade do corpo com relação a um ponto, ou plano, de referência. Dessa forma, pode-se dizer que estimação da posição é um subgrupo da navegação (GROVES, 2013).

Inicialmente, os VANTs eram equipados com sensores para estimação da posição e orientação, em que geralmente eram empregados o GPS e o INS (KANELLAKIS; NIKOLAKOPOULOS, 2017). Como esses sistemas são suscetíveis a erros, a fusão dos dados se tornou uma alternativa promissora para aumentar a acurácia e aumentar a confiabilidade, além de estender as possibilidades para sistemas alternativos, como os baseados em sensores imageadores. Nesse sentido, foram propostas diferentes abordagens de estimação da posição baseadas em técnicas de visão computacional (BALAMURUGAN et al., 2016; KANELLAKIS; NIKOLAKOPOULOS, 2017; AL-KAFF et al., 2018). As técnicas de estimação da posição, na qual estão contidas as abordagens baseadas em visão computacional, podem ser divididas em dois grandes grupos

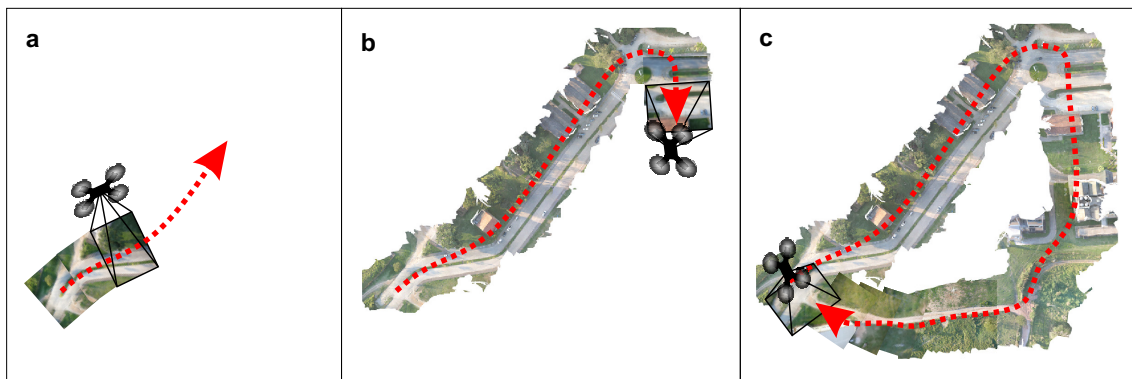
(GROVES, 2013):

- *Navegação estimada.* A expressão original é *dead reckoning*, provavelmente derivada de *deduced reckoning*, e pode ser definida como o cálculo da posição atual de um corpo baseado no seu avanço a partir de uma posição anterior. Por isso, pode-se dizer que, na navegação estimada, obtém-se a posição relativa, e não absoluta, de um veículo. Neste grupo são incluídas técnicas de estimação da posição que buscam calcular a velocidade e orientação realizadas pelo corpo, como a odometria por exemplo, em que o número de rotações realizadas pela roda de um veículo é contabilizado e multiplicado pelo seu diâmetro. Quando são utilizadas técnicas de visão computacional para se obter o movimento relativo de um corpo, o sistema de estimação da posição é também chamado de odometria visual. A maior desvantagem das técnicas deste grupo é o rápido aumento da imprecisão da posição estimada em função dos erros que se acumulam a cada novo movimento executado pelo veículo.
- *Posicionamento absoluto ou posicionamento fixo.* Esta categoria engloba sistemas de estimação da posição em que o cálculo é realizado a partir da identificação de informações externas ao corpo. Tais informações podem ser sinais artificiais ou estruturas naturais presentes no ambiente. Além disso, pode ser necessário comparar essas informações com informações preexistentes salvas em um banco de dados para inferir a posição do veículo. Na área marítima, por exemplo, a chamada navegação visual é definida como a estimação da posição de uma embarcação a partir do reconhecimento visual de estruturas presentes na costa terrestre (MIGUENS, 2019). Apesar de não apresentar acumulação de erros parciais no cálculo da posição, como ocorre com o *dead reckoning*, este grupo depende de informações prévias sobre o ambiente para que seja possível reconhecê-lo. Além disso, pode ocorrer de as informações externas serem propositalmente modificadas ou erroneamente reconhecidas ao serem comparadas com as informações internas preexistentes, levando a uma estimação falsa da posição do veículo.

Quando instrumentos externos não estão disponíveis para auxiliar um veículo na estimação da sua posição, os dois grupos, aqui chamados de navegação estimada e posicionamento fixo, podem ser entendidos como subgrupos de outro grupo maior, denominado *Simultaneous Localization and Mapping* (SLAM), que, em tradução livre, significa localização e mapeamento simultâneos (CADENA et al., 2016). Como o

nome indica, a ideia é, ao mesmo tempo, estimar a posição e realizar o mapeamento do ambiente no qual o veículo está inserido. No SLAM, um veículo que não possui acesso a informações prévias sobre um ambiente pode iniciar sua exploração utilizando técnicas de *dead reckoning* ao mesmo tempo que constrói um mapa coletando informações do ambiente. Quando o veículo retorna a algum local já visitado, as novas informações coletadas podem ser comparadas ao mapa criado anteriormente por meio de alguma técnica de posicionamento fixo. Nos casos em que um local é revisitado, ocorre o chamado fechamento de circuito (do inglês, *loop closure*), uma característica chave do SLAM. Caso não ocorra fechamento de circuitos, o SLAM se resume à navegação estimada. E caso um mapa sobre o ambiente esteja disponível antes de iniciar a navegação, a navegação estimada se torna uma ferramenta auxiliar para tornar o sistema mais robusto e menos suscetível a falhas. Como o fechamento de circuitos se dá pelo reconhecimento de uma parte do ambiente ao ser revisitado pelo veículo, ele também pode ser chamado de reconhecimento de regiões (CADENA et al., 2016). Um exemplo de como o SLAM funciona, em que o ambiente é sensorado por imagens aéreas do solo, pode ser visto na Figura 2.6.

Figura 2.6 - Ilustração do funcionamento do SLAM com imagens aéreas.



As imagens foram capturadas por uma câmera acoplada a um *drone* direcionada para o solo. A aeronave inicia a trajetória usando técnicas de odometria visual como está ilustrado em (a). Em seguida, em (b), as imagens são processadas para formar um mapa do trajeto já sensorado à medida que novas imagens são capturadas. Em (c) ocorre o reconhecimento de uma região já visitada, caracterizando o fechamento do circuito.

Fonte: Produção do autor.

### 2.2.3.1 Trabalhos relacionados

Para veículos aéreos, a necessidade de voltar a uma região já sobrevoada para que ocorra o fechamento do circuito pode exigir um custo extra de tempo de voo que muitas aeronaves não dispõem, uma vez que o custo de energia para se sustentar no ar é dispendioso e *drones* de pequeno porte são alimentados por baterias elétricas. Dessa forma, é interessante que um mapa esteja disponível previamente ao voo, tornando desnecessário o retorno a uma região já sobrevoada. Diversos trabalhos com esse objetivo foram desenvolvidos ao longo das últimas décadas, como pode ser observado na Tabela 2.3.

Tabela 2.3 - Trabalhos de estimação da posição de VANTs por meio de imagens a partir do reconhecimento de regiões.

#	Técnicas	Algoritmos	Referência
1	OV e RR	CCNZ e Hausdorff	(SIM et al., 2002)
2	RR	Filtro de Gabor e RNA	(SHIGUEMORI et al., 2007)
3	RR	Operador de Harris	(SAMADZADEGAN et al., 2007)
4	OV e RR	KLT e CC	(CONTE; DOHERTY, 2008)
5	OV e RR	KLT e CCNZ	(CONTE; DOHERTY, 2009)
6	RR	Histograma	(LINDSTEN et al., 2010)
7	RR	SIFT	(CESETTI et al., 2011)
8	RR	SIFT	(SAMADZADEGAN; ABDI, 2012)
9	RR	CCNZ	(LIU et al., 2015)
10	RR	CC	(BRAGA et al., 2016)
11	RR	SURF	(ABDI et al., 2016)
12	RR	Correlação de Fases	(WAN et al., 2016)
13	RR	SIFT, SURF e outros	(BLAZHKO et al., 2017)
14	RR	SURF	(STEPANOV; TISHCHENKO, 2017)

Por limitação de espaço, as técnicas de Odometria Visual e Reconhecimento de Regiões são chamadas de OV e RR respectivamente. As demais siglas são definidas no texto.

A Tabela 2.3 foi criada a partir de artigos de revisões sobre outros trabalhos na área de estimação da posição de VANTs utilizando visão computacional (BALAMURUGAN et al., 2016; KANELLAKIS; NIKOLAKOPOULOS, 2017; AL-KAFF et al., 2018). A partir desses trabalhos, foi utilizada a técnica da bola de neve, em que novas referências são selecionadas a partir das listas de citações dos trabalhos anteriores. Os novos trabalhos escolhidos servem, por sua vez, como fonte de citações para que outros trabalhos sejam encontrados. O processo pode ser repetido até que não se encontre nenhum novo trabalho relevante na área. Como consequência do processo, foram selecionados 14 trabalhos que utilizam o reconhecimento de regiões para estimar a



posição de VANTs.

Todos os trabalhos utilizam informações de imagens óticas capturadas por sensores imageadores monoculares a bordo de *drones*. A única exceção é o trabalho de número 1, que utiliza também o modelo digital de elevação da região sobrevoada para realizar a estimação da posição. Além disso, alguns trabalhos apresentam também técnicas de odometria visual como forma de aumentar o desempenho da estimação da posição de VANTs. Todos os trabalhos apresentam técnicas que contam com informações sobre o ambiente, como mapas ou imagens georreferenciadas, disponíveis antes do início do voo.

1. [Sim et al. \(2002\)](#) utilizam duas informações complementares sobre o ambiente sobrevoado: o modelo digital de elevação e imagens óticas. O trabalho também faz uso de duas técnicas de estimação da posição diferentes: a odometria visual e o reconhecimento de regiões. Na primeira, o algoritmo de *blockmatching* é empregado, enquanto na segunda, para calcular a posição absoluta do veículo, são utilizadas a correlação cruzada normalizada de média zero (ver Capítulo 3), representada na Tabela 2.3 como CCNZ, nas informações providas pelo modelo digital de elevação e a medida de Hausdorff (*Robust-Oriented Hausdorff Measure*) nas imagens óticas. No final, é proposta uma integração envolvendo as duas técnicas para reduzir o erro da posição estimada.
2. [Shiguemori et al. \(2007\)](#) apresentam uma forma de navegação baseado em pontos de referência naturais presentes no solo. Para isso, é necessário que uma estação de controle planeje antecipadamente e escolha os pontos de referência observáveis pela aeronave durante o percurso da mesma. Um banco de filtros Gabor é empregado para extração das características dos marcos. As características são então utilizadas por duas redes neurais artificiais diferentes para classificar as regiões escolhidas.
3. [Samadzadegan et al. \(2007\)](#) descrevem a ideia de utilizar imagens capturadas em voo para, posteriormente, compará-las a outra imagem que serve como referência para estimar a posição da aeronave. Para isso, é utilizado o algoritmo baseado em pontos de interesse de Harris, ou operador de Harris.
4. [Conte e Doherty \(2008\)](#) abordam a estimação da posição por meio da odometria visual e também por meio do reconhecimento de regiões. No primeiro caso, utiliza o algoritmo Kanade-Lucas-Tomasi (KLT) para de-

tectar pontos de interesse no solo e calcular a posição relativa da aeronave. Como o cálculo da posição relativa acumula erros ao longo do voo, é necessária uma técnica de posicionamento fixo para reestabelecer a posição correta. Para isso, é proposto o uso de extratores de bordas e Correlação Cruzada (CC) para realizar o registro de imagens entre as imagens capturadas durante o voo e uma imagem georreferenciada obtida previamente.

5. [Conte e Doherty \(2009\)](#) é uma versão modificada do artigo publicado em 2008, [Conte e Doherty \(2008\)](#). A primeira mudança é a utilização da correlação cruzada normalizada de média zero ao realizar o registro de imagens entre as imagens capturadas em voo e a referência. Além disso, não são mais utilizados extratores de bordas antes de realizar o cálculo da correlação. A segunda modificação é o uso de odometria visual para obtenção da posição relativa do veículo.
6. [Lindsten et al. \(2010\)](#) propõem o uso de histogramas de classe calculados a partir de imagens capturadas por VANTs para comparar com mapas pré-classificados. A comparação é feita por meio da informação contida nos histogramas, que, por sua vez, são calculados a partir das características presentes nas imagens aéreas divididas em zonas circulares. A principal vantagem de utilizar histogramas é que estes não sofrem influência da rotação nem da luminosidade entre a imagem referência e a imagem capturada.
7. [Cesetti et al. \(2011\)](#) apresentam o uso de imagens de satélite para realizar o cálculo da estimação da posição de aeronaves. Para realizar o registro de imagens, é utilizado o algoritmo *Scale-Invariant Feature Transform* (SIFT) para detectar pontos de interesse do ambiente.
8. [Samadzadegan e Abdi \(2012\)](#) é baseado no trabalho anterior do mesmo autor publicado anteriormente, [Samadzadegan et al. \(2007\)](#). A ideia é realizar o registro de imagens para estimar a posição absoluta de uma aeronave durante o voo utilizando o algoritmo SIFT. Assim como no trabalho original, as imagens capturadas em voo devem ser comparadas a uma referência obtida previamente.
9. [Liu et al. \(2015\)](#) propõem o uso da energia da imagem baseado no filtro de Sobel para reduzir o efeito da luminosidade e auxiliar no cálculo da estimação da posição. O cálculo é realizado entre imagens capturadas em voo e imagens referência, sendo que o algoritmo utilizado para realizar o registro de imagens é a CCNZ.

10. Braga et al. (2016) utilizam extratores de borda clássicos de Canny e Sobel para comparar com o extrator de bordas proposto pelos autores. A ideia da extração de bordas é remover os detalhes contidos nas imagens aéreas para obter um melhor casamento entre o mapa utilizado como referência e as imagens capturadas pelo *drone* em tempo real. O casamento entre os dois é realizado utilizando a CC.
11. Abdi et al. (2016) propõem o uso do algoritmo *Speeded-Up Robust Features* (SURF) para extrair pontos de interesse de uma imagem referência e, assim, utilizar as posições dos pontos localizados no solo. Com as equações de colinearidade, calcula a atitude, posição e altura da aeronave. Os experimentos são feitos a partir da imagem de um satélite, que foi artificialmente modificada para simular as imagens aéreas.
12. Wan et al. (2016) apresentam a ideia de remover o efeito da luminosidade nas imagens. Para isso, propõem o uso do algoritmo de correlação de fases para realizar o registro de imagens capturadas em voo com imagens previamente obtidas de satélites.
13. Blazhko et al. (2017) utilizam diferentes algoritmos baseados em pontos de interesse para efetuar o registro de imagens entre imagens capturadas por *drones* e imagens utilizadas como referência. O trabalho apresenta um estudo comparativo entre os algoritmos testados levando em consideração o tempo de processamento e a quantidade de pontos de interesse detectados. Além disso, apresenta uma discussão sobre o uso de imagens capturadas em diferentes estações do ano, afirmando que as diferenças entre elas podem ser a causa de problemas na realização do registro de imagens.
14. Stepanov e Tishchenko (2017) propõem o uso de pontos de interesse naturais localizados no solo para estimar a posição de aeronaves. Para isso, utilizam o algoritmo SURF em imagens de satélite para extrair esses pontos e utilizá-los posteriormente durante o voo. Os experimentos são realizados por meio de imagens sintéticas obtidas a partir de imagens de satélite que simulam voos de VANTs.

Todos os trabalhos aqui listados apresentam metodologias diferentes para realizar a estimação da posição de VANTs por meio de imagens. Porém, nenhum deles discute uma metodologia contendo a adaptação como forma de otimizar o sistema de estimação da posição por imagens. Além disso, os casos de uso nos quais tais metodologias

foram testadas apresentam diferentes características, o que revela a importância de um sistema que possa ser adaptado para cada caso.

### **2.3 Considerações finais sobre este capítulo**

O presente capítulo apresentou, de forma sucinta, uma RSL sobre SRDs. Tal revisão resultou em uma agenda de pesquisa, na qual se destaca, entre outros itens, a possibilidade de utilização da reconfiguração como forma de prover um sistema com a característica da autoadaptação. Além disso, também foram discutidos aspectos gerais sobre VANTs, como nomenclatura, legislação, definição e pesquisas desenvolvidas na área de estimação da posição por imagens. No próximo capítulo, será abordada e discutida a metodologia da estimação da posição por imagens utilizada neste trabalho, bem como fundamentos necessários para outros capítulos.

### 3 REGISTRO DE IMAGENS E APRENDIZADO DE MÁQUINA

Neste trabalho, a estimação da posição de VANTs é realizada pela técnica de reconhecimento de regiões por meio do registro automático de imagens. A ideia central da técnica jaz no reconhecimento da região na qual o VANT está sobrevoando ao associá-la a um mapa, que serve como referência para localizar a posição do veículo. A técnica envolve uma série de etapas, em que algoritmos de processamento de imagens são utilizados, assim como informações da aeronave durante o percurso de voo. Os algoritmos de cada etapa dessa técnica são aqui definidos com base nos trabalhos de Silva (2016) e Braga (2018), que a investigaram e desenvolveram. Dessa forma, o presente capítulo é dedicado à definição dos algoritmos empregados por tal técnica de estimação da posição. Além disso, são apresentados os conceitos das metodologias de aprendizado de máquina utilizadas neste trabalho.

Os parâmetros de entrada necessários para correta execução da técnica podem ser vistos na Tabela 3.1. Tais parâmetros foram divididos em 2 grupos: estáticos e dinâmicos. O primeiro grupo inclui valores que se modificam ao longo do voo realizado pela aeronave, enquanto o segundo inclui valores que permanecem inalterados dentro de um mesmo percurso. Os parâmetros do primeiro grupo são formados pelas imagens capturadas, e pelas atitude e altitude da aeronave no instante de captura das mesmas. Enquanto os valores de entrada do segundo grupo são constituídos pela imagem usada como referência e sua resolução espacial, e o ângulo de visada instantâneo da câmera acoplada ao VANT.

Tabela 3.1 - Parâmetros de entrada necessário para estimação da posição por imagens.

<b>Parâmetros de entrada</b>	
<b>Dinâmicos</b>	<b>Estáticos</b>
Imagem	Imagem referência
Atitude da aeronave	Resolução espacial da imagem referência
Altitude da aeronave	Campo de visada instantâneo da câmera

De maneira geral, o objetivo do reconhecimento de regiões por meio de imagens é utilizar imagens obtidas por meio de sensores imageadores a bordo da aeronave para, após uma série de processamentos, compará-las a uma imagem referência e, assim, estimar a posição da aeronave. A imagem referência, i.e., a imagem que será utilizada como mapa é uma imagem georreferenciada, que pode ser formada, por exemplo, por um mosaico de imagens ou por uma imagem de satélite. É importante

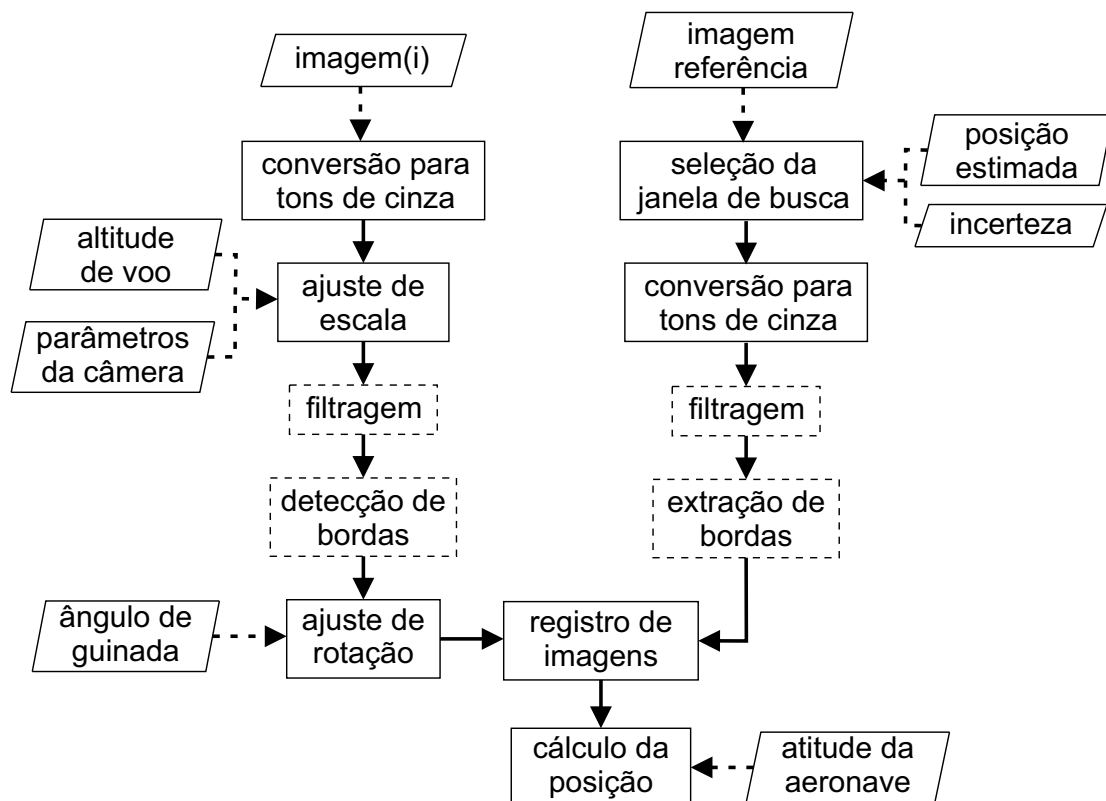
que a imagem referência cubra toda a área na qual o *drone* irá realizar o sobrevoo, de forma que este não visite áreas que se encontram além dos limites do mapa de referência.

Como a imagem georreferenciada e a imagem capturada no momento do voo são obtidas sob diferentes condições, elas estão submetidas a diferentes características e, por isso, etapas de processamento são executadas com o intuito de atenuar as diferenças. Apesar de ter sido empregada em voos noturnos (SILVA, 2016), a técnica utilizada aqui foi desenvolvida para voos diurnos em que a posição do sensor imageador está em nadir, i.e., orientado para o centro da Terra. Na Figura 3.1 pode ser visto fluxograma envolvendo todas as etapas necessárias para execução da técnica de reconhecimento de regiões por meio do registro de imagens.

Para uma imagem capturada em um instante  $i$  durante o voo, a primeira etapa é formada pela sua conversão em uma nova imagem em tons de cinza, de forma que os algoritmos subsequentes possam operar corretamente. Logo em seguida, uma sequência de etapas é realizada com o objetivo de melhorar a semelhança entre a imagem capturada e a referência. Entre elas, duas são obrigatórias: o ajuste de escala e o ajuste de rotação, que se fazem necessárias porque os algoritmos usados na etapa do registro de imagens são invariantes a escala e rotação e apresentam resultados ruins caso o ângulo de rotação e a escala sejam muito diferentes entre as imagens as quais estão sendo comparadas (GONZALEZ; WOODS, 2008). As etapas de filtragem e detecção de bordas são opcionais, porém buscam eliminar características nas imagens que possam influenciar negativamente a etapa do registro de imagens. Em cada etapa, um ou mais algoritmos podem ser utilizados, compondo um conjunto de possibilidades que afetam diretamente os resultados na estimação da posição.

Da mesma forma que a imagem capturada durante o voo é processada ao longo de diferentes etapas, a imagem referência é afetada pelos mesmos processamentos, com exceção da correção de escala e rotação, pois essas informações são utilizadas como referência. A única etapa adicional é a seleção da janela de busca. Nessa etapa, a partir da posição estimada de onde o *drone* está e de uma medida de incerteza, uma porção da imagem referência é selecionada para realização do registro de imagens. Dessa forma, reduz-se a chance de erro por meio de reconhecimento espúrio de regiões semelhantes e também o esforço computacional das etapas subsequentes. O cálculo do tamanho da área de busca em dado instante  $i$  é dado pela Equação 3.1.

Figura 3.1 - Fluxograma da técnica de estimação da posição baseado em reconhecimento de regiões.



A imagem( $i$ ) representa a imagem obtida no instante  $i$  durante o voo. As etapas de processamento são representadas por retângulos, em que os de linha contínuas representam etapas obrigatórias e os de linha tracejada representam etapas opcionais. Os parâmetros de entrada são representados por paralelogramos e se conectam às etapas de processamento por meio de linhas tracejadas.

Fonte: Produção do autor.

$$J(i) = \begin{bmatrix} lar(f_i) + inc/res(f_i) \\ alt(f_i) + inc/res(f_i) \end{bmatrix}, \quad (3.1)$$

em que  $f_i$  é a imagem aérea capturada após o ajuste de escala,  $lar(f_i)$  e  $alt(f_i)$  é a largura e altura da imagem respectivamente,  $res(f_i)$  é a sua resolução espacial, e  $inc$  é a incerteza em metros. A partir da posição estimada da aeronave, pode-se calcular a região onde a janela estará centralizada.

Nesse contexto, o presente capítulo se divide em 6 seções, que se dedicam a detalhar

os algoritmos presentes em cada umas das etapas que compõe a técnica de estimação da posição baseada em reconhecimento de regiões. Os aspectos básicos da imagem digital, que formam a base para entendimento de tais seções, são apresentados no Apêndice B. Na primeira Seção, 3.1, o algoritmo de conversão para imagens em tons de cinza é apresentado. Os algoritmos responsáveis pelas transformadas geométricas das imagens, que incluem os algoritmos de ajuste de escala e de rotação são apresentados na Seção 3.2. Nas seções seguintes, 3.3 e 3.4, são discutidos os algoritmos utilizados para filtrar e extrair as bordas das imagens. Os algoritmos utilizados para realizar o registro de imagens são apresentados na Seção 3.5. Na Seção 3.6, é analisado o cálculo da posição da aeronave com base na atitude da mesma durante a captura da imagem. Por último, na Seção 3.7 são abordados os fundamentos teóricos das técnicas de aprendizado de máquina utilizados nos capítulos subsequentes.

### 3.1 Conversão para escala de cinza

Uma das etapas da metodologia de estimação da posição por meio do registro de imagens é, sempre que necessário, a transformação da imagem capturada pelo VANT em uma imagem monocromática. Assim, a imagem monocromática, ou imagem em escala de cinza, pode ser representada por uma única função bidimensional,  $f(x, y)$ , cujos valores são finitos e discretos. A partir desta, pode-se então aplicar os algoritmos necessários para o cálculo da posição do VANT.

A transformação de uma imagem colorida para uma imagem em tons de cinza implica na perda de informação da mesma, uma vez que elimina suas informações sobre cores. A perda, porém, apresenta benefícios, pois simplifica as etapas de processamento subsequentes da metodologia de estimação da posição por meio de registro de imagens e reduz o esforço computacional dos algoritmos empregados (HUGHES et al., 2013; GONZALEZ; WOODS, 2008).

Uma das formas de se converter uma imagem colorida para uma imagem monocromática é por meio da composição ponderada da intensidade de cada uma das bandas da primeira. Caso a imagem esteja no padrão RGB, as bandas vermelho, verde e azul devem ser misturadas para se obter a nova imagem em tons de cinza. A imagem monocromática foi obtida a partir da banda Y do modelo de cores YIQ e pode ser calculada como

$$f_m(x, y) = 0,299 \cdot f_R(x, y) + 0,597 \cdot f_G(x, y) + 0,114 \cdot f_B(x, y), \quad (3.2)$$



em que  $f_m$  é a imagem monocromática, e  $f_R$ ,  $f_G$  e  $f_B$  é a banda vermelha, verde e azul da imagem colorida respectivamente. A porcentagem utilizada em cada uma das bandas para compor a imagem monocromática está relacionada à sensibilidade que o olho humano possui do espectro da luz (HUGHES et al., 2013).

### 3.2 Transformações geométricas

As transformações geométricas são utilizadas para mitigar os efeitos provocados pelo movimento da câmera quando embarcado no *drone*. As imagens capturadas durante o voo por VANTs estão suscetíveis a diferentes tipos de distorções, sejam elas consequência do revelado da superfície imageada, do conjunto óptico das lentes da câmera utilizada ou do movimento da aeronave. Para que a técnica de registro de imagens por meio da correlação cruzada funcione corretamente, a distorção de escala e rotação devem ser corrigidas. Ambas podem ser realizadas aplicando transformações geométricas a imagens digitais.

As transformações geométricas de uma imagem monocromática podem ser realizadas em dois passos. Primeiramente, as coordenadas de cada *pixel*,  $(x, y)$ , da nova imagem devem ser calculadas. Em seguida, uma intensidade,  $f_m(x, y)$ , devem ser atribuídos a cada um dos *pixels* obtidos pela etapa anterior. No caso específico das transformações de escala e rotação, a seguinte equação pode ser utilizada (GONZALEZ; WOODS, 2008; HARTLEY; ZISSERMAN, 2004)

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} T_{1,1} & T_{1,2} \\ T_{2,1} & T_{2,2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad (3.3)$$

em que  $(x, y)$  são as coordenadas da imagem original e  $(x', y')$  são as coordenadas da imagem após a transformação. A transformação pode então ser realizada ao se determinar os valores de  $T_{1,1}$ ,  $T_{1,2}$ ,  $T_{2,1}$  e  $T_{2,2}$ .

Porém, ao se realizar a transformação da Equação 3.3, pode ocorrer de algumas coordenadas  $(x', y')$  ficarem sem valores, ocasionando lacunas na nova imagem. Assim como pode ocorrer de determinadas coordenadas serem mapeadas para o mesmo ponto, criando sobreposição de valores. Para evitar esses problemas, pode-se realizar a transformação inversa (GONZALEZ; WOODS, 2008)

$$\begin{bmatrix} x \\ y \end{bmatrix} = T^{-1} \begin{bmatrix} x' \\ y' \end{bmatrix}. \quad (3.4)$$

A partir da transformada inversa é possível obter as coordenadas da nova imagem. A intensidade, ou brilho, dos *pixels* pode ser então calculada escolhendo as coordenadas mais próximas na imagem original. Essa técnica é chamada de vizinho mais próximo (do inglês, *nearest neighbor*) e apresenta menor custo computacional quando comparada às técnicas que utilizam interpolação. Além disso, ela mantém inalterado o valor dos *pixels* da imagem original (GONZALEZ; WOODS, 2008; HUGHES et al., 2013).

### 3.2.1 Ajuste de escala

O ajuste de escala pode ser visto como uma transformada geométrica de redimensionamento da imagem, i.e., uma transformação que modifica a resolução espacial da imagem (GONZALEZ; WOODS, 2008). Como o sensor imageador está em posição nadir, a resolução espacial será aqui tratada como a porção do solo representada por um *pixel* e, por isso, será equivalente ao *Ground Sample Distance* (GSD), que pode ser traduzido como distância da amostra do solo.

O ajuste de escala é responsável por igualar o GSD da imagem capturada pelo VANT durante o voo e o GSD da imagem referência. Em geral, os sensores imageadores utilizados e as condições de voo tornam a resolução espacial das imagens diferentes. O GSD de cada uma das imagens pode ser calculado usando a equação B.2, que depende dos parâmetros intrínsecos da câmera e da distância entre a lente e o objeto imageado. Distorções que dependem de outros fatores (e.g., aberrações ópticas e irregularidades presentes no relevo do terreno) não são aqui consideradas.

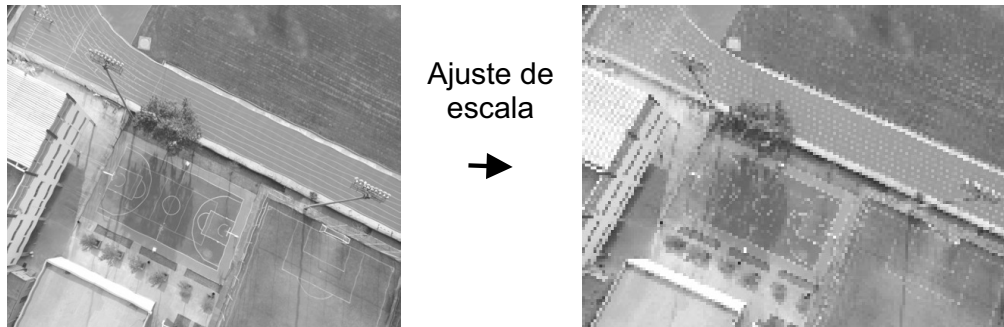
Nesse contexto, a distância entre a câmera e a superfície imageada pode ser entendida como a distância do *drone* ao solo, i.e., sua altitude. Esta, por sua vez, pode ser obtida por meio dos sensores a bordo da aeronave, como um barômetro. Os parâmetros intrínsecos da câmera podem ser obtidos por meio de um teste de calibração ou junto ao representante do produto. Após calcular o GSD das duas imagens (sendo uma capturada em voo e a outra a referência), pode-se obter os parâmetros necessários para redimensionar a imagem por meio da matriz (GONZALEZ; WOODS, 2008; HARTLEY; ZISSERMAN, 2004)

$$T = \begin{bmatrix} T_x & 0 \\ 0 & T_y \end{bmatrix}, \quad (3.5)$$

em que  $T_x$  e  $T_y$  são calculados pela divisão entre o GSD da imagem capturada e a imagem referência em cada um dos eixos. A imagem na qual as transformações

serão aplicadas é aquela que possui menor GSD. Na Figura 3.2 pode ser vista uma imagem capturada por um VANT com GSD de aproximadamente  $14 \text{ cm}/\text{pixel}$  (à esquerda) e a mesma imagem após ser redimensionada para um GSD de  $50 \text{ cm}/\text{pixel}$  (à direita).

Figura 3.2 - Imagem capturada por VANT antes e depois de sofrer ajuste de escala.



A imagem aérea original (à esquerda) foi capturada durante um voo realizado por um VANT à cerca de 70 metros do solo, possui tamanho de  $520 \times 427 \text{ pixels}$  e GSD de aproximadamente  $14 \text{ cm}/\text{pixel}$ . A mesma imagem com GSD de  $50 \text{ cm}/\text{pixel}$  (à direita) e tamanho de  $150 \times 112 \text{ pixels}$ . O FOV das duas imagens é aproximadamente o mesmo.

Fonte: Produção do autor.

### 3.2.2 Ajuste de rotação

Além das imagens apresentarem aproximadamente o mesmo GSD, elas precisam estar orientadas aproximadamente no mesmo ângulo. A imagem referência é, por padrão, orientada para o norte. As imagens capturadas em voo, porém, não possuem uma orientação única. Esse problema pode ser corrigido caso seja utilizado um *gimbal*, o que às vezes não acontece. Dessa forma, o ajuste de rotação se faz necessário para corrigir a orientação das imagens.

O ângulo de guinada do *drone* no momento de captura da imagem pode ser obtido por meio de seus sensores, como uma bússola eletrônica. A partir desse, pode-se obter o ângulo de guinada da câmera instalada no corpo da aeronave. Usando esse parâmetro e as Equações 3.3 e 3.4, é possível realizar o ajuste de rotação da imagem por meio da matriz de transformação

$$T = \begin{bmatrix} \cos(\kappa) & -\text{sen}(\kappa) \\ \text{sen}(\kappa) & \cos(\kappa) \end{bmatrix}, \quad (3.6)$$

em que  $\kappa$  é o ângulo de guinada câmera no momento de captura da imagem. A Figura 3.3 exemplifica o ajuste de rotação, em que uma imagem aérea obtida por uma câmera a bordo de um *drone* (à esquerda) é rotacionada usando o processo descrito acima (à direita). Um detalhe importante é a borda preta gerada após a rotação da imagem e que não faz parte da mesma. Isso ocorre como consequência do preenchimento (ou *padding*, no inglês) de valores que não fazem parte da imagem original para manter o formato retangular da imagem.

Figura 3.3 - Imagem capturada por VANT antes e depois de sofrer ajuste de rotação.

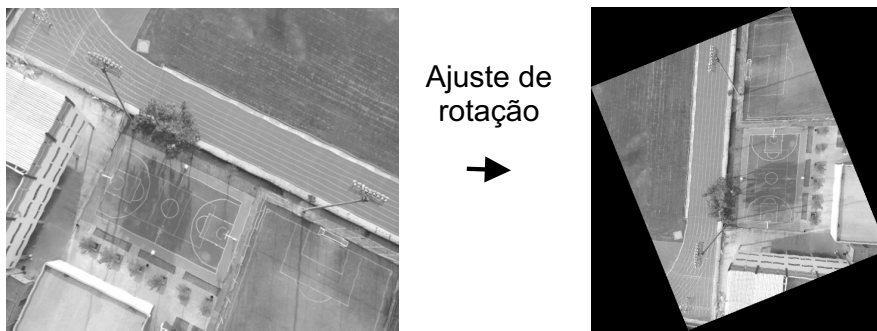


Imagem aérea (à esquerda) capturada durante um voo realizado por um VANT e após ser rotacionada em um ângulo de  $112^\circ$  (à direita). A imagem original possui tamanho de  $520 \times 427 \text{ pixels}$ , enquanto a imagem rotacionada possui tamanho de  $591 \times 642 \text{ pixels}$ . O acréscimo no tamanho da imagem se dá pelo preenchimento de zeros em decorrência da transformação e gera a região preta ao redor da imagem.

Fonte: Produção do autor.

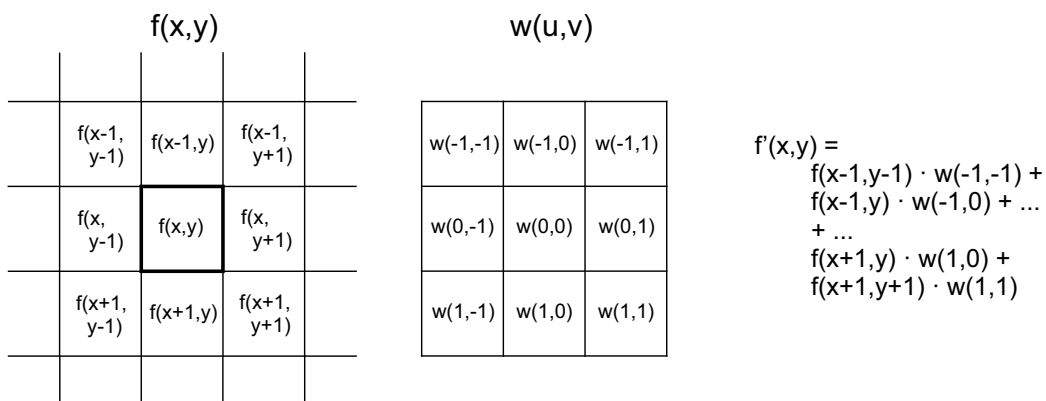
### 3.3 Filtros no Domínio do Espaço

Filtros podem ser entendidos como operações que alteram a intensidade dos *pixels* em imagens para produzir novas imagens. O termo “filtro” está associado ao domínio da frequência, pois é utilizado para descrever mecanismos que permitem ou rejeitam a passagem de determinadas frequências de um sinal. Quando essas manipulações são realizadas no domínio do espaço, recebem o nome de filtros espaciais, ou filtros no domínio do espaço (GONZALEZ; WOODS, 2008; HUGHES et al., 2013).

De forma geral, a filtragem busca melhorar o aspecto das imagens manipulando suas informações de forma a deixá-la mais apropriada para determinada aplicação (GONZALEZ; WOODS, 2008). A manipulação ocorre por meio da modificação de um *pixel* por meio de uma operação que envolve os valores do próprio *pixel* e seus vizinhos. Dessa forma, filtros espaciais envolvem um conjunto de valores, que servem como entrada, para se obter um novo valor. Caso a operação realizada na imagem seja linear, o filtro é denominado de filtro linear. Caso contrário, o filtro é denominado de não linear.

Filtros espaciais lineares operam como a soma de produtos entre uma imagem e uma janela, que também pode ser chamada de núcleo (do inglês, *kernel*), máscara, e padrão (do inglês, *template*). Essa operação ocorre entre os valores do *kernel* e a intensidade dos *pixels* da imagem na vizinhança apropriada e recebe o nome de convolução (GONZALEZ; WOODS, 2008; HUGHES et al., 2013). O valor dos coeficientes da janela define a natureza do filtro e o efeito que este causará na imagem. A mecânica da convolução é exemplificada na Figura 3.7 com um núcleo de tamanho  $3 \times 3$ .

Figura 3.4 - Convolução entre uma imagem e uma janela de tamanho  $3 \times 3$ .



A convolução entre a imagem  $f$  e a janela  $w$  envolve a multiplicação de todos os 9 elementos da janela pelos respectivos elementos da imagem. O resultado final é obtido somando-se todas as multiplicações.

Fonte: Produção do autor.

Como pode ser visto na Figura 3.4, a operação de convolução envolve a multiplicação entre os elementos da janela,  $w(u, v)$ , e os elementos da imagem,  $f(x, y)$ . O valor do

novo *pixel*,  $f'(x, y)$ , pode ser calculado centralizando a janela nas coordenadas  $(x, y)$  e aplicando a convolução por meio da equação (GONZALEZ; WOODS, 2008; HUGHES et al., 2013)

$$f'(x, y) = \sum_{u=-a}^a \sum_{v=-b}^b f(x + s, y + t) \cdot w(u, v), \quad (3.7)$$

em que  $f'(x, y)$  é o valor do novo *pixel*, e  $a$  e  $b$  representam o tamanho da janela de forma que a multiplicação envolva todos os elementos da mesma. Aqui, o formato da janela dos filtros é assumido como retangular de tamanho  $2a + 1$  e  $2b + 1$ , i.e., tamanhos horizontal e vertical ímpares. Apesar de ser possível trabalhar com janelas de lados pares, a utilização de lados de tamanho ímpar facilita a indexação e permite que as operações sejam realizadas com valores inteiros (GONZALEZ; WOODS, 2008).

Para gerar uma nova imagem,  $f'$ , a operação de convolução deve ser aplicada a todos os *pixels* da imagem original,  $f$ . Para isso, todas as coordenadas  $(x, y)$  devem ser visitadas. Quando as coordenadas estão próximas das laterais da imagem, parte da janela fica “para fora”, pois não há valores na imagem que possam ser associados aos valores da janela. Para solucionar esse problema, pode-se preencher as lacunas com zeros ou outros valores (GONZALEZ; WOODS, 2008). Aqui, a técnica escolhida replica os valores da borda mais próxima para os valores mais afastados exigidos pela janela.

Nesse contexto, pode-se dizer que filtros espaciais lineares envolvem a convolução de uma imagem com uma janela. Modificar os valores da janela implica em modificar o efeito que o filtro causa na imagem. E modificar o tamanho do filtro implica em modificar a intensidade com que esses efeitos são aplicados. Nas Subseções 3.3.1 e 3.3.2 são discutidos os filtros lineares da média e de Gauss, respectivamente, e seus efeitos em imagens aéreas. Na Subseção 3.3.3 é apresentado o filtro não linear da mediana.

### 3.3.1 Filtro da Média

O filtro da média pode ser obtido ao se definir uma janela cujos coeficientes possuam todos o mesmo valor e uma constante responsável pela normalização, obtida por meio da soma dos coeficientes da janela (GONZALEZ; WOODS, 2008; HUGHES et al., 2013). Quando a escolha dos coeficientes é 1, o que geralmente ocorre, a constante de normalização passa a ser o número de elementos do filtro. Ao se realizar a convolução

(ver Equação 3.7) de uma imagem com tal janela, o valor de cada *pixel* da nova imagem passa a ser a média dos *pixels* da imagem original e da sua vizinhança.

O filtro da média também recebe o nome de filtro de suavização, pois reduz as transições de intensidade dentro da imagem. Nesse sentido, o filtro da média pode ser utilizado para reduzir o impacto de detalhes irrelevantes em uma imagem (GONZALEZ; WOODS, 2008). Na Figura 3.5 pode ser visto um exemplo do efeito causado pelo filtro da média com uma janela de tamanho  $11 \times 11$ .

Figura 3.5 - Exemplo do efeito do filtro da média em uma imagem aérea.



A imagem aérea capturada por uma câmera a bordo de um VANT (à esquerda) foi suavizada com o filtro da média com uma janela de tamanho  $11 \times 11$ , gerando a imagem à direita.

Fonte: Produção do autor.

### 3.3.2 Filtro de Gauss

O filtro de Gauss é definido de forma semelhante ao filtro da média, com uma janela seguida por um termo de normalização. Os coeficientes da janela, porém, não são todos iguais, mas seguem a distribuição dada por (GONZALEZ; WOODS, 2008)

$$w(u, v) = Ke^{-\frac{u^2+v^2}{2\sigma^2}} \quad (3.8)$$

em que  $u$  e  $v$  são valores inteiros e correspondem aos índices da janela,  $K$  é uma constante e  $\sigma$  é o desvio padrão. A escolha de  $K$  se torna irrelevante com o termo de normalização da janela. E a escolha do valor de *sigma* deve ser feito de acordo com o tamanho da janela para evitar que janelas de diferentes tamanhos produzam o mesmo resultado (GONZALEZ; WOODS, 2008).

Assim como o filtro da média, o filtro Gaussiano produz a suavização da imagem. A

diferença se dá em como essa suavização é feita. Enquanto o filtro de média cria uma suavização linear na qual há diferenças de intensidade na imagem, o filtro Gaussiano apresenta uma suavização mais forte e mais uniforme (GONZALEZ; WOODS, 2008). A Figura 3.6 ilustra o efeito causado pelo filtro de Gauss ao ser aplicado em uma imagem capturada por uma câmera acoplada a um VANT.

Figura 3.6 - Exemplo do efeito do filtro de Gauss em uma imagem aérea.



À imagem aérea capturada por uma câmera a bordo de um VANT (à esquerda) foi aplicado um filtro Gaussiano com uma janela de tamanho  $11 \times 11$  e  $\sigma = 2$ , gerando a imagem à direita.

Fonte: Produção do autor.

### 3.3.3 Filtro da Mediana

Diferentemente do filtro da média e de Gauss, o filtro da mediana não é linear. O cálculo realizado envolve a ordenação de um conjunto de *pixels* de uma região da imagem e a seleção do valor central do conjunto para compor o novo *pixel* (GONZALEZ; WOODS, 2008). Embora o processo de convolução não seja usado, a janela pode ser utilizada como forma de visualização do cálculo. Ao percorrer uma imagem com uma janela de tamanho arbitrário, cada novo *pixel* recebe o valor da mediana do conjunto de *pixels* envolvido pela janela da imagem original.

O filtro da mediana é efetivo na eliminação de ruídos conhecidos como sal-e-pimenta (do inglês, *salt-and-pepper*), que podem ser descritos como pontos brancos e pretos presentes na imagem. Na Figura 3.7 pode ser visto uma imagem aérea capturada por um VANT, a mesma imagem com ruído do tipo sal-e-pimenta artificialmente adicionado e a imagem resultante após a utilização do filtro da mediana na imagem com ruído.



Figura 3.7 - Exemplo do efeito do filtro da mediana em uma imagem aérea com ruído.



À imagem aérea capturada por uma câmera acoplada a um VANT (à esquerda) foi adicionado, de forma artificial, o ruído do tipo sal-e-pimenta, gerando uma nova imagem com pontos brancos e pretos (ao centro), que foi filtrada com o filtro da mediana de tamanho  $3 \times 3$ , gerando a imagem à direita.

Fonte: Produção do autor.

### 3.4 Detecção de Bordas

Algoritmos de segmentação de imagens são baseados em uma de duas características básicas presentes em imagens: a continuidade e a descontinuidade (GONZALEZ; WOODS, 2008). Este último assume que a imagem possui diferenças entre regiões com intensidade suficientemente grandes para que seus limites possam ser detectados. Tais limites definem as regiões presentes dentro de uma imagem e são chamados de bordas ou contornos.

Quando a intensidade de um *pixel* muda abruptamente, ela é chamada de *pixel* de borda, indicando que naquela parte específica da imagem existe uma transição entre regiões. Conjuntos interligados de *pixels* de bordas são chamados de segmentos de bordas ou apenas de bordas. A operação executada em imagens para identificar as bordas é chamada de detecção de bordas ou extração de bordas (GONZALEZ; WOODS, 2008).

Como o contorno de uma região da imagem é demarcado pela diferença de intensidade dos seus *pixels*, o processo de detecção de bordas pode ser realizado por meio de derivadas de primeira e segunda ordem (GONZALEZ; WOODS, 2008). É possível calcular as derivadas de uma imagem digital, i.e., de um sinal discreto, utilizando o método das diferenças finitas. Escolhendo a forma da diferença central, pode-se definir

$$\frac{\partial f(x, y)}{\partial x} = \frac{f(x + 1, y) - f(x - 1, y)}{2}, \quad (3.9)$$

para a derivada no eixo  $x$ , i.e., o eixo horizontal. O mesmo vale para o eixo vertical, desde que seja feita a devida alteração para a variável  $y$ . O cálculo das derivadas numéricas pode ser então realizado por meio da convolução (ver Seção 3.3) da imagem com uma janela (GONZALEZ; WOODS, 2008). A partir das derivadas, é possível calcular o gradiente da imagem, definido como o vetor

$$\nabla f(x, y) \equiv \begin{bmatrix} g_x(x, y) \\ g_y(x, y) \end{bmatrix} = \begin{bmatrix} \frac{\partial f(x,y)}{\partial x} \\ \frac{\partial f(x,y)}{\partial y} \end{bmatrix}. \quad (3.10)$$

E, a partir do gradiente, a força (ou magnitude) pode ser calculada pela norma Euclidiana do vetor,

$$\|\nabla f(x, y)\| = \sqrt{g_x^2(x, y) + g_y^2(x, y)}. \quad (3.11)$$

Por último, para definir se a magnitude do gradiente é uma borda, pode-se aplicar um limiar, que atribui 255 (branco) a todos os valores acima desse limiar e zero (preto) aos demais valores. Essa etapa também é chamada de binarização, pois a imagem passa a aceitar apenas dois valores, tornando-se uma imagem binária. A escolha do limiar pode ser realizada por métodos automáticos que fazem uso do histograma da imagem (HUGHES et al., 2013).

### 3.4.1 Operador de Sobel

O operador de Sobel, ou operador de Sobel-Feldman, foi proposto no final dos anos 1960 por Irwin Sobel e Gary Feldman como uma forma de calcular o gradiente de uma imagem (SOBEL; FELDMAN, 1973). A operação pode ser definida em duas etapas. Primeiro, deve-se realizar a convolução de uma imagem com duas janelas diferentes para calcular as derivadas aproximadas das variações de intensidade dos *pixels* nos dois eixos da imagem. A janela que calcula as variações da componente horizontal pode ser definida como (SOBEL; FELDMAN, 1973; GONZALEZ; WOODS, 2008)

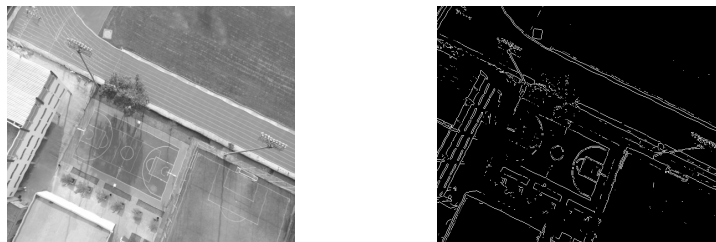
$$g_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad (3.12)$$

e a janela que calcula as variações da componente vertical como

$$g_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}. \quad (3.13)$$

Na etapa seguinte, as duas componentes são unificadas para calcular a magnitude do gradiente da imagem por meio da Equação 3.11. O resultado é uma imagem contendo a demarcação das regiões onde ocorre mudança de intensidade dos *pixels*. Por último, pode-se aplicar um limiar para tornar a imagem binária. Na Figura 3.8 pode ser vista uma imagem capturada por um VANT antes e após a utilização do operador de Sobel para detecção de bordas.

Figura 3.8 - Exemplo do efeito do operador de Sobel-Feldman em uma imagem aérea.



Fonte: Produção do autor.

### 3.4.2 Operador de Canny

O algoritmo para detecção de bordas proposto por John Canny em 1986 (CANNY, 1986), conhecido também como operador de Canny, é mais complexo que o operador de Sobel e busca atingir 3 objetivos básicos (GONZALEZ; WOODS, 2008):

- Alto índice de acerto, em que o maior número de bordas possível deve ser encontrado sem obter falsos positivos;
- As coordenadas indicadas como bordas devem ser precisas, i.e., a distância entre o ponto marcado como borda e a borda verdadeira deve ser mínimo;
- As bordas detectadas devem ser únicas para cada ponto de correspondência na imagem. O detector de bordas não deve indicar mais de um ponto como borda onde só deveria existir um único ponto.

Apesar de ser difícil de atingir os 3 objetivos, o operador de Canny busca satisfazê-los da melhor maneira possível. Nesse sentido, esse operador é um algoritmo sofisticado, que pode ser dividido em 5 etapas (CANNY, 1986; GONZALEZ; WOODS, 2008). Na primeira etapa a imagem é filtrada com um filtro Gaussiano (ver Subseção 3.3.2). Em seguida o gradiente da imagem (Equação 3.10) é calculado utilizando alguma forma de derivada discreta. O terceiro passo é o cálculo da direção, ou ângulo, do gradiente, que pode ser obtido por meio da equação

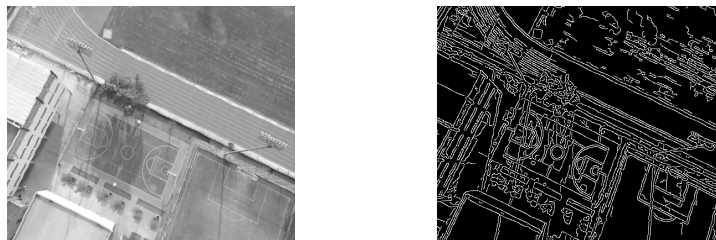
$$\alpha(x, y) = \tan^{-1} \left[ \frac{g_x(x, y)}{g_y(x, y)} \right], \quad (3.14)$$

em que  $g_x$  e  $g_y$  foram calculados no segundo passo. Na etapa seguinte, o gradiente da imagem e a orientação do gradiente são utilizados para reduzir a espessura das bordas. O processo, chamado de supressão não máxima, pode ser resumido em duas partes:

- O ângulo do gradiente de um dado ponto é utilizado para se obter os dois *pixels* mais próximos que estão localizados ao longo daquele mesmo ângulo;
- A intensidade do gradiente dos vizinhos são verificados. Caso algum deles possua intensidade maior, o ponto é suprimido. Caso contrário, o ponto é mantido.

Por último, o quinto passo se resume em aplicar um limiar para definir quais pontos são bordas e quais não são. No caso específico do operador de Canny, dois limiares são aplicados, sendo um superior e outro inferior. Dessa forma, pode-se eliminar as bordas espúrias (ou falsos positivos). Na Figura 3.9 podem ser vistos os efeitos do método de Canny em uma imagem aérea capturada por um *drone*.

Figura 3.9 - Exemplo do efeito do operador de Canny em uma imagem aérea.



Fonte: Produção do autor.

### 3.4.3 Redes Neurais Artificiais

Uma alternativa aos operadores tradicionais, como Canny e Sobel, é o uso de RNAs para detecção de bordas. As RNAs podem ser vistas como filtros não lineares quando comparadas aos filtros espaciais (ver Seção 3.3), em que ocorre a convolução da imagem com uma janela (BRAGA, 2018). A ideia básica por trás das RNAs é o treinamento a partir de dados de entrada predefinidos, que visam aumentar seu desempenho em dada tarefa (ver Seção 3.7.3). No caso da detecção de bordas, o objetivo do treinamento é aumentar a acurácia de detecção das mesmas.

A RNA já treinada e testada foi utilizada para compor o conjunto de técnicas de detecção de bordas, sendo a arquitetura e os parâmetros de treinamento desenvolvidos em outros trabalhos (BRAGA et al., 2016; BRAGA, 2018). Uma MLP foi treinada com o algoritmo de retropropagação de erro, no qual classifica padrões binários de bordas. O treinamento envolveu um conjunto total de 26 padrões diferentes, sendo 24 de bordas e 2 de não bordas. Os padrões definidos para o treinamento da MLP são possíveis combinações de matrizes binárias com tamanho  $3 \times 3$ , uma vez que matrizes maiores implicaria em conjuntos de treinamento muitos extensos devido ao número de combinações possíveis (BRAGA, 2018).

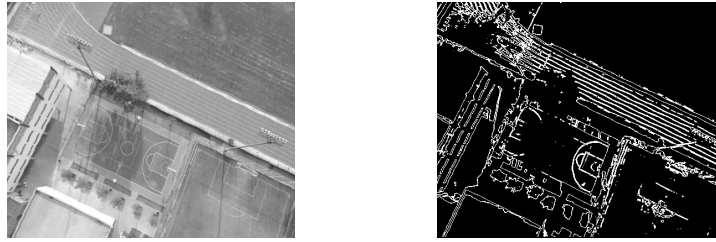
A arquitetura da MLP é composta por 9 neurônios artificiais de entrada (um para cada célula da matriz  $3 \times 3$ ), um neurônio artificial de saída, que classifica se a entrada é uma borda ou uma não borda, e 18 neurônios na camada escondida. A arquitetura foi concebida por meio de uma meta-heurística chamada de Colisão de Múltiplas Partículas (do inglês, *Multiple Particle Collision Algorithm*) (BRAGA, 2018).

A utilização da MLP como extrator de bordas ocorre de maneira semelhante ao processo de convolução (ver Seção 3.3), em que a janela percorre a imagem para classificar os *pixels* em borda ou não borda. No caso da MLP, a janela possui tamanho de  $3 \times 3$  *pixels*, totalizando as 9 entradas aceitas pela rede neural, que faz a classificação. O processo de binarização da imagem é feito utilizando o método de Otsu (BRAGA, 2018). O resultado do processo de detecção de bordas com a MLP descrita pode ser visto na Figura 3.10.

### 3.4.4 Congruência de Fase

A congruência de fase é uma medida adimensional que reflete os comportamento de um sinal no domínio da frequência. A decomposição de uma onda quadrada utilizando séries de Fourier, por exemplo, revela que as ondas senoidais estão em

Figura 3.10 - Exemplo do efeito da detecção de bordas realizado por uma RNA.



Fonte: Produção do autor.

fase, i.e., em máxima congruência, no exato ponto de crescimento e decrescimento da onda quadrada. Nesse sentido, as informações de uma imagem no domínio da frequência podem ser utilizados para detectar onde suas componentes estão em fase e, conseqüentemente, determinar a posição das bordas da imagem (KOVESI, 2000).

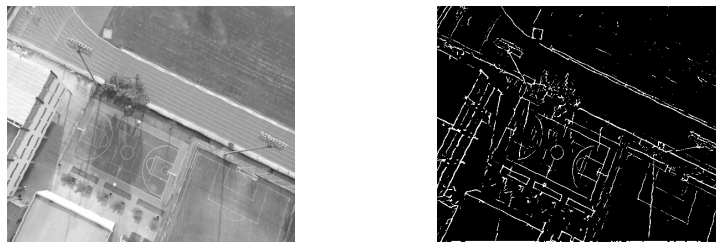
Apesar de ser uma técnica que apresenta excelentes resultados, seu problema crucial é a sensibilidade a ruídos presentes no sinal. Além disso, o esforço computacional desta técnica é consideravelmente maior quando comparado aos operadores tradicionais, como o operador de Sobel. Neste trabalho, a técnica de detecção de bordas por congruência de fase segue as diretrizes propostas por Kovesi (2000) e envolve 4 passos, resumidos a seguir:

- Como a congruência de fase é diretamente proporcional à função de energia local de um sinal, este pode ser utilizado para encontrar os pontos de máxima congruência. Para isso, pode-se realizar a convolução do sinal com filtros de escalas variadas para se obter vetores de resposta. A congruência de fase pode então ser obtida a partir do somatório das escalas dos vetores. O filtro de *Log-Gabor* foi escolhido por ser um modelo conveniente para o problema.
- A congruência de fase é suscetível a ruídos presentes no sinal e, por esse motivo, deve possuir algum mecanismo para corrigir o problema. Ao supor que o ruído é aditivo, o espectro de potência do ruído é constante e as bordas só ocorrem em pontos isolados da imagem, é possível identificar o nível de ruído que influencia a energia do sinal. E, uma vez calculado, é possível usá-lo para suprimir o efeito do ruído no cálculo da congruência de fase.

- A detecção de bordas só terá sentido se a congruência de fase possuir diversas frequências. Caso só haja uma frequência, não é possível avaliar onde ela está em concordância com outras frequências. Para evitar que certas frequências influenciem a detecção da borda, elas devem ser ponderadas. A função sigmoide foi escolhida para a ponderação, com um parâmetro responsável pelo fator de corte. Dessa forma, respostas espúrias podem ser evitadas e a localização das bordas se torna mais apurada.
- Por último, o cálculo deve ser generalizado para duas dimensões, que pode ser realizado aplicando os filtros unidimensionais em diferentes orientações e, posteriormente, combinar suas respostas.

Dessa forma, é possível aplicar a congruência de fase para detectar bordas em uma imagem digital. Um exemplo do efeito da congruência de fase na detecção de bordas de uma imagem aérea pode ser visto na Figura 3.11.

Figura 3.11 - Exemplo da técnica de congruência de fase na detecção de bordas.



Fonte: Produção do autor.

### 3.5 Registro de imagens

A tarefa de comparar duas imagens que podem ter sido capturadas em momentos diferentes, em posições diferentes e com sensores diferentes pode ser realizada por meio do registro de imagens. Nesse sentido, registro de imagens pode ser entendido como a transformação que permite que as informações de uma imagem sejam comparadas a outra. O processo consiste em definir uma das imagens como referência e aplicar transformações geométricas na outra de forma que ambas fiquem alinhadas (BROWN, 1992). Aqui, vale lembrar que as transformações geométricas de escala e rotação foram realizadas em etapas anteriores, o que define esta etapa como responsável pelo cálculo da matriz de translação entre as duas imagens.

Uma forma de utilizar informações presentes em imagens para calcular suas correspondências consiste em escolher uma região de uma imagem (chamada *template*) e usá-la como modelo para rastrear sua posição em outra imagem usando algum tipo de técnica de busca, em que, neste trabalho, são utilizadas técnicas baseadas em correlação cruzada. Essa forma é indicada para sequências de imagens com pouca ou nenhuma alteração da visão da câmera durante a captura (SZELISKI, 2011).

Dessa forma, um *template* pode ser definido como uma região contida em uma imagem, ou até uma imagem inteira, que possui algum tipo de informação capaz de identificá-la. O *template* pode então ser utilizado para estimar a matriz de translação que tornam as duas imagens desalinhadas. O processo ocorre de maneira similar ao da convolução (ver Seção 3.3), em que uma janela é aplicada à imagem referência de forma a gerar uma nova imagem, sendo a janela o próprio *template* e a nova imagem o resultado da correlação. Aqui, o *template* será a imagem capturada em voo por uma câmera acoplada a um VANT, enquanto a imagem de busca, ou imagem referência, será a imagem georreferenciada utilizada como mapa da região sobrevoada. Nesse sentido, a matriz de translação entre as duas imagens pode ser estimada pela técnica de correlação cruzada e afins.

### 3.5.1 Correlação Cruzada

A correlação é utilizada desde os anos 1970 para rastrear e localizar regiões em imagens (HANNAH, 1974). Ao longo do tempo sofreu modificações, como a redução do custo computacional (LEWIS, 1995) e o suporte para uso de máscaras (PADFIELD, 2010). Apesar de ser uma técnica muito importante na área de processamento de imagens, a correlação não é tolerante a variação de escala e rotação, i.e., as imagens submetidas ao seu uso devem possuir aproximadamente a mesma escala e estar orientadas na mesma direção.

A correlação cruzada (CC) pode ser descrita como uma operação matemática *pixel* a *pixel* entre um *template* e uma imagem e está diretamente relacionada à distância Euclidiana. Esta, por sua vez, pode ser definida, entre dois objetos bidimensionais (e.g., duas imagens), como o somatório do quadrado das diferenças dos valores no espaço selecionado (LEWIS, 1995), ou, matematicamente, como

$$d_{f,t}^2(u, v) = \sum_{x,y} [f(x, y) - t(x - u, y - v)]^2, \quad (3.15)$$

em que  $f$  é a imagem na qual ocorre a busca e  $(x, y)$  o espaço de busca da imagem



sob o *template*  $t$ , localizado na posição  $(u, v)$ . A Equação 3.15 pode ser expandida para

$$d_{f,t}^2(u, v) = \sum_{x,y} [f^2(x, y) - 2f(x, y)t(x - u, y - v) + t^2(x - u, y - v)]. \quad (3.16)$$

Supondo que os termos  $\sum f^2(x, y)$  e  $\sum t^2(x - u, y - v)$  são constantes, então o restante da equação

$$cc(u, v) = \sum_{x,y} [f(x, y)t(x - u, y - v)] \quad (3.17)$$

pode ser considerado uma medida de similaridade entre a imagem e o *template*, i.e., quanto maior o valor de  $cc(u, v)$ , maior a semelhança entre o *template* e a imagem em que ocorre a busca na posição  $(u, v)$ .

Embora a Equação 3.17 seja capaz de definir a CC matematicamente, no quesito de implementação ainda é necessário definir quais são os valores de  $x$  e  $y$ . Como correlação é calculada para todos os elementos sobrepostos entre o *template* e a imagem de busca, dado um *template* de dimensão  $M$  por  $N$  e a imagem de busca  $P$  por  $Q$ , a matriz de correlação terá tamanho  $M + P - 1$  por  $N + Q - 1$ . Levando em conta essa sobreposição, a Equação 3.17 deve ser modificada para

$$cc(u, v) = \sum_{x=u-M+1}^u \sum_{y=v-N+1}^v [f(x, y)t(x - u + M, y - v + N)], \quad (3.18)$$

em que  $f(x, y) = 0$  para  $x \geq P + 1$  ou  $x \leq 0$  e  $y \geq Q + 1$  ou  $y \leq 0$ . Por questões de simplicidade e legibilidade, a Equação 3.17 será utilizada para derivar as técnicas subsequentes.

Uma vez calculada a matriz de correlação, é esperado que a posição correta da região de interesse seja indicada pelo seu maior valor. Porém, isso nem sempre acontece. Às vezes, ao contrário do que foi assumido, o termo  $\sum f^2(x, y)$  não é constante, i.e., a imagem em que ocorre a busca pode apresentar uma região com mais brilho (*pixels* com valor mais alto) do que as outras regiões. Dessa forma, o termo  $\sum f^2(x, y)$  poderá influenciar o cálculo e indicar a posição do *template* errada. Um outro problema da Equação 3.17 é que os valores de  $cc(u, v)$  não são normalizados, o que pode dificultar a comparação dos resultados entre imagens de diferentes tamanhos.

Embora apresente algumas limitações, a CC é uma técnica simples e pode ser facilmente programada. No domínio do tempo, o uso de múltiplos laços aninhados para realizar as multiplicações entre as duas imagens eleva o custo computacional. Nesse caso, a alternativa é a implementação no domínio da frequência por meio do uso da transformada discreta de *Fourier*, que pode ser realizada pela convolução entre o *template* e a imagem de busca (GONZALEZ; WOODS, 2008). No domínio da frequência, a CC pode ser calculada como

$$\text{conv}(f, t') = \mathcal{F}^{-1}\{\mathcal{F}(f)\mathcal{F}(t')\}, \quad (3.19)$$

em que  $\mathcal{F}$  é a transformada de *Fourier*,  $f$  é a imagem e  $t'$  é o *template* rotacionado em  $180^\circ$ .

### 3.5.2 Correlação Cruzada Normalizada

De acordo com (NAKHMANI; TANNENBAUM, 2013), a correlação cruzada normalizada (CCN) pode ser definida por

$$\text{ccn}(u, v) = \frac{\sum_{x,y} [f(x, y)t(x - u, y - v)]}{\sqrt[2]{\sum_{x,y} f(x, y)^2 \sum_{x,y} t(x - u, y - v)^2}}, \quad (3.20)$$

que é baseada na Equação 3.17, porém com um novo termo no denominador responsável pela normalização. O resultado da equação é um valor entre 0 e 1, sendo 1 a maior similaridade possível entre *template* e imagem.

Embora a correlação normalizada seja uma sofisticação da CC, ela não soluciona o problema do brilho da imagem, uma vez que o termo  $\sum f^2(x, y)$  ainda não é levado em consideração no cálculo. Dessa forma, uma nova técnica foi proposta, sendo denominada de correlação cruzada normalizada de média zero (CCNZ) (NAKHMANI; TANNENBAUM, 2013), ou também de simplesmente de correlação cruzada normalizada (LEWIS, 1995; GONZALEZ; WOODS, 2008; PADFIELD, 2010; SZELISKI, 2011). Neste trabalho, a nomenclatura usada será diferente para os dois tipos de correlação, sendo a CCNZ definida como

$$\text{ccnz}(u, v) = \frac{\sum_{x,y} \{[f(x, y) - \bar{f}_{u,v}][t(x - u, y - v) - \bar{t}]\}}{\sqrt[2]{\sum_{x,y} [f(x, y) - \bar{f}_{u,v}]^2 \sum_{x,y} [t(x - u, y - v) - \bar{t}]^2}}, \quad (3.21)$$

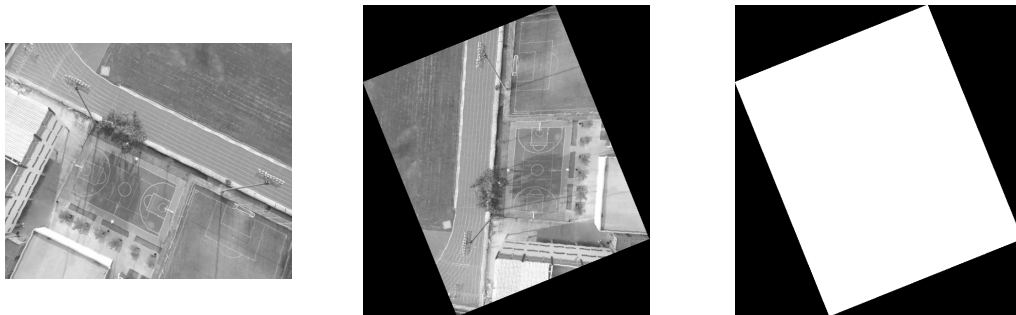
em que  $\bar{f}_{u,v}$  é a média no espaço da imagem de busca sob  $t$  e  $\bar{t}$  é a média do *template*. A principal vantagem de se utilizar a Equação 3.21 para calcular a medida de similaridade é que ela leva em consideração a mudança de brilho da imagem. Quando o valor médio dos *pixels* da imagem e do *template* são incluídas na equação, o cálculo passa a considerar apenas a diferença entre os valores e não seu valores absolutos. Além disso, o resultado da equação se torna um valor entre  $-1$  e  $1$ , sendo  $-1$  completamente anticorrelacionado e  $1$  completamente correlacionado.

Apesar de solucionar o problema da diferença de brilho entre as imagens, a CCNZ possui um custo computacional extra quando comparado à CC. Para reduzi-lo, foi proposta a utilização de somas parciais e transformadas de *Fourier* (LEWIS, 1995).

### 3.5.3 O uso de máscaras

O uso de máscaras em imagens se faz necessário em muitas aplicações do mundo real (PADFIELD, 2010). Em processamento de imagens, as máscaras podem ser definidas como matrizes lógicas bidimensionais responsáveis por demarcar a região da imagem a ser utilizada, em que 1 representa um valor de interesse e 0 um valor a ser ignorado. Um exemplo de utilização de uma máscara pode ser visto na Figura 3.12.

Figura 3.12 - Exemplo do uso de máscara para demarcar a região válida de uma imagem.



A imagem aérea capturada por um VANT (à esquerda) ao ser rotacionada em  $112^\circ$  (imagem do meio) ocasiona a criação de regiões pretas como consequência do *padding*. Para evitar esses valores, uma máscara binária (à direita) pode ser utilizada para demarcar a região válida da imagem.

Fonte: Produção do autor.

Ao rotacionar a imagem original (imagem mais à esquerda da Figura 3.12), obtém-se uma nova imagem, rotacionada e acrescida de uma borda (imagem do meio da

Figura 3.12). Esta borda é resultado do “enchimento” lateral para que a imagem mantenha sua forma retangular e, por isso, é chamada de preenchimento ou, em inglês, *padding*. Por não fazer parte da imagem original, o preenchimento não pode ser considerado nos cálculos e, por isso, é necessária a utilização de uma máscara que indique quais valores devem ser utilizados e quais devem ser ignorados (imagem mais à direita da Figura 3.12).

No cálculo da CC (Equação 3.17), todos os valores inválidos da imagem também devem ser desconsiderados. Para isso, pode-se observar a máscara da imagem,  $M_f$ , e a máscara do *template*,  $M_t$ . Tomando  $M = M_f \cap M_t$  para todos os valores de  $(u, v)$ , a Equação 3.17 pode ser reescrita da seguinte forma

$$cc(u, v) = \sum_{x, y \in M} [f(x, y)t(x - u, y - v)] \quad (3.22)$$

e a Equação 3.21 como

$$ccnz(u, v) = \frac{\sum_{x, y \in M} \{[f(x, y) - \bar{f}_{x, y}][t(x - u, y - v) - \bar{t}_{x, y}]\}}{\sqrt{\sum_{x, y \in M} [f(x, y) - \bar{f}_{x, y}]^2 \sum_{x, y \in M} [t(x - u, y - v) - \bar{t}_{x, y}]^2}}. \quad (3.23)$$

Implementar as Equações 3.22 e 3.23 no domínio temporal é trivial, porém o número de operações torna o esforço computacional do processo muito alto.

### 3.5.4 Correlação Cruzada com Máscara

No domínio da frequência, a CC com máscara pode ser implementada da mesma maneira que a CC sem máscara. A única diferença é que todos os valores a serem ignorados demarcados pela máscara devem ser zerados antes de iniciar o cálculo da convolução.

Por outro lado, a implementação com máscara da versão da CCN e da CCNZ é mais complexa. Isto se dá por 2 fatores. Primeiro, os valores do *template* variam de acordo com a região de interesse da imagem de busca como consequência da sobreposição entre as máscaras. O mesmo vale para os valores no espaço da imagem de busca,  $\bar{f}_{x, y}$ . Segundo, ao não utilizar o *padding*, o número de valores utilizados no cálculo pode não ser suficiente para se atingir um resultado estatisticamente confiável (PADFIELD, 2010).

O primeiro problema pode ser resolvido ao manipular a Equação 3.23 de forma a calcular seus somatórios utilizando convoluções (ver Apêndice C). Dessa forma, com o custo total de 6 convoluções, o numerador e o denominador podem ser calculados para todos os valores de  $(u, v)$ , o que coloca o custo computacional da versão da CCNZ com máscara no mesmo nível de complexidade da versão sem máscara (PADFIELD, 2010).

O segundo problema é a consequência do uso de máscaras, em que o *padding* não é contabilizado no cálculo. Dessa forma, pode ocorrer da quantidade de elementos resultante da intersecção entre as máscaras não ser suficiente para se obter um resultado estatisticamente confiável. Para solucionar esse problema, é necessário aplicar uma função que relacione o resultado da correlação e a quantidade de valores utilizados no cálculo. Um exemplo é aplicar uma função de corte, em que os valores abaixo de um limiar  $\delta$  são zerados, i.e.,

$$ccnz^*(u, v) = \begin{cases} ccnz(u, v), & \text{para } q(u, v) \geq \delta \\ 0, & \text{para } q(u, v) < \delta \end{cases} \quad (3.24)$$

em que  $q$  é a quantidade de elementos envolvidos no cálculo da convolução, e que pode ser calculada pela Equação C.7 do Apêndice C.

Aplicando o mesmo raciocínio para a CCN (ver Equação 3.20), pode-se chegar a conclusão de que é possível resolvê-la com o uso de 3 convoluções. É mais uma convolução extra para calcular a quantidade de elementos na intersecção das máscaras e aplicar o limiar para eliminar os resultados estatisticamente não confiáveis. Dessa forma, é possível calcular a CCN com máscara no domínio da frequência utilizando um total de 4 convoluções.

Por último, é possível calcular a distância Euclidiana (Equação 3.15) com máscara no domínio da frequência. A Equação expandida, 3.16, possui 3 somatórios que podem ser resolvidos com 3 convoluções. Adicionando a convolução responsável por calcular o número de elementos na intersecção das máscaras, o cálculo da distância Euclidiana com máscara no domínio da frequência pode ser resolvido com um total de 4 convoluções.

### 3.6 Cálculo da posição

A matriz de translação entre a imagem capturada pelo *drone* e a imagem referência pode ser usada para calcular a posição da imagem capturada pela câmera acoplada ao VANT. Porém, a posição da imagem e a posição do VANT podem ser diferentes, pois, embora a câmera esteja em posição nadir, está presa à estrutura da aeronave e, conseqüentemente, sofre os movimentos realizados por esta. Esse problema pode ser resolvido com o uso de um *gimbal*, que é um acessório responsável por estabilizar a câmera de forma a eliminar os movimentos indesejáveis durante a captura das imagens. Porém, muitos *drones* não estão equipados com *gimbal*, o que torna essa etapa da técnica necessária.

A atitude de um corpo no espaço tridimensional pode ser definida por 3 ângulos perpendiculares entre si (GROVES, 2013; KRAUS, 2007). A rotação do corpo pode então ser calculada combinando as rotações individuais de cada um dos eixos. Escolhendo  $\omega$  como ângulo de arfagem,  $\phi$  como rolagem e  $\kappa$  como guinada, pode-se definir a rotação de um corpo como (KRAUS, 2007)

$$R_\omega = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\omega) & -\text{sen}(\omega) \\ 0 & \text{sen}(\omega) & \cos(\omega) \end{bmatrix}; \quad (3.25)$$

$$R_\phi = \begin{bmatrix} \cos(\phi) & 0 & \text{sen}(\phi) \\ 0 & 1 & 0 \\ -\text{sen}(\phi) & 0 & \cos(\phi) \end{bmatrix}; \quad (3.26)$$

$$R_\kappa = \begin{bmatrix} \cos(\kappa) & -\text{sen}(\kappa) & 0 \\ \text{sen}(\kappa) & \cos(\kappa) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (3.27)$$

A rotação de um corpo em torno dos 3 eixos pode ser calculada pela combinação

$$R = R_\omega R_\phi R_\kappa = \begin{bmatrix} R_{1,1} & R_{1,2} & R_{1,3} \\ R_{2,1} & R_{2,2} & R_{2,3} \\ R_{3,1} & R_{3,2} & R_{3,3} \end{bmatrix}. \quad (3.28)$$

Para se obter o deslocamento entre o ponto projetado pela câmera e a posição do

VANT, basta utilizar os elementos da matriz da Equação 3.28

$$\begin{bmatrix} Des_x \\ Des_y \end{bmatrix} = \begin{bmatrix} H \cdot R_{1,3}/R_{3,3} \\ H \cdot R_{2,3}/R_{3,3} \end{bmatrix} \quad (3.29)$$

em que  $Des_x$  e  $Des_y$  são os deslocamentos no eixo  $x$  e  $y$  respectivamente, e  $H$  é a altitude da aeronave.

### 3.7 Aprendizado de máquina

O ato de classificar é utilizado em diversas atividades humanas e permite a tomada de decisões frente a uma situação nova baseada em informações passadas. Além disso, a análise (e mineração) de dados permite que novas informações sejam encontradas, assim como padrões e tendências que podem não estar inicialmente visíveis. Esse processo é denominado de “aprender com os dados” (HASTIE et al., 2009). De forma geral, pode-se dividir o processo de aprendizado em dois: dado um montante de dados, deseja-se agrupá-los em conjuntos com características semelhantes; ou, dado um montante de dados, deseja-se construir uma regra capaz de classificar uma nova informação em uma das classes preexistentes. Enquanto o primeiro é conhecido como Aprendizagem não Supervisionada, o segundo é chamado de Aprendizagem Supervisionada (HAYKIN, 2009).

Diferentes abordagens surgiram ao longo dos anos buscando atingir tais expectativas (HASTIE et al., 2009). Segundo Kotsiantis (2007), os classificadores podem ser agrupados em diferentes tipos de técnicas, como as baseadas em lógica, em redes neurais e em estatística. Aqui, o termo aprendizado de máquina (do inglês, *Machine Learning*) será utilizado para incluir qualquer técnica de classificação independentemente de sua vertente.

Três técnicas de aprendizado de máquina utilizadas nos capítulos subsequentes são apresentadas ao longo desta seção. Na Subseção 3.7.1, é apresentado o conceito, de forma geral, da árvore de classificação. Em seguida, na Subseção 3.7.2, são discutidos os aspectos básicos das Máquinas de Vetores de Suporte. Por último, na Subseção 3.7.3 é apresentada a estrutura genérica das RNAs, bem como a estrutura que caracteriza um Perceptron de Múltiplas Camadas.

### 3.7.1 Árvore de classificação

Classificadores baseados em árvores são fáceis e rápidos de serem construídos e resultam em modelos muitas vezes facilmente interpretáveis por seres humanos. Além disso, eles aceitam misturas de dados numéricos e categóricos, e apresentam alta resistência à inclusão de variáveis de entrada irrelevantes, i.e., variáveis que não contribuem de maneira significativa para uma melhor classificação dos dados. Por outro lado, árvores de classificação raramente proveem alta acurácia de classificação quando comparada a outras técnicas mais sofisticadas (HASTIE et al., 2009).

Uma árvore é um grafo direcionado acíclico, o que confere a característica de possuir apenas um caminho entre qualquer par de nós. O nó do qual partem os caminhos que alcançam todos os demais nós é chamado de raiz. O nó raiz, ou apenas raiz, é caracterizado por não possuir parentes, i.e., não possuir nenhum outro nó cujas arestas estão direcionadas a ele. Qualquer outro nó da árvore possui ao menos um parente e nós que não possuem descendentes são chamados de folhas (BISHOP, 2006).

Métodos de classificação baseados em árvore dividem o espaço amostral em subgrupos por meio de regras de decisão. Cada subgrupo do espaço amostral pode ser subdividido em novos grupos em um processo recorrente. Esse processo de subdivisões ocorre até que alguma regra de parada seja satisfeita, tal como a classificação correta de todo espaço amostral ou o alcance de um comprimento máximo da árvore (HASTIE et al., 2009).

Um exemplo de árvore de classificação binária e a divisão causada no espaço amostras bidimensional pode ser visto na Figura 3.13.

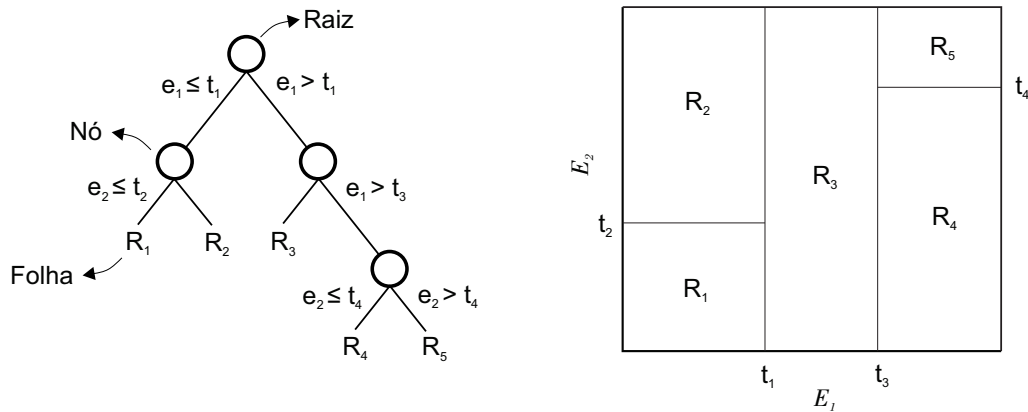
A parte mais difícil no treinamento de uma árvore de classificação se dá na escolha das regras que separam as amostras. As regras devem definir quais são os pontos de separação que melhor dividem as amostras em grupos. Uma forma de definir tais regras é por meio da medida de impureza dos nós, que contabiliza a proporção de amostras por classe para uma dada região (HASTIE et al., 2009).

Tomando o conjunto amostral  $\{(e_i, d_i)\}_{i=1}^N$ , em que  $e_i$  é um vetor de observações e  $d_i$  é a classe a qual pertence, a proporção de observações pertencentes a uma classe  $k$  em um nó  $m$  da árvore é dada por (HASTIE et al., 2009)

$$p_{mk} = \frac{1}{N_m} \sum_{e_i \in R_m, d_i=k} 1, \quad (3.30)$$



Figura 3.13 - Exemplo de árvore de classificação binária.



A árvore de classificação (à esquerda) é composta por 4 nós e 5 folhas e particiona o espaço amostral em 5 regiões diferentes, conforme ilustrado no quadro da direita.

Fonte: Adaptado de Hastie et al. (2009).

em que  $R_m$  é a região do espaço amostral contendo  $N_m$  amostras. Diferentes medidas de impureza podem ser calculadas a partir da Equação 3.30, tal como o índice de Gini:

$$Gini = \sum_{k=1}^K p_{mk}(1 - p_{mk}), \quad (3.31)$$

em que  $K$  é o total de classes do conjunto amostral do nó  $m$ . Por último, basta ponderar as medidas de impureza pelo número de amostras de cada nó descendente após a divisão. As regras de classificação são então escolhidas de forma a minimizar a medida de impureza (HASTIE et al., 2009).

### 3.7.2 Máquinas de Vetor de Suporte

Máquinas de Vetor de Suporte (*Support Vector Machines* - SVM) são classificadores cuja regra de decisão está baseada na construção de um subespaço afim plano, i.e., um hiperplano, que busca maximizar a margem de separação do espaço amostral de treinamento. SVMs são baseados no conceito de vetores de suporte, que formam a base para a construção da regra de decisão e também o princípio básico por trás do *kernel* da técnica. Além disso, os vetores de suporte têm papel fundamental na etapa de treinamento da técnica, que realiza o ajuste do hiperplano por meio de um problema de otimização convexo (BISHOP, 2006; HAYKIN, 2009).

Para compreender o funcionamento da SVM, supõe-se, inicialmente, um conjunto amostral  $\{(e_i, d_i)\}_{i=1}^N$ , em que  $e_i$  é um vetor contendo as características de entrada da  $i$ -ésima amostra e  $d_i$  é a classe de saída desejada para a mesma amostra. Supondo um subconjunto linearmente separável com duas classes,  $d_i = 1$  ou  $d_i = -1$ , a equação que separa os dois padrões mediante um hiperplano é dada por

$$w \cdot e + v = 0, \quad (3.32)$$

em que  $w$  é o vetor de pesos,  $e$  é o vetor de entrada,  $w \cdot e$  é o produto escalar entre os dois vetores e  $v$  é um viés. A ideia fundamental por trás da técnica de SVM é obter um hiperplano ótimo que separe as amostras com a maior margem possível, i.e., definir  $w_0$  e  $v_0$  de modo a maximizar a separação entre as classes do espaço amostral.

O subconjunto de pontos do espaço amostral que se localiza exatamente no limite da margem do hiperplano é chamado de vetores de suporte e são responsáveis por definir os parâmetros do hiperplano ótimo (HAYKIN, 2009). Todas as amostras restantes são irrelevantes para o processo. Os vetores de suporte são formados pelas amostras que estão o mais próximo possível do hiperplano ótimo e, além de formarem um subconjunto de pontos difíceis de serem classificados, determinam de maneira fundamental a localização do plano que separará as classes (BISHOP, 2006).

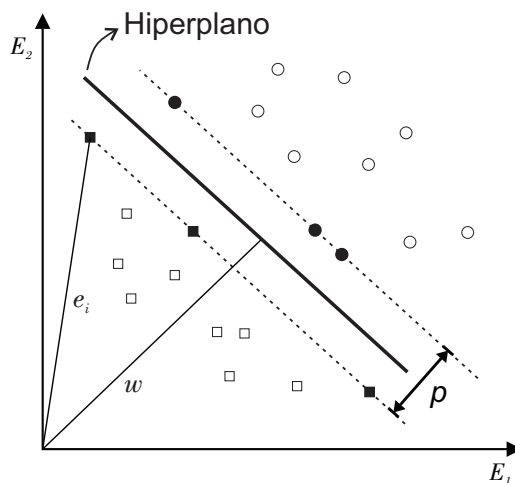
Escolhendo  $p$  como o valor ótimo para a margem que separará as classes do conjunto amostral, pode-se demonstrar que (HAYKIN, 2009)

$$p = \frac{2}{\|w_0\|}. \quad (3.33)$$

Nesse sentido, para maximizar  $p$ , deve-se minimizar a norma euclidiana do vetor de pesos,  $w_0$ . Um exemplo, em duas dimensões, de hiperplano dividindo um conjunto de dados linearmente separável pode ser visto na Figura 3.14.

A partir das Equações 3.32 e 3.33, pode-se formalizar o problema como um problema de otimização convexo, o que torna qualquer solução local do problema uma solução também global (BISHOP, 2006). A solução é dada por meio dos multiplicadores de Lagrange (HAYKIN, 2009)

Figura 3.14 - Exemplo de hiperplano em um conjunto de dados linearmente separável.



As duas classes linearmente separáveis do subconjunto amostral são ilustradas por quadrados e círculos. Em um espaço bidimensional, o hiperplano é representado por uma reta. A margem é delimitada pelas linhas tracejadas com distância  $p$  uma da outra, que tocam os vetores de suporte, representados pelos círculos e quadrados sólidos.

Fonte: Adaptado de [Haykin \(2009\)](#).

$$w_0 = \sum_{i=1}^{N_L} \alpha_{0,i} d_i e_i, \quad (3.34)$$

em que  $N_L$  é o número de vetores de suporte para os quais os multiplicadores de Lagrange ótimos,  $\alpha_{0,i}$ , são não nulos. E o viés,  $v_0$ , é dado por

$$v_0 = 1 - w_0 e_v, \quad (3.35)$$

em que  $e_v$  são os vetores de suporte do conjunto amostral nos quais os multiplicadores de Lagrange,  $\alpha_{0,i}$ , são não nulos.

Por último, é possível realizar uma transformação não linear a partir do vetor de entrada para um outro espaço e, assim, calcular o hiperplano ótimo nesse espaço. Tomando  $\{\varphi_j(e)\}_{j=1}^{\infty}$  como um conjunto de funções não lineares que transformam o espaço do conjunto de entrada, o hiperplano que separa as classes no novo espaço pode ser calculado por

$$w_0 = \sum_{i=1}^{N_L} \alpha_i d_i \phi(e_i), \quad (3.36)$$

em que  $\phi(e_i) = [\varphi_1(e_1), \varphi_2(e_2), \dots]^T$  e  $N_L$  é o número de vetores de suporte. A equação pode então ser reescrita como (HAYKIN, 2009)

$$\sum_{i=1}^{N_L} \alpha_i d_i \phi(e_i) \phi(e) = 0, \quad (3.37)$$

em que o produto  $\phi(e_i)\phi(e)$  pode ser representado por  $k(e, e_i)$  e é chamado de *kernel*. O *kernel* é uma função que calcula o produto interno entre dois valores no espaço de entrada após sofrerem a transformação  $\phi$  para um novo espaço. A importância do *kernel* em SVMs está relacionada ao fato de que um conjunto de entrada linearmente inseparável no espaço original pode se tornar linearmente separável em um novo espaço (BISHOP, 2006). Há diferentes tipos de *kernel* que podem ser utilizados em SVMs (HAYKIN, 2009), entre eles, o polinomial,

$$k_p(e, e_i) = (e^T e_i + 1)^\beta, \quad (3.38)$$

em que o grau do polinômio,  $\beta$ , é uma configuração de entrada e pode ser definido pelo usuário. E a função de base radial,

$$k_{rbf}(e, e_i) = \exp\left(-\frac{1}{2\gamma^2} \|e - e_i\|^2\right), \quad (3.39)$$

em que  $\gamma$  é um parâmetro de entrada.

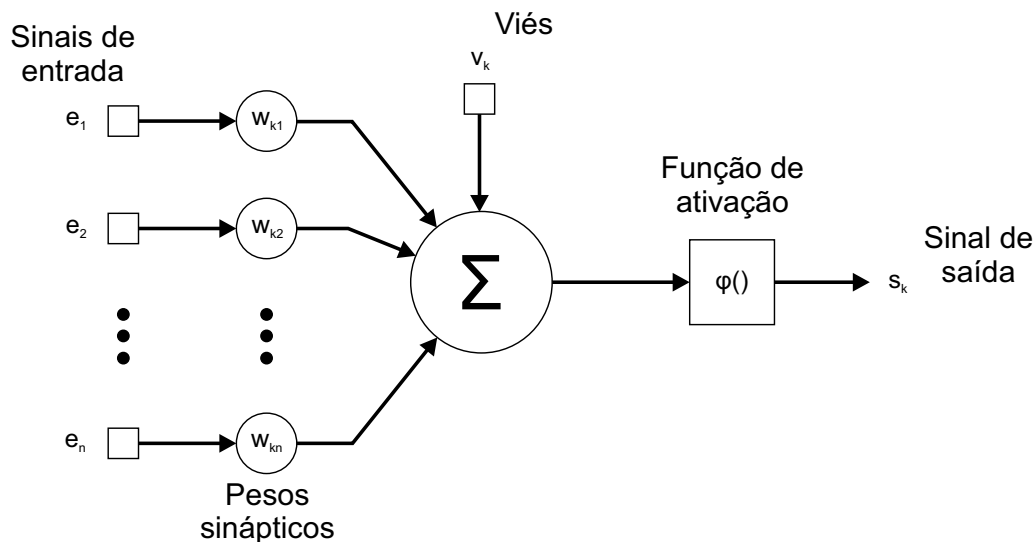
Para problemas linearmente inseparáveis, é possível obter a solução do problema de maneira semelhante, utilizando também multiplicadores de Lagrange. A principal diferença jaz na definição de margem, que se torna flexível, e permite que amostras sejam toleradas no espaço entre os vetores de suporte e o hiperplano, assim como classificações erradas. O problema passa, então, a ser um problema de otimização em que se busca minimizar a probabilidade de classificações erradas. Para isso, é necessário definir parâmetros de entrada responsáveis pela escolha entre o aumento da complexidade do modelo e o esforço por minimizar o número de amostras separadas incorretamente na amostra de treinamento (BISHOP, 2006; HAYKIN, 2009).

### 3.7.3 Redes Neurais Artificiais

RNAs são capazes de modelar, com as restrições e vantagens de uma máquina, a forma como o cérebro humano realiza uma tarefa particular. A construção da RNA cede lugar ao procedimento de aprendizado, cujo objetivo é modificar os pesos sinápticos da rede a fim de treiná-la para um propósito específico. Dessa forma, as RNAs podem ser entendidas como um processador distribuído formado por uma única unidade de processamento que se repete, o neurônio artificial, que é capaz não apenas de armazenar conhecimento, mas também de utilizá-lo em prol de uma tarefa (HAYKIN, 2009).

Inspirado no neurônio biológico humano, o neurônio artificial busca imitar suas funções por meio de somas e multiplicações de sinais. Aqui, vale ressaltar que a complexa organização contendo diferentes níveis de estruturas que o cérebro humano possui não é reproduzida por uma RNA. O neurônio é uma unidade de processamento, que também armazena informação, e é a base para o funcionamento de diferentes tipos de RNAs que podem ser concebidos a partir dela (HAYKIN, 2009). A Figura 3.15 ilustra o modelo de um neurônio artificial não linear.

Figura 3.15 - Modelo não linear de um neurônio artificial.



O neurônio artificial possui basicamente 3 elementos: um conjunto de pesos sinápticos, um adicionador responsável por somar os sinais de entrada e uma função de ativação.

Fonte: Adaptado de Haykin (2009).

O neurônio da Figura 3.15 pode ser representado matematicamente por meio das equações (BISHOP, 2006; HAYKIN, 2009)

$$u_k = \sum_{i=1}^n w_{ki} e_i, \quad (3.40)$$

e

$$d_k = \varphi(u_k + v_k), \quad (3.41)$$

em que, para dado neurônio  $k$ ,  $e_1, e_2, \dots, e_i$  são os sinais de entrada,  $w_{k1}, w_{k2}, \dots, w_{ki}$  são os pesos dos respectivos sinais de entrada,  $v_k$  é o viés,  $\varphi()$  é a função de ativação e  $d_k$  é o sinal de saída do neurônio. O viés,  $v_k$ , possui a capacidade de aumentar ou diminuir o valor de entrada da função de ativação. A função de ativação, por sua vez, determina o valor de saída do neurônio a partir da combinação da soma linear dos sinais,  $u_k$ , e do viés,  $v_k$ . Existem diferentes funções que podem ser utilizadas para realizar a ativação do neurônio (HAYKIN, 2009). Um exemplo de função de ativação é a função de Heaviside, dada por

$$\varphi(h_k) = \begin{cases} 1, & h_k \geq 0; \\ 0, & h_k < 0, \end{cases} \quad (3.42)$$

em que  $h_k = u_k + v_k$ . Também pode-se citar a função logística, dada por

$$\varphi(h_k) = \frac{1}{1 + \exp(-ah_k)}, \quad (3.43)$$

em que  $a$  é o parâmetro que define a inclinação da curva. Quando  $a \rightarrow \infty$ , a função logística se iguala a função de Heaviside. Além das citadas, outras funções podem ser utilizadas, tais como a função identidade e a tangente hiperbólica (BISHOP, 2006; HAYKIN, 2009).

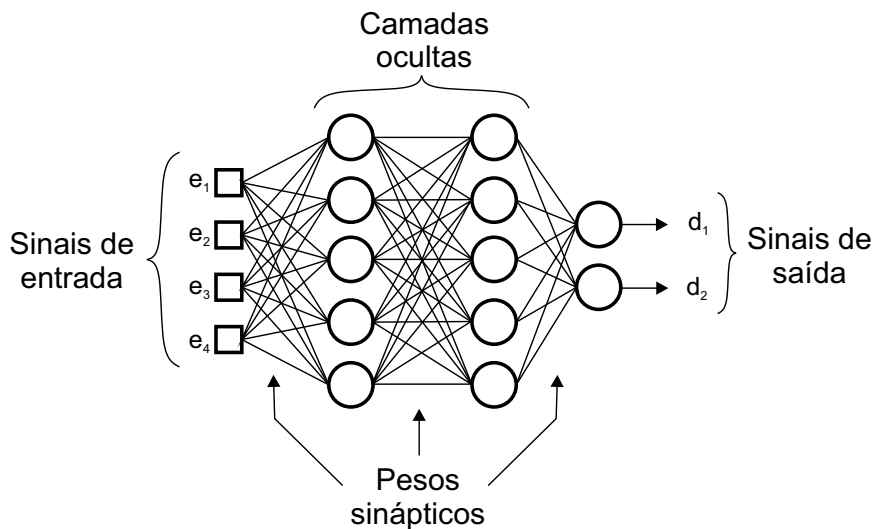
A partir do neurônio artificial, que é a unidade básica a ser utilizada na construção de redes neurais, pode-se construir estruturas com diferentes características. A estrutura das redes influencia o tipo de algoritmo de aprendizado que deverá ser empregado para que sejam treinadas (HAYKIN, 2009).

### 3.7.3.1 Perceptron de Múltiplas Camadas

Uma das redes neurais amplamente conhecida é o Perceptron de Múltiplas Camadas (do inglês, *MultiLayer Perceptron* - MLP) (BISHOP, 2006). As redes que possuem uma ou mais de uma camada oculta de neurônios são chamadas de multicamadas. A camada oculta se refere a toda camada que não pode ser vista diretamente pelos sinais de entrada nem pelos sinais de saída. Com a adição de pelo menos uma camada intermediária, i.e., oculta, espera-se que a influência adicionada contribua de maneira positiva no desempenho da rede (HAYKIN, 2009).

Além das múltiplas camadas, as características básicas de uma MLP exigem que o modelo de cada neurônio da rede possua uma função de ativação não linear diferenciável. Outra característica é o alto grau de conectividade entre os neurônios, cuja extensão é determinada pelos pesos sinápticos da rede (HAYKIN, 2009). Um exemplo de perceptron multicamadas pode ser visto na Figura 3.16.

Figura 3.16 - Estrutura de um perceptron multicamadas.



A MLP possui um total de 2 camadas ocultas, com 5 neurônios em cada uma. A quantidade de camadas ocultas pode variar, assim como a quantidade de neurônios em cada uma delas.

Fonte: Adaptado de Haykin (2009).

Apesar das características da MLP apresentarem reverses, como o fato de tornar o processo de aprendizado difícil de interpretar, é possível obter os pesos sinápticos da

rede por meio de um mecanismo de aprendizado automático. Um método popular para treinamento é o algoritmo de retropropagação, que pode ser dividido em duas etapas. Primeiro, é realizado um processo denominado de processamento direto, em que os pesos sinápticos são fixados e os sinais de entrada são propagados pelas camadas da rede até alcançarem a saída. Na segunda fase, um processamento denominado processamento reverso é realizado, em que um sinal de erro é produzido por meio da comparação da saída desejada com a saída obtida para, então, ser propagado no sentido inverso da rede. Nessa fase, a propagação do erro é responsável por, camada por camada, realizar ajustes sucessivos nos pesos sinápticos da rede. O treinamento ocorre utilizando múltiplas iterações das duas etapas anteriores, chamadas épocas de treinamento (HAYKIN, 2009).

Nesse sentido, a estrutura da MLP pode ser definida pelo usuário, que estipula o número de camadas ocultas, bem como a quantidade de neurônios em cada uma delas, assim como a função de ativação da rede. Os valores dos pesos sinápticos, por sua vez, são obtidos por meio de um algoritmo de treinamento, que, utilizando a técnica de retropropagação do erro, ajusta seus pesos de forma a aumentar a acurácia da classificação da rede. Os parâmetros do algoritmo de treinamento, como o número de épocas de treinamento e a taxa de aprendizado, também podem ser definidas pelo usuário (HAYKIN, 2009).

### **3.8 Considerações finais sobre este capítulo**

O presente capítulo abordou a técnica de estimação da posição de VANTs por meio de imagens. Esta é composta por uma sequência de etapas, em que podem ser empregados diferentes algoritmos de processamento de imagens. Além disso, foram também discutidas as metodologias de aprendizado de máquina que serão utilizadas posteriormente. Ambos tópicos serão necessários nos capítulos subsequentes como partes integrantes da metodologia autoadaptativa, que será abordada no próximo capítulo.



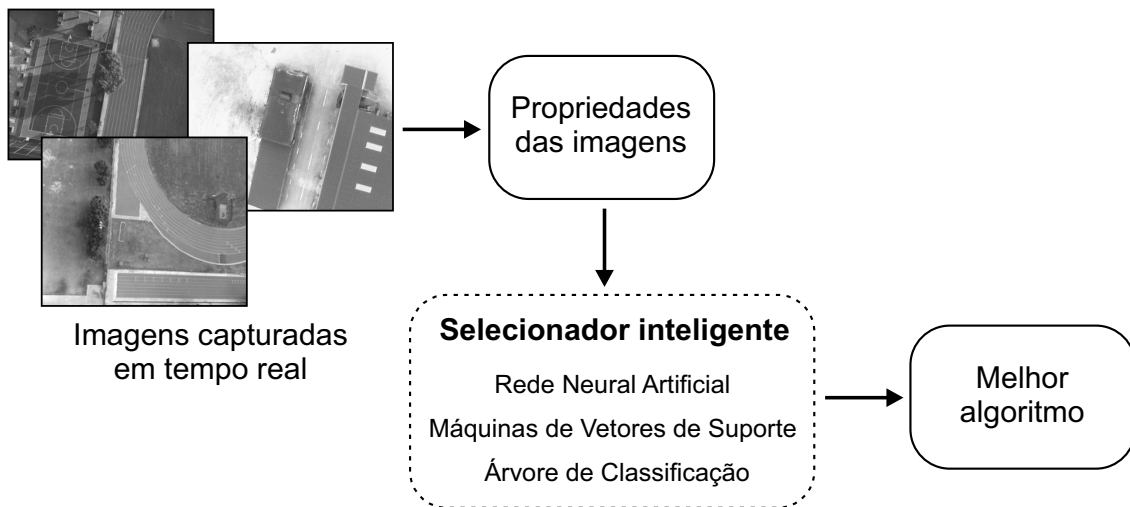
## 4 ABORDAGEM AUTOADAPTATIVA PARA ESTIMAÇÃO DA POSIÇÃO POR IMAGENS

Há diferentes opções de técnicas e abordagens que podem ser empregadas na estimação da posição de VANTs por meio de imagens (ver Seção 2.2.3). As possibilidades aumentam ainda mais se forem consideradas técnicas baseadas em informações providas por outros tipos de sensores (CADENA et al., 2016). Todas essas técnicas buscam resolver o mesmo problema, ou a mesma classe de problemas, que, de forma geral, é estimar a posição de um corpo no espaço. Cada alternativa envolve um conjunto de etapas a serem processadas, assim como um conjunto de sinais providos por sensores, responsáveis por interpretar as informações do ambiente. A sequência de passos, que se inicia com a manipulação dos dados de entrada, como as informações sensorizadas do ambiente, e que se estende até se obter um resultado, os dados de saída, é chamada de algoritmo. Cada algoritmo, definido por sua sequência única de etapas, é caracterizado por um desempenho próprio, mas não único. Nesse sentido, considerar uma única possibilidade de algoritmo frente à diversidade de situações que um problema real pode eventualmente apresentar é buscar sua solução levando em consideração uma única abordagem, independentemente do desempenho que esta apresente.

À medida que algoritmos eram criados para solucionar problemas preexistentes, a escolha do mais propício a ser utilizado em conjuntos específicos de problemas tornou uma ideia cada vez mais clara: não há algoritmo universal que resolva de maneira ideal todo um domínio de problemas (SMITH-MILES, 2009). Obviamente, quando há mais de um algoritmo apto a resolver o mesmo problema, opta-se pelo melhor, i.e., aquele que apresenta o melhor desempenho. Porém, quando o problema muda, o algoritmo a ser utilizado pode mudar também. Nesse contexto, Rice (1976) propôs o modelo abstrato de seleção de algoritmos, no qual busca responder a seguinte pergunta: dado um conjunto de algoritmos capazes de resolver uma classe de problemas, como selecionar o melhor algoritmo?

Uma visão geral da abordagem autoadaptativa é apresentada na Figura 4.1. Como pode ser visto, a ideia central da adaptação está em escolher um algoritmo específico para cada imagem capturada, sendo este o que apresenta melhor desempenho dentre um conjunto de algoritmos preestabelecido. A escolha é feita por meio de um decisor, que é, aqui, composto por técnicas de aprendizado de máquina. Dessa forma, ao adaptar os algoritmos para cada caso, espera-se aumentar o desempenho do sistema como um todo.

Figura 4.1 - Visão geral da abordagem autoadaptativa.



As propriedades extraídas das imagens capturadas por VANTs são utilizadas para que o selecionador inteligente decida qual é o melhor algoritmo a ser empregado para aquele caso específico.

Fonte: Produção do autor.

Nesse sentido, o presente capítulo é dedicado ao modelo de seleção de algoritmos, que é a base teórica na qual se fundamenta o processo de autoadaptação analisado nesta tese, assim como a adequação do mesmo ao problema de estimação da posição de VANTs baseado em imagens. A organização do capítulo se dá da seguinte forma. Na Seção 4.1 é discutido o modelo abstrato para seleção de algoritmos. Na seção seguinte, 4.2, é apresentada a forma como o modelo foi adaptado dentro das diretrizes do problema de estimação da posição de VANTs baseado em imagens. Por último, na Seção 4.3 são apresentadas as considerações finais do capítulo.

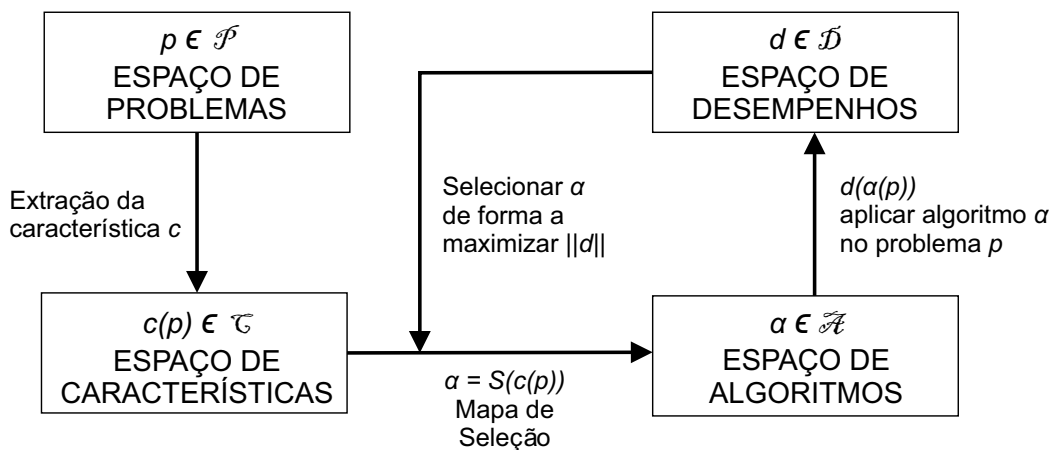
#### 4.1 O problema da seleção de algoritmos

Em 1976, John Rice propôs uma resposta à pergunta de como escolher o melhor algoritmo disponível para uma dada instância de um problema (RICE, 1976). Como resultado, Rice desenvolveu um modelo abstrato de seleção de algoritmos, que conta com 4 espaços teóricos que podem ser adaptados a diferentes problemas reais.

O problema da seleção de algoritmos surge quando há mais de uma possibilidade de escolha, dentro de um conjunto de possibilidades, para se resolver um problema particular. Como a escolha implica diretamente na forma como o problema será re-

solvido, i.e., o algoritmo que será utilizado, deve-se escolher aquela que apresenta melhor desempenho, independentemente de como este último seja representado. A escolha do algoritmo pode ser guiada a partir do aspecto, das características e detalhes, que o problema apresenta. Esse é o raciocínio básico que está por trás do modelo abstrato de Rice, ilustrado na Figura 4.2.

Figura 4.2 - Modelo abstrato para seleção de algoritmos proposto por John Rice.



Cada instância  $p$  do espaço de problemas possui um conjunto de características,  $\{c_1(p), c_2(p), \dots\}$ , que compõe o espaço de características. Estas são associadas a um algoritmo  $\alpha$ , pertencente ao espaço de algoritmos, por meio do mapa de seleção,  $S(c(p))$ . Cada algoritmo apresenta um desempenho  $d$ , que varia de acordo com  $p$ . A ideia central do modelo é associar  $p$  ao algoritmo  $\alpha$  de forma a maximizar  $d$ . O modelo original proposto por Rice (1976) não considera o ciclo formado pelo desempenho  $d$  como fator para definição do mapa de seleção. Por esse motivo, o modelo foi alterado para o proposto por Smith-Miles (2009).

Fonte: Adaptado de Smith-Miles (2009).

Cada um dos 4 espaços do modelo abstrato da seleção de algoritmos está descrito a seguir (RICE, 1976):

- Espaço de problemas: este espaço é composto pelo conjunto de problemas e suas instâncias, o que o torna grande e muito diverso. Aqui, são incluídos não apenas problemas do mundo real, mas também as instâncias dos problemas, que apresentam características particulares do mesmo. As instâncias são importantes para definir as particularidades do problema que causam incrementos ou decrementos no desempenho dos algoritmos.

- Espaço de características: aqui são recolhidas as propriedades que caracterizam as instâncias do problema, i.e., formas de descrever as particularidades dessas instâncias. Este espaço é extremamente importante para que seja possível associar as instâncias do problema a um algoritmo específico.
- Espaço de algoritmos: este espaço reúne o conjunto contendo os algoritmos passíveis de solucionar as instâncias do problema. Assim como o espaço de problemas, é grande e variado. Para Rice, algoritmos idênticos configurados com parâmetros diferentes não são contabilizados como algoritmos distintos.
- Espaço de desempenhos: neste espaço são consideradas as métricas que tornam um algoritmo mais elegível que outro. Aqui, desempenho não pode ser considerado apenas como eficiência computacional ou uso de memória, mas como qualquer medida capaz de classificar algoritmos, inclusive medidas abstratas, como complexidade de implementação e suscetibilidade a erros. O conjunto de desempenhos pode ser composto por mais de um métrica, o que torna necessária a ponderação das métricas para eleger um único melhor algoritmo. Para resolver essa situação, Rice propôs um modelo mais sofisticado contendo também um espaço de critérios composto por pesos que deveriam ser atribuídos a cada um dos desempenhos.

Por ter sido concebido na forma de modelo abstrato e genérico, ele pode ser adaptado a diferentes problemas do mundo real. Um exemplo é o problema da seleção de algoritmos de quadratura para calcular a área de uma função matemática. O espaço de problemas passa então a ser definido como o conjunto de funções de uma variável, cujas instâncias podem ser  $f(x) = x^2$  e  $f(x) = \text{sen}(x)$ . O espaço de características, na busca por descrever a forma e o aspecto das funções, pode incluir a presença de singularidades, descontinuidades, e comportamentos repetitivos. O espaço de algoritmos, por sua vez, é composto por técnicas para calcular a quadratura, como o método de Simpson e a regra de Gauss. O espaço de desempenhos pode ser formado pelo montante de trabalho, i.e., número de passos necessários para se chegar ao resultado, uso de memória, simplicidade de uso, e complexidade de implementação. Por último, deve ser realizado o mapeamento entre o espaço de características e o espaço de algoritmos, em que, hipoteticamente, poderia ser o seguinte: o método de Simpson é indicado para funções que possuem comportamento repetitivo, como a senoidal, pois reduz o montante de trabalho realizado; a regra de Gauss é indicada para funções que não possuem comportamento senoidal, como a quadrática  $f(x) = x^2$ , pelo efeito contrário do mesmo motivo.

Nesse contexto, o processo de mapeamento entre o espaço de características e o espaço de algoritmos é tarefa de crucial importância no problema de seleção de algoritmos. Sem um mapeamento adequado, não é possível associar as instâncias do problema, definida por meio de suas características, ao algoritmo desejado. Uma das chaves para esse problema é escolher propriedades não apenas simples de serem calculadas, mas também que sejam capazes de descrever corretamente as instâncias do problema para que estas possam ser associadas aos algoritmos, i.e., para um domínio particular de problemas, as propriedades devem correlacionar as instâncias deste com o desempenho obtido pelos algoritmos.

Além da definição do espaço de características, ainda é necessário realizar o procedimento chamado de mapeamento ou mapa de seleção. Este pode ser considerado o componente mais importante no problema de seleção de algoritmos, pois é responsável por escolher os algoritmos a serem utilizados em cada caso. O mapa de seleção pode ser visto como um algoritmo, em que é necessário determinar seus parâmetros para que seja capaz de realizar, da melhor maneira possível, a correta associação entre o espaço de características e o espaço de algoritmos. A escolha do algoritmo adequado para realizar o mapeamento é um problema de seleção de algoritmos, o que revela a recursividade do modelo de Rice (RICE, 1976; SMITH-MILES, 2009).

Para a comunidade de aprendizado de máquina, o problema da seleção de algoritmos pode ser entendido como um problema de classificação. Além disso, a escolha dos parâmetros da técnica de classificação também se caracteriza como um problema de seleção de algoritmos, denominada meta-aprendizado. O modelo de Rice serve ainda como base para a exploração e desenvolvimento de outras áreas, como a de sistemas autoadaptativos (SMITH-MILES, 2009).

## **4.2 Estratégia autoadaptativa para estimação da posição por imagens**

A partir do modelo de seleção de algoritmos proposto por Rice (ver Seção 4.1), é possível construir um sistema autoadaptativo, em que a adaptação se dá pela escolha do algoritmo mais propício para cada situação. Para adequar o modelo à estimação da posição de VANTs por imagens, os 4 espaços deverão ser definidos. Uma vez definidos, a etapa restante é realizar a correta associação entre as características do problema e as combinações de algoritmos propostas para resolvê-lo, denominado, no modelo de Rice, de mapa de seleção. Para isso, técnicas de aprendizado de máquina foram escolhidas como forma de realizar a classificação entre o espaço de características e o espaço de algoritmos.

Nesse contexto, a presente seção se encontra organizada em 5 subseções, em que a primeira, 4.2.1, é dedicada à definição do espaço de problemas, a segunda, 4.2.2, ao espaço de características, a 4.2.3 ao espaço de algoritmos e a 4.2.4 ao espaço de desempenhos. Por último, na Subseção 4.2.5 são apresentadas as técnicas de aprendizado de máquina a serem utilizadas como mapa de seleção.

#### 4.2.1 Espaço de problemas

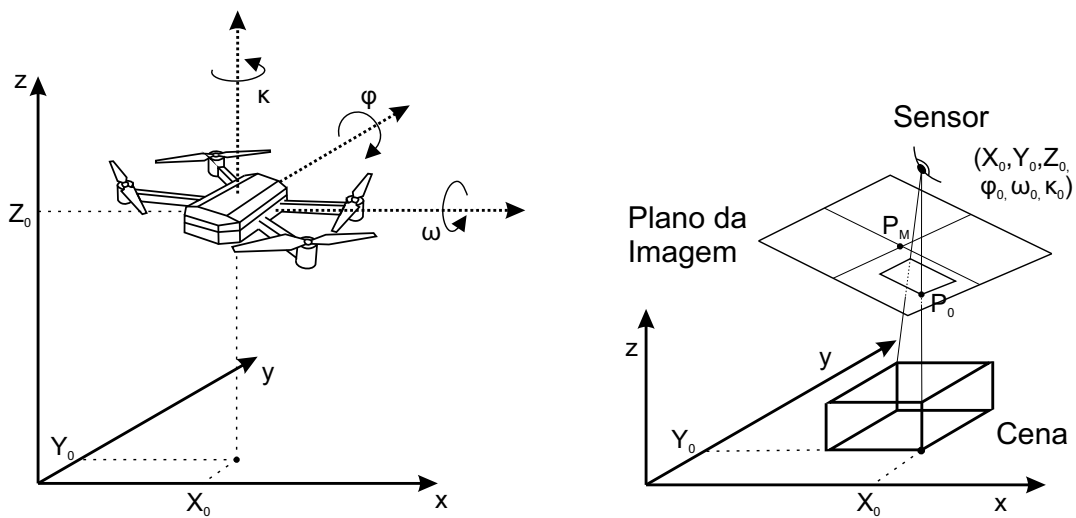
VANTs, assim como outros veículos que se movimentam pelo ar, possuem liberdade nos 3 eixos do espaço. Como consequência, tais veículos podem se movimentar livremente sobrevoando diferentes regiões, e.g., espaços florestais, montanhas e cidades, assim como podem também explorar locais em que há limitação física de espaço, e.g., interior de prédios, cavernas e antros. As características presentes em cada um dos grupos descritos acima se dá basicamente pela presença de obstáculos ao redor, e principalmente acima, do *drone*, diferenciando ambientes fechados, ou ambientes internos, de ambientes abertos, também chamados de ambientes externos. Neste trabalho, apenas este último grupo, i.e., ambientes abertos, será utilizado para exploração da técnica autoadaptativa de estimação da posição baseado em imagens.

Além disso, há uma variedade de sensores que podem ser utilizados na estimação da posição de VANTs, como câmeras monoculares e estéreo, sensores baseados em pulsos de laser, sonares e unidades inerciais (GROVES, 2013). Neste trabalho, foram escolhidas câmeras monoculares que operam na faixa visível do espectro eletromagnético. Além disso, por se tratar de ambientes abertos, as câmeras devem estar orientadas para o solo, em visada nadir, como requisito para que a técnica de estimação da posição opere corretamente. Idealmente, o ângulo de orientação do sensor está próximo do zero em relação à nadir. Porém, em condições nas quais se encontra acoplado diretamente à fuselagem do veículo, o movimento realizado por este influencia o ângulo de visada da câmera, modificando a imagem capturada. A Figura 4.3 ilustra as características de posição e orientação de VANTs e também da câmera a ele acoplada.

A posição de um corpo no espaço é definida por 3 coordenadas ( $X_0, Y_0, Z_0$ , na Figura 4.3), que definem sua posição nos eixos do plano,  $x$  e  $y$ , e no eixo vertical,  $z$ , respectivamente. É conveniente adotar a Terra como referencial para as coordenadas espaciais por meio da equivalência entre o eixo vertical e altitude, i.e., distância entre o corpo e a superfície do solo, e os demais como longitude e latitude, no qual o referencial é o centro da Terra (GROVES, 2013). A atitude do veículo, por sua vez, é definida por 3 ângulos, chamados arfagem (do inglês, *pitch*), rolagem (do inglês,

*roll*) e guinada (do inglês, *yaw*), que resultam do movimento horário seguindo a regra da mão direita em torno dos eixos  $x$ ,  $y$  e  $z$ , respectivamente (KRAUS, 2007). A influência dos ângulos no cálculo da posição do sensor pode ser corrigida a partir das equações de colinearidade (ver Capítulo 3, Seção 3.6).

Figura 4.3 - Posição e orientação de VANTs no espaço e efeito em imagens.



O estado de um corpo no espaço (à esquerda) pode ser representado por 6 graus de liberdade, sendo 3 ângulos e 3 coordenadas do espaço (GROVES, 2013). O ângulo de visada da câmera no momento de captura da cena altera a composição da imagem, conforme representado no esquema à direita. Nesta ilustração, a projeção da posição do sensor ( $X_0, Y_0$ ) não coincide com o ponto central da imagem,  $P_M$ , mas com o ponto arbitrário  $P_0$ .

Fonte: Produção do autor.

Dessa forma, o espaço de problemas é aqui definido como a estimação da posição das coordenadas geográficas, latitude e longitude, ao longo do percurso de voo de um VANT. A altitude e os ângulos do veículo não são parte do problema, embora também possam ser estimados por técnicas baseadas em visão computacional. As instâncias do problema são compostas por imagens capturadas por câmeras monoculares a bordo de *drones* orientadas para o solo, somadas a demais informações de orientação do veículo.

### 4.2.2 Espaço de características

Como o espaço de algoritmos está diretamente relacionado a efeitos que o ambiente pode causar em imagens e, conseqüentemente, impactar no desempenho dos algoritmos, este espaço é composto por um conjunto de propriedades que buscam caracterizar as imagens. Tais propriedades podem ser divididas em 3 tipos: baseadas em tons, ou intensidade do brilho da imagem, baseadas em textura, nas quais a distribuição espacial do brilho da imagem é levada em consideração, e baseadas no posicionamento da câmera no instante da sua captura. As propriedades contidas nos dois primeiros grupos são calculadas por meio de operações matemáticas a partir da informação contida nas imagens, enquanto as propriedades integrantes do último grupo são obtidas por meio de sensores a bordo do *drone*. É importante ressaltar que as propriedades aqui definidas não apenas descrevem características presentes em imagens aéreas, mas também são rapidamente computáveis. Dessa forma, espera-se, de maneira eficiente, extrair características de instâncias do problema que sejam suficientemente relevantes para associar o espaço de problemas ao espaço de algoritmos.

O espaço de características é composto, ao todo, por 14 características, sendo 4 propriedades da imagem baseadas em tons: média, desvio padrão, entropia e contraste (diferença entre intensidade máxima e mínima da imagem); 7 propriedades da imagem baseadas em textura: autocorrelação, dissimilaridade, homogeneidade, energia, contraste, diferença de variância e entropia; e 3 propriedades da aeronave: altura em relação ao solo, ângulo de rolagem e ângulo de arfagem no momento da captura da imagem. Enquanto as propriedades do posicionamento da câmera podem ser obtidas junto aos sensores a bordo da aeronave (ver Apêndice B), as demais podem ser calculadas a partir da intensidade dos elementos gráficos da imagem em escala de cinza (GONZALEZ; WOODS, 2008; HARALICK et al., 1973). Embora algumas propriedades possuam o mesmo nome (e.g., entropia), o cálculo realizado para obtê-las é diferente, resultando em propriedades também diferentes.

As propriedades do primeiro grupo, i.e., baseadas em tons, podem ser calculadas a partir do histograma da imagem, em que se ignora a organização espacial dos seus valores (GONZALEZ; WOODS, 2008). Nesse sentido, a média do brilho da imagem é dada por

$$avg_h(f) = \frac{\sum_x \sum_y f(x, y)}{xy}, \quad (4.1)$$



em que  $f$  é a imagem e  $x, y$  suas coordenadas (ver Apêndice B). O desvio padrão pode ser obtido pela equação

$$std_h(f) = \sqrt{\frac{1}{xy-1} \sum_x \sum_y [f(x, y) - avg(f)]^2}. \quad (4.2)$$

A entropia pode ser calculada por meio da equação

$$etp_h(f) = \sum \{hist_n(f) \cdot \log_2[hist_n(f)]\}, \quad (4.3)$$

em que  $hist_n()$  é o histograma normalizado composto por 256 intervalos. A diferença entre o máximo e o mínimo da imagem são dados por

$$dif_h(f) = max(f) - min(f), \quad (4.4)$$

em que  $max()$  e  $min()$  representam, respectivamente o maior e menor valores do histograma da imagem, i.e., o maior e menor valores da intensidade de brilho da imagem. As propriedades baseadas em textura, em que a informação estrutural do arranjo das superfícies pode ser obtida, são calculadas a partir da Matriz de Coocorrência de Níveis de Cinza (do inglês, *Gray-Level Co-occurrence Matrix* - GLCM) (HARALICK et al., 1973). Para tornar as propriedades invariáveis à rotação e à quantização de níveis de cinza, a matriz de coocorrência normalizada foi calculada. O cálculo da matriz normalizada pode ser encontrado no Apêndice D.

Tomando  $\tilde{P}$  como a GLCM,  $C$  como o número de escalas de cinza da imagem quantizada e  $\Delta = 1/C$ , as medidas podem ser calculadas por meio das fórmulas representadas na Tabela 4.1 (LÖFSTEDT et al., 2019).

### 4.2.3 Espaço de algoritmos

A escolha dos algoritmos se baseia nos trabalhos de Goltz (2011), Silva (2016) e Braga (2018), que fazem parte de uma sequência de esforços do grupo PITER<sup>1</sup> para desenvolvimento de tecnologias de navegação por imagens para VANTs. Basicamente, a técnica consiste em comparar imagens capturadas em voo a uma referência, que foi obtida previamente e que contém informações sobre toda região do

---

<sup>1</sup>Processamento de Imagens em TEmpo Real (PITER) é um grupo vinculado ao Instituto de Estudos Avançados (IEAv).

Tabela 4.1 - Texturas de Haralick invariantes a tons de cinza e equações correspondentes.

Definição	Equação
Autocorrelação	$\sum_{i=1}^C \sum_{j=1}^C \left(\frac{i}{C} \cdot \frac{j}{C}\right) \tilde{P}(i, j) \Delta^2$
Energia	$\sum_{i=1}^C \sum_{j=1}^C \tilde{P}(i, j)^2 \Delta^2$
Homogeneidade	$\sum_{i=1}^C \sum_{j=1}^C \frac{\tilde{P}(i, j)}{1 + (i/C - j/C)^2} \Delta^2$
Entropia	$- \sum_{i=1}^C \sum_{j=1}^C \tilde{P}(i, j) \log[\tilde{P}(i, j)] \Delta^2$
Dissimilaridade	$\sum_{i=1}^C \sum_{j=1}^C \left(\frac{i}{C} - \frac{j}{C}\right) \tilde{P}(i, j) \Delta^2$
Contraste	$\sum_{i=1}^C \sum_{j=1}^C \left(\frac{i}{C} - \frac{j}{C}\right)^2 \tilde{P}(i, j) \Delta^2$
Diferença de variância	$\sum_{k=0}^{C-1} \left(\frac{k+1}{C} - \tilde{\mu}_{x-y}\right)^2 \tilde{P}_{x-y}(k) \Delta$
$\tilde{\mu}_{x-y}$	$\sum_{k=0}^{C-1} \frac{k+1}{C} \tilde{P}_{x-y}(k) \Delta$
$\tilde{P}_{x-y}(k)$	$\sum_{i=1}^C \sum_{j=1}^C \Big _{ i-j =k} \tilde{P}(i, j) \Delta$

vo. Por serem imagens diferentes, que podem ter sido capturadas por sensores diferentes, ou em momentos diferentes, ou ainda em ângulos diferentes, a técnica utiliza algoritmos de processamento de imagens que buscam tornar a cena mais verossímil, eliminando possíveis ruídos ou objetos transitórios que possam ocasionalmente estar presentes. A combinação dos algoritmos pode melhorar ou piorar a acurácia da técnica, a depender do efeito causado por cada um deles. A descrição detalhada de cada etapa da técnica, assim como os algoritmos utilizados em cada uma delas, pode ser encontrada no Capítulo 3.

Nesse contexto, o espaço de algoritmos é composto por combinações de algoritmos

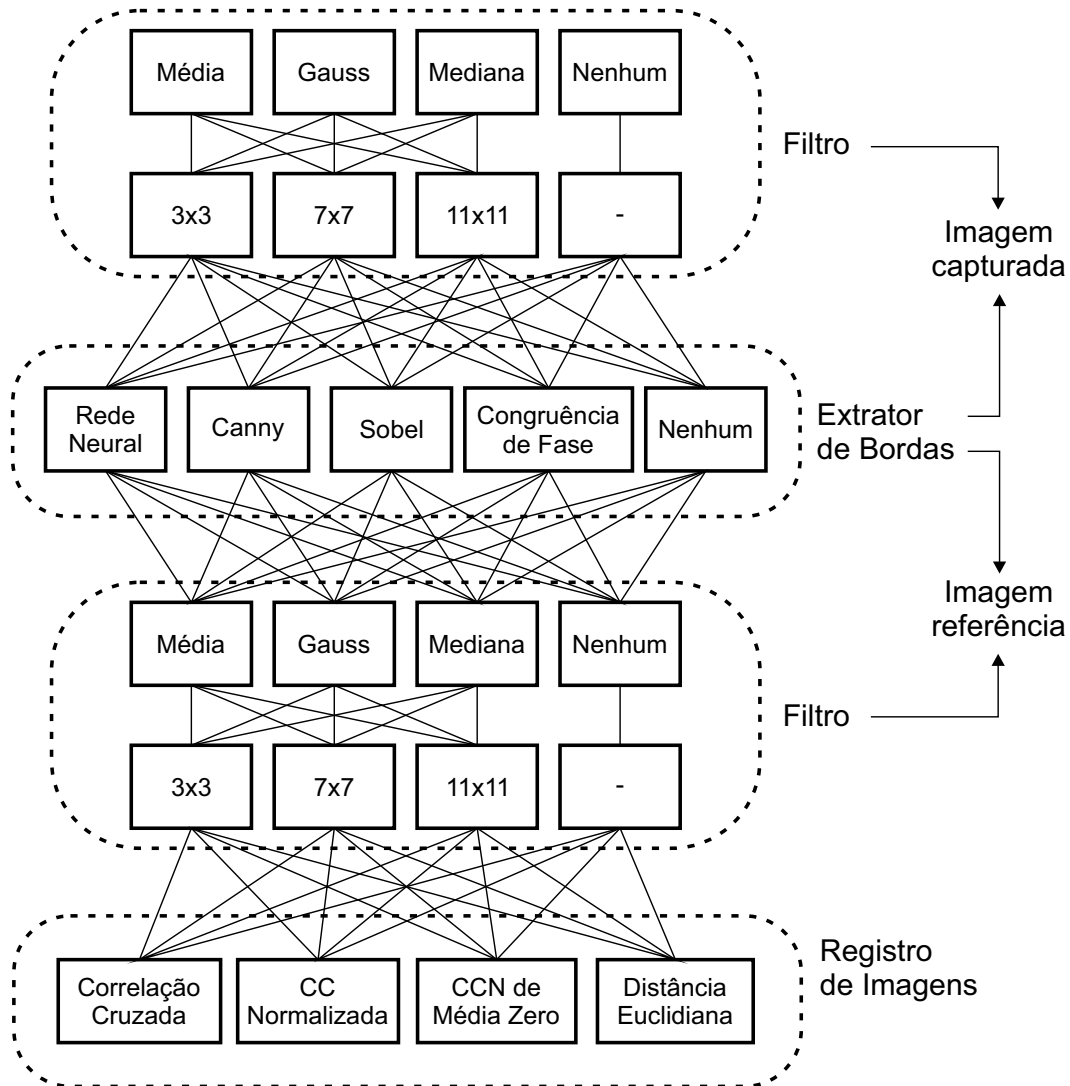
de processamento de imagens, que incluem filtros, extratores de bordas e técnicas baseadas de correlação cruzada. Aqui, cabe a desambiguação: este espaço é composto por algoritmos que, por sua vez, são formados por combinações de outros algoritmos. Nesse sentido, o termo utilizado doravante para espaço de algoritmos será espaço de combinações de algoritmos, ou apenas combinações. Dessa forma, o objetivo do modelo é selecionar a melhor combinação de algoritmos dado um conjunto de características do problema. A Figura 4.4 ilustra as possibilidades que se originam a partir das diferentes combinações entre os algoritmos utilizados na técnica de estimação da posição por imagens.

Rice, ao propor o modelo abstrato de seleção de algoritmos, definiu o espaço de algoritmos como sendo composto apenas por processos diferentes de resolução de problemas, o que faria com que os parâmetros escolhidos por estes não fossem reconhecidos como algoritmos diferentes. Aqui, porém, a escolha de parâmetros pode ser também entendida como um problema de seleção de algoritmos, embora, no significado estrito do termo, seja a mesma metodologia de resolução contendo configurações diferentes (SMITH-MILES, 2009). Dessa forma, cada composição entre a técnica de filtragem e o tamanho do *kernel* é vista como um algoritmo diferente. O total de combinações possíveis entre filtros, extrator de bordas e técnica de registro de imagens é igual a 2000 e, dessa forma, todas essas 2000 possibilidades foram consideradas.

#### 4.2.4 Espaço de desempenhos

Na estimação da posição por imagens, a principal forma de medir a eficácia de uma combinação de algoritmos é mediante sua acurácia, i.e., o quão próxima a posição estimada corresponde à posição real do veículo. A posição estimada de uma instância do problema pode ser calculada por meio de uma combinação específica de algoritmos e resultará em um valor que, sujeito a erros, está diretamente relacionado à combinação escolhida. A posição real do veículo, por sua vez, pode ser obtida a partir de instrumentos a bordo do mesmo. Neste trabalho, o espaço de desempenho será constituído pela diferença entre a posição estimada pela técnica baseada em imagens e a posição prevista pelo GNSS do VANT. Embora a posição provida pelo GNSS também esteja suscetível a erros, ela será utilizada como verdade fundamental (do inglês, *Ground Truth*). Vale ressaltar que o erro associado ao GNSS varia entre os equipamentos e, por isso, não é o mesmo para todos os experimentos utilizados neste trabalho. Ainda assim, ele é suficientemente pequeno para ser utilizado como verdade fundamental.

Figura 4.4 - Combinações de algoritmos possíveis para estimação da posição por imagens.



As combinações são formadas a partir da aplicação, ou não, de um filtro e um detector de bordas na imagem capturada pelo VANT para a imagem referência, sendo o detector de bordas o mesmo para ambas imagens. Cada filtro é formado pela combinação de 3 opções diferentes (Média, Gauss e Mediana) e por 3 opções de *kernels* (3x3, 7x7 e 11x11). O conjunto de detectores de bordas é formado por 4 possibilidades: Rede neural, Canny, Sobel e Congruência de fases. Por último, um algoritmo de registro de imagens deve ser escolhido entre 4 diferentes possibilidades: Correlação Cruzada, Correlação Cruzada Normalizada, Correlação Cruzada Normalizada de Média Zero e Distância Euclidiana. A descrição de todos os algoritmos aqui citados pode ser encontrada no Capítulo 3.

Fonte: Produção do autor.

A partir da imagem georreferenciada, o valor das coordenadas cartográficas obtido pelos sinais do GNSS pode ser transcrito para um par de coordenadas da imagem. A

distância entre estes e a posição estimada podem então ser utilizadas para calcular o erro a partir da equação

$$erro_p(f) = \sqrt{(x_{img} - x_{gnss})^2 + (y_{img} - y_{gnss})^2}, \quad (4.5)$$

em que  $(x_{img}, y_{img})$  são as coordenadas da imagem  $f$  obtidas por meio do método de estimação da posição por imagens e  $(x_{gnss}, y_{gnss})$  são as coordenadas da imagem fornecidas pelo GNSS. Para obter a distância, ou erro, em metros, basta dividir a resultado da Equação 4.5 pela resolução espacial da imagem,  $res(f)$ ,

$$erro_m(f) = erro_{pxl}(f)/res(f). \quad (4.6)$$

#### 4.2.5 Aprendizado de Máquina como mapa de seleção

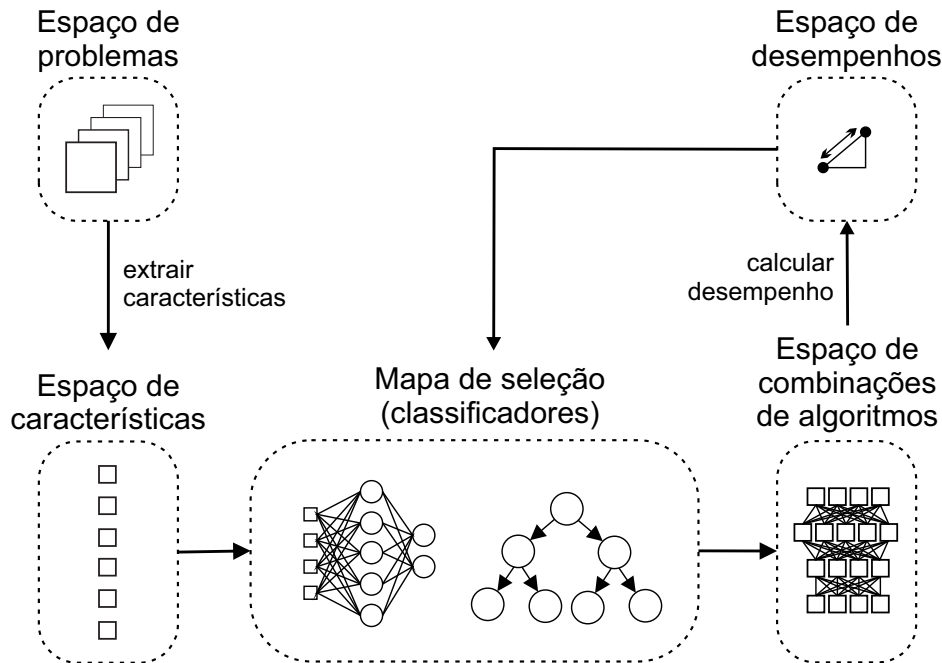
Em linhas gerais, o objetivo da abordagem autoadaptativa para estimação da posição pode ser definida como sendo a correta associação de combinações de algoritmos a partir de um determinado conjunto de propriedades da imagem. Para realizar essa escolha, que corresponde ao mapa de seleção no modelo abstrato de seleção de algoritmos, deve-se adotar um mecanismo de classificação. Tal mecanismo proverá as regras para o correto endereçamento, a partir do espaço de características, dos conjuntos de combinações contidos no espaço de algoritmos. Dessa forma, o problema de seleção de algoritmos pode ser dividido em dois: primeiro, definir qual a melhor combinação entre o espaço de algoritmos e o espaço de características; segundo, estabelecer qual técnica de classificação apresenta melhor acurácia ao associar os espaços citados. Um esquema em que o modelo abstrato de Rice foi modificado para incluir classificadores no lugar do mapa de seleção é apresentado na Figura 4.5.

Nesse contexto, três técnicas de aprendizado de máquina foram escolhidas para compor o mapa de seleção. Como cada técnica possui suas próprias peculiaridades, é esperado que cada uma obtenha uma acurácia diferente em termos de somas de classificações corretas e erradas. As técnicas são a Árvore de Classificação (ver Subseção 3.7.1), a Máquina de Vetores de Suporte (ver Subseção 3.7.2) e o Perceptron de Múltiplas Camadas (ver Subseção 3.7.3).

Aqui, pode-se dizer que o núcleo inteligente do sistema autoadaptativo é formado por uma técnica de aprendizado de máquina, que é responsável por classificar as imagens capturadas pelo VANT e associá-la a uma determinada combinação de

algoritmos. Nesse sentido, há diferentes maneiras de classificar as imagens e associá-las a uma dada combinação. Tal associação é fundamental para que a abordagem autoadaptativa funcione corretamente e, assim, proveja o sistema com inteligência suficiente para realizar as reconfigurações adequadas.

Figura 4.5 - Esquema em que classificadores são utilizados na seleção de algoritmos.



O esquema ilustra a forma como técnicas de classificação podem ser utilizadas como o mapa de seleção do problema de seleção de algoritmos. A partir de instâncias do problema é possível extrair características que servirão como informações para selecionar as técnicas. Nesse sentido, o espaço de características atua como entrada para os classificadores, que associam um conjunto de características a uma combinação de algoritmos. O treinamento é realizado por meio da observação do desempenho de cada combinação.

Fonte: Produção do autor.

### 4.3 Considerações finais sobre este capítulo

Neste capítulo, foi apresentada a abordagem autoadaptativa para estimação da posição de VANTs por imagens. Tal abordagem é baseada no modelo abstrato de seleção de algoritmos proposto por John Rice, composto por 4 espaços e um mapa de seleção. Os espaços, assim como o mapa de seleção, são fundamentais para o desenvolvimento do sistema autoadaptativo. O mapa de seleção, porém, pode ainda ser considerado o núcleo inteligente do sistema, pois é o responsável por decidir a

reconfiguração, ou adaptação, deste. No próximo capítulo, serão apresentados os experimentos dos quais conjuntos de dados foram obtidos para posterior análise da metodologia autoadaptativa.





## 5 AVALIAÇÕES EXPERIMENTAIS

Diferentes experimentos foram selecionados e organizados a fim de servirem de base para testar a metodologia proposta. Cada experimento envolveu o sobrevoo de uma região específica por um VANT, cujo controle foi planejado e realizado por uma equipe em solo. Em todos experimentos, o VANT foi equipado com uma câmera em visada nadir para a captura das imagens. Após a realização do experimento, as informações da aeronave, tal como sua altura de voo, e as imagens capturadas pela câmera embarcada foram recolhidos, processados e sincronizados. Dessa forma, o conjunto de dados resultante contém todas as informações necessárias para executar a metodologia proposta.

Os experimentos foram organizados de acordo com a localização geográfica onde foram obtidos. Em cada local, o número de experimentos realizados varia, assim como a aeronave utilizada e seus sensores, a área sobrevoada, as condições climáticas e a imagem referência utilizada. Na Tabela 5.1 podem ser vistos o número de experimentos selecionados bem como a respectiva cidade e país onde o voo foi realizado.

Tabela 5.1 - Número de experimentos realizados e suas respectivas localizações.

País	Cidade	Experimentos
Brasil	São José dos Campos	1
Suécia	Revinge	1
México	Monterrey	3

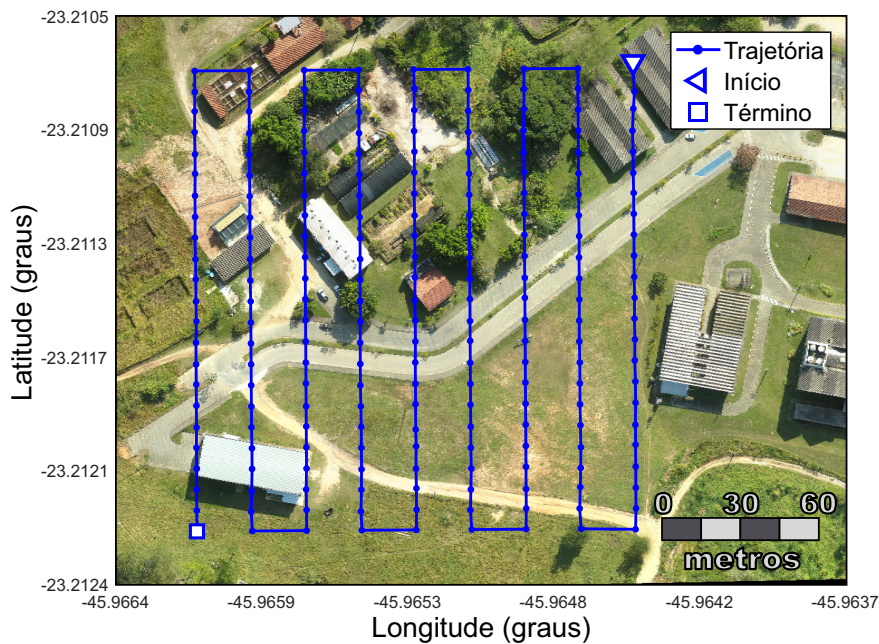
Nesse contexto, o presente capítulo se divide em 3 seções, uma para cada região onde foram realizados os experimentos. Na primeira seção, 5.1, são definidas as condições em que foi obtido o conjunto de dados por meio do voo realizado na cidade de São José dos Campos, no Brasil. Em seguida, a Seção 5.2 é dedicada ao experimento realizado na cidade de Revinge, na Suécia. Por último, na Seção 5.3, são tratados os detalhes técnicos dos experimentos realizado em Monterrey, no México.

### 5.1 Experimento realizado no Brasil

No Brasil, apenas um experimento foi realizado. A execução do experimento foi feita pelo grupo PITER, vinculado ao Instituto de Estudos Avançados. O voo foi realizado na cidade de São José dos campos ( $23^{\circ}10' S$ ,  $45^{\circ}53' O$ ) e envolveu um quadricóptero de pequeno porte.

No experimento, um VANT modelo *Phantom 4*, produzido pela DJI, sobrevoou o campus da Universidade do Vale do Paraíba (UNIVAP) a uma altura de 32 m. Como pode ser visto na Figura 5.1, a trajetória do VANT, em forma de zigue-zague, soma um total de 1,75 km.

Figura 5.1 - Trajetória percorrida por um VANT durante o experimento em São José dos Campos, Brasil.



Os círculos azuis indicam a posição do VANT no momento da captura das imagens e são interligados por segmentos de retas. A posição da aeronave é dada pelos dados do GNSS a bordo da mesma.

Fonte: Produção do autor.

Por se tratar de um *Phantom 4*, não houve necessidade de instalar uma câmera para registrar as imagens, pois o mesmo possui um sistema de vídeo acoplado. Esse sistema conta com um *gimbal* ativo de 3 eixos e uma câmera digital com resolução para imagens RGB de  $4000 \times 3000$  pixels, e campos de visão horizontal de  $98,33^\circ$  e vertical de  $73,75^\circ$ . Durante o experimento, à cerca de 32 m de altura, as imagens capturadas possuem GSD de aproximadamente 1,37 cm. No total, 207 imagens foram capturadas.

Além disso, o *drone* possui um sistema no qual as informações de voo são armazenadas em arquivos e registradas em sua memória interna. A partir desses arquivos foi

possível extrair as informações do VANT durante todo o percurso. Os dados foram posteriormente selecionados e associados às imagens capturadas no experimento. Dessa forma, a atitude da câmera assim como informações de altura e posicionamento geográfico do veículo foram adicionadas à base de dados com as imagens.

A imagem utilizada como referência foi obtida por outro VANT, que sobrevoou a região propositalmente em um dia diferente para garantir que as condições climáticas fossem diferentes. O mosaico de imagens foi processado e georreferenciado para se obter a imagem da região. A resolução espacial da imagem referência é de  $2,1\text{ cm}$  e cobre todo o trajeto percorrido pelo VANT.

## 5.2 Experimento realizado na Suécia

O conjunto de imagens registradas na Suécia foi obtido a partir de um experimento executado pelo Departamento de Ciência da Informação e da Computação (*Department of Computer and Information Science*) da Universidade de Linköping (CONTE; DOHERTY, 2008; BRAGA et al., 2016). O experimento foi realizado em uma região utilizada para treinamento de resgate de emergências e situa-se na cidade de Revinge, Suécia ( $55^{\circ}43' \text{ N}$ ,  $13^{\circ}28' \text{ L}$ ).

O voo foi feito por um VANT do modelo R-MAX, fabricado pela *Yamaha Motor Company*. O Yamaha R-MAX é um helicóptero com motor de combustão interna e impulsionado por dois rotores, um principal e um lateral. A aeronave percorreu uma distância de  $916\text{ m}$  a uma velocidade média de  $3\text{ m/s}$  e altura aproximada de  $60\text{ m}$ .

Uma câmera digital foi fixada diretamente à aeronave em ângulo perpendicular ao solo (visada nadir). Por esse motivo, a atitude da aeronave está diretamente relacionada à atitude da câmera. O sensor imageador utilizado possui campo de visão horizontal de  $43.62^{\circ}$  e vertical de  $31.85^{\circ}$  e é capaz de capturar imagens monocromáticas de  $360 \times 288\text{ pixels}$  com GSD de  $12\text{ cm/pixel}$ . Ao todo, 1593 imagens foram capturadas a uma frequência de  $25\text{ Hz}$ . Na Figura 5.2 pode ser vista a trajetória percorrida pela aeronave durante a execução do experimento.

Além da câmera acoplada ao *drone*, outros sensores complementares fizeram parte dos dispositivos presentes na aeronave. Entre eles, um barômetro, 3 giroscópios e 3 acelerômetros (CONTE; DOHERTY, 2009). A fim de obter a atitude da câmera e a altura do VANT durante a captura das imagens, os dados obtidos pelos sensores foram processados e sincronizados manualmente após o voo (CONTE; DOHERTY, 2008). A imagem referência associada a este experimento cobre toda a área sobrevoada pelo

VANT, é monocromática e possui GSD de  $50\text{ cm/pixel}$ .

Figura 5.2 - Trajetória percorrida por um VANT durante a realização do experimento em Revinge, Suécia.



A posição do VANT no momento da captura das imagens é indicada pelos círculos azuis presentes na imagem e foram obtidos por meio do GNSS a bordo da aeronave.

Fonte: Produção do autor.

### 5.3 Experimentos realizados no México

No México, diversos experimentos foram realizados em cooperação com o Centro de Investigação em Ciências Físico Matemáticas da *Univerdad Autónoma de Nuevo León*. Todos os voos foram realizados na cidade de Monterrey, mais precisamente no campus da universidade ( $25^{\circ}43' \text{ N}$ ,  $100^{\circ}19' \text{ O}$ ). De todos os experimentos realizados, alguns apresentaram falhas e não puderam ser processados posteriormente para serem utilizados neste trabalho. Ao todo, 3 experimentos foram executados com êxito.

Os experimentos foram realizados por um hexacóptero construído pela equipe da universidade, também responsável pela realização dos voos. O VANT é impulsionado por 6 rotores e alimentado por duas baterias de lítio. Conta com uma placa *pixhawk* instalada no centro do corpo do VANT e responsável por gerenciar suas funções

de estabilização e controle. Além disso, a *pixhawk* armazena informações de voo de sensores embarcados na própria placa, que incluem um barômetro e sensores inerciais. O hexacóptero pode ser visto na Figura 5.3.

Os dados foram processados e sincronizados após a conclusão dos experimentos para se obter a atitude, altura e posicionamento geográfico do VANT no instante de captura de cada imagem. A sincronia se fez necessária porque o sistema da câmera e o sistema de controle do VANT operavam de maneira independente. Dessa forma, o relógio de ambos sistemas foi utilizado como referência para que as informações fossem associadas às imagens.

Figura 5.3 - Hexacóptero utilizado na realização dos experimentos em Monterrey, México.



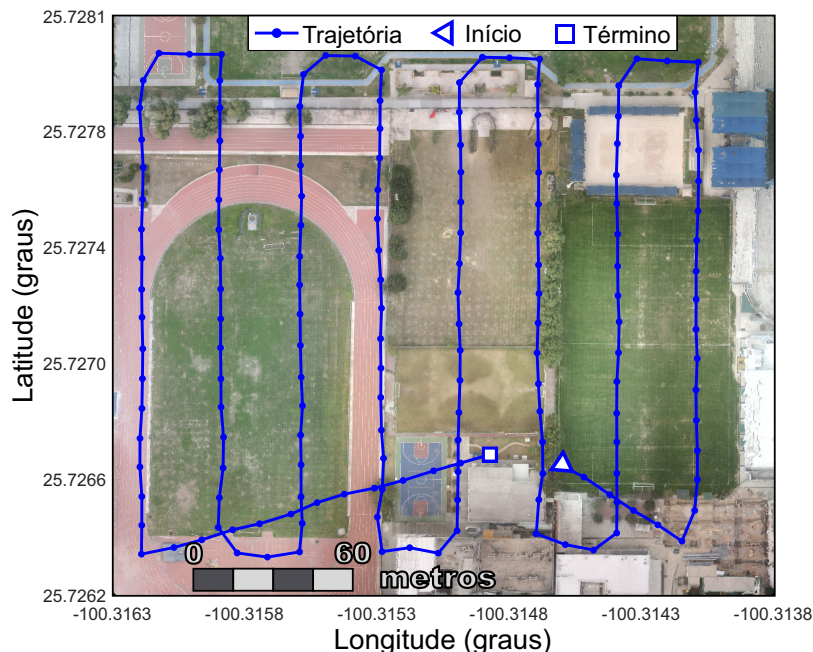
Fonte: Produção do autor.

No primeiro experimento, o VANT percorreu uma distância de 1,86 *km* a uma altura aproximada de 70 *m*. A trajetória foi composta por linhas paralelas com orientação norte-sul, gerando um formato em zigue-zague (ver Figura 5.4). A câmera utilizada no experimento foi uma Sequoia, da empresa *Parrot*, acoplada ao VANT em visada nadir. A câmera Sequoia é capaz de capturar imagens RGB com resolução de  $4608 \times 3456$  *pixels* com campos de visão horizontal de  $63,9^\circ$  e vertical de  $50,1^\circ$ . As imagens capturadas neste experimento tem GSD de, aproximadamente, 1,9 *cm*. Ao todo, 165 imagens foram capturadas a uma frequência de 1 *Hz*.

A imagem referência foi obtida por meio de um mosaico de imagens capturadas pelo mesmo *drone*, porém em um dia diferente e sob outras condições de voo (3º experimento, ver Figura 5.6). A resolução espacial da imagem referência é de 2,2 *cm*

e cobre toda a área sobrevoada pelo hexacóptero durante o experimento.

Figura 5.4 - Trajetória percorrida por um VANT durante o primeiro experimento em Monterrey, México.



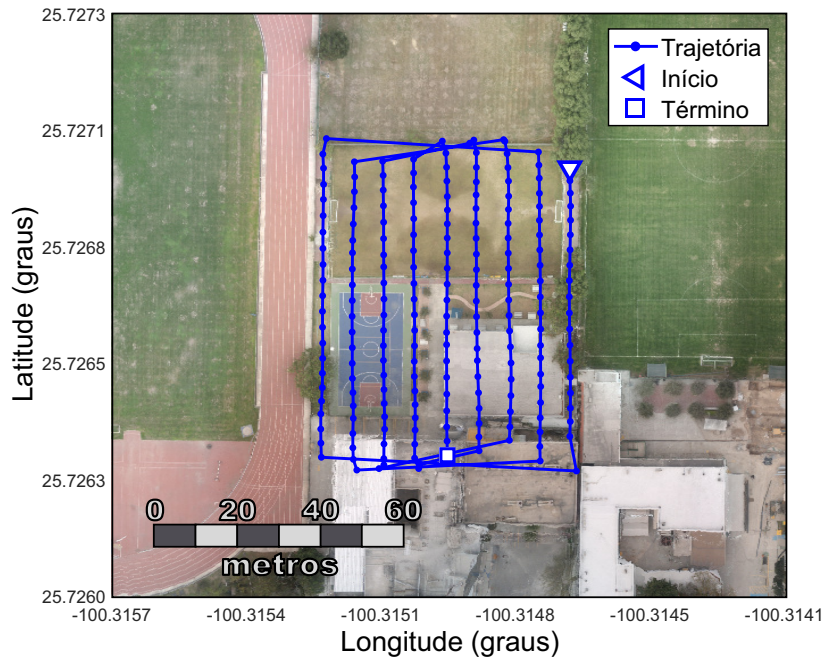
Os círculos azuis indicam a posição do VANT, obtida por meio do GNSS da aeronave, no momento da captura das imagens. O modelo da trajetória é em zig-zague composto por linhas paralelas apontando para o norte geográfico.

Fonte: Produção do autor.

No segundo experimento, o VANT percorreu uma distância total de cerca de 1 *km* a uma altura aproximada de 45 *m*. A trajetória foi planejada antes do início do voo e é composta por retângulos circunscritos entre si, formando uma espécie de espiral (ver Figura 5.5). Uma câmera digital foi acoplada a sua estrutura em ângulo perpendicular ao solo. A câmera modelo ILCE-5100L, fabricada pela Sony, captura imagens RGB de  $6000 \times 4000$  *pixels*, cujo campo de visão horizontal de  $83,37^\circ$  e vertical de  $55,58^\circ$  permite uma resolução espacial do solo, à 45 *m* de altura, de 1,1 *cm*. Ao todo, 198 imagens foram capturadas no experimento em localidades definidas durante o planejamento da rota. A câmera foi conectada à placa *pixhawk* por um cabo responsável por emitir o sinal de captura da imagem. Dessa forma, as imagens foram capturadas em posições geográficas específicas ignorando a distância temporal entre elas. A Figura 5.5 ilustra a trajetória percorrida pelo VANT durante

a captura das imagens, bem como sua posição no momento de captura das imagens. A imagem referência é a mesma do experimento anterior e foi obtida por meio de um mosaico de imagens capturadas no terceiro experimento.

Figura 5.5 - Trajetória percorrida por um VANT durante o segundo experimento em Monterrey, México.

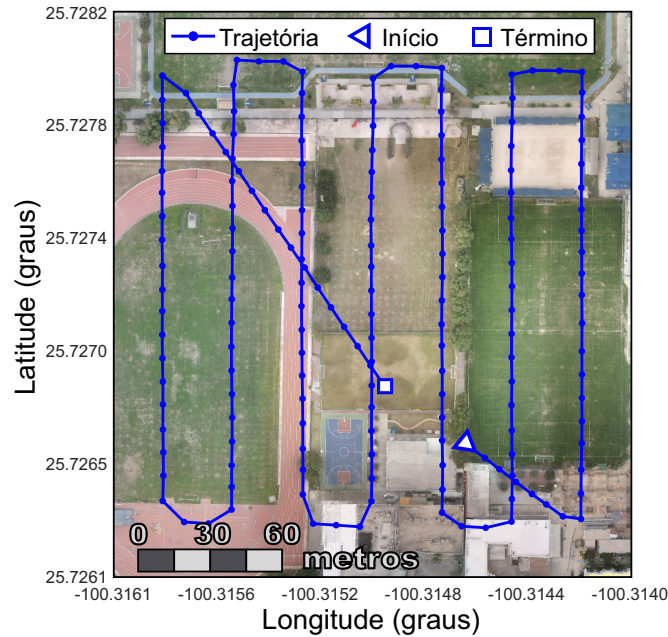


Os círculos azuis indicam a posição do VANT, obtida por meio do GNSS a bordo da aeronave, no momento da captura das imagens. A trajetória segue um modelo semelhante a uma espiral, em que são executados movimentos circunscritos entre si.

Fonte: Produção do autor.

O terceiro experimento foi realizado com o objetivo de se obter uma imagem referência da área sobrevoada pelos outros experimentos. Para isso, foi definida uma trajetória em zigue-zague com sobreposição parcial das imagens (ver Figura 5.6). Ao todo, o VANT percorreu  $1,72 \text{ km}$  a uma altura de aproximadamente  $80 \text{ m}$ . A câmera utilizada para registrar as imagens foi a mesma do experimento anterior, uma Sequoia, que permitiu a captura de imagens com resolução espacial de cerca de  $2,2 \text{ cm}$ . No total, foram obtidas 175 imagens a uma frequência de  $1 \text{ Hz}$ . As imagens foram posteriormente utilizadas para construção de um mosaico georreferenciado a fim de ser utilizado como referência.

Figura 5.6 - Trajetória percorrida por um VANT durante o terceiro experimento em Monterrey, México.



Os círculos azuis indicam a posição do VANT, obtida por meio do GNSS da aeronave, no momento da captura das imagens. A trajetória foi planejada de forma que as imagens capturadas possuam sobreposição parcial.

Fonte: Produção do autor.

#### 5.4 Considerações finais sobre este capítulo

No presente capítulo, foram descritos os experimentos que resultaram em conjuntos de dados a serem utilizados para testar a metodologia autoadaptativa proposta. Cada experimento contou com características específicas, tais como o VANT utilizado no voo, o sensor imageador embarcado na aeronave e o terreno sobrevoado, que se revelam importantes fatores para que seja possível testar a técnica de estimação de posição por imagens proposta. Nesse sentido, o próximo capítulo apresenta os resultados obtidos a partir dos conjuntos de dados aqui descritos.



## 6 RESULTADOS

Neste trabalho, a análise da metodologia autoadaptativa, apresentada no Capítulo 4, se dará por meio da investigação dos resultados alcançados a partir dos conjuntos de imagens obtidos nos experimentos descritos no Capítulo 5. Nesse sentido, diferentes configurações de algoritmos e técnicas, assim como duas métricas de erro são apresentadas de forma a proporcionar uma base para que os resultados possam ser comparados entre si.

As métricas de erro são baseadas no espaço de desempenhos da abordagem autoadaptativa (ver Seção 4.2.4) e compreendem a distância entre a posição do VANT estimada pela técnica de reconhecimento de regiões por imagens e a posição provida pelo GNSS do veículo. Por serem métricas que tem como objetivo avaliar o desempenho dos experimentos como um todo, o cálculo se dá em cima do conjunto de imagens obtido em cada voo do referido experimento. Tomando o conjunto de imagens como  $\{f_i\}_{i=1}^N$ , a primeira métrica pode ser definida como

$$EM = \frac{1}{N} \sum_{i=1}^N erro_m(f_i), \quad (6.1)$$

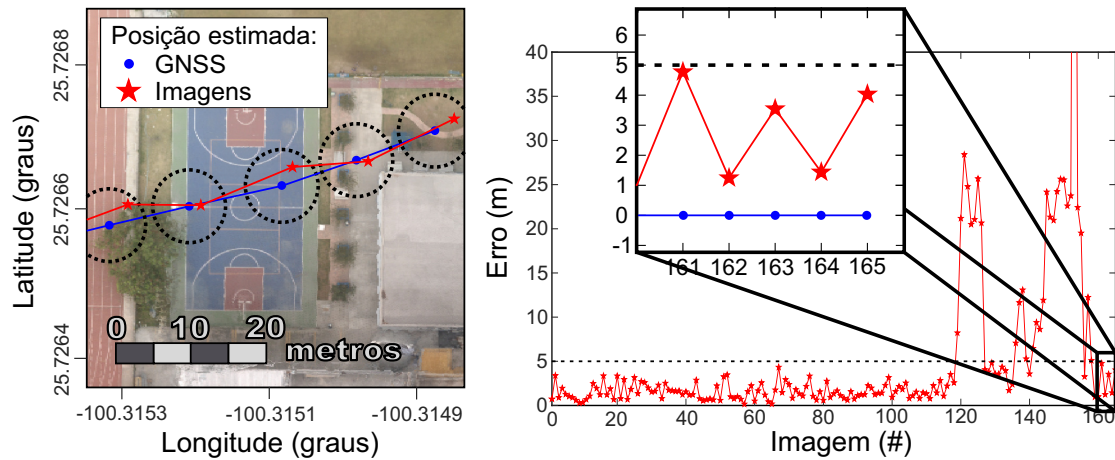
em que a função  $erro_{mtr}()$  é a distância entre a posição do VANT estimada por imagens e a posição estimada pelo GNSS dada pela Equação 4.6. A segunda métrica pode ser calculada como

$$TBC = \frac{100}{N} \sum_{i=1}^N \begin{cases} 1, & \text{se } erro_m(f_i) < \epsilon; \\ 0, & \text{caso contrário,} \end{cases} \quad (6.2)$$

em que  $\epsilon$  define, em metros, o limite máximo de erro tolerável. Essa métrica é denominada de Taxa de Boa Correspondência (TBC) e compreende a porcentagem das estimações das posições cujos erros são inferiores ao limite  $\epsilon$ .

Neste trabalho, de acordo com o que foi proposto por Braga (2018),  $\epsilon$  foi definido como 5 metros. Dessa forma, espera-se analisar o desempenho de cada conjunto de imagens (i.e., de cada experimento) considerando não apenas a média do erro, mas também a quantidade de boas correspondências dos conjuntos. Uma ilustração sobre o erro, obtido a partir da comparação entre as estimações realizadas pelo GNSS e pela técnica de reconhecimento de regiões por imagens, pode ser encontrada na Figura 6.1.

Figura 6.1 - Análise comparativa entre as posições estimadas pelo GNSS e pela técnica de reconhecimento por imagens.



O quadro à esquerda ilustra a posição do VANT estimada com base nos dados do GNSS da aeronave (círculos em azul) e a posição estimada por meio da técnica de reconhecimento de regiões por imagens (estrelas em vermelho). Os círculos com a linha tracejada possuem raio de 5 metros, e delimitam as estimações da posição que se caracterizam como TBC. O quadro maior à direita e sua ampliação ilustram a distância entre a posição obtida pelo GNSS e a posição calculada pela técnica de reconhecimento de regiões, que pode ser entendida como o erro da posição real da aeronave em relação a posição estimada por meio da técnica baseada em imagens. As 5 posições estimadas do quadro da esquerda equivalem às 5 posições do quadro menor da direita, que é uma ampliação dos últimos 5 valores do quadro maior.

Fonte: Produção do autor.

O desempenho da técnica de estimação da posição por imagens, dado pelo erro, foi calculado para cada uma das 2000 combinações de algoritmos (ver Subseção 4.2.3) em cada um dos 5 conjuntos de imagens obtidos por meio dos experimentos. Como forma de eliminar a influência da resolução espacial da imagem referência e para diminuir o esforço computacional no cômputo dos resultados, elas foram redimensionadas para terem o mesmo GSD de  $50 \text{ cm/pxl}$ . Da mesma forma, as imagens capturadas por VANTs foram redimensionadas para, à mesma altitude, registrarem imagens com o mesmo GSD. Assim, pode-se garantir que as eventuais diferenças no desempenho dos resultados foram causadas por fatores que não estão associados à resolução dos sensores imageadores.

Além disso, os parâmetros de entrada de algoritmos que não puderam ser calculados automaticamente foram definidos a partir das suas próprias características ou com base na literatura científica. Como apresentado na Seção 4.2.3, três tamanhos de

janelas foram utilizados para os filtros espaciais (com dimensões  $3 \times 3$ ,  $7 \times 7$  e  $11 \times 11$ ), sendo contabilizados como algoritmos próprios. Com base no tamanho da janela dos filtros, o desvio padrão do filtro de Gauss,  $\sigma$  (ver Equação 3.8), foi definido como 3. Os parâmetros do detector de bordas baseado em congruência de fases foram escolhidos de acordo com os padrões definidos no trabalho do autor que desenvolveu o método, Kovesi (2000). Para os algoritmos do registro de imagens que utilizam máscara, o limiar  $\delta$  (ver Equação 3.24) foi escolhido de forma a aceitar apenas valores que cobrem pelo menos 90% da imagem capturada pelo VANT. E, por último, a incerteza da janela de busca (ver Equação 3.1) foi definida como 20 metros.

Com base nesses parâmetros, o desempenho de cada conjunto de dados foi analisado a partir dos resultados obtidos utilizando a abordagem não adaptativa e a abordagem autoadaptativa. No primeiro caso, é possível calcular os resultados de forma direta. A abordagem autoadaptativa, por sua vez, requer um mapa de seleção (ver Seção 4.2.5), responsável por associar as características das imagens (ver Seção 4.2.2) às combinações de algoritmos. Para realizar tal associação, 3 técnicas de aprendizado de máquina foram empregadas: árvore de classificação, SVM e MLP. Todos os conjuntos de imagens foram utilizados para treinamento, sendo este realizado a partir da validação cruzada em 10 etapas. Os parâmetros e a arquitetura de cada uma das técnicas foi obtido por meio de tentativas e erros.

O treinamento foi feito da seguinte forma. A partir das combinações de algoritmos e seus devidos parâmetros, foi possível calcular o desempenho de cada uma das imagens obtidas a partir dos experimentos descritos no Capítulo 5. Estas, por sua vez, foram agrupadas em regiões a partir de um processo iterativo em que se observou a acurácia na classificação das técnicas, as características presentes nas imagens e o desempenho das combinações de algoritmos. Dessa forma, cada conjunto de dados foi dividido em subconjuntos (ou regiões), compostas por imagens com características similares. Em seguida, para cada região era eleita uma combinação de algoritmos com base no desempenho da mesma. Assim, as técnicas de aprendizado de máquina poderiam ser treinadas para associar as características de um subconjunto de imagens a uma combinação de algoritmos eleita como ideal para aquele caso. Os parâmetros de treinamento das técnicas e as arquiteturas das mesmas foram obtidos por meio de tentativa e erro observando a acurácia na classificação dos dados, em que se utilizou a validação cruzada com 10 etapas. O processo foi realizado no ambiente de trabalho *Orange*, versão 3.26.0 (DEMŠAR et al., 2013). Os parâmetros de treinamento das técnicas podem ser vistos na Tabela 6.1.

Tabela 6.1 - Configurações de entrada utilizados no treinamento das técnicas de aprendizado de máquina.

Técnica	Parâmetro	Valor
Árvore	Limite para divisão dos nós	90% de classificação
SVM	Função <i>kernel</i>	<i>RBF</i> , $\gamma = 0,07$
	Custo para amostras classificadas erradas	5
MLP	Número de camadas escondidas	3
	Número de neurônios	90, 60, 30
	Função de ativação	Tangente hiperbólica
	Número de épocas de treinamento	4000

A árvore de classificação foi treinada de forma a balancear a complexidade do modelo e a acurácia na classificação das amostras. No modelo final, a árvore contou com 225 nós e 113 folhas. O nó raiz classifica o conjunto amostral inicial a partir do valor da textura de Haralick de autocorrelação, o que demonstra a importância da medida de textura na classificação das imagens. A função *kernel* da técnica SVM foi definida como RBF (ver Equação 3.39), sendo o parâmetro da função,  $\gamma$ , definido como 0,07. O parâmetro custo, responsável por aumentar ou diminuir a largura da margem do hiperplano, foi definido como 5. A arquitetura da MLP contou com 3 camadas escondidas, sendo a primeira composta por 90 neurônios, a segunda por 60 e a terceira por 30, e a função de ativação foi definida como tangente hiperbólica.

Os conjuntos de treinamento foram organizados em planilhas Excel, que pe um formato suportado no ambiente de trabalho *Orange*. Cada linha fazia referência a uma imagem e as colunas eram utilizadas para estruturar as características extraídas das imagens. No Apêndice E pode ser visto um exemplo de tal planilha. A partir dessas, as técnicas de aprendizado de máquina foram treinadas.

A acurácia das técnicas de aprendizado de máquina de cada subconjunto de imagens pode ser obtida a partir das matrizes de confusão das mesmas. Essas podem ser encontradas no Apêndice F. Como todos os conjuntos de imagens foram utilizados para estimar os parâmetros das técnicas de aprendizado de máquina, não foi possível separar um grupo específico de imagens para avaliar o desempenho da abordagem autoadaptativa. Nesse sentido, foi realizada uma simulação com base nos erros e acertos de cada técnica a partir da distribuição de predições de cada uma das classes. A simulação foi realizada sorteando-se um número aleatório para cada imagem do conjunto de imagens em análise. Esse número seria responsável por classificar a imagem em uma das classes, respeitando a probabilidade de cada classe

dada pela matriz de confusão. Assim, foi possível simular o número de classificações corretas e incorretas para cada conjunto de imagens, bem como o melhor e pior caso de cada conjunto. Por último, ao incluir as classificações incorretas, foi possível calcular o desempenho da abordagem autoadaptativa e compará-la às abordagens não adaptativas. Cada simulação foi repetida 1000 vezes.

Nesse contexto, o presente capítulo se encontra organizado da seguinte forma. Na primeira seção, 6.1, são apresentados e discutidos os resultados referentes ao conjunto de dados obtido no Brasil. Em seguida, na Seção 6.2, são apresentados os resultados obtidos a partir dos dados do experimento realizado na Suécia. Na Seção 6.3 são descritos os resultados referentes aos conjuntos de dados decorrentes dos experimentos produzidos no México. Uma análise geral envolvendo todos os resultados discutidos anteriormente é apresentada na Seção 6.4. Por último, na Seção 6.5 são expostas as considerações finais do capítulo.

## 6.1 Resultados do experimento realizado no Brasil

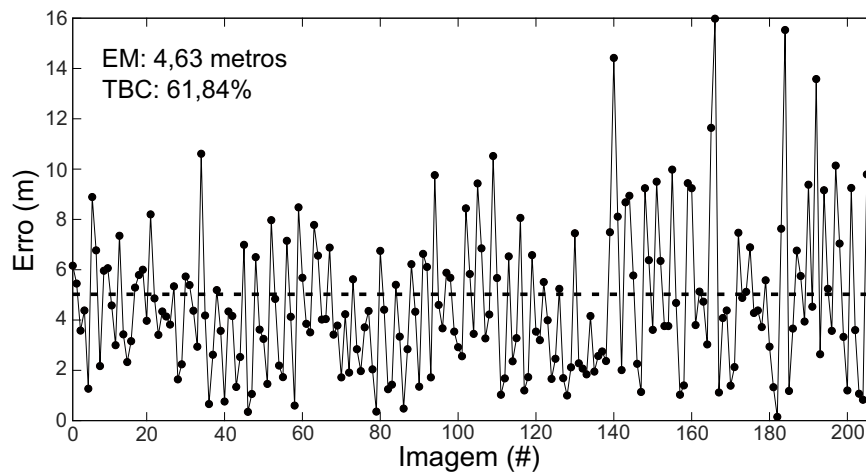
Nesta seção, são apresentados os resultados referentes ao conjunto de dados obtidos por meio do experimento realizado no Brasil. Esse conjunto conta com um total de 207 imagens, que foram processadas para cada uma das 2000 combinações de algoritmos. Assim, cada imagem do experimento pode se associar a 2000 combinações diferentes de algoritmos, que influenciam diretamente o desempenho da técnica de estimação da posição de VANTs. Nesse sentido, as análises aqui apresentadas buscam explorar essas possíveis associações entre imagens e combinações de algoritmos.

Primeiramente, os resultados foram obtidos por meio da abordagem não adaptativa, em que uma única combinação de algoritmos é utilizada para todo o conjunto de imagens. Nesse sentido, existem duas mil possibilidades não adaptativas, uma para cada combinação, sendo aquelas que apresentaram melhores métricas de erro selecionadas para análise. Para o conjunto de imagens analisado nesta seção, a combinação que apresentou menor EM também apresentou maior TBC. O erro da estimação da posição de cada uma das 207 imagens que compõe o conjunto pode ser visto na Figura 6.2.

Como pode ser visto na Figura 6.2, as estimações de posição apresentam erros cujas medidas de boas correspondências são intercaladas por medidas cujo erro é maior do que 5 metros. Esse fato revela a ausência de uma região específica e bem delimitada sobrevoada pelo VANT durante a execução do percurso. Além das características do ambiente de sobrevoo, também é importante lembrar que a distância em relação ao

solo da aeronave nesse experimento é de aproximadamente 32 metros, o mais baixo de todos os experimentos. Essa pode ser a causa responsável pela rápida alteração entre cenários diferentes de sobrevoo e também a causa para a rápida alternância nas medidas de erro.

Figura 6.2 - Erro da estimação da posição utilizando a abordagem não adaptativa aplicada ao conjunto de imagens do experimento realizado no Brasil.



A distribuição do erro ao longo da trajetória revela a ausência de grupos contíguos de imagens com erros similares, e alta variância. A linha tracejada representa o limite máximo de erro tolerável para medidas de boa correspondência.

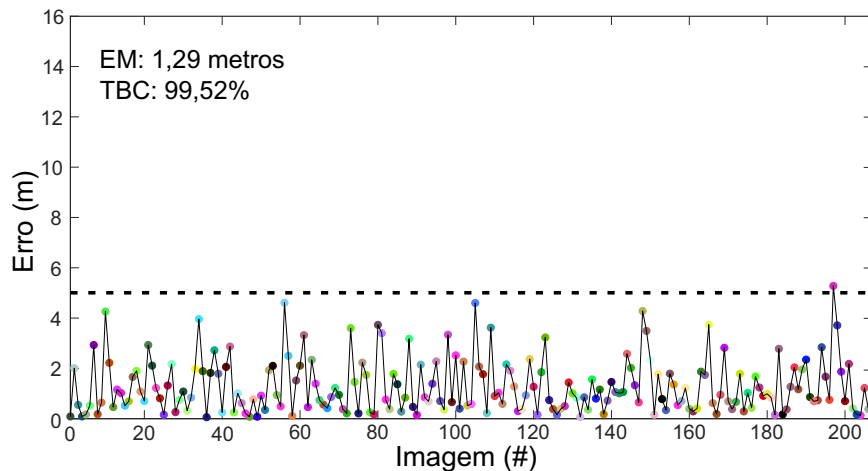
Fonte: Produção do autor.

Para o caso em análise, a combinação de algoritmos baseada na abordagem não adaptativa com menor erro possui EM de 4,63 metros e TBC de 61,84%. Os algoritmos utilizados em tal combinação são o filtro da média com janela  $7 \times 7$  aplicado à imagem capturada pelo VANT, o filtro da mediana com janela  $11 \times 11$  aplicado à imagem referência, e a CCNZ como medida de similaridade. Nenhum detector de bordas se fez necessário para essa combinação. Com base nos algoritmos utilizados, pode-se afirmar que a filtragem das imagens é importante para remover eventuais ruídos que prejudicam o cálculo da medida de similaridade. A ausência da etapa de detecção de bordas pode ser justificada pelo fato de que a medida de similaridade utilizada (CCNZ) é capaz de remover os efeitos da diferença de brilho entre as imagens, mas mantém inalterada a informação contida nas mesmas.

Em seguida, foi realizada uma análise ideal da abordagem adaptativa, na qual o mapeamento é realizado diretamente entre as instâncias do problema e as combina-

ções de algoritmos que apresentam melhor desempenho. Esse é um caso extremo, pois a associação entre imagens e combinações de algoritmos é feita de maneira artificial. Para o conjunto de dados aqui analisado, cada uma das 207 imagens foi associada à melhor combinação de algoritmos disponível. Os resultados dessa análise são exibidos na Figura 6.3.

Figura 6.3 - Erro da estimação da posição utilizando a abordagem adaptativa ideal aplicada ao conjunto de imagens do experimento realizado no Brasil.



A linha tracejada representa o limite máximo de erro tolerável para medidas de boa correspondência. Círculos de mesma cor representam estimações realizadas com uma mesma combinação de algoritmos.

Fonte: Produção do autor.

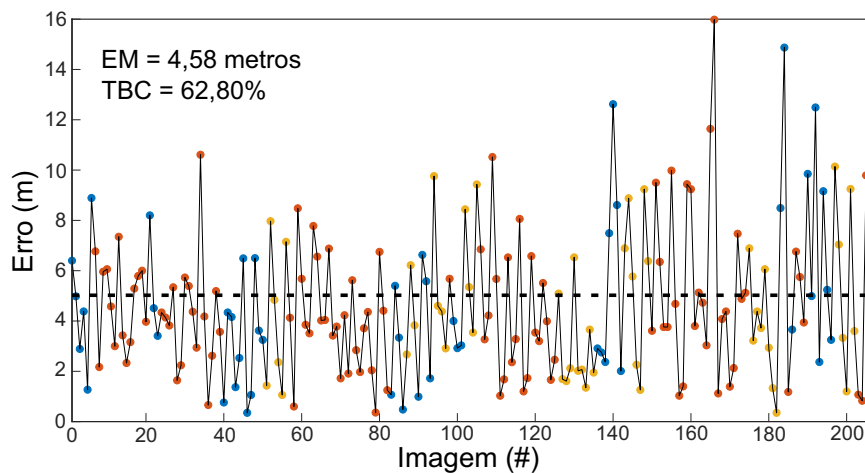
Conforme ilustrado na Figura 6.3, a ampla maioria das medidas ficou abaixo do limite de erro dos 5 metros, impactando nas métricas de erro, em que  $EM = 1,29$  metro e  $TBC = 99,51\%$ . Ao comparar tais resultados com a abordagem não adaptativa, pode-se notar uma diminuição do EM e aumento do TBC. Assim, pode-se afirmar que o emprego de diferentes combinações de algoritmos na estimação da posição de um VANT ao longo da sua trajetória se mostra útil para diminuir o erro da mesma.

Essa estratégia pode ser entendida como ideal, pois associa teoricamente a melhor combinação de algoritmos disponível à cada instância do problema. Dessa forma, o melhor desempenho possível é atingido, seja considerando o conjunto de imagens como um todo, seja considerando apenas imagens individuais. Porém, nem sempre é possível realizar o mapeamento direto entre as instâncias do problema e o espaço

de combinações de algoritmos, pois o mapa de seleção não pode ser construído na prática. Aqui, o mapa de seleção é composto por técnicas de aprendizado de máquina, o que equivale afirmar que não é possível treinar essas técnicas para realizar uma associação confiável entre as imagens capturadas em voo e as combinações de algoritmos. No conjunto de 207 imagens em análise, as técnicas de aprendizado de máquina deveriam ser capazes de classificá-las entre 161 classes diferentes (formadas pelas combinações de algoritmos), sendo 125 dessas classes composta por apenas uma imagem.

Nesse sentido, outra estratégia para abordagem adaptativa foi utilizada. Tal estratégia consiste em eleger subconjuntos de imagens para serem tratados como uma única classe, i.e., uma mesma combinação de algoritmos é empregada para todo um subconjunto de imagens. Assim, o desempenho individual de cada imagem do subconjunto é sacrificado em detrimento ao desempenho do grupo. Para o conjunto de imagens em análise, 3 subconjuntos foram definidos com base nas características das imagens. Os resultados dessa estratégia podem ser vistos na Figura 6.4.

Figura 6.4 - Erro da estimação da posição utilizando a abordagem adaptativa por regiões aplicada ao conjunto de imagens do experimento realizado no Brasil.



A linha tracejada representa o limite máximo de erro tolerável para medidas de boa correspondência. Círculos de mesma cor representam estimações realizadas com uma mesma combinação de algoritmos. Círculos com cor azul, vermelha e amarela representam a região A, B e C respectivamente.

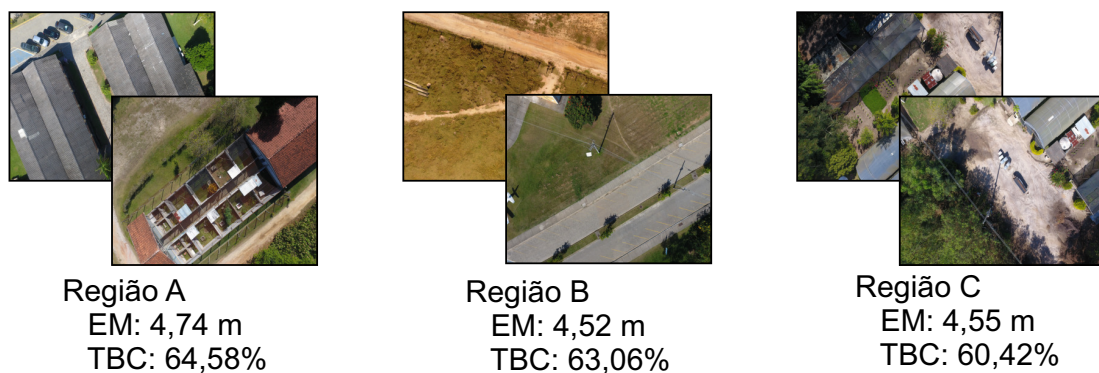
Fonte: Produção do autor.

Para cada dos 3 subconjuntos de imagens, foi atribuída a combinação de algoritmos



com melhor desempenho. Tais subconjuntos foram escolhidos a partir de uma inspeção visual das imagens, em que se buscou reunir as imagens com características semelhantes em um mesmo grupo, e da análise dos desempenho desses grupos, em que se buscou minimizar o erro. Assim, espera-se alcançar uma alta acurácia na classificação e também um ganho no desempenho dos resultados. Exemplos de imagens aéreas caracterizando as regiões de sobrevoo podem ser vistas na Figura 6.5.

Figura 6.5 - Exemplos de imagens dos subconjuntos referente ao experimento realizado no Brasil.



Duas imagens de cada subconjunto foram escolhidas como forma de ilustrar as características de cada região, cujas métricas de erro aparecem abaixo das mesmas.

Fonte: Produção do autor.

Como pode ser visto na Figura 6.5, as métricas de erro apresentam valores próximos independente da região. Além disso, as imagens possuem detalhes, tais como estradas e edificações, que auxiliam a técnica de estimação da posição. Essas características, somadas a outras intrínsecas ao experimento, tornam difícil a escolha de cenários adequados e bem delimitados para que a abordagem autoadaptativa apresente melhora no desempenho. Isso também pode ser entendido ao observar a Figura 6.4, cujo perfil das medidas de erro é semelhante ao perfil da abordagem não adaptativa (ver Figura 6.2). Nesse sentido, as métricas de EM e TBC apresentaram um incremento muito pequeno, de 4,63 metros para 4,57 metros e de 61,84% para 62,80% respectivamente. Aqui, porém, é importante salientar que a análise do erro não leva em consideração as eventuais classificações erradas das técnicas de aprendizado de máquina.

Embora o desempenho da presente análise não apresente o mesmo ganho que a estratégia adaptativa ideal, é possível treinar técnicas de aprendizado de máquina

que associem grupos de imagens a combinações de algoritmos mantendo um alto índice de acerto na classificação. As matrizes de confusão das técnicas podem ser vistas no Apêndice F. Na Tabela 6.2 podem ser vistas a acurácia de cada técnica por cenário.

Tabela 6.2 - Acurácia das técnicas de aprendizado de máquina por região referente ao experimento realizado no Brasil.

	Regiões		
	A	B	C
MLP	87,50%	96,40%	95,83%
SVM	79,17%	96,40%	93,75%
Árvore	89,58%	94,59%	79,17%

Dessa forma, a metodologia autoadaptativa por regiões pode ser testada e analisada. Para calcular a eficácia dessa abordagem, a taxa de acerto de cada técnica de aprendizado de máquina foi obtida a partir da matriz de confusão das referidas técnicas (ver Tabela 6.2). Em seguida, essa taxa foi utilizada para obter a probabilidade de acerto e erro e, assim, calcular o impacto que as classificações espúrias causam no desempenho da abordagem. Tal solução foi proposta como forma de resolver a falta de um conjunto de imagens para ser utilizado exclusivamente como validação dos classificadores e, assim, permitir que os resultados fossem comparados à abordagem não adaptativa. Tal solução foi executada 1000 vezes para cada técnica de aprendizado de máquina, nas quais as imagens classificadas foram sorteadas aleatoriamente observando a probabilidade de acerto e erro relativa a cada classe. Os resultados são exibidos na Tabela 6.3.

Tabela 6.3 - Resultados de EM e TBC ao considerar a taxa de acerto dos classificadores do conjunto de imagens do experimento realizado no Brasil.

	EM (metros)			TBC		
	Máximo	Médio	Mínimo	Máximo	Médio	Mínimo
MLP	4,76	4,60	4,51	63,77%	62,65%	61,35%
SVM	4,77	4,61	4,51	63,77%	62,60%	60,87%
Árvore	4,75	4,57	4,41	64,73%	62,89%	60,87%

Como apresentado na Tabela 6.3, ao incluir a taxa de classificações espúrias de cada técnica de aprendizado de máquina e simular o processo repetidas vezes, as métricas

de erro apresentam uma variação em torno de um valor médio. Isso ocorre porque cada imagem está associada a uma dada combinação de algoritmos, que produz um desempenho específico. Quando essa combinação é alterada, o desempenho também se altera, sendo este, às vezes, melhor do que o inicialmente previsto. Da mesma forma, o desempenho também pode se revelar pior.

Nesse sentido, a classificação imperfeita das técnicas resulta em um espectro de valores com uma margem para mais e para menos em torno de um valor médio. A margem está relacionada ao índice de acerto relativo a cada técnica e o valor médio está relacionado ao resultado de uma classificação 100% correta, que, para o caso em análise, tem  $EM = 4,58$  metros e  $TBC = 62,80\%$ . Por isso, quanto maior a acurácia de um classificador, mais estável serão seus resultados, pois menos alterações ocorrerão nas medidas de erro.

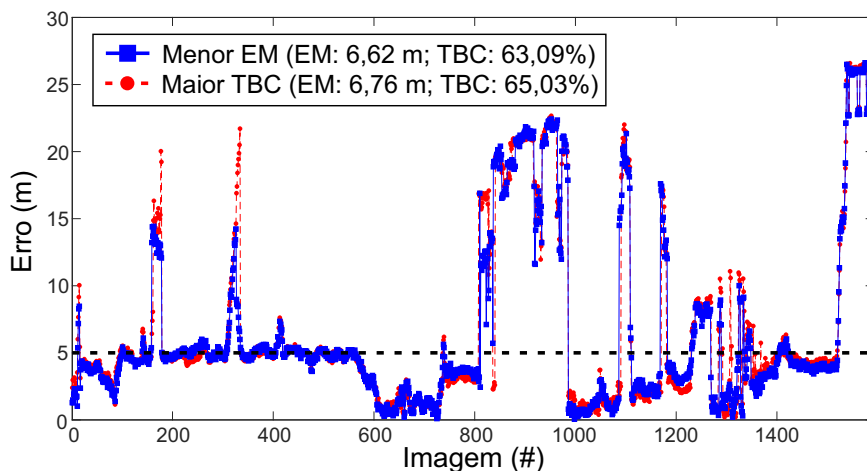
Como apresentado na Tabela 6.3, os resultados da abordagem autoadaptativa por regiões não apresentam ganho significativo quando comparados à abordagem não adaptativa. Ao contrário, no pior caso ainda apresentam uma pequena piora no desempenho. Isso pode ser consequência das características do experimento, que dificultaram a definição de grupos específicos de imagens e suas correspondentes associações às melhores combinações de algoritmos disponíveis. Talvez o número de subconjuntos de imagens, 3, tenha sido insuficiente para atribuir a devida combinação de algoritmos para cada imagem, e um número maior deveria ter sido utilizado. Porém, nesse caso, seria difícil obter técnicas de aprendizado de máquina aptas a classificar corretamente as imagens, como ocorre com a abordagem adaptativa ideal. De uma forma ou de outra, o resultado obtido nesta seção é relevante para a discussão final apresentada na Seção 6.4.

## 6.2 Resultados do experimento realizado na Suécia

A mesma estrutura de análise dos resultados apresentados na seção anterior, 6.1, é seguida na presente seção. Porém, o conjunto de imagens aqui analisado foi obtido a partir de um experimento realizado na Suécia e conta com um total de 1593 imagens (ver seção 5.2).

A primeira análise foi realizada com base nos resultados obtidos por meio da abordagem não adaptativa, na qual uma mesma combinação de algoritmos é utilizada para todas as imagens. A distribuição do erro ao longo da trajetória das duas abordagens não adaptativas que apresentaram menor EM e maior TBC podem ser vistas na Figura 6.6.

Figura 6.6 - Erro da estimação da posição utilizando a abordagem não adaptativa aplicada ao conjunto de imagens do experimento realizado na Suécia.



Os quadrados azuis ligados por uma linha contínua exibem as medidas de erro da abordagem não adaptativa com menor EM. Os círculos vermelhos ligados por uma linha tracejada demarcam as medidas de erro da abordagem não adaptativa com maior TBC. A distribuição das medidas de erro revela a presença de grupos de imagens cujos erros são similares, e.g., o grupo que compreende as imagens de número 800 a 1000, cujo erro médio é de aproximadamente 20 metros. A linha tracejada representa o limite máximo de erro tolerável para medidas de boa correspondência.

Fonte: Produção do autor.

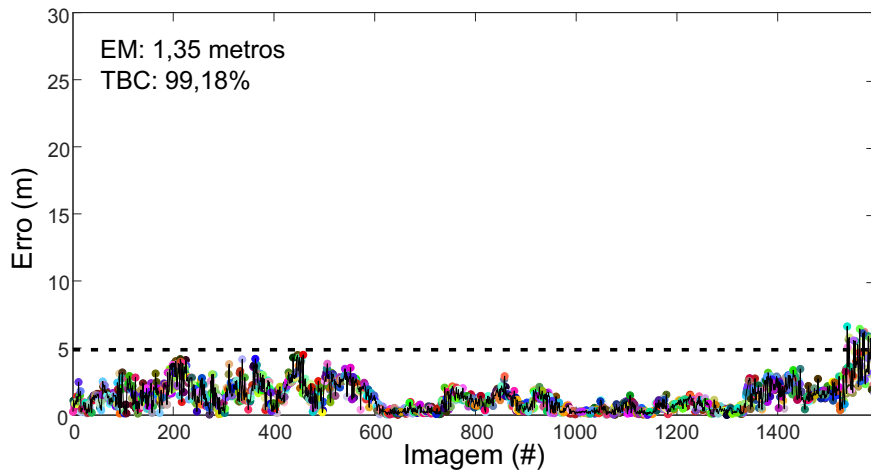
Conforme ilustrado na Figura 6.6, as duas abordagens não adaptativas que apresentaram menor EM e maior TBC possuem um comportamento similar e estão em grande parte sobrepostas. É interessante notar que, embora o resultado da abordagem não adaptativa do experimento anterior (ver Figura 6.2) apresenta TBC similar aos resultados do experimento aqui em análise, o comportamento das medidas de erro são bem diferentes. No experimento em análise, as medidas de erro apresentam um comportamento mais constante, em que grupos de imagens possuem erro próximo dos 5 metros, ora apresentam erro em torno dos 15 metros. Isso pode ser explicado pelas características do experimento e também do terreno sobrevoado, que impactam diretamente nas estimações da posição.

A combinação de algoritmos com menor EM é composta por um filtro de média com tamanho  $11 \times 11$ , aplicado à imagem referência, e a CCNZ como medida de similaridade. A combinação com maior TBC é composta pelos mesmos algoritmos, mas com uma janela de tamanho  $3 \times 3$ . Isso explica a proximidade das duas distribuições das medidas de erro. Além disso, é interessante perceber que o tamanho

da janela do filtro utilizado influencia as métricas de erro, sendo a janela grande responsável por diminuir o erro médio e a pequena por aumentar as medidas de boa correspondência.

Na sequência, foram obtidos resultados por meio da abordagem adaptativa ideal, em que cada imagem é associada diretamente à melhor combinação de algoritmos disponível. A distribuição das medidas de erro para o conjunto de imagens obtido no experimento realizado na Suécia pode ser vista na Figura 6.7.

Figura 6.7 - Erro da estimação da posição utilizando a abordagem adaptativa ideal aplicada ao conjunto de imagens do experimento realizado na Suécia.



A linha tracejada representa o limite máximo de erro tolerável para medidas de boa correspondência. Círculos de mesma cor representam estimações realizadas com uma mesma combinação de algoritmos.

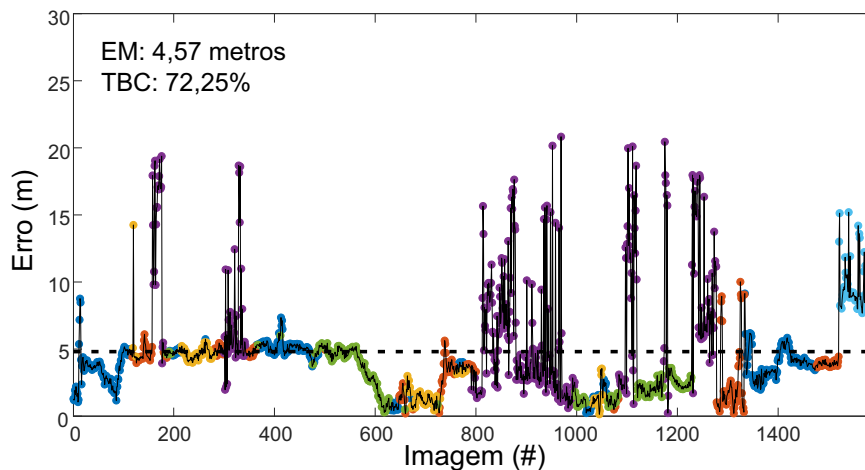
Fonte: Produção do autor.

Na abordagem ideal, as medidas de erro, em sua quase totalidade, ficaram abaixo do limite tolerável de erro, sendo caracterizadas como medidas de boa correspondência. Nesse sentido, as métricas de erro, com  $EM = 1,35$  metros e  $TBC = 99,18\%$ , são as melhores possíveis dado o espaço de combinações de algoritmos disponível neste trabalho. Porém, utilizando as técnicas de aprendizado de máquina apresentadas neste trabalho, não foi possível associar individualmente as imagens às devidas combinações de forma acurada.

Dessa forma, foi utilizada a estratégia por cenários, na qual subconjuntos de imagens com características semelhantes são associadas à melhor combinação disponível para

todo o grupo. Para o conjunto de imagens obtido no experimento realizado na Suécia, 6 regiões foram definidas para serem tratadas como subconjuntos de imagens. A distribuição das medidas de erro levando em conta as regiões definidas pode ser vista na Figura 6.8.

Figura 6.8 - Erro da estimação da posição utilizando a abordagem adaptativa por regiões aplicada ao conjunto de imagens do experimento realizado na Suécia.



A linha tracejada representa o limite máximo de erro tolerável para medidas de boa correspondência. Círculos de mesma cor representam estimações realizadas com uma mesma combinação de algoritmos. Círculos com cor azul, laranja, amarela, roxa, verde, azul claro representam as regiões A, B, C, D, E e F respectivamente.

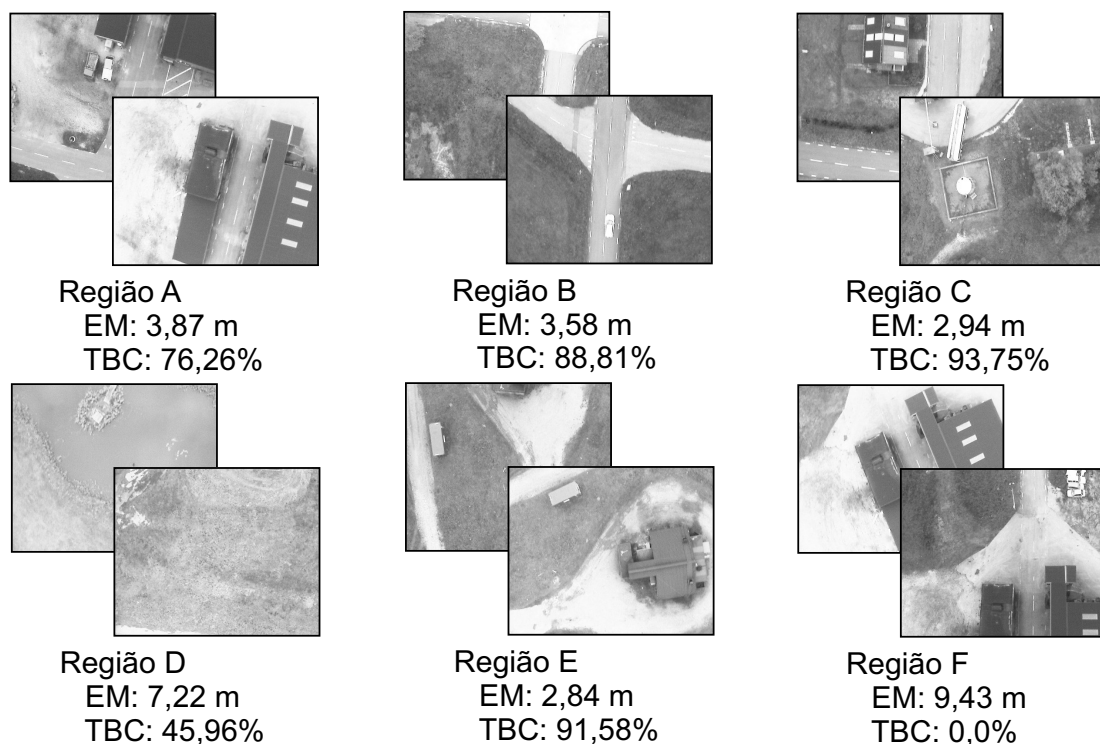
Fonte: Produção do autor.

Para cada uma das 6 regiões foi atribuída a combinação de algoritmos com melhor desempenho. A definição dos cenários para o conjunto de imagens em análise foi mais simples se comparado ao conjunto de imagens do experimento passado. Isso se deu principalmente porque o perfil das medidas de erro delimitou diferentes regiões de sobrevoo (ver Figura 6.6). Além disso, as características das imagens puderam ser associadas a tais regiões, como ilustrado na Figura 6.9, que apresenta duas imagens exemplo para cada região assim como suas respectivas métricas de erro.

As regiões A, B, C e E (respectivas regiões de cor azul, laranja, amarela e verde da Figura 6.8) são áreas de sobrevoo que apresentam métricas de erro com bom desempenho. Por outro lado, as regiões D e F (representadas pelas cores roxo e azul claro na Figura 6.8) apresentam métricas com desempenho pior. A primeira destas duas últimas, a região D, é composta por imagens com poucos detalhes, uma vez que

compreende cenários de lagos e gramados com pouca ou nenhuma estrutura urbana. O subconjunto F, por sua vez, é formado por imagens com estruturas semelhantes à região A, porém com métricas de erro piores. Como esse cenário é caracterizado por uma série de manobras realizadas pelo VANT antes do pouso, pode-se afirmar que os movimentos da aeronave causaram distorções nas imagens, o que pode ter sido a causa para a piora dos resultados.

Figura 6.9 - Exemplos de imagens dos subconjuntos referente ao experimento realizado na Suécia.



Duas imagens de cada subconjunto foram escolhidas como forma de ilustrar as características de cada região, cujas métricas de erro aparecem abaixo das mesmas.

Fonte: Produção do autor.

Nesse contexto, embora a combinação de algoritmos com melhor desempenho tenha sido atribuída à cada subconjunto de imagens, alguns apresentaram métricas de erro insatisfatórias, e.g., a região D. Isso revela a ausência de uma combinação adequada para tais subconjuntos, que terminam por prejudicar o desempenho de todo o conjunto. Além disso, é importante observar que as regiões A, B, C e E estão associadas à combinações de algoritmos que não utilizam detectores de bordas. Por outro lado,

as regiões D e F utilizam, respectivamente, o operador de Canny e o extrator baseado em congruência de fase. Dessa forma, pode-se afirmar que a detecção de bordas auxilia a estimação da posição em cenários com essas características.

Conforme exibido na Figura 6.4, ao definir as regiões e atribuir a melhor combinação para cada cenário, as métricas de erro apresentam melhora quando comparadas à da abordagem não adaptativa. No caso do EM, ocorre uma melhora de 6,62 metros para 4,57 metros e, para a TBC, de 65,03% para 72,25%. Esses valores, porém, não consideram a taxa de acurácia de cada técnica de aprendizado de máquina e, conseqüentemente, as eventuais classificações erradas. Nesse sentido, para averiguar a eficácia dessa estratégia, as taxas de acerto de cada técnica de aprendizado de máquina foram consideradas no cálculo das métricas de erro. A acurácia das técnicas associada a cada região foi obtida junto a matriz de confusão (ver Apêndice F). A acurácia de cada técnica pode ser vista na Tabela 6.4.

Tabela 6.4 - Acurácia das técnicas de aprendizado de máquina por região referente ao experimento realizado na Suécia.

	Região					
	A	B	C	D	E	F
MLP	93,70%	89,18%	92,36%	94,99%	93,77%	98,63%
SVM	90,34%	85,45%	87,50%	88,58%	87,55%	94,52%
Árvore	85,29%	83,96%	84,03%	92,76%	83,88%	84,93%

O processo, repetido 1000 vezes, classifica cada imagem do conjunto de forma aleatória levando em consideração a taxa de acerto e erro de cada uma das técnicas de aprendizado de máquina. Dessa forma, a estratégia autoadaptativa pode ser testada considerando a influência dos acertos e erros de cada classificador. Os resultados são exibidos na Tabela 6.5.

Tabela 6.5 - Resultados de EM e TBC ao considerar a taxa de acerto dos classificadores do conjunto de imagens do experimento realizado na Suécia.

	EM (metros)			TBC		
	Máximo	Médio	Mínimo	Máximo	Médio	Mínimo
MLP	4,85	4,73	4,63	72,00%	71,04%	70,06%
SVM	5,05	4,89	4,76	71,37%	70,13%	68,99%
Árvore	5,15	4,94	4,76	71,25%	69,91%	68,36%



Na Tabela 6.5 podem ser vistos o melhor e o pior caso das duas métricas de erro ao incluir as classificações espúrias de cada técnica de aprendizado de máquina. É interessante comparar tais resultados com a acurácia de cada técnica (ver Tabela 6.4) e perceber que o desempenho está diretamente relacionado à acurácia das técnicas, i.e., quanto maior a acurácia, melhor o desempenho. Dessa forma, é importante notar que o desempenho de cada técnica varia, sendo o melhor desempenho dado pela MLP, seguida pela SVM, e a árvore de classificação na última posição. A explicação se dá pelo fato de que, à medida que a acurácia das técnicas diminui, a quantidade de imagens classificadas erroneamente aumenta, causando diminuição no desempenho das métricas de erro. Por esse motivo, caso fosse considerado, hipoteticamente, classificações 100% corretas, o EM seria 4,57 metros e a TBC 72,25%, um resultado não alcançado nem no melhor caso da MLP. Por último, vale comparar esses resultados aos resultados das abordagens não adaptativas (ver Figura 6.6) para perceber um incremento no desempenho de ambas as métricas de erro independente do classificador utilizado.

Nesse contexto, a abordagem autoadaptativa por regiões apresentou ganho no sentido de diminuir o erro da estimação da posição por imagens ao se considerar a trajetória completa realizada pelo VANT. No experimento aqui em análise, pode-se afirmar que as regiões de sobrevoo puderam ser mais facilmente identificadas do que o experimento passado (ver Seção 6.1), o que facilitou a escolha dessas regiões.

Por último, ao analisar os resultados da abordagem adaptativa ideal, pode-se afirmar que as métricas de erro ainda podem ser aprimoradas. Porém, é difícil delimitar regiões de sobrevoo de forma a diminuir o erro e ainda manter uma alta acurácia de classificação. É importante notar que não é sempre que a melhor combinação de algoritmos de uma região apresenta um desempenho com alta TBC. Na região D, por exemplo, a TBC foi de 45% por ser uma região cujas imagens apresentaram poucos detalhes que auxiliavam a técnica de estimação da posição. Nesses casos, pode ser interessante ampliar o espaço de combinação de algoritmos com novas técnicas. Assim, os mesmos classificadores poderiam ser utilizados, pois as regiões continuariam sendo as mesmas, mas o desempenho destas seriam diferentes.

### **6.3 Resultados dos experimentos realizados no México**

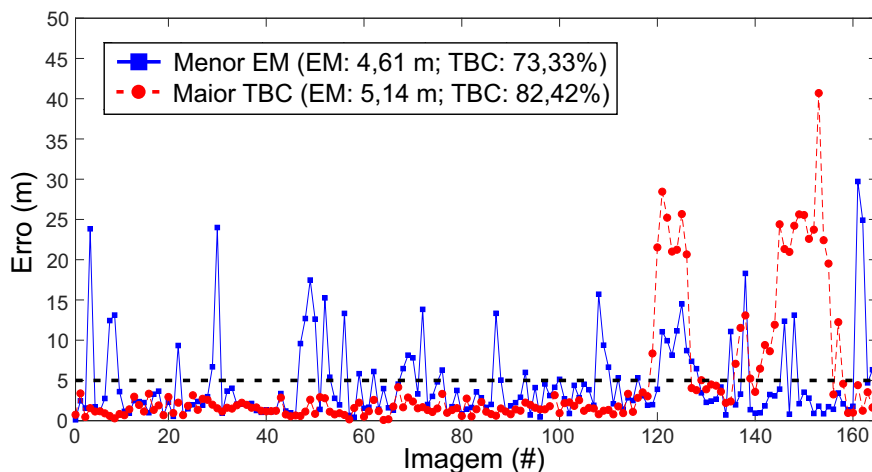
No México, foram realizados 2 experimentos diferentes com a finalidade de obter conjuntos de imagens aéreas registradas por VANTs (ver seção 5.3). Por esse motivo, a presente seção se encontra dividida em 2 subseções, uma para cada experimento. Em ambos os casos, a estrutura de análise dos resultados é idêntica e segue a mesma

organização apresentada nas seções anteriores.

### 6.3.1 Primeiro experimento

O conjunto de imagens obtido no primeiro experimento realizado no México conta com um total de 165 arquivos. A primeira análise realizada nesse conjunto é baseada nos resultados obtidos por meio abordagem não adaptativa, em que a melhor combinação de algoritmos é escolhida considerando o conjunto de imagens como um todo. A distribuição das medidas de erro ao longo da trajetória do VANT pode ser vista na Figura 6.10.

Figura 6.10 - Erro da estimação da posição utilizando a abordagem não adaptativa aplicada ao conjunto de imagens do primeiro experimento realizado no México.



Os quadrados azuis ligados por uma linha contínua exibem as medidas de erro da abordagem não adaptativa com menor EM. Os círculos vermelhos ligados por uma linha tracejada demarcam as medidas de erro da abordagem não adaptativa com maior TBC. A linha tracejada preta representa o limite máximo de erro tolerável para medidas de boa correspondência.

Fonte: Produção do autor.

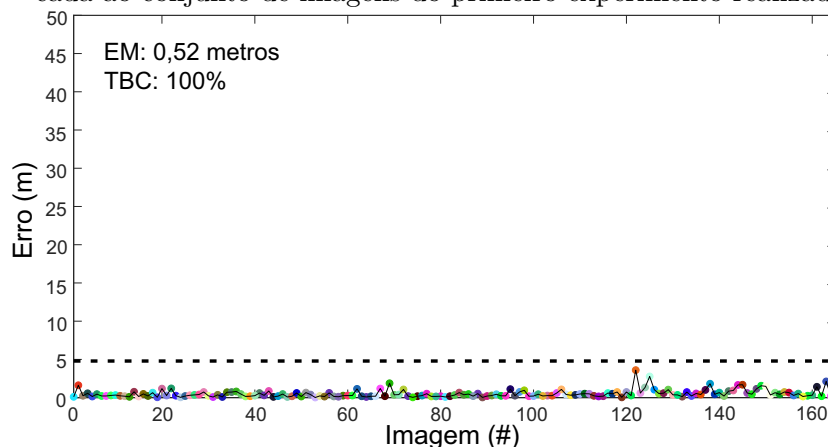
Como pode ser visto na Figura 6.10, as duas abordagens não adaptativas com melhor desempenho considerando as métricas EM e TBC apresentam comportamentos diferentes. No caso do menor EM, medidas que não podem ser consideradas de boa correspondência surgem esporadicamente ao longo de toda trajetória. No caso da combinação que apresentou maior TBC, tais medidas só aparecem em regiões bem definidas ao final da trajetória. Nesse sentido, pode-se afirmar que as características do terreno sobrevoado causam direta influência nas duas combinações de algoritmos

aqui discutidas.

A combinação com menor EM é composta pelo filtro de Gauss com janela  $3 \times 3$  aplicado à imagem capturada pelo VANT, o filtro da mediana de mesmo tamanho aplicado à imagem referência e o extrator de bordas baseado em congruência de fase utilizado em ambas imagens. A medida de similaridade utilizada é a CC. No caso da combinação com maior TBC, o mesmo filtro é aplicada à imagem capturada pelo VANT, porém, na imagem referência, foi utilizado o filtro da média com janela de tamanho  $3 \times 3$ . Nenhum extrator de bordas foi aplicado às imagens e a medida de similaridade empregada foi a CCNZ. Dessa forma, pode-se afirmar que a primeira combinação de algoritmos utilizada (com extração de bordas) influenciou o comportamento das medidas de erro, pois, ao eliminar diversos detalhes presentes nas imagens, terminou por diminuir o desempenho da medida em diversas ocasiões ao longo do voo, mas também ajudou a aumentar o mesmo nas regiões localizadas ao final da trajetória. Por outro lado, na combinação com maior TBC, ao contar com todos os detalhes presente nas imagens, as medidas de erro se mantiveram aproximadamente constantes e abaixo do limite de erro tolerável até alcançarem as regiões ao final da trajetória, onde o erro aumentou drasticamente.

Em seguida, os resultados referentes à abordagem adaptativa ideal foram calculados. Nessa estratégia, cada imagem é associada à melhor combinação de algoritmos, sendo tal associação realizada de maneira hipotética. A distribuição das medidas de erro para essa estratégia podem ser vistas na Figura 6.11.

Figura 6.11 - Erro da estimação da posição utilizando a abordagem adaptativa ideal aplicada ao conjunto de imagens do primeiro experimento realizado no México.

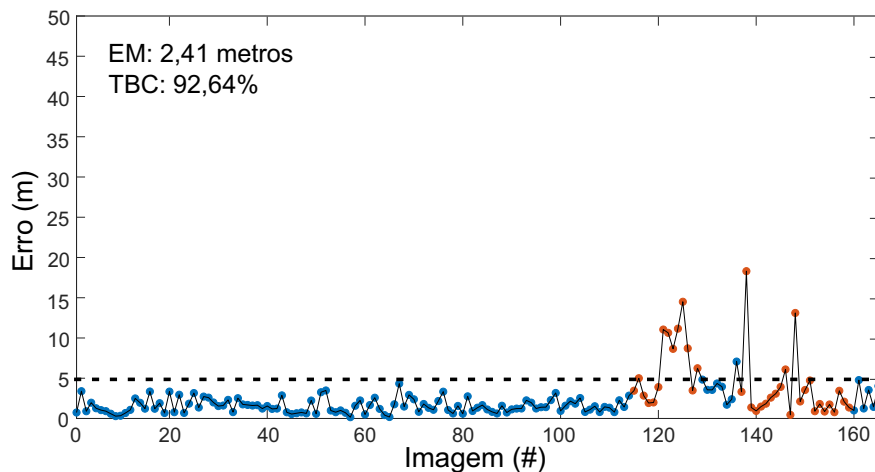


A linha tracejada representa o limite máximo de erro tolerável para medidas de boa correspondência. Círculos de mesma cor representam estimações realizadas com uma mesma combinação de algoritmos.

Fonte: Produção do autor.

A estratégia ideal apresenta um perfil de distribuição das medidas de erro muito próximo ao zero. Nesse caso, 100% das medidas ficaram abaixo do limite de 5 metros e foram classificadas como medidas de boa correspondência. Infelizmente, a tarefa de associar, individualmente, as imagens à melhor combinação de algoritmos disponível é muito complexa para que as técnicas de aprendizado de máquina analisadas neste trabalho apresentassem um nível de acurácia satisfatório. Nesse sentido, a estratégia por cenários, ou subconjuntos de imagens, foi utilizada. Para o caso em análise, apenas 2 subconjuntos de imagens foram escolhidos. O resultado das medidas de erro considerando os 2 cenários pode ser visto na Figura 6.12.

Figura 6.12 - Erro da estimação da posição utilizando a abordagem adaptativa por regiões aplicada ao conjunto de imagens do primeiro experimento realizado no México.



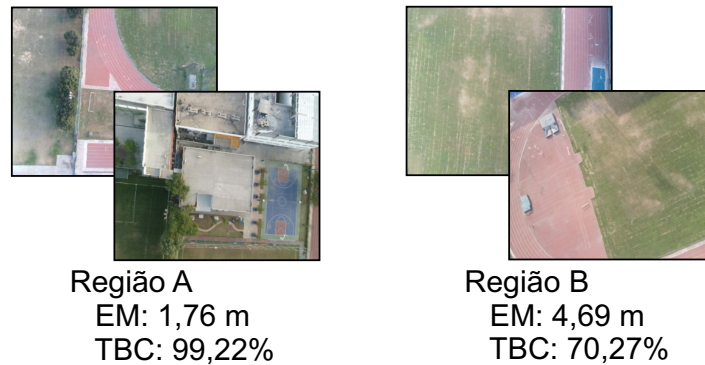
A linha tracejada representa o limite máximo de erro tolerável para medidas de boa correspondência. Círculos de mesma cor representam estimações realizadas com uma mesma combinação de algoritmos. Círculos com cor azul e laranja representam as regiões A e B respectivamente.

Fonte: Produção do autor.

As duas regiões foram escolhidas com base na abordagem não adaptativa, em que a combinação com melhor desempenho foi elegida para cada caso. É interessante notar que outros cenários poderiam ser escolhidos, fragmentando a região A (representada pela cor azul na Figura 6.12). Porém, o ganho seria pequeno, uma vez que a distribuição das medidas de erro estão próximas de zero. Além disso, a complexidade para classificar as imagens nas devidas regiões aumentaria, o que influenciaria o desempenho final da abordagem autoadaptativa por regiões. Exemplo de imagens

de ambas regiões podem ser vistas na Figura 6.13.

Figura 6.13 - Exemplos de imagens dos subconjuntos referente ao primeiro experimento realizado no México.



Duas imagens de cada subconjunto foram escolhidas como forma de ilustrar as características de cada região. As métricas de erro de cada região estão abaixo das imagens.

Fonte: Produção do autor.

As combinações de algoritmos utilizadas na abordagem autoadaptativa por regiões é semelhante às combinações empregadas na abordagem não adaptativa. A única diferença se deu na manipulação dos filtros, em que na região A não foi utilizado nenhum filtro, e na região B foi utilizado o filtro de Gauss com tamanho  $3 \times 3$ . A mesma explicação apresentada na abordagem anterior pode ser usada para explicar a presente análise. A diferença na composição dos filtros, entretanto, pode ser justificada pela especialização das combinações de algoritmos nos referidos subconjuntos de imagens. Por último, vale salientar que a combinação de algoritmos associada à região B (representada pela cor laranja na Figura 6.12) aumenta o desempenho do conjunto de imagens, mas não elimina completamente as medidas que não foram classificadas como boa correspondência.

Porém, as métricas de erro consideradas até aqui não levam em consideração a acurácia da classificação das técnicas de aprendizado de máquina. A taxa de acerto foi obtida a partir da matriz de confusão (ver Apêndice F) e pode ser vista, para cada região, na Tabela 6.6.

Nesse contexto, a acurácia de cada técnica de aprendizado de máquina foi considerada no cálculo do desempenho da abordagem autoadaptativa por regiões. O cálculo foi repetido 1000 vezes em um processo aleatório que respeitava a probabilidade de

Tabela 6.6 - Acurácia das técnicas de aprendizado de máquina por região referente ao primeiro experimento realizado no México.

	Região	
	A	B
MLP	100%	94,60%
SVM	99,22%	94,60%
Árvore	99,22%	100%

acerto e erro das técnicas. Por último, as métricas de erro foram estimadas, assim como o pior e melhor caso entre as 1000 repetições. Os resultados podem ser vistos na Tabela 6.7.

Tabela 6.7 - Resultados de EM e TBC ao considerar a taxa de acerto dos classificadores do conjunto de imagens do primeiro experimento realizado no México.

	EM (metros)			TBC		
	Máximo	Médio	Mínimo	Máximo	Médio	Mínimo
MLP	3,46	2,55	2,36	93,94%	92,21%	89,70%
SVM	3,57	2,59	2,37	93,94%	91,99%	89,70%
Árvore	2,66	2,43	2,40	92,73%	92,56%	90,91%

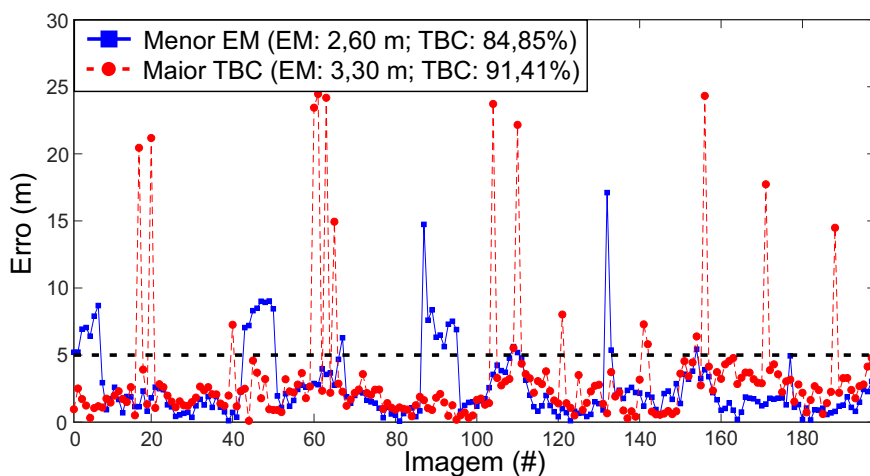
Ao incluir as classificações espúrias de cada técnica de aprendizado de máquina, as métricas de erro apresentam uma variação em torno de um valor médio. Isso ocorre por conta das imagens associadas à combinação de algoritmos incorreta, ocasionando, na maior parte dos casos, influência negativa no desempenho. A variação no desempenho está diretamente associada à acurácia da técnica de aprendizado de máquina. Aqui, a árvore de classificação apresentou maior acurácia que as demais técnicas, o que resultou em melhores resultados quando comparado às demais técnicas.

Porém, no caso médio, as métricas de erro estão próximas umas das outras, com EM de aproximadamente 2,5 metros e TBC em torno dos 92%. Quando esses valores são comparados à abordagem não adaptativa, pode-se notar um aumento significativo, principalmente quando a média do erro é considerada. Mesmo tomando os piores casos de cada técnica, o ganho ainda é positivo. Dessa forma, pode-se considerar que a abordagem autoadaptativa por regiões apresentou desempenho superior quando comparada à abordagem não adaptativa.

### 6.3.2 Segundo experimento

O segundo experimento realizado no México conta com um conjunto de 198 imagens. Primeiramente, a análise foi realizada utilizando uma única combinação de algoritmos para todo o conjunto, sendo, por isso, chamada de abordagem não adaptativa. O perfil das medidas de erro para cada imagem ao longo da trajetória do VANT pode ser visto na Figura 6.14.

Figura 6.14 - Erro da estimação da posição utilizando a abordagem não adaptativa aplicada ao conjunto de imagens do segundo experimento realizado no México.



Os quadrados azuis ligados por uma linha contínua exibem as medidas de erro da abordagem não adaptativa com menor EM. Os círculos vermelhos ligados por uma linha tracejada demarcam as medidas de erro da abordagem não adaptativa com maior TBC. A linha tracejada preta representa o limite máximo de erro tolerável para medidas de boa correspondência.

Fonte: Produção do autor.

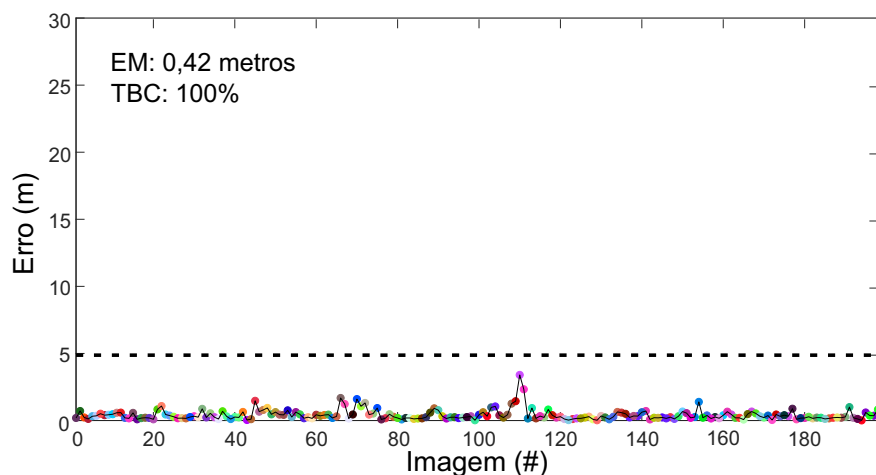
Ao observar a Figura 6.14 pode-se notar que a maior parte das medidas de erro estão abaixo da linha dos 5 metros, que separa as medidas de boa correspondência das outras medidas. Porém, em diversos pontos ao longo da trajetória, as combinações apresentam medidas com erro que podem alcançar 25 metros. A principal explicação para esse comportamento se dá ao observar a trajetória do VANT, que tem formato de espiral e sobrevoa as mesmas regiões várias vezes ao longo do voo. Nesse sentido, pode-se afirmar que as características do cenário impactam no desempenho das medidas de erro e, toda vez que o VANT sobrevoa tais regiões, o desempenho piora.

A combinação com menor EM é composta apenas pelo filtro da média com janela de

tamanho  $7 \times 7$  aplicado à imagem referência e a CCNZ como medida de similaridade. É interessante notar que essa combinação possui bom desempenho na maior parte da trajetória, salvo em pequenos grupos de imagens, nos quais o erro fica próximo dos 10 metros. A combinação com maior TBC, por sua vez, é formada por dois filtros, o de Gauss e o da média, ambos com janela de tamanho  $3 \times 3$ , sendo o primeiro aplicado à imagem capturada pelo VANT e o segundo à imagem referência, o detector de bordas baseado em congruência de fase e a CCN. O comportamento é similar ao caso anterior, em que grupos de imagens apresentam comportamento diferente do restante da trajetória e chegam a alcançar erros de 25 metros. Aqui, porém, os grupos de imagens não equivalem aos mesmos da combinação anterior, o que pode ser justificado pela adição da etapa de detecção de bordas. Esta elimina detalhes das imagens que podem ser úteis em alguns casos, mas que terminam diminuindo o desempenho da técnica em outros casos.

Na sequência, a abordagem adaptativa ideal, na qual cada imagem é individualmente associada à melhor combinação de algoritmos disponível, foi analisada. Vale lembrar que a associação é feita sem a utilização de nenhuma técnica de aprendizado de máquina. O perfil de distribuição das medidas de erro para essa abordagem pode ser visto na Figura 6.15.

Figura 6.15 - Erro da estimação da posição utilizando a abordagem adaptativa ideal aplicada ao conjunto de imagens do segundo experimento realizado no México.



A linha tracejada representa o limite máximo de erro tolerável para medidas de boa correspondência. Círculos de mesma cor representam estimações realizadas com uma mesma combinação de algoritmos.

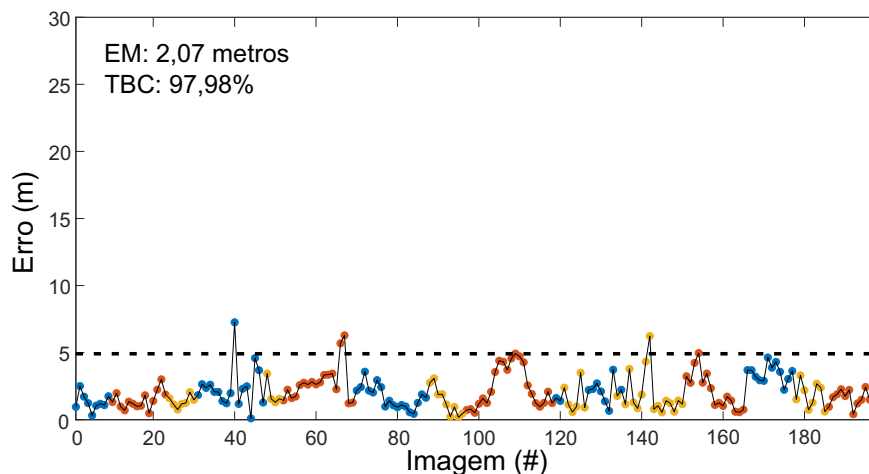
Fonte: Produção do autor.



Como pode ser visto na Figura 6.11, todas as medidas de erro foram classificadas como medidas de boa correspondência. Além disso, a média do erro é de aproximadamente 42 centímetros. Essas informações revelam o potencial da adaptação na estimação da posição por imagens. Porém, assim como nas análises dos experimentos anteriores, a estratégia ideal é de difícil efetivação, pois é difícil associar, com alta acurácia, cada imagem à combinação com melhor desempenho.

Dessa forma, a estratégia baseada em regiões, ou subconjunto de imagens, foi adotada. Para o conjunto de imagens analisado nesta seção, 3 regiões foram definidas. A distribuição das medidas de erro considerando tais regiões pode ser vista na Figura 6.16

Figura 6.16 - Erro da estimação da posição utilizando a abordagem adaptativa por regiões aplicada ao conjunto de imagens do segundo experimento realizado no México.



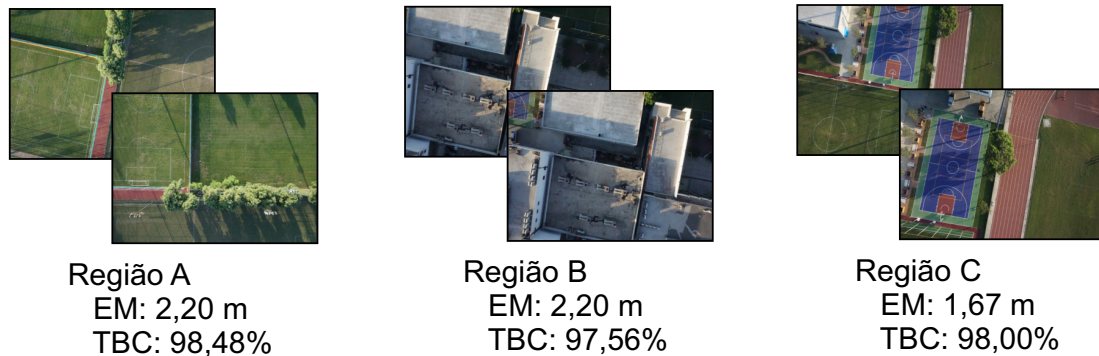
A linha tracejada representa o limite máximo de erro tolerável para medidas de boa correspondência. Círculos de mesma cor representam estimações realizadas com uma mesma combinação de algoritmos. Círculos de cor azul, vermelho e amarelo representam as regiões A, B e C respectivamente.

Fonte: Produção do autor.

Como ilustrado na Figura 6.16, as imagens foram divididas em 3 subconjuntos distintos. Para cada subconjunto, as combinações com melhor desempenho foram escolhidas. As regiões foram delimitadas com base na características das mesmas e também com base no perfil das medidas de erro da estratégia não adaptativa (ver Figura 6.14). Também é interessante notar que as regiões se repetem, pois a traje-

tória realizada pelo VANT tem formato de espiral, o que faz com que as mesmas regiões sejam revisitadas. Exemplos de imagens aéreas dessas regiões podem ser vistos na Figura 6.17.

Figura 6.17 - Exemplos de imagens dos subconjuntos referente ao segundo experimento realizado no México.



Duas imagens de cada subconjunto foram escolhidas como forma de ilustrar as características de cada região. As métricas de erro de cada região estão abaixo das imagens.

Fonte: Produção do autor.

Uma combinação de algoritmos específica foi utilizada em cada região sobrevoada pelo VANT. Nessas combinações, é interessante ressaltar o fato de que o detector de bordas baseado em congruência de fase foi utilizado nas regiões A e C. E, na região B, nenhum extrator de bordas se fez necessário. Com base nos detalhes presentes nas imagens da Figura 6.17, pode-se dizer que as bordas presentes nas imagens das regiões A e C são retas e ocupam quase a totalidade da imagem. O mesmo não ocorre nas imagens da região B, cujas bordas são mais complexas e se repartem diversas vezes, acompanhando as curvas dos telhados das edificações presentes. Essa é, provavelmente, a justificativa para os diferentes desempenhos apresentados por tais combinações. E também explica o comportamento das duas combinações com melhor desempenho da estratégia não adaptativa.

Como já explicado nas análises dos experimentos anteriores, as métricas de erro aqui obtidas não consideram a taxa de acerto e erro das técnicas de aprendizado de máquina. A acurácia de tais técnicas foi calculada a partir da matriz de confusão das mesmas (ver apêndice F). Na Tabela 6.8 podem ser vistas as acurácias de cada uma das técnicas para cada uma das regiões definidas para o conjunto de imagens do experimento em análise.

Tabela 6.8 - Acurácia das técnicas de aprendizado de máquina por região referente ao segundo experimento realizado no México.

	Região		
	A	B	C
MLP	95,45%	97,56%	96,00%
SVM	98,48%	97,56%	86,00%
Árvore	86,36%	93,90%	74,00%

Para considerar a acurácia das técnicas de aprendizado de máquina no cálculo das métricas de erro, um processo iterativo foi realizado. Este processo foi repetido 1000 vezes e levou em consideração a taxa de acerto de cada técnica. A partir das repetições, o pior e melhor caso de cada métrica de erro foi obtida. Os resultados podem ser vistos na Tabela 6.9.

Tabela 6.9 - Resultados de EM e TBC ao considerar a taxa de acerto dos classificadores do conjunto de imagens do segundo experimento realizado no México.

	EM (metros)			TBC		
	Máximo	Médio	Mínimo	Máximo	Médio	Mínimo
MLP	2,53	2,15	2,01	98,99%	97,36%	94,95%
SVM	2,49	2,14	2,01	98,99%	97,27%	94,44%
Árvore	2,75	2,30	1,99	98,99%	95,80%	92,42%

Incluindo a acurácia na classificação de cada região, pode-se notar que esta influencia as métricas de erro. Esta influência se explica pelas classificações erradas, que ocasionam uma piora ou, às vezes, melhora nos resultados. Nesse sentido, o pior e melhor caso de cada métrica pode ser analisado, assim como a média de todas as repetições. Comparando a Tabela 6.8 com a Tabela 6.9, pode-se observar que a variação no desempenho das métricas de erro está associada à acurácia de cada técnica de aprendizado de máquina. Aqui, a MLP e a SVM possuem maior acurácia que a árvore de classificação, o que ocasionou em melhores resultados quando comparado à esta última.

No caso médio, independente da técnica de aprendizado de máquina utilizada, ambas as métricas apresentaram ganho quando comparados à estratégia não adaptativa. O único caso que apresenta piora é o pior caso da árvore de classificação, cujo EM = 2,75 metros não supera os 2,60 metros da abordagem não adaptativa. Em todos os demais casos, a abordagem adaptativa por regiões apresenta ganho de desempenho.

## 6.4 Análise dos resultados

Neste capítulo, quatro conjuntos de dados obtidos por quatro experimentos diferentes foram analisados e discutidos. Cada experimento possui suas próprias características, tais como terreno sobrevoado, trajetória realizada pelo VANT, altitude da aeronave e sensor imageador a bordo desta. Todas essas diferenças contribuíram para os resultados encontrados, que, levando em consideração a abordagem não adaptativa, apresentaram diferentes combinações de algoritmos com melhor desempenho. Nesse sentido, pode-se afirmar que as peculiaridades de cada experimento demandam por uma combinação específica, que apresenta melhor desempenho que as demais.

Porém, embora experimentos distintos estejam relacionados a combinações de algoritmos também distintas, uma segunda análise foi realizada. Nessa, cada imagem foi associada à melhor combinação disponível no espaço de algoritmos. Essa abordagem extrema, denominada ideal, foi realizada para verificar o potencial de ganho quando comparado à abordagem não adaptativa. Em todos os casos, os resultados apresentaram desempenho superior. Porém, tal estratégia é de difícil execução, pois a associação entre as imagens e as combinações de algoritmos desejada é muito complexa.

Nesse sentido, uma terceira análise foi realizada. Nessa, subconjuntos de imagens foram associadas às combinações que apresentaram o melhor desempenho para todo o grupo. Tais grupos foram divididos com base nas características presentes nas imagens, e, por isso, estão atreladas às regiões de sobrevoo. Dessa forma, cada combinação de algoritmos foi associada a uma região de sobrevoo específica, cujas imagens possuem similaridades entre si.

Para realizar a associação entre os subconjuntos de imagens e as combinações de algoritmos, foram utilizadas técnicas de aprendizado de máquina. Em específico, uma MLP, uma SVM e uma árvore de classificação. Por serem técnicas diferentes, cada uma delas apresentou uma acurácia diferente. A taxa de acerto das três técnicas citadas por ser vista na Tabela 6.10.

Como pode ser visto na Tabela 6.10, a técnica que apresentou maior acurácia considerando todos os quatro conjuntos de imagens foi a MLP. Isso pode ser explicado pelo fato de que a MLP é capaz de dividir o espaço de características de forma não linear, ao contrário do que a árvore de classificação é capaz de fazer. Esse fato se confirma ao analisar a acurácia da SVM, que também divide o espaço de forma

Tabela 6.10 - Acurácia de classificação das técnicas de aprendizado de máquina considerando todos os conjuntos de imagens.

Técnica	Acurácia
MLP	94,1%
SVM	90,2%
Árvore	87,7%

linear, com hiperplanos, mas que é capaz de modificá-lo por meio da função *kernel*. Além disso, pode-se afirmar que a taxa de acerto dos classificadores não é única para cada conjunto de imagens, uma vez que, no primeiro experimento realizado no México, a árvore de classificação obteve maior acurácia que os demais.

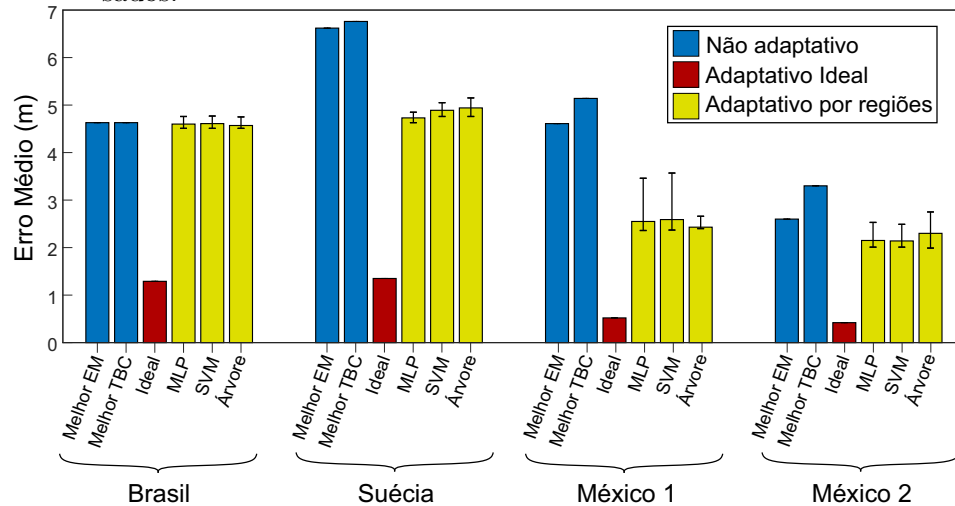
De qualquer forma, a taxa de acerto dos classificadores implica em resultados cujos valores apresentam uma margem em torno de um ponto médio. Nesse sentido, a acurácia dos classificadores tem impacto direto na estabilidade dos resultados, pois cada associação errada faz com que a imagem seja processada por uma combinação de algoritmos que não foi inicialmente prevista. Caso existisse um classificador perfeito, todas as imagens seriam associadas corretamente e não haveria variação nos resultados.

Nesse contexto, o desempenho da abordagem autoadaptativa por regiões apresenta variação em torno de um pior e um melhor caso. O mesmo não acontece com a abordagem não adaptativa, pois uma única combinação de algoritmos é utilizada para todo o conjunto de imagens. Da mesma forma, a abordagem adaptativa ideal também não apresenta variação, pois a associação entre imagens e combinações foi feita de maneira direta, sem considerar possíveis erros de classificação. O desempenho das três análises de cada conjunto de imagens pode ser visto nas Figuras 6.18 e 6.19.

Como pode ser visto na Figura 6.18, a abordagem adaptativa por regiões apresenta menor EM do que a abordagem não adaptativa, mas pior desempenho que a abordagem adaptativa ideal. Da mesma forma, o comportamento da TBC se dá de maneira oposta, o que representa um pior desempenho das abordagens não adaptativas (ver Figura 6.19). É importante lembrar que as abordagens não adaptativas são duas: a que apresentou menor EM e a com maior TBC. Ainda assim, a estratégia autoadaptativa por regiões apresenta maior desempenho do que ambas não adaptativas. O único caso em que não ocorre aumento significativo no desempenho é o conjunto de dados relativo ao experimento realizado no Brasil. Como discutido anteriormente, as características do terreno sobrevoado aliado à baixa altitude do VANT podem ter

contribuído para que a estratégia autoadaptativa não tivesse êxito. Para esse caso específico, pode-se afirmar que não foi possível obter aumento no desempenho do sistema recorrendo à reconfiguração deste, pois, ou as características do meio não eram suficientes para justificar a adaptação do sistema, ou os cenários definidos para a reconfiguração não permitiram ganhos significativos de desempenho.

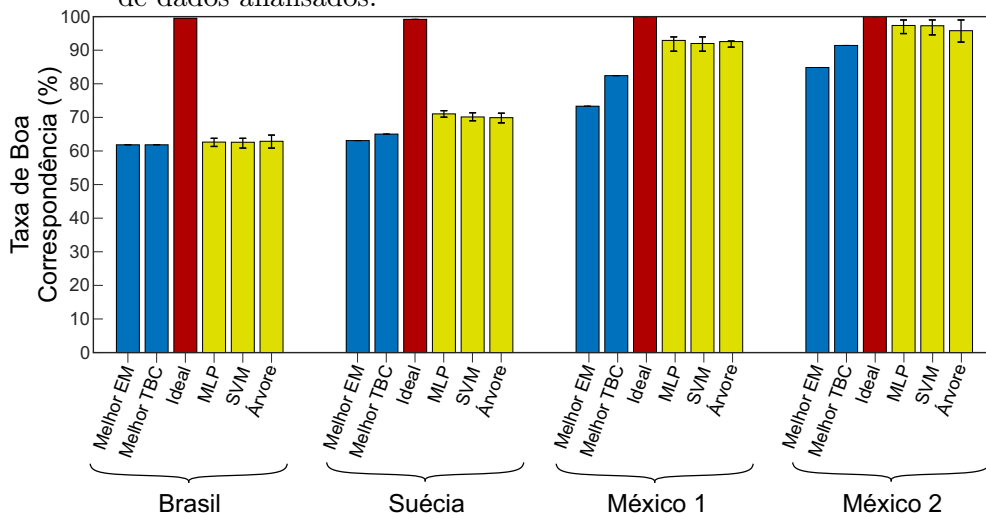
Figura 6.18 - Erro médio das três abordagens de cada um dos conjuntos de dados analisados.



O EM de cada uma das 5 abordagens é apresentado para cada conjunto de dados analisado. As barras de erro aparecem apenas nos casos da abordagem autoadaptativa por regiões, sendo delimitadas pelo pior e melhor valor.

Fonte: Produção do autor.

Figura 6.19 - Taxa de boa correspondência das três abordagens de cada um dos conjuntos de dados analisados.

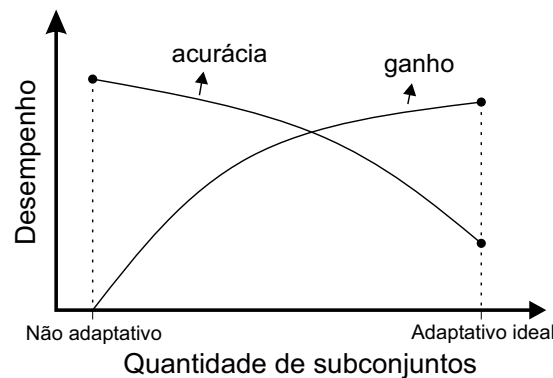


A TBC de cada uma das 5 abordagens é apresentado para cada conjunto de dados analisado. As barras de erro aparecem apenas nos casos da abordagem autoadaptativa por regiões, sendo delimitadas pelo pior e melhor valor. A legenda para as cores das colunas pode ser vista na Figura 6.18.

Fonte: Produção do autor.

Nesse contexto, podem-se definir duas características principais que devem ser consideradas na adaptação. Primeiro, a acurácia, ou taxa de acerto, dos classificadores. Aqui, é importante lembrar que existe um classificador perfeito, e ele ocorre toda vez que uma única combinação é utilizada para todas as imagens, i.e., ele é dado pela abordagem não adaptativa. Segundo, o aumento no desempenho do sistema quando ocorre a adaptação. Nesse caso, o melhor ganho possível é dado pela abordagem adaptativa ideal, em que cada instância do problema está associada ao melhor algoritmo disponível. Entre a abordagem não adaptativa e a abordagem adaptativa ideal, que podem ser considerados casos extremos, existe uma variedade de situações possíveis que podem apresentar ganhos expressivos no desempenho e que também podem satisfazer critérios mínimos de acerto nas classificações. A Figura 6.20 ilustra o comportamento das duas curvas em relação ao número de subconjuntos utilizados para dividir o conjunto inicial de imagens.

Figura 6.20 - Comportamento do desempenho e acurácia em relação ao aumento do número de subconjunto utilizados.



As curvas representam a acurácia da classificação e o ganho teórico obtido ao se associar corretamente combinações de algoritmos à subconjuntos de dados à medida que estes aumentam. Quando não ocorre adaptação, a acurácia é máxima (pois há apenas um subconjunto), mas o ganho é mínimo. Por outro lado, quando o número de subconjuntos é máximo (i.e., uma imagem para cada subconjunto), o ganho é ideal, mas a acurácia é mínima.

Fonte: Produção do autor.

É importante salientar que as curvas apresentadas na Figura 6.20 podem se alterar toda vez que um dos espaços do modelo de Rice for modificado (ver Seção 4.1). Por exemplo, ao trocar as combinações de algoritmos utilizadas neste trabalho por novas combinações, a curva de desempenho teria seu comportamento provavelmente

alterado. Da mesma forma, se outras técnicas de aprendizado de máquina fossem utilizadas, a acurácia destas também se alteraria. Por último, pode-se afirmar que a autoadaptação acarreta em ganhos no desempenho do sistema quando ambas curvas são satisfeitas.

Nesse contexto, pode-se afirmar que a hipótese proposta nesta tese de doutorado (ver Seção 1.2) foi validada. A abordagem autoadaptativa para seleção automática de combinações de algoritmos é capaz de aumentar o desempenho do sistema no sentido de diminuir o erro na estimação da posição do VANT. Isso se dá pelo emprego especializado de algoritmos em cada ambiente específico de sobrevoo, diferentemente do que ocorre com a abordagem não adaptativa, em que uma única combinação de algoritmos é utilizada para todos os ambientes.

## **6.5 Considerações finais sobre este capítulo**

No presente capítulo, foram discutidas as análises dos resultados obtidos a partir de 4 conjuntos de dados diferentes. Cada conjunto, por apresentar características intrínsecas resultantes da forma como foram obtidos, manifestou resultados únicos no sentido de que diferentes combinações de algoritmos foram utilizadas para cada um deles. Além disso, a abordagem autoadaptativa revelou ganhos no desempenho de um mesmo conjunto quando comparado à abordagem não adaptativa, validando a hipótese proposta. Por último, foi apresentada uma análise geral tendo como base todos os resultados obtidos, em que se propõe uma curva do comportamento esperado da metodologia autoadaptativa. O próximo capítulo apresenta as considerações finais deste trabalho, bem como os trabalhos futuros que podem ser desenvolvidos a partir dos resultados aqui alcançados.



## 7 CONCLUSÕES

Neste trabalho, foi analisada a possibilidade de se utilizar uma abordagem autoadaptativa para aumentar o desempenho da técnica de estimação da posição de VANTs por imagens. Para isso, um modelo abstrato de seleção de algoritmos foi utilizado com o objetivo de auxiliar na organização do problema. Nesse sentido, quatro espaços foram definidos: o espaço de problemas, o espaço de características, o espaço de combinações de algoritmos e o espaço de desempenhos. Por último, o mapa de seleção, responsável por realizar a associação entre o espaço de problemas e o espaço de algoritmos, foi também definido, sendo este composto por técnicas de aprendizado de máquina.

A análise da abordagem autoadaptativa se deu ao selecionar imagens do espaço de problemas e associá-las à combinações pertencentes ao espaço de algoritmos. Como um dos componentes da abordagem autoadaptativa, um processo de Caracterização foi realizado, onde 2000 combinações de técnicas de processamento digital de imagens e diferentes cenários foram considerados. Acredita-se que essa Caracterização é uma das maiores contribuições desse trabalho, pois a mesma se baseia na filosofia de se realizar um processo investigativo exaustivo de métodos e técnicas de processamento digital de imagens já existentes, para resolver o problema proposto nessa tese de doutorado.

Três estratégias diferentes foram apresentadas, sendo: a abordagem não adaptativa, em que uma única combinação é empregada em todo o conjunto de imagens; a abordagem adaptativa ideal, em que ignora-se a acurácia das técnicas de aprendizado de máquina para que o melhor desempenho possível seja alcançado; e, a autoadaptativa, em que se busca um balanço entre a taxa de acerto dos classificadores e o ganho no desempenho ao se considerar a reconfiguração dos algoritmos. Esta última apresentou ganho em três dos quatro conjuntos de dados analisados.

Nesse contexto, pode-se afirmar que a abordagem autoadaptativa por cenários é de potencial uso para aumentar o desempenho da técnica de estimação da posição de VANTs por imagens, dado que a mesma obteve um desempenho superior se comparada à abordagem não adaptativa. Isso ocorre porque cada cenário possui características intrínsecas e, por isso, apresenta demanda por algoritmos especializados para funcionar de maneira mais adequada. Nesse sentido, pode-se afirmar que, quanto mais ambientes diferentes são imageados durante o sobrevoo, mais se faz necessária a adaptação do sistema e, assim, maior o impacto positivo que a abordagem autoadaptativa poderá apresentar nos resultados.

Por outro lado, é importante analisar se o emprego de tal abordagem é viável dada as condições do ambiente na qual o problema estará imerso. Nesse sentido, o modelo abstrato de seleção de algoritmos ajuda a organizar e explorar as possibilidades de cada situação. Por exemplo, caso a região seja similar no sentido de que um único algoritmo seja capaz de estimar a posição de maneira acurada durante toda a missão, a autoadaptação se faria desnecessária.

Além disso, o modelo pode ser utilizado para realizar a adaptação de maneira *offline*, i.e., antes de iniciar o sobrevoo em uma região. Nesse caso, a sequência correta de algoritmos seria escolhida e embarcada no VANT antes do início da missão. Assim, elimina-se o tempo de processamento da reconfiguração durante o voo, mas também se elimina a capacidade do sistema de se autoadaptar. Além disso, é necessário conhecer previamente o ambiente que será sobrevoado. A escolha dependerá do projetista do sistema, que deverá optar por aquele que mais atende aos requisitos do problema dado o ambiente em que se encontra.

## 7.1 Trabalho futuros

Nesse contexto, são propostos novos trabalhos que buscam ampliar as pesquisas no contexto da autoadaptação aplicada à estimação da posição de VANTs por imagens. Entre eles, pode-se citar:

- A exploração de uma forma automática de selecionar subconjuntos de imagens de forma a otimizar a acurácia dos classificadores e, também, o ganho no desempenho ocasionado pela reconfiguração;
- O aumento do espaço de algoritmos adicionando novas técnicas de estimação de posição, como a odometria visual, que, assim como a técnica por reconhecimento de regiões, pode se beneficiar com a característica da adaptação;
- A utilização da abordagem autoadaptativa para detectar regiões que não são propícias para estimar a posição por meio de imagens para acionar um sistema de falha segura, ou um sistema de estimação da posição independente;
- A análise da possibilidade de uso de um segundo nível de adaptação de algoritmos, no qual as técnicas de aprendizado de máquina são escolhidas de acordo com ambiente em que estão;

- A implementação da metodologia autoadaptativa em sistemas embarcados e sua posterior aplicação em um VANT com a finalidade de observar a viabilidade da mesma quando testada em tempo real;
- O uso da técnica de *data augmentation* para auxiliar no treinamento das técnicas de aprendizado de máquina como forma de aumentar a quantidade de exemplos de imagens de treinamento;
- A análise para verificar a possibilidade de uso, em nível de simulação e de implementação em sistemas embarcados, de redes neurais artificiais mais sofisticadas para classificação de imagens;
- A exploração de novas formas de aumentar o espaço de desempenhos, em que se destaca a medição do tempo de processamento necessário para execução de cada algoritmo em estudo, assim como o tempo adicional necessário para a readaptação do sistema;
- A utilização da teoria de erros para comparar as incertezas originadas por diferentes técnicas de estimação da posição e posterior análise criteriosa dos resultados obtidos.



## REFERÊNCIAS BIBLIOGRÁFICAS

ABDI, G.; SAMADZADEGAN, F.; KURZ, F. Pose estimation of unmanned aerial vehicles based on a vision-aided multi-sensor fusion. **ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, p. 193–199, 2016. ISSN 2194-9034. 7, 30, 33

ABNOUS, R.; ZENG, C.; CHOWDHURY, S. Dynamics and control design of a blended wing-body transitioning UAV. In: **AIAA/ISSMO MULTIDISCIPLINARY ANALYSIS AND OPTIMIZATION CONFERENCE**. [S.l.: s.n.], 2017. p. 1–16. 2

AGÊNCIA NACIONAL DE AVIAÇÃO CIVIL. **Regulamento brasileiro da aviação civil especial**. [S.l.: s.n.], 2017. 1–26 p. 23

AL-KAFF, A.; MARTÍN, D.; GARCÍA, F.; ESCALERA, A. de la; ARMINGOL, J. M. Survey of computer vision algorithms and applications for unmanned aerial vehicles. **Expert Systems with Applications**, v. 92, p. 447–463, 2018. ISSN 09574174. 6, 10, 11, 24, 25, 27, 30

ANDERSON, J.; KALRA, N.; STANLEY, K. D.; SORENSEN, P.; SAMARAS, C.; OLUWATOLA, T. **Autonomous vehicle technology: a guide for policymakers**. [S.l.: s.n.], 2014. 214 p. ISBN 9780833083982. 22

ANTHONY, R.; RETTBERG, A.; CHEN, D.; JAHNICH, I.; BOER, G. de; EKELIN, C. A dynamically reconfigurable automotive control system architecture. **IFAC Proceedings Volumes**, v. 41, n. 2, p. 9308–9313, 2008. ISSN 14746670. 19

AYKIN, M. D.; NEGAHDARIPOUR, S. On feature matching and image registration for two-dimensional forward-scan sonar imaging. **Journal of Field Robotics**, v. 30, n. 4, p. 602–623, jul 2013. ISSN 15564959. 143

BALAMURUGAN, G.; VALARMATHI, J.; NAIDU, V. P. S. Survey on UAV navigation in GPS denied environments. In: **INTERNATIONAL CONFERENCE ON SIGNAL PROCESSING, COMMUNICATION, POWER AND EMBEDDED SYSTEM**. [S.l.]: IEEE, 2016. p. 198–204. ISBN 978-1-5090-4620-1. ISSN 0957-8811. 27, 30

BISHOP, C. M. **Pattern recognition and machine learning**. [S.l.]: Springer, 2006. ISBN 978-0387-31073-2. 62, 63, 64, 66, 68, 69

BLAZHKO, V.; KALINOVSKY, A.; KOVALEV, V. Unmanned Aerial Vehicle (UAV): back to base without satellite navigation. In: KRASNOPROSHIN, V. V.; ABLAMEYKO, S. V. (Ed.). **PATTERN RECOGNITION AND INFORMATION PROCESSING**. Cham: Springer, 2017. p. 140–149. ISBN 978-3-319-54220-1. 30, 33

BOBDA, C. **Introduction to reconfigurable computing**. [S.l.]: Springer, 2007. 1–359 p. ISSN 1660-2110. ISBN 978-1-4020-6088-5. 21

BRADY, H. Drone saves drowning swimmers for first time ever. **National Geographic**, 2018. Disponível em: <<https://www.nationalgeographic.com/adventure/lists/surfing/lifeguard-drone-rescue-australia-spd/>>. 4

BRAGA, J. R. G. **Navegação autônoma de VANT por imagens Lidar**. 307 p. Tese (Doutorado em Computação Aplicada) — Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2018. 6, 7, 11, 35, 51, 79, 95

BRAGA, J. R. G.; VELHO, H. F. C.; CONTE, G.; DOHERTY, P.; SHIGUEMORI, E. H. An image matching system for autonomous UAV navigation based on neural network. In: **INTERNATIONAL CONFERENCE ON CONTROL, AUTOMATION, ROBOTICS AND VISION**. [S.l.]: IEEE, 2016. p. 1–6. ISBN 978-1-5090-3549-6. 30, 33, 51, 89

BROWN, L. G. A survey of image registration techniques. **ACM Computing Surveys**, v. 24, n. 4, p. 325–376, 1992. ISSN 0360-0300. 53

CADENA, C.; CARLONE, L.; CARRILLO, H.; LATIF, Y.; SCARAMUZZA, D.; NEIRA, J.; REID, I.; LEONARD, J. J. Past, present, and future of simultaneous localization and mapping: toward the robust-perception age. **IEEE Transactions on Robotics**, v. 32, n. 6, p. 1309–1332, dec 2016. ISSN 1552-3098. 6, 28, 29, 71

CANNY, J. A computational approach to edge detection. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, n. 6, p. 679–698, nov 1986. ISSN 0162-8828. 49, 50

CESETTI, A.; FRONTONI, E.; MANCINI, A.; ASCANI, A.; ZINGARETTI, P.; LONGHI, S. A visual global positioning system for unmanned aerial vehicles used in photogrammetric applications. **Journal of Intelligent & Robotic Systems**, v. 61, n. 1-4, p. 157–168, 2011. ISSN 0921-0296. 30, 32

CHOWDHURY, S.; MALDONADO, V.; PATEL, R. Conceptual design of a multi-ability reconfigurable unmanned aerial vehicle (UAV) through a synergy of

3D CAD and modular platform planning. In: **AIAA/ISSMO MULTIDISCIPLINARY ANALYSIS AND OPTIMIZATION CONFERENCE**. [S.l.: s.n.], 2014. p. 16–20. ISBN 9781624102837. 16, 20

CLAUS, C.; STECHELE, W. Autovision—reconfigurable hardware acceleration for video-based driver assistance. In: PLATZNER, M.; TEICH, J.; WEHN, N. (Ed.). **Dynamically reconfigurable systems**. Dordrecht: Springer, 2010. p. 375–394. ISBN 978-90-481-3484-7. 19

CONTE, G.; DOHERTY, P. An integrated UAV navigation system based on aerial image matching. In: **IEEE AEROSPACE CONFERENCE**. [S.l.]: IEEE, 2008. p. 1–10. ISBN 978-1-4244-1487-1. ISSN 1095-323X. 30, 31, 32, 89

\_\_\_\_\_. Vision-based unmanned aerial vehicle navigation using geo-referenced information. **EURASIP Journal on Advances in Signal Processing**, n. 1, p. 387308, 2009. ISSN 1687-6180. 7, 30, 32, 89

CORKE, P. **Robotics, vision and control**. Berlin, Heidelberg: Springer, 2011. (Springer Tracts in Advanced Robotics, v. 73). ISBN 978-3-642-20143-1. 4

DALAMAGKIDIS, K.; VALAVANIS, K. P.; PIEGL, L. A. **On integrating nnmanned aircraft systems into the national airspace system**. Dordrecht: Springer, 2012. ISBN 978-94-007-2478-5. 24, 26

DEMŠAR, J.; CURK, T.; ERJAVEC, A.; GORUP, Č.; HOČEVAR, T.; MILUTINOVIČ, M.; MOŽINA, M.; POLAJNAR, M.; TOPLAK, M.; STARIČ, A.; ŠTAJDOHAR, M.; UMEK, L.; ŽAGAR, L.; ŽBONTAR, J.; ŽITNIK, M.; ZUPAN, B. Orange: data mining toolbox in python. **Journal of Machine Learning Research**, v. 14, p. 2349–2353, 2013. Disponível em: <<http://jmlr.org/papers/v14/demsar13a.html>>. 97

ESTRIN, G. Reconfigurable computer origins: the UCLA fixed-plus-variable (F+V) structure computer. **IEEE Annals of the History of Computing**, v. 24, n. 4, p. 3–9, oct 2002. ISSN 1058-6180. 20, 21

FEDERAL AVIATION ADMINISTRATION. **Interim operational approval guidance 08-01**. [S.l.: s.n.], 2008. 1–18 p. 26

FORNARI, G.; SANTIAGO JÚNIOR, V. A. Dynamically reconfigurable systems: a systematic literature review. **Journal of Intelligent & Robotic Systems**, v. 95, n. 3-4, p. 829–849, sep 2019. ISSN 0921-0296. 7, 13

- GALINDO, J. C. F. **Estudo de métodos de extração de características aplicados ao problema da estimação da posição de um vant em navegação autônoma com visão computacional**. 80 p. Dissertação (Mestrado) — Universidade Federal de Minas Gerais, Belo Horizonte, 2016. 11
- GAO, B.; HU, G.; GAO, S.; ZHONG, Y.; GU, C. Multi-sensor optimal data fusion for INS/GNSS/CNS integration based on unscented Kalman filter. **International Journal of Control, Automation and Systems**, v. 16, n. 1, p. 129–140, feb 2018. ISSN 1598-6446. 6
- GOLTZ, G. A. M. **Redes neurais artificiais em imagens para estimação da posição de um vant**. 117 p. Dissertação (Mestrado em Computação Aplicada) — Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2011. 79
- GONZALEZ, R. C.; WOODS, R. E. **Digital image processing**. New Jersey: Pearson Prentice Hall, 2008. 977 p. ISBN 978-0-13-168728-8. 36, 38, 39, 40, 42, 43, 44, 45, 46, 47, 48, 49, 50, 56, 78, 143, 145
- GROVES, P. D. **Principles of GNSS, inertial, and multi-sensor integrated navigation systems**. 2. ed. [S.l.]: Artech House, 2013. ISBN 978-1-60807-005-3. 5, 6, 27, 28, 60, 76, 77
- HANNAH, M. J. **Computer matching of areas in stereo images**. AAI7427032. Tese (Doutorado) — Stanford University, Stanford, CA, USA, 1974. 54
- HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. Textural features for image classification. **IEEE Transactions on Systems, Man, and Cybernetics**, n. 6, p. 610–621, nov 1973. ISSN 0018-9472. 78, 79, 151, 152
- HARTLEY, R.; ZISSERMAN, A. **Multiple view geometry in computer vision**. [S.l.]: Cambridge University Press, 2004. ISBN 9780521540513. 39, 40
- HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The elements of statistical learning: data mining, inference, and prediction**. 2. ed. [S.l.]: Springer, 2009. (Springer series in statistics). ISBN 0387848576,9780387848570. 61, 62, 63
- HAYKIN, S. **Neural networks and learning machines**. [S.l.]: Pearson Prentice Hall, 2009. ISBN 13:978-0-13-147139-9. 61, 63, 64, 65, 66, 67, 68, 69, 70
- HEREDIA, G.; DURAN, A.; OLLERO, A. Modeling and simulation of the HADA reconfigurable UAV. **Journal of Intelligent and Robotic Systems: Theory and Applications**, v. 65, n. 1-4, p. 115–122, 2012. ISSN 09210296. 2



HUGHES, J. F.; DAM, A. van; MCGUIRE, M.; SKLAR, D. F.; FOLEY, J. D.; FEINER, S. K.; AKELEY, K. **Computer graphics: principles and practice**. 3. ed. [S.l.]: Addison-Wesley Professional, 2013. ISBN 0321399528,9780321399526. 38, 39, 40, 42, 43, 44, 48, 143

JAIMES, B. R. A. **Estratégias para aumentar a robustez de estimação de posição geográfica em vants através de imagens**. 141 p. Dissertação (Mestrado) — Universidade Federal de Minas Gerais, Belo Horizonte, 2016. 11

JIANG, W.; LI, Y.; RIZOS, C. Improved decentralized multi-sensor navigation system for airborne applications. **GPS Solutions**, v. 22, n. 3, p. 78, jul 2018. ISSN 1080-5370. 5

JÓZWIAK, L.; NEDJAH, N. Modern architectures for embedded reconfigurable systems - a survey. **Journal of Circuits, Systems and Computers**, v. 18, n. 02, p. 209–254, apr 2009. ISSN 0218-1266. 17, 20

KANELLAKIS, C.; NIKOLAKOPOULOS, G. Survey on computer vision for UAVs: current developments and trends. **Journal of Intelligent & Robotic Systems**, p. 1–28, jan 2017. ISSN 0921-0296. 5, 6, 11, 22, 24, 26, 27, 30

KESSAL, L.; ABEL, N.; DEMIGNY, D. Real-time image processing with dynamically reconfigurable architecture. **Real-Time Imaging**, v. 9, n. 5, p. 297–313, 2003. ISSN 10772014. 20

KOFMAN, W.; HERIQUE, A.; GOUTAIL, J.-P.; HAGFORS, T.; WILLIAMS, I. P.; NIELSEN, E.; BARRIOT, J.-P.; BARBIN, Y.; ELACHI, C.; EDENHOFER, P.; LEVASSEUR-REGOURD, A.-C.; PLETTEMEIER, D.; PICARDI, G.; SEU, R.; SVEDHEM, V. The Comet Nucleus Sounding Experiment by Radiowave Transmission (CONSERT): a short description of the instrument and of the commissioning stages. **Space Science Reviews**, v. 128, n. 1-4, p. 413–432, may 2007. ISSN 0038-6308. 22

KOTSIANTIS, S. B. Supervised machine learning: a review of classification techniques. In: **CONFERENCE ON EMERGING ARTIFICIAL INTELLIGENCE APPLICATIONS IN COMPUTER ENGINEERING**. Netherlands: IOS Press, 2007. p. 3–24. ISBN 978-1-58603-780-2. 61

KOVESI, P. Phase congruency: a low-level image invariant. **Psychological Research**, v. 64, n. 2, p. 136–148, 2000. ISSN 14302772. 52, 97

- KRAUS, K. **Photogrammetry: geometry from images and laser scans**. [S.l.]: Walter de Gruyter, 2007. ISBN 978-3-11-019007-6. 60, 77
- KRUPITZER, C.; ROTH, F. M.; VANSYCKEL, S.; SCHIELE, G.; BECKER, C. A survey on engineering approaches for self-adaptive systems. **Pervasive and Mobile Computing**, v. 17, n. PB, p. 184–206, 2015. ISSN 15741192. 1, 2, 16
- LEWIS, J. P. Fast normalized cross-correlation. **Vision Interface**, 1995. 54, 56, 57
- LINDSTEN, F.; CALLMER, J.; OHLSSON, H.; TÖRNQVIST, D.; SCHÖN, T. B.; GUSTAFSSON, F. Geo-referencing for UAV navigation using environmental classification. In: **INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION**. [S.l.]: IEEE, 2010. p. 1420–1425. ISBN 978-1-4244-5038-1. ISSN 10504729. 30, 32
- LIU, X.; WANG, H.; FU, D.; YU, Q.; GUO, P.; LEI, Z.; SHANG, Y. An area-based position and attitude estimation for unmanned aerial vehicle navigation. **Science China Technological Sciences**, v. 58, n. 5, p. 916–926, 2015. ISSN 1674-7321. 30, 32
- LÖFSTEDT, T.; BRYNOLFSSON, P.; ASKLUND, T.; NYHOLM, T.; GARPEBRING, A. Gray-level invariant Haralick texture features. **PLOS ONE**, v. 14, n. 2, p. e0212110, feb 2019. ISSN 1932-6203. 79, 152
- LYKE, J. C.; CHRISTODOULOU, C. G.; VERA, G. A.; EDWARDS, A. H. An introduction to reconfigurable systems. **Proceedings of the IEEE**, v. 103, n. 3, p. 291–317, 2015. ISSN 00189219. 1, 4, 20
- MAURETTE, M. Mars rover autonomous navigation. **Autonomous Robots**, v. 14, n. 2, p. 199–208, 2003. 22
- MCCLUNEY, W. R. **Introduction to radiometry and photometry**. Norwood: Artech House, 1994. 417 p. ISBN 0-89006-678-7. 143, 144
- MIGUENS, A. P. **Navegação: a ciência e a arte. Volume 1 - navegação costeira, estimada e em águas restritas**. [s.n.], 2019. 519 p. Disponível em: <<https://www.marinha.mil.br/dhn/?q=pt-br/npublicacoes>>. 28
- MÖLLER, L.; SOARES, R.; CARVALHO, E.; GREHS, I.; CALAZANS, N.; MORAES, F. Infrastructure for dynamic reconfigurable systems: choices and trade-offs. In: **ANNUAL SYMPOSIUM ON INTEGRATED CIRCUITS AND SYSTEMS DESIGN**. New York: ACM, 2006. p. 44–49. ISBN 1-59593-479-0. 20

NAKHMANI, A.; TANNENBAUM, A. A new distance measure based on generalized Image Normalized Cross-Correlation for robust video tracking and image recognition. **Pattern Recognition Letters**, v. 34, n. 3, p. 315–321, feb 2013. ISSN 01678655. 56

NONAMI, K. Prospect and recent research & development for civil use autonomous unmanned aircraft as UAV and MAV. **Journal of System Design and Dynamics**, v. 1, n. 2, p. 120–128, 2007. 4

NOVO, E. M. L. d. M. **Sensoriamento remoto: princípios e aplicações**. São Paulo: Blucher, 2010. 388 p. ISBN 978-85-212-0540-1. 145

OTTO, A.; AGATZ, N.; CAMPBELL, J.; GOLDEN, B.; PESCH, E. Optimization approaches for civil applications of unmanned aerial vehicles (UAVs) or aerial drones: a survey. **Networks**, v. 72, n. 4, p. 411–458, dec 2018. ISSN 00283045. 4, 22

PADFIELD, D. Masked FFT registration. In: **IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION**. [S.l.]: IEEE, 2010. p. 2918–2925. ISBN 978-1-4244-6984-0. 54, 56, 57, 58, 59, 148

RICE, J. R. The algorithm selection problem. **Advances in Computers**, v. 15, n. C, p. 65–118, 1976. ISSN 00652458. 71, 72, 73, 75

SAE International. **Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles**. [s.n.], June 2018. 35 p. Disponível em: <[https://www.sae.org/standards/content/j3016\\_201806/](https://www.sae.org/standards/content/j3016_201806/)>. 2, 3

SALEHIE, M.; TAHVILDARI, L. Self-adaptive software: landscape and research challenges. **ACM Transactions on Autonomous and Adaptive Systems**, v. 4, n. 2, p. 1–42, may 2009. ISSN 15564665. 2

SAMADZADEGAN, F.; ABDI, G. Autonomous navigation of unmanned aerial vehicles based on multi-sensor data fusion. In: **IRANIAN CONFERENCE ON ELECTRICAL ENGINEERING**. [S.l.]: IEEE, 2012. p. 868–873. ISBN 978-1-4673-1149-6. 30, 32

SAMADZADEGAN, F.; HAHN, M.; SAEEDI, S. Position estimation of aerial vehicle based on a vision aided navigation system. **Proceedings of visualization and exploration of geospatial data - Stuttgart**, 2007. 30, 31, 32

- ŠAŠAK, J.; GALLAY, M.; KAŇUK, J.; HOFIERKA, J.; MINÁR, J. Combined use of terrestrial laser scanning and UAV photogrammetry in mapping alpine terrain. **Remote Sensing**, v. 11, n. 18, p. 2154, sep 2019. ISSN 2072-4292. [6](#)
- SHIGUEMORI, E. H.; MARTINS, M. P.; MONTEIRO, M. V. T. Landmarks recognition for autonomous aerial navigation by neural networks and Gabor transform. In: ASTOLA, J. T.; EGIAZARIAN, K. O.; DOUGHERTY, E. R. (Ed.). **IMAGE PROCESSING: ALGORITHMS AND SYSTEMS V**. [S.l.: s.n.], 2007. v. 6497. [30](#), [31](#)
- SILVA FILHO, P. F. F. **Automatic landmark recognition in aerial images for the autonomous navigation system of unmanned aerial vehicles**. 158 p. Dissertação (Mestrado em Engenharia Eletrônica e Computação) — Instituto Tecnológico de Aeronáutica, São José dos Campos, 2016. [11](#)
- SILVA, W. da. **Navegação autônoma de VANT em período noturno com imagens infravermelho termal**. 99 p. Dissertação (Mestrado em Computação Aplicada) — Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2016. [6](#), [7](#), [11](#), [35](#), [36](#), [79](#)
- SIM, D.-G.; PARK, R.-H.; KIM, R.-C.; LEE, S. U.; KIM, I.-C. Integrated position estimation using aerial image sequences. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 24, n. 1, p. 1–18, 2002. ISSN 01628828. [30](#), [31](#)
- SMITH-MILES, K. A. Cross-disciplinary perspectives on meta-learning for algorithm selection. **ACM Computing Surveys**, v. 41, n. 1, p. 1–25, jan 2009. ISSN 0360-0300. [71](#), [73](#), [75](#), [81](#)
- SOBEL, I.; FELDMAN, G. A 3x3 isotropic gradient operator for image processing. **Pattern Classification and Scene Analysis**, p. 271–272, 1973. [48](#)
- STEPANOV, D.; TISHCHENKO, I. Methods of visual navigation of the UAV flying over the nonplanar district. In: KIM, J.-H.; KARRAY, F.; JO, J.; SINCAK, P.; MYUNG, H. (Ed.). **ROBOT INTELLIGENCE TECHNOLOGY AND APPLICATIONS**. Cham: Springer, 2017. p. 551–561. ISBN 978-3-319-31293-4. [7](#), [30](#), [33](#)
- SZELISKI, R. **Computer vision: algorithms and applications**. London: Springer, 2011. 812 p. (Texts in Computer Science). ISSN 10636919. ISBN 978-1-84882-934-3. [27](#), [54](#), [56](#), [143](#), [144](#), [145](#)

- TANG, B.-h.; ZENG, W.-j.; ZHOU, Z.-x. The application of multi-source information fusion technology in vehicle integrated navigation system. In: **INTERNATIONAL CONFERENCE ON BIOINFORMATICS AND COMPUTATIONAL BIOLOGY**. New York, NY, USA: ACM, 2018. p. 107–113. ISBN 9781450363488. 5
- WAGNER, M.; ZÖBEL, D.; MEROTH, A. Re-configuration in SOA-based adaptive driver assistance systems. **SIGBED Review**, v. 11, n. 3, p. 30–35, 2014. ISSN 1551-3688. 16, 19, 20
- WAKEFIELD, J. Life-saving kidney delivered by drone. **BBC**, 2019. Disponível em: <<https://www.bbc.com/news/technology-48132595>>. 4
- WAN, X.; LIU, J.; YAN, H.; MORGAN, G. L. Illumination-invariant image matching for autonomous UAV localisation based on optical sensing. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 119, p. 198–213, 2016. ISSN 09242716. 30, 33
- WANG, X. W.; CHEN, W. N.; WANG, Y.; YOU, H. J.; PENG, C. L. The design and implementation of hardware task configuration management unit on dynamically reconfigurable SoC. In: **INTERNATIONAL CONFERENCE ON EMBEDDED SOFTWARE AND SYSTEMS**. [S.l.: s.n.], 2009. p. 179–184. ISBN 9780769536781. 20
- YAN, R.-j.; PANG, S.; SUN, H.-b.; PANG, Y.-j. Development and missions of unmanned surface vehicle. **Journal of Marine Science and Application**, v. 9, n. 4, p. 451–457, 2010. ISSN 1993-5048. 22
- YUH, J. Design and control of autonomous underwater robots: a survey. **Autonomous Robots**, v. 8, n. 1, p. 7–24, 2000. ISSN 1573-7527. 22
- ZELLER, M.; PREHOFER, C.; KREFFT, D.; WEISS, G. Towards runtime adaptation in AUTOSAR. **SIGBED Review**, v. 10, n. 4, p. 17–20, 2013. ISSN 1551-3688. 19
- ZONG, C.; MONDAL, S.; HALL, D. A.; JAFARI, R. Digitally assisted analog front-end power management strategy via dynamic reconfigurability for robust heart rate monitoring. **SIGBED Review**, v. 12, n. 3, p. 36–39, 2015. ISSN 1551-3688. 16



## APÊNDICE A - *STRING* DE BUSCA DA RSL

As buscas foram feitas utilizando-se os próprios mecanismos de busca disponíveis nos portais das respectivas bases de dados. As *strings* são divididas em dois módulos ligados pelo conectivo *AND*. O primeiro módulo contém sinônimos de SRDs, enquanto o segundo módulo contém descrições de infraestruturas de sistemas computacionais (e.g. hardware, software). As palavras-chave sobre processamento de imagens foram adicionadas com o objetivo secundário de selecionar artigos que trouxessem sistemas reconfiguráveis como possíveis soluções para esse domínio. Como é de conhecimento, processamento de imagem é muito relevante no contexto aeroespacial como um todo, por exemplo, seja se for tomado o segmento espacial de sistemas como satélites (câmeras embarcadas em satélites), assim como as atividades de distribuição de imagens obtidas por tais satélites. A versão final da *string* de busca para as bases *Scopus*, *ACM*, *Web of Science* e *Science Direct* é a seguinte:

((*“Dynamic reconfigurable system”* OR *“Dynamically reconfigurable system”* OR *“Dynamically self reconfigurable system”* OR *“Dynamic reconfigurable embedded system”* OR *“Dynamic reconfigurable real-time system”* OR *“Reconfigurable embedded system”* OR *“Dynamic adaptive system”* OR *“Dynamic adaptable system”* OR *“Dynamic self adaptable system”* OR *“Reconfigurable real-time system”* OR *“Partial reconfigurable system”*) AND (*“hardware device”* OR *“system on chip”* OR *“system-on-a-chip”* OR *“system-on-chip”* OR *“embedded software”* OR *“embedded system”* OR *“image processing”* OR *“image processor”* OR *“image analysis”* OR *“classification of images”* OR *“processing of images”* OR *“image-processing”*)).

Como a base de dados *IEEE* permite no máximo 15 palavras-chave na busca, a *string* original foi modificada para a seguinte versão:

((*“Dynamic reconfigurable system”*) OR (*“Dynamically reconfigurable system”*) OR (*“Dynamically self reconfigurable system”*) OR (*“Reconfigurable embedded system”*) OR (*“Partial reconfigurable system”*)) AND ((*“hardware device”*) OR (*“system on chip”*) OR (*“system-on-a-chip”*) OR (*“system-on-chip”*) OR (*“embedded software”*) OR (*“embedded system”*) OR (*“image processing”*)).





## APÊNDICE B - ASPECTOS BÁSICOS DA IMAGEM DIGITAL

Segundo Gonzalez e Woods (2008), uma imagem pode ser definida como “uma função bidimensional,  $f(x, y)$ , em que  $x$  e  $y$  são coordenadas (planas) espaciais e a amplitude de  $f$ , em qualquer par de coordenadas  $(x, y)$ , é chamado de intensidade ou nível de cinza da imagem naquele ponto”. Cada ponto que compõe a imagem, cuja localização se dá pelo par de coordenadas  $x, y$ , é chamado de elemento da imagem (do inglês, *picture element*), ou *pixel*. Em outras palavras, uma imagem é uma matriz de *pixels*. Quando as coordenadas  $(x, y)$  e a intensidade dos *pixels* ( $f$ ) são finitos e discretos, a imagem é considerada digital (GONZALEZ; WOODS, 2008; HUGHES et al., 2013).

Na maioria dos casos, a composição da imagem se dá pela interpretação de bandas do espectro eletromagnético, que serve como fonte de informação. Porém, em alguns casos, a composição da imagem se dá por outros meios, como, por exemplo, sensores de ultrassom, que podem ser usados para gerar imagens a partir da detecção de ondas sonoras refletidas por objetos presentes em um ambiente (AYKIN; NEGAHDARIPOUR, 2013). Aqui, apenas o primeiro grupo será discutido.

Além disso, a complexidade de detalhes que compõe o sistema de captura de imagens, como composição das lentes, tipos de filmes em fotografias analógicas, ruídos eletrônicos inerentes à fotografia digital, filtros de banda e sensibilidade radiométrica, fogem do escopo deste trabalho (SZELISKI, 2011; GONZALEZ; WOODS, 2008). Nesta seção, os aspectos básicos da formação de imagens digitais no âmbito da fotometria serão apresentados para que sirvam como fundamento para compreensão das seções subsequentes.

A fotometria é o estudo que busca descrever e medir a propagação da radiação eletromagnética quando esta é perceptível ao olho humano. A fração do espectro eletromagnético que pode ser detectado pelo olho é chamada de luz e, apesar de não possuir limites precisos por variar de pessoa para pessoa, está localizada na região do espectro cujo comprimento de onda compreende 360 *nm* até 800 *nm* (MCCLUNEY, 1994; GONZALEZ; WOODS, 2008). É essa região do espectro que engloba todo o conjunto de cores percebido pelos seres humanos, que varia do violeta até o vermelho, passando pelo azul, verde, amarelo e laranja.

Quando a luz não possui cor, é chamada de monocromática (ou acromática) e possui como único atributo, a intensidade (MCCLUNEY, 1994; GONZALEZ; WOODS, 2008). Como a percepção da intensidade da luz monocromática se dá por meio da variação entre o preto e o branco, incluindo tons de cinza, ela também é denominada de “níveis

de cinza”. Nesse sentido, uma imagem monocromática também pode ser chamada de imagem em tons de cinza ou imagem em escala de cinza. Aqui, vale lembrar que imagens em preto e branco são diferentes de imagens monocromáticas, pois possuem apenas duas intensidades, representadas pelo preto e pelo branco, o que lhes confere o nome de imagens binárias.

Imagens coloridas podem então ser obtidas por meio da combinação de bandas específicas do espectro eletromagnético (MCCLUNEY, 1994). Uma representação muito utilizada é a RGB (*Red, Green, Blue*), que foi padronizada levando em consideração a percepção humana de cores. O padrão RGB descreve o comprimento de onda específico das cores primárias vermelho, verde e azul e a proporção de cada uma delas na composição de novas cores. Para gerar imagens coloridas, sensores presentes em câmeras digitais usam filtros que capturam a luz em faixas específicas do espectro eletromagnético, registrando o ambiente fotografado em múltiplas bandas monocromáticas, que mais tarde são processadas para formar uma imagem colorida (SZELISKI, 2011). Na Figura B.1 pode ser vista a decomposição de uma imagem colorida capturada por um VANT em suas bandas vermelho, verde e azul, em que cada banda é constituída por uma matriz de valores discretos,  $f(x, y)$ , que representa a intensidade (ou brilho) da imagem naquele local.

Figura B.1 - Estrutura de uma imagem RGB em suas bandas, vermelho, verde e azul.



Uma imagem digital colorida pode ser vista como uma combinação de imagens monocromáticas de bandas específicas do espectro eletromagnético. As imagens monocromáticas, por sua vez, podem ser definidas como matrizes finitas e discretas, em que cada célula representa a intensidade de luz num ponto.

Fonte: Produção do próprio autor.

De forma geral, uma câmera digital é um sistema capaz de capturar a luz proveniente de um ambiente e direcioná-la para sensores capazes de traduzir essa informação em sinais elétricos (NOVO, 2010). Enquanto a primeira parte está relacionada à parte óptica presente na câmera, a segunda compreende o uso de materiais sensíveis à luz e o processo de digitalização dessa informação. Na prática, o processo de aquisição de imagens digitais envolve uma série de detalhes adicionais que não serão considerados aqui.

Um modelo comumente empregado em câmeras digitais é a utilização de um chip composto por células fotossensíveis, em que cada célula é responsável por capturar uma pequena porção da luz do ambiente (SZELISKI, 2011; GONZALEZ; WOODS, 2008). Um exemplo desse tipo de célula é o CCD (*Charged Coupled Device*), que pode compor chips de  $2000 \times 2000$  detetores ou até maiores. Cada um desses detetores é responsável por um elemento que compõe a imagem, i.e., um *pixel*. A parte óptica da câmera, geralmente composta por um conjunto de lentes, é então utilizada para direcionar os feixes de luz na matriz de fotosensores. Dessa forma, a luz proveniente do ambiente pode ser corretamente direcionada para cada elemento fotossensível a fim de capturar a imagem. Na Figura B.2 pode ser visto um modelo simplificado em duas dimensões do funcionamento de uma câmera digital.

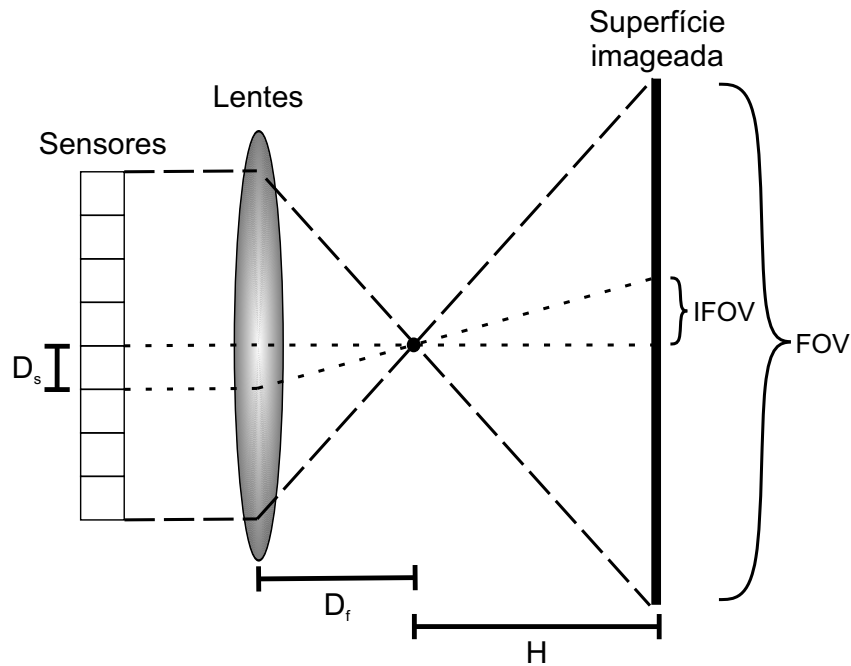
O alcance dos ângulos que um único elemento sensor é capaz de imagear é chamado *Instantaneous Field of View* (IFOV) e pode ser traduzido como campo de visada instantâneo. Na imagem digital, o IFOV é o menor elemento visível, i.e., um pixel. O conjunto de todos os detetores forma o campo de visada da câmera (do inglês, *Field of View* - FOV) e corresponde a totalidade dos *pixels* presente na imagem (NOVO, 2010).

O IFOV e o FOV podem ser calculados usando trigonometria (NOVO, 2010). Na Figura B.2, um conjunto de lentes com distância focal  $D_f$  é responsável por direcionar a luz proveniente de uma superfície imageada a uma distância  $H$  em um chip fotossensível, composto por detetores de dimensão  $D_s$ . Dessa forma, o IFOV, em radianos, pode ser calculado com a fórmula

$$IFOV_r = \frac{D_s}{D_f}, \quad (\text{B.1})$$

sendo  $D_s$  a dimensão do elemento sensor e  $D_f$  a distância focal. O IFOV pode também ser calculado em unidade de medida de comprimento, desde que a distância

Figura B.2 - Modelo simplificado de uma câmera digital em duas dimensões.



O modelo representa a entrada de luz de uma dada superfície imageada, que passa pelo conjunto óptico do sistema (lentes), até atingir os sensores fotossensíveis. A dimensão do detector,  $D_s$ , associado à distância focal  $D_f$  e à distância  $H$  da superfície imageada controlam o campo de visada instantâneo (IFOV) do sensor. O conjunto de todos os detectores forma o campo de visada do sensor (FOV).

Fonte: Produção do próprio autor.

até o objeto imageado,  $H$ , seja adicionada à Equação B.1,

$$IFOV_m = \frac{D_s}{D_f} H. \quad (B.2)$$

O FOV pode ser obtido a partir da multiplicação do IFOV pela quantidade de detectores presentes no sensor imageador.

## APÊNDICE C - CÁLCULO DA CORRELAÇÃO CRUZADA NORMALIZADA COM MÁSCARA

Sendo  $f$  uma imagem e  $t$  um *template*, a CCNZ com máscara pode ser calculada por meio de convoluções. Para isso, as equações originais devem ser reescritas como

$$ccnz(u, v) = \frac{\sum_{x,y \in M} \{[f(x, y) - \bar{f}_{x,y}][t(x - u, y - v) - \bar{t}_{x,y}]\}}{\sqrt[2]{\sum_{x,y \in M} [f(x, y) - \bar{f}_{x,y}]^2} \sqrt[2]{\sum_{x,y \in M} [t(x - u, y - v) - \bar{t}_{x,y}]^2}}, \quad (C.1)$$

em que  $M_f$  é a máscara da imagem,  $M_t$  é a máscara do *template* e  $M = M_f \cap M_t$ , é reescrita como

$$ccnz(u, v) = \frac{NUM_{ccnz(u,v)}}{\sqrt[2]{DEN1_{ccnz(u,v)}} \sqrt[2]{DEN2_{ccnz(u,v)}}}. \quad (C.2)$$

Expandindo os termos do numerador da Equação C.2,  $NUM_{ccnz(u,v)}$ , obtém-se

$$\begin{aligned} NUM_{ccnz(u,v)} &= \sum_{x,y \in M} \{[f(x, y) - \bar{f}_{x,y}][t(x - u, y - v) - \bar{t}_{x,y}]\} \\ NUM_{ccnz(u,v)} &= \sum_{x,y \in M} [f(x, y)t(x - u, y - v) - f(x, y)\bar{t}_{x,y} \\ &\quad - \bar{f}_{x,y}t(x - u, y - v) + \bar{f}_{x,y}\bar{t}_{x,y}] \\ NUM_{ccnz(u,v)} &= \sum_{x,y \in M} f(x, y)t(x - u, y - v) - \sum_{x,y \in M} f(x, y)\bar{t}_{x,y} \\ &\quad - \sum_{x,y \in M} \bar{f}_{x,y}t(x - u, y - v) + \sum_{x,y \in M} \bar{f}_{x,y}\bar{t}_{x,y}. \end{aligned}$$

Como os dois últimos termos são equivalentes,

$$NUM_{ccnz(u,v)} = \sum_{x,y \in M} f(x, y)t(x - u, y - v) - \sum_{x,y \in M} f(x, y)\bar{t}_{x,y}, \quad (C.3)$$

em que

$$\bar{t}_{x,y} = \frac{\sum_{x,y \in M} t(x-u, y-v)}{\sum_{x,y \in M} 1}.$$

Aplicando o mesmo raciocínio na primeira parte do denominador,  $DEN1_{ccnz(u,v)}$ , obtém-se

$$DEN1_{ccnz}(u, v) = \sum_{x,y \in M} [f(x, y) + \bar{f}_{x,y}]^2$$

$$DEN1_{ccnz}(u, v) = \sum_{x,y \in M} [f(x, y)^2 - 2f(x, y)\bar{f}_{x,y} + \bar{f}_{x,y}^2]$$

$$DEN1_{ccnz}(u, v) = \sum_{x,y \in M} f(x, y)^2 - 2 \sum_{x,y \in M} f(x, y)\bar{f}_{x,y} + \sum_{x,y \in M} \bar{f}_{x,y}^2.$$

Como os dois últimos somatórios são equivalentes,

$$DEN1_{ccnz}(u, v) = \sum_{x,y \in M} f(x, y)^2 - \sum_{x,y \in M} f(x, y)\bar{f}_{x,y}, \quad (C.4)$$

em que

$$\bar{f}_{x,y} = \frac{\sum_{x,y \in M} f(x, y)}{\sum_{x,y \in M} 1}.$$

Seguindo os mesmos passos para a segunda parte do denominador,  $DEN2_{ccnz(u,v)}$ , obtém-se

$$DEN2_{ccnz}(u, v) = \sum_{x,y \in M} t(x-u, y-v)^2 - \sum_{x,y \in M} t(x-u, y-v)\bar{t}_{x,y}. \quad (C.5)$$

Todos os somatórios das Equações C.3, C.4 e C.5 podem ser calculados por convolução (PADFIELD, 2010). Para isso, todos os valores  $(x, y) \notin M_f$  devem ser previamente

zerados na imagem  $f$  e todos os valores  $(x, y) \notin M_t$  devem ser zerados no *template*  $t$ . Dessa forma, tem-se que

$$\sum_{x,y \in M} f(x, y)t(x - u, y - v) = \text{conv}(f, t) = \mathcal{F}^{-1}\{\mathcal{F}(f)\mathcal{F}(t)\}, \quad (\text{C.6})$$

e

$$\sum_{x,y \in M} 1 = \text{conv}(M_f, M'_t) = \mathcal{F}^{-1}\{\mathcal{F}(M_f)\mathcal{F}(M'_t)\}, \quad (\text{C.7})$$

em que  $M'_t$  é igual a  $M_t$  rotacionado  $180^\circ$ . Da mesma forma,

$$\sum_{x,y \in M} f(x, y) = \text{conv}(f, M'_t) = \mathcal{F}^{-1}\{\mathcal{F}(f)\mathcal{F}(M'_t)\}, \quad (\text{C.8})$$

e

$$\sum_{x,y \in M} f(x, y)^2 = \text{conv}(f \cdot f, M'_t) = \mathcal{F}^{-1}\{\mathcal{F}(f \cdot f)\mathcal{F}(M'_t)\}. \quad (\text{C.9})$$

Para calcular os somatórios do *template*, basta utilizar as Equações C.8 e C.9 com a devida substituição das variáveis. Dessa forma, o numerador e o denominador podem ser calculados para todos os valores de  $(u, v)$ .

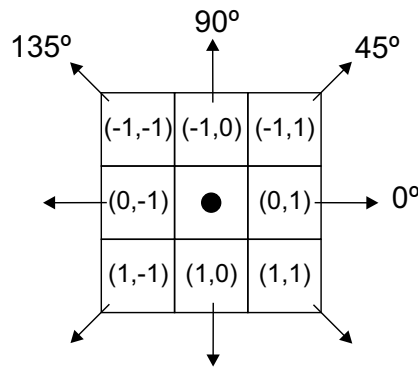




## APÊNDICE D - CÁLCULO DA MATRIZ DE COCORRÊNCIA

A matriz de coocorrência de níveis de cinza é o elemento base utilizado para calcular medidas de texturas de imagens monocromáticas conhecidas como texturas de Haralick (HARALICK et al., 1973). O cálculo da matriz se dá a partir de valores vizinhos presentes na imagem, que, por conta da sua organização espacial, permitem analisar a relação de dependência entre regiões da imagem. Por esse motivo, a matriz de coocorrência é também chamada de matriz de dependência espacial de tons de cinza (HARALICK et al., 1973). A relação de vizinhança entre valores de uma imagem se dá por uma distância e um ângulo. A Figura D.1 ilustra os 8 vizinhos possíveis para distância de um elemento gráfico.

Figura D.1 - Vizinhos mais próximos para uma distância de 1 elemento gráfico.



Fonte: Adaptado de Haralick et al. (1973).

A matriz de coocorrência é então calculada analisando a intensidade do valor do elemento gráfico da imagem e seu vizinho. Para isso, é necessário que a imagem seja primeiro quantizada a níveis de cinza na faixa de valores desejada, i.e., para dada imagem  $f$  de tamanho  $X \times Y$ , e níveis de cinza na faixa  $[c_{min}, c_{max}]$ , a nova imagem quantizada,  $f_q$ , deverá ter tamanho  $X \times Y$  com níveis de cinza na faixa  $[1, C]$ . A partir da imagem quantizada é possível calcular a matriz de coocorrência,  $O \in \mathbb{N}^{C \times C}$ , através da soma do número de vezes em que cada par de níveis de cinza ocorrem como vizinhos. O vizinho pode ser definido como um vetor bidimensional,  $(d_x, d_y)$ , sendo  $d_x, d_y \in \mathbb{Z}$ , que compreendem a distância horizontal e vertical entre os elementos vizinhos da imagem. Dessa forma, cada elemento da matriz é calculado como

$$O(i, j) = \sum_x^X \sum_y^Y \begin{cases} 1, & \text{se } f(x, y) = i \text{ e } f(x + d_x, y + d_y) = j; \\ 0, & \text{caso contrário,} \end{cases} \quad (\text{D.1})$$

em que os valores periféricos da imagem, i.e., suas bordas, são desconsideradas do cálculo. Um exemplo para uma imagem monocromática de tamanho  $4 \times 4$  e quantizada para 4 tons de cinza, ou 4 intensidades de brilho, é apresentado na Figura D.2.

Figura D.2 - Exemplo do cálculo da matriz de coocorrência para direções de  $0^\circ$  e  $90^\circ$ .

			$0^\circ$ ( $d_x=0, d_y=1$ )		$90^\circ$ ( $d_x=-1, d_y=0$ )																																																
$f =$	<table border="1" style="border-collapse: collapse; text-align: center;"> <tr><td>0</td><td>0</td><td>1</td><td>1</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>1</td></tr> <tr><td>0</td><td>2</td><td>2</td><td>2</td></tr> <tr><td>2</td><td>2</td><td>3</td><td>3</td></tr> </table>	0	0	1	1	0	0	1	1	0	2	2	2	2	2	3	3	$O =$	<table border="1" style="border-collapse: collapse; text-align: center;"> <tr><td>2</td><td>2</td><td>1</td><td>0</td></tr> <tr><td>0</td><td>2</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>3</td><td>1</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td></tr> </table>	2	2	1	0	0	2	0	0	0	0	3	1	0	0	0	0	$O =$	<table border="1" style="border-collapse: collapse; text-align: center;"> <tr><td>3</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>2</td><td>0</td><td>0</td></tr> <tr><td>2</td><td>2</td><td>1</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>2</td><td>0</td></tr> </table>	3	0	0	0	0	2	0	0	2	2	1	0	0	0	2	0
0	0	1	1																																																		
0	0	1	1																																																		
0	2	2	2																																																		
2	2	3	3																																																		
2	2	1	0																																																		
0	2	0	0																																																		
0	0	3	1																																																		
0	0	0	0																																																		
3	0	0	0																																																		
0	2	0	0																																																		
2	2	1	0																																																		
0	0	2	0																																																		

Fonte: Adaptado de Haralick et al. (1973).

Para que as dependências angulares não influenciem nas medidas, que são obtidas a partir da matriz de coocorrência, é indicado que se utiliza a média das matrizes em diferentes direções (HARALICK et al., 1973). Dessa forma, a matriz se torna invariante a rotação e as medidas, obtidas por meio dela, também.

Por outro lado, os níveis de cinza da imagem quantizada ainda precisam ser determinados como parâmetro de entrada. Para eliminar os efeitos dessa escolha, podem-se calcular medidas de texturas invariantes a escala de cinza (LÖFSTEDT et al., 2019). Nesse sentido, a matriz de coocorrência deve ser normalizada da seguinte forma:

$$\tilde{P} = \frac{O}{\sum_{i=1}^C \sum_{j=1}^C O(i, j) \Delta_i \Delta_j}, \quad (\text{D.2})$$

em que  $\Delta_i = \Delta_j = \Delta = 1/C$ . A partir da matriz normalizada,  $\tilde{P}$ , é possível calcular as medidas de textura propostas por Haralick et al. (1973) com pequenas alterações nas equações originais (LÖFSTEDT et al., 2019).

## APÊNDICE E - ESTRUTURA DAS PLANILHAS DE TREINAMENTO

Para treinar as técnicas de aprendizado de máquina, as características das imagens (ver Seção 4.2.2) foram extraídas e organizadas em planilhas. A estrutura destas é da seguinte forma: cada linha, com exceção da primeira, corresponde a uma imagem; e, cada coluna é uma característica da referida imagem. A primeira linha é reservada para os nomes das características. Um exemplo, extraído das imagens do conjunto da Suécia, pode ser visto na Tabela E.1.

Tabela E.1 - Exemplo de planilha contendo as características das imagens.

Autocorrelation	Dissimilarity	Energy	Entropy	Homogeneity
0.4491	0.0235	46.4776	-3.0438	0.9980
0.4632	0.0216	52.2760	-3.1596	0.9985
0.4725	0.0244	50.1682	-3.1247	0.9977
0.4996	0.0250	51.9007	-3.0900	0.9979
0.5060	0.0224	57.0809	-3.1667	0.9985
0.4946	0.0252	57.5094	-3.1046	0.9976
0.4857	0.0257	55.8720	-3.0749	0.9977
0.4850	0.0230	56.9543	-3.1231	0.9985
0.5068	0.0232	60.4219	-3.1697	0.9985



## APÊNDICE F - MATRIZ DE CONFUSÃO DAS TÉCNICAS DE APRENDIZADO DE MÁQUINA

As matrizes de confusão deste apêndice revelam a quantidade de classificações predita pelas técnicas (eixo horizontal) versus a quantidade de classificações verdadeiras (eixo vertical). As linhas da matriz revelam a porcentagem de imagens preditas correta e erroneamente em relação ao total de imagens pertencentes à classe da mesma linha. Dessa forma, a soma dos valores de uma linha sempre totaliza 100% e, na diagonal principal, estão localizadas a porcentagem de predições corretas de cada classe.

Nesse sentido, as classes, ou subconjunto de imagens, estão associadas a regiões de sobrevoo específicas, cujas imagens possuem características que determinam a combinação de algoritmos específica que deve ser utilizada. Assim, pode-se associar diretamente as regiões de sobrevoo aos algoritmos que devem ser empregados nessas regiões. Por motivos de legibilidade, a composição de cada combinação foi omitida das Tabelas F.1, F.2 e F.3, e substituída por uma sigla. A definição das siglas pode ser encontrada na Tabela F.4, que descreve quais algoritmos foram utilizados em cada uma das referidas classes.

Tabela F.1 - Matriz de confusão em porcentagem da árvore de classificação.

	Predito												
	C01	C02	C03	C04	C05	C06	C07	C08	C09	C10	C11	C12	C13
<b>C01</b>	84,9	0	0	9,6	0	0	0	0	0	0	5,5	0	0
<b>C02</b>	0	89,6	0	0	0	0	0	0	0	0	0	10,4	0
<b>C03</b>	0,9	0,9	94,6	0	0	0	0	0,9	0	0	0	2,7	0
<b>C04</b>	1,5	0	0	85,3	0	0	0,6	0	4,0	0	5,0	0	3,6
<b>C05</b>	0	0	0	0	100	0	0	0	0	0	0	0	0
<b>C06</b>	0	0	0	0	0	86,3	0	6,1	0	7,6	0	0	0
<b>C07</b>	0,6	0	0	1,4	0	0	92,7	0	3,1	0	2,2	0	0
<b>C08</b>	0	0	0	0	1,2	2,4	0	94,0	0	2,4	0	0	0
<b>C09</b>	0,4	0	0	6,6	0	0	7,0	0	83,8	0	2,2	0	0
<b>C10</b>	0	0	0	0	0	8,0	0	18,0	0	74,0	0	0	0
<b>C11</b>	0,4	0	0	5,2	0	0	4,1	0	1,5	0	83,9	0	4,9
<b>C12</b>	0	8,3	12,5	0	0	0	0	0	0	0	0	79,2	0
<b>C13</b>	0,7	0	0	6,2	0,4	0	0	0	0	0	1,5	0	91,2

Tabela F.2 - Matriz de confusão em porcentagem da SVM.

	Predito												
	C01	C02	C03	C04	C05	C06	C07	C08	C09	C10	C11	C12	C13
Verdade													
C01	94,5	0	0	0	0	0	0	0	0	0	5,5	0	0
C02	0	79,2	8,3	0	0	0	0	0	0	0	0	12,5	0
C03	0	0,9	96,4	0	0	0	0	0	0	0	0	2,7	0
C04	0,8	0	0	90,4	0	0	0	0	3,6	0	2,9	0	2,3
C05	0	0	0	0	94,6	0	0	0	0	0	0	0	5,4
C06	0	0	0	0	0	98,5	0	0	0	1,5	0	0	0
C07	0	0	0	2,2	0	0	88,6	0	4,5	0	4,7	0	0
C08	0	0	0	0	0	1,2	0	97,6	0	1,2	0	0	0
C09	0	0	0	5,1	0	0	6,6	0	87,2	0	1,1	0	0
C10	0	0	0	0	0	4,0	0	10,0	0	86,0	0	0	0
C11	0	0	0	5,2	0	0	4,5	0	3,0	0	85,8	0	1,5
C12	0	2,1	4,2	0	0	0	0	0	0	0	0	93,7	0
C13	0	0	0	5,5	0,4	0	0	0	0	0	1,1	0	93,0

Tabela F.3 - Matriz de confusão em porcentagem da MLP.

	Predito												
	C01	C02	C03	C04	C05	C06	C07	C08	C09	C10	C11	C12	C13
Verdade													
C01	98,6	0	0	0	0	0	0	0	0	0	1,4	0	0
C02	0	87,5	4,2	0	0	0	0	0	0	0	0	8,3	0
C03	0	1,8	96,4	0	0	0	0	0	0	0	0	1,8	0
C04	0,2	0	0	93,7	0	0	0	0	1,9	0	2,3	0	1,9
C05	0	0	0	0	94,6	0	0	0	0	0	0	0	5,4
C06	0	0	0	0	0	95,5	0	1,5	0	3,0	0	0	0
C07	0	0	0	0,6	0	0	95,0	0	2,5	0	1,9	0	0
C08	0	0	0	0	0	1,2	0	97,6	0	1,2	0	0	0
C09	0	0	0	2,9	0	0	2,9	0	93,8	0	0,4	0	0
C10	0	0	0	0	0	0	0	4,0	0	96,0	0	0	0
C11	0	0	0	3,0	0	0	4,1	0	2,2	0	89,2	0	1,5
C12	0	0	4,2	0	0	0	0	0	0	0	0	95,8	0
C13	0,4	0	0	2,6	0	0	0	0,4	0	0	1,1	0	95,5

Tabela F.4 - Composição das combinações de algoritmos de cada subconjunto de imagens.

	<b>Imagem Capturada</b>		<b>Imagem Referência</b>		<b>Registro de</b>
	<b>Filtro</b>	<b>Borda</b>	<b>Filtro</b>	<b>Borda</b>	<b>Imagens</b>
<b>C01</b>	Média 11x11	Cong. Fases	Média 7x7	Cong. Fases	CCN
<b>C02</b>	Média 3x3	Nenhum	Mediana 11x11	Nenhum	CCNZ
<b>C03</b>	Média 7x7	Nenhum	Mediana 11x11	Nenhum	CCNZ
<b>C04</b>	Gauss 11x11	Nenhum	Média 11x11	Nenhum	CCNZ
<b>C05</b>	Gauss 3x3	Cong. Fases	Gauss 3x3	Cong. Fases	CC
<b>C06</b>	Gauss 3x3	Cong. Fases	Gauss 3x3	Cong. Fases	CCN
<b>C07</b>	Gauss 7x7	Canny	Média 7x7	Canny	CCNZ
<b>C08</b>	Mediana 3x3	Nenhum	Mediana 3x3	Nenhum	CCN
<b>C09</b>	Mediana 3x3	Nenhum	Nenhum	Nenhum	CCNZ
<b>C10</b>	Mediana 3x3	Cong. Fases	Nenhum	Cong. Fases	CC
<b>C11</b>	Nenhum	Nenhum	Média 11x11	Nenhum	CCNZ
<b>C12</b>	Nenhum	Nenhum	Mediana 11x11	Nenhum	CCN
<b>C13</b>	Nenhum	Nenhum	Nenhum	Nenhum	CCNZ





## PUBLICAÇÕES TÉCNICO-CIENTÍFICAS EDITADAS PELO INPE

### **Teses e Dissertações (TDI)**

Teses e Dissertações apresentadas nos Cursos de Pós-Graduação do INPE.

### **Manuais Técnicos (MAN)**

São publicações de caráter técnico que incluem normas, procedimentos, instruções e orientações.

### **Notas Técnico-Científicas (NTC)**

Incluem resultados preliminares de pesquisa, descrição de equipamentos, descrição e ou documentação de programas de computador, descrição de sistemas e experimentos, apresentação de testes, dados, atlas, e documentação de projetos de engenharia.

### **Relatórios de Pesquisa (RPQ)**

Reportam resultados ou progressos de pesquisas tanto de natureza técnica quanto científica, cujo nível seja compatível com o de uma publicação em periódico nacional ou internacional.

### **Propostas e Relatórios de Projetos (PRP)**

São propostas de projetos técnico-científicos e relatórios de acompanhamento de projetos, atividades e convênios.

### **Publicações Didáticas (PUD)**

Incluem apostilas, notas de aula e manuais didáticos.

### **Publicações Seriadas**

São os seriados técnico-científicos: boletins, periódicos, anuários e anais de eventos (simpósios e congressos). Contam destas publicações o Internacional Standard Serial Number (ISSN), que é um código único e definitivo para identificação de títulos de seriados.

### **Programas de Computador (PDC)**

São a seqüência de instruções ou códigos, expressos em uma linguagem de programação compilada ou interpretada, a ser executada por um computador para alcançar um determinado objetivo. Aceitam-se tanto programas fonte quanto os executáveis.

### **Pré-publicações (PRE)**

Todos os artigos publicados em periódicos, anais e como capítulos de livros.