BRAYAN RENE ACEVEDO JAIMES

ESTRATÉGIAS PARA AUMENTAR A ROBUSTEZ DE ESTIMAÇÃO DE POSIÇÃO GEOGRÁFICA EM VANTS ATRAVÉS DE IMAGENS

Belo Horizonte, Minas Gerais
Dezembro de 2016

Universidade Federal de Minas Gerais
Escola de Engenharia
Programa de Pós-Graduação em Engenharia Elétrica

ESTRATÉGIAS PARA AUMENTAR A ROBUSTEZ DE ESTIMAÇÃO DE POSIÇÃO GEOGRÁFICA EM VANTS ATRAVÉS DE IMAGENS

Dissertação apresentada ao Curso de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do grau de Mestre em Engenharia Elétrica.

BRAYAN RENE ACEVEDO JAIMES

Belo Horizonte, Minas Gerais
Dezembro de 2016

Resumo

Nesta dissertação são propostas distintas abordagens para melhorar o processo de estimação de posição geográfica através de imagens capturadas por VANTs (Veículos Aéreos Não Tripulados). Em primeiro lugar, visando fazer mais robusta a navegação autônoma destas aeronaves sem a necessidade de utilizar o sinal do GPS (*Global Position System*), duas novas abordagens de casamento de imagens e com baixo tempo de processamento foram desenvolvidas. A primeira utiliza um extrator de bordas *canny* adaptativo e a segunda utiliza limiarização. Com essas técnicas é possível dar solução aos problemas de sobrestimação de bordas e de inclusão de ruído que afetam a comparação das imagens e, consequentemente, a estimação de posição.

Em segundo lugar, duas abordagens foram desenvolvidas para fazer correção de distorção projetiva, ajuste de escala, rotação e translação em imagens aéreas capturadas por VANTs quando a posição da câmera não é perpendicular a Terra. A primeira delas é uma técnica que toma como base o conhecimento prévio dos ângulos de inclinação do VANT, fornecidos pelos sensores inerciais da aeronave para obter a matriz homográfica e corrigir a imagem. Estes ángulos compõem os parâmetros de rotação da matriz homográfica, que é também composta por outras matrizes concatenadas que representam parâmetros intrínsecos da câmera e de translação na imagem. A segunda abordagem apresenta uma robusta correção de distorção projetiva e espectral nas imagens aéreas. Esta técnica baseia-se no casamento de pontos característicos extraídos entre a imagem capturada pelo VANT e a imagem georeferenciada, utilizando os algoritmos SURF e MSAC para estimar os parâmetros que compõem a matriz homográfica e, desta forma, corrigir a imagem.

A avaliação das abordagens propostas leva a consideração diferentes tipos de terreno (vegetação, urbano e rodovia) na aplicação dos testes. Também, é considerada a avaliação com imagens de diferentes sensores com distorção de perspectiva, escala, rotação e translação. As métricas de avaliação foram o erro da distância média na estimação de posição e tempo de processamento. Para a correção de distorção de pers-

pectiva métricas como, o número de pontos característicos extraídos em cada uma das imagens, o número de casamentos estimados entre imagens, eficiência, *recall*, precisão e o tempo de processamento foram considerados. Os resultados obtidos ao longo dos diferentes testes aplicados nas técnicas são promissórios, apresentam baixo tempo de processamento e indicam que podem ser usadas em condições reais de voo.

Abstract

In this thesis, different approaches to improve the geographical position estimation process through UAV (Unmanned Aerial Vehicles) images are proposed. In first place, two new template matching approaches with low processing time were developed aiming to make a more robust autonomous navigation of the aircrafts without the need to use GPS (Global Position System) signal. The first uses an adaptive Canny edge detector and the second one uses thresholding. With these techniques, it is possible to solve the edge overestimation and the noise inclusion that affect the image comparison and, consequently, the position estimation.

In second place, two approaches to correct projective distortion, scale adjustment, rotation and translation in UAV images were developed when the camera position is not perpendicular to earth. The first of them is a technique that uses the (previous) knowledge of UAV tilt angles provided by the aircraft inertial sensors to obtain the homographic matrix and correct the image. These angles compose the rotation parameters of the homographic matrix that is also composed by other concatenated matrices that representing the camera intrinsic parameters and the image translation. The second approach presents a robust correction of projective and spectral distortions in images captured by UAVs. This technique is based in the keypoints matching extracted between the UAV image and the georeferenced one. It also uses the SURF and MSAC algorithms in order to estimate the parameters that compose the homographic matrix and, thus, the image is corrected.

The evaluation of the proposed approaches considered different land types (forest, urban and highway) in the tests application. Also, the evaluation with images obtained from different sensors with distortion of perspective, scale, rotation, and translation is considered. The evaluation metrics were the mean distance error in the position estimation and the processing time. Now, for the perspective distortion correction, metrics like the keypoints number extracted on each image, the estimated matching number between images, efficiency, recall, precision and processing time were conside-

red. The obtained results throughout the different tests applied in the techniques are promissory, have low processing time and indicate that they can be used in real flight conditions.

Dedico este trabalho,

A Deus pela saúde, força e constância fornecida.

Aos meus pais, Maria Jaimes e Gustavo Acevedo por todo o apoio que me deram neste aprendizagem, são a motivação na minha vida.

As minhas irmãs, Nury e Leidy por me apoiar e me ajudar neste processo.

A minha família por sempre permanecer junta, apesar da distância.

Brayan Rene Acevedo Jaimes

Agradecimentos

Ao meu orientador, Professor Frank Sill Torres, pelos valiosos conselhos e discussões, pela paciência e disponibilidade. Sempre acompanhou de perto todo o meu processo.

Ao meu co-orientador, Professor Cristiano Leite de Castro, pelo apoio e tranquilidade que me transmitiu nos momentos mais críticos do meu trabalho.

Ao Professor Antônio de Pádua Braga, pela oportunidade de fazer parte de seu equipe de trabalho no Laboratório de Inteligencia Computacional - LITC. Um Excelente professor.

Ao Professor Elcio Hideiti Shiguemori, pela valiosa assessoria no meu trabalho e conhecimentos dados nas conversas semanais do grupo PITER. Excelente pessoa e professor.

Aos meus amigos e colegas do LITC, pelas trocas de ideias, pelo apoio e o excelente ambiente de trabalho.

Aos meus amigos Brasileiros e Colombianos, pela amizade e colaboração ao longo deste tempo convivido.

A Maria Vitoria Africano, pela incondicional companhia e apoio constante em todo meu processo no Brasil. Por todos os bons momentos que compartilhamos, o aprendizagem e conhecimento mútuo adquirido.

À CAPES pelo apoio financeiro.

Sumário

1 Introdução **1**

1.1 Motivação 1

1.2 Objetivos e Contribuições do Trabalho 4

1.3 Organização da Dissertação 5

2 Estado da arte **6**

2.1 Estimação de posição em VANTs via Casamento de Imagens 6

2.1.1 Casamento de Imagens por Correlação Espacial 9

2.1.2 Métodos de Extração de Bordas 10

2.1.3 Algoritmo Detector de Bordas Canny 12

2.1.4 Trabalhos Relacionados 14

2.2 Correção de Distorções Geométricas em Imagens Aéreas Capturadas por VANTs 16

2.2.1 Tipos de Correções Geométricas 18

2.2.2 Modelo de Calibração da Câmera 19

2.2.3 Transformadas Geométricas no Plano (2D) 22

2.2.4 Transformação Linear Direta - DLT 25

2.3 SURF 26

2.3.1 Detecção de Pontos de Interesse 27

2.3.2 Localização de Pontos de Interesse 28

2.3.3 Trabalhos Relacionados 30

3 Métodos de Estimação de Posição em VANTs via Casamento de Imagens **33**

3.1 Detector de Bordas Canny Adaptativo 33

3.2 Cálculo de Correlação por Limiarização de Imagens 40

3.3 Resumo do Capítulo 44

4 Correção de Distorção Projetiva em Imagens Capturadas por VANTs Utilizando um Método Paramétrico **46**

4.1 Imagens Aéreas com Visada não Nadir . . . 47
4.2 Visão Geral Técnica . . . 47
4.3 Cálculo da Matriz de Rotação e Matriz de Parâmetros Intrínsecos da Câmera . . . 48
4.4 Análise de Homografia e Transformação Projetiva . . . 49
4.5 Resumo do Capítulo . . . 50

5 Correção de Distorção Projetiva e Espectral em Imagens Capturadas por VANTs Utilizando um Método não Paramétrico 52
5.1 Visão Geral Técnica . . . 53
5.2 Pré-processamento . . . 53
5.2.1 Equalização de Histograma . . . 53
5.2.2 Realce de Bordas e Ajuste de Intensidade . . . 55
5.3 Extração de Características e Casamento de Pontos Característicos com SURF . . . 58
5.3.1 Remoção de Outliers com MSAC . . . 60
5.4 Estimação de Homografia a Partir do Casamento de Pontos Característicos 62
5.5 Resumo do Capítulo . . . 63

6 Resultados Experimentais e Análises 64
6.1 Área de Trabalho e Software . . . 64
6.2 Casamento de Imagens . . . 65
6.3 Correção de Distorção Projetiva com Método Paramétrico . . . 70
6.4 Correção de Distorção Projetiva com Método não Paramétrico . . . 73
6.4.1 Imagens de Igual Sensor . . . 74
6.4.2 Imagens de Diferente Sensor . . . 80
6.4.3 Caracterização dos Algoritmos SURF e MSAC . . . 84
6.4.4 Estimação de Posição para Imagens de Diferente Sensor . . . 85
6.5 Análise do Vídeo . . . 88

7 Conclusões e Trabalhos Futuros 94
7.1 Conclusões . . . 94
7.1.1 Quanto à Estimação de Posição via Casamento de Imagens . . . 94
7.1.2 Quanto à Correção de Distorção Projetiva . . . 94
7.1.3 Quanto ao Teste do Vídeo . . . 97
7.2 Publicações em Eventos Científicos . . . 97
7.3 Trabalhos Futuros . . . 98

Referências Bibliográficas 99

Anexo 1 **108**

Anexo 2 **111**

Anexo 3 **114**

Anexo 4 **115**

A4.1 Teste de Eficiência em Imagens de Diferente Sensor 115

A4.2 Teste de *Recall* e *Precision* em Imagens de Diferente Sensor 116

Lista de Figuras

2.1 Diagrama de estimação de posição geográfica com a técnica Casamento de Imagens. 8
2.2 Arranjo para a obtenção da correlação de $f(x, y)$ e $w(x, y)$ de (Gonzalez and Woods, 2008). 10
2.3 Processo de correlação entre as imagens. 11
2.4 Exemplo do detector de bordas *canny* aplicado em uma imagem aérea capturada por VANT. 14
2.5 Exemplo do detector de bordas *Canny* aplicado em uma imagem de satélite. 14
2.6 Exemplo de imagem com visada nadir e com perspectiva. 17
2.7 Deslocamento na imagem devido ao movimento de rotação da Terra. . . . 18
2.8 Efeitos das distorções geométricas causadas pelo movimento do VANT (instabilidade). 18
2.9 Distorções causadas pela lente da câmera abordo do VANT. 19
2.10 Modelo da câmera composto por parâmetros intrínsecos e extrínsecos . . . 20
2.11 Transformação de plano projetivo 22
2.12 Filtro tipo caixa. Derivada parcial da segunda ordem da função gaussiana em imagens discretas na direção $y - (L_{yy})$ (2.12(a)) e na direção $xy - (L_{xy})$ (2.12(b)). A Figura 2.12(c) e 2.12(d) representam a aproximação com o filtro tipo caixa das Figuras 2.12(a) 2.12(b). 28
2.13 Espaço-escala implementado em SURF. 29
2.14 Filtros $Haar - wavelet$ para obter as respostas em x 2.14(a) e y 2.14(b). . 29
2.15 Comportamento do descritor para três padrões de imagem com intensidades diferentes dentro de uma sub-região. 30
2.16 Não casamento de pontos característicos devido à diferenças de contraste. . 30

3.1 Imagen capturadas pelo VANT: (a) Classe Vegetação, (b) Classe Urbano, (c) Classe Rodovia. 34
3.2 Detecção de bordas com *canny*: (a) Classe Vegetação, (b) Classe Urbano, (c) Classe Rodovia. 34

3.3 Resultado da estimação de posição quando existe sobrestimação de bordas em imagem de classe Vegetação (fig 3.2(a)). 35
3.4 Desvio padrão da imagem. 37
3.5 Histograma de uma Sub-imagem após de ser calculada a diferença dos pixeis entre colunas. 38
3.6 Obtenção do limiar final resultante do análise feito sobre a imagem *template*. Variacao do limiar (linha azul), média do limiar (linha vermelha). 38
3.7 Extração de bordas com *canny* sobre a imagem de classe Vegetação, aplicando diferentes valores de *sigma*. 39
3.8 Obtenção do *sigma*. 40
3.9 Resultado da estimação de posição aplicando o método de detecção de bordas com *canny* adaptativo, aplicada em imagem da classe Vegetação (fig 3.2(a)). 41
3.10 Obtenção do limiar α. 42
3.11 Resultado da estimação de posição utilizando o método de correlação por limiarização de imagens, aplicada em imagem da classe Vegetação (fig 3.2(a)). 43
3.12 Diagrama de estimação de posição geográfica com a técnica Casamento de Imagens incluído os métodos propostos neste capítulo. 45

4.1 Tipos de Visadas em VANTs. 47
4.2 Processo Geral de Correção da Imagem 48
4.3 Diagrama de estimação de posição geográfica com a técnica Casamento de Imagens incluído o método proposto neste capítulo. 51

5.1 Arquitetura para correção de distorção projetiva e espectral em imagens aéreas. 54
5.2 Diagrama esquemático de equalização de histograma. 55
5.3 (a) Imagem georeferenciada em tons de cinza (b) Imagem capturada pelo VANT em tons de cinza; (c) Imagem georeferenciada depois de ser aplicado o filtro Laplaciano (d) Imagem capturada pelo VANT depois de ser aplicado o filtro Laplaciano; (e) Imagem georeferenciada depois da subtração com a imagem filtrada (f) Imagem capturada pelo VANT depois da subtração com a imagem filtrada. 57

5.4 Pré-processamento: (a) Imagem capturada pelo VANT original em tons de cinza e (c) Seu histograma correspondente; (b) Imagem capturada pelo VANT depois do Pré-processamento e (d) Seu histograma correspondente; (e) Imagem georeferenciada original em tons de cinza e (g) Seu histograma correspondente; (f) Imagem georeferenciada depois do pré-processamento e (h) Seu histograma correspondente. 59
5.5 Detecção e casamento de pontos característicos com SURF e MSAC: (a) Imagens de teste já pré-processadas; (b) Detecção de pontos característicos na imagem georeferenciada (esquerda) e na imagem do VANT (direita); (c) Localização dos pontos característicos na imagem georeferenciada (esquerda) e na imagem do VANT (direita); (d) Casamento de pontos característicos com SURF entre a imagem georeferenciada (esquerda) e a imagem do VANT (direita); (e) Casamento de pontos característicos entre a imagem georeferenciada (esquerda) e a imagem do VANT (direita) usando SURF+MSAC. 61
5.6 Diagrama de estimação de posição geográfica com a técnica Casamento de Imagens incluído o método proposto neste Capítulo. 63

6.1 Rota planejada para avaliação dos testes. 66
6.2 Imagens capturadas por VANT de diferente sensor (experimento 1). 68
6.3 Imagens capturadas por VANT de igual sensor (experimento 2). 69
6.4 Imagem com distorção (esquerda) - Imagem corrigida (direita). 71
6.5 Imagem com distorção (esquerda) - Imagem corrigida (direita). 71
6.6 Imagem com distorção (esquerda) - Imagem corrigida (direita). 71
6.7 Imagem com distorção (esquerda) - Imagem corrigida (direita). 72
6.8 Imagem com distorção (esquerda) - Imagem corrigida (direita). 72
6.9 Imagem com distorção (esquerda) - Imagem corrigida (direita). 72
6.10 Imagem urbana. 75
6.11 Imagem urbana (Mineirão). 75
6.12 Imagem urbana com presença de gramado. 76
6.13 Imagem Vegetação. 76
6.14 Imagem urbana com rodovia. 77
6.15 Fábrica. 77
6.16 Vegetação Rio. 78
6.17 (a) Relação de número de casamentos com e sem pré-processamento; (b) Relação de tempos de processamento com e sem pré-processamento; (c) Relação de casamento vs. tempo de processamento. 79

6.18 O resultado do método proposto para o caso de teste Airport: (a) Imagem georeferenciada; (b) Imagem capturada pelo VANT; (c) A saída do método proposto. 81
6.19 O resultado do método proposto para o caso de teste Stadium: (a) Imagem georeferenciada; (b) Imagem capturada pelo VANT; (c) A saída do método proposto. 81
6.20 O resultado do método proposto para o caso de teste Tanks: (a) Imagem georeferenciada; (b) Imagem capturada pelo VANT; (c) A saída do método proposto. 82
6.21 O resultado do método proposto para o caso de teste Bush: (a) Imagem georeferenciada; (b) Imagem capturada pelo VANT; (c) A saída do método proposto. . 82
6.22 O resultado do método proposto para o caso de teste Highway: (a) Imagem georeferenciada; (b) Imagem capturada pelo VANT; (c) A saída do método proposto. 83
6.23 Seleção de sub-regiões para cada imagem de teste. 86
6.24 Fluxograma de Reconhecimento de Padrões e estimação de posição com a técnica Casamento de Imagens de diferente sensor. 87
6.25 sub-regiões encontrados utilizando *template matching* em cada imagem. . . 89
6.26 Imagem capturada pelo VANT na região de São Carlos-SP. 89
6.27 Resultado da estimação de posição para o trecho 1. 90
6.28 Distância média calculada para o trecho 1. 91
6.29 Resultado da estimação de posição para o trecho 2. 92
6.30 Distância média calculada para o trecho 2. 92

A1.1 Mapa georreferenciado obtido do satélite *Quickbird*, com uma resolução espacial de 0.6 m sobre a região de São José dos Campos no estado de São Paulo - Brasil. 108
A1.2 Imagens georeferenciadas e imagens capturadas por VANT com diferença de luminosidade e presença de ruido (experimento 1). 109
A1.3 Imagens georeferenciadas e Imagens capturadas por VANT de igual sensor (experimento 2). 110
A2.1 Imagens capturadas por VANT sobre a região de São Carlos, SP - Brasil. . 111
A2.2 Imagens capturadas por VANT sobre a região de São Carlos, SP - Brasil. . 112
A2.3 Imagens capturadas por VANT sobre a região de São Carlos, SP - Brasil. . 113
A3.1 Mapa georreferenciado sobre a região de São José dos Campos no estado de São Paulo - Brasil. 114
A4.1 Número de pontos vs. variação do *threshold*. 116
A4.2 Número de casamentos vs. variação do *threshold*. 116
A4.3 Eficiência Vs. variação do *threshold*. 117
A4.4 Distância máxima em pixeis permissível para o casamento, t_d. 117
A4.5 Resultado de *Recall* nas imagens de teste. 120

A4.6 Resultado da média e desvio padrão para *Recall*. 120
A4.7 Resultado de *Precision* nas imagens de teste. 121
A4.8 Resultado da média e desvio padrão para *Precision*. 122
A4.9 Gráfico de *Recall* vs. $1 - Precision$. 123

Lista de Tabelas

2.1 Resumo dos diferentes tipos de transformações geométricas 25

6.1 Resultados comparativos de desempenho nos diferentes métodos aplicados para a estimação de posição em imagens de diferente sensor. 68
6.2 Resultados comparativos de desempenho nos diferentes métodos aplicados para a estimação de posição em imagens de igual sensor. 69
6.3 Ângulos (Yaw, Pitch, Roll) de cada imagem com distorção e seus respectivos tempos de processamento na correção. Latitude e Longitude mostra a posição da imagem aérea que foi capturada para comparar com Google Maps. 73
6.4 Resultados de extração e casamento de pontos característicos com e sem pré-processamento em imagens com igual sensor. Tempo de processamento para cada caso avaliado. 79
6.5 Média de pontos característicos extraídos das imagens. Média, valor mínimo e máximo de casamento de pontos característicos nas imagens. Tempos de processamento obtidos de cada método em cada caso de teste depois de 30 execuções. 84
6.6 Média de pontos característicos extraídos das imagens. Média, valor mínimo e máximo de casamento de pontos característicos nas imagens. Tempos de processamento obtidos de cada método em cada caso de teste depois de 30 execuções. Caracterizados os algoritmos SURF e MSAC e encontrar o melhor ponto de operação. 85
6.7 Estimativa de posição entre as sub-regiões da imagem capturada pelo VANT e a imagem georeferenciada utilizando a técnica de Casamento de Imagens - *template matching*. 88
6.8 Desempenho da aplicação de Correção de perspectiva vs. rotação para estimação de posição na rota 1 91
6.9 Desempenho da aplicação de Correção de perspectiva vs. rotação para estimação de posição na rota 2 92

A4.1 Matriz de confusão para um classificador binário. 118

A4.2 Média de pontos característicos extraídos das imagens. Média, valor mínimo e máximo de casamento de pontos característicos nas imagens. Tempos de processamento obtidos de cada método em cada caso de teste depois de 30 execuções. Caracterizados os algoritmos SURF e MSAC e encontrar o melhor ponto de operação. 123

Lista de Siglas

ASIFT	Affine-SIFT
BLOB	Binary Large OBject - Objeto Grade Binário
DLT	Direct Linear Transform - Transformação Linear Direta
DoG	Diferença de Gaussiana
FN	Falsos Negativos
FP	Falsos Positivos
GIS	Geographic Information System
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
HOG	Histogram of Oriented Gradients
IEAV	Instituto de Estudos Avançados da Aeronáutica
INS	Inertial Navigation System
LIDAR	Laser Imaging Detection and Ranging
M-PCA	Multiple Particle Collision Algorithm
MSAC	M-estimator Sample Concensus
NAV	Navegação Autonoma de VANTs (Simulador de voo)
PCA-SIFT	Principal Components Analysis-SIFT
PITER	Processamento de Imagens em Tempo Real
RANSAC	RANdom SAmple Consensus
RBF	Rede com funções de Base Rasial
RGB	Red, Green, Blue image
RNA	Redes Neurais Artificiais
RNC	Rede Neural Celular
RPMC	Rede Perceptron de Múltiplas Camadas
SIFT	Scale Invariant Feature Transform
SURF	Speeded-Up Robust Transform
SVD	Singular Value Decomposition
TCP-IP	Protocolo de Control de Transmisión-Protocolo de Internet
TIN	Triangulated Irregular Networks
TN	Verdadeiros negativos
TP	Verdadeiros positivos
UR-SIFT	Uniform Robust SIFT
VANT	Veículo Aéreo Não Tripulado

Lista de Simbolos

n	Número de pixeis
c	Correlação
f	Imagem a ser processada
g	Máscara de convolução
h	Matriz resultante após aplicação da convolução entre matrizes
$f_c(x)$	Resposta de impulso do filtro
$G(x)$	Borda unidimensional
n_o	Amplitude de ruído na imagem
X_{max}	Distancia entre máximos adjacentes
ϵ	Orientação da transformação
t_x	Fator de translação em x
t_y	Fator de translação em y
R	Matriz de rotação
s	Fator de escala
H_s	Transformada de Similaridade
H_A	Transformada Affim
H_p	Transformada Projetiva
H	Matriz Homográfica
σ	Desvio padrão utilizado em filtro gaussiano
w	Peso atribuído
α	Limiar calculado
L	Imagem de salida limiarizada
A	Matriz de parâmetros intrínsecos
f_{cam}	Distância focal da câmera,
k_u	Fator de escala horizontal
k_v	Fator de escala vertical
u_o, v_o	Coordenadas do ponto central da câmara

ϕ	Ângulo de rolagem
θ	Ângulo de arfagem
ψ	Ângulo de guinada
$R_{\phi,\theta,\psi}$	Matriz de rotação nos ângulos rolagem, arfagem e guinada
P_p	Matriz de projeção perspectiva
t_p	Posição da visão da cena
n_p	Vetor normal ao plano da cena
L_p	Níveis de intensidade discretos na imagem
n_k 51	Número de pixeis com ik nível de intensidade
f_i	Função de densidade de probabilidade
F_k	Função de densidade cumulativa
K	Numero de hipoteses em algoritmo MSAC
t_o	Threshold do algoritmo MSAC
f_{custo}	Função de custo
X	Pontos característicos de referência
X'	Pontos característicos de destino
R_T	Raio médio da circunferência da terra
lat	Latitude
lon	Longitude
D	Distância de separação entre a posição estimada e a real
Var	Variância
DP	Desvio padrão entre a distância real e estimada
$P1$	Ponto característico extraído na imagem 1
$P2$	Ponto real característico extraído na imagem 2
$P2'$	Ponto característico extraído na imagem 2
t_d	Distância máxima em pixeis permissível

Capítulo 1

Introdução

1.1 Motivação

O constante desenvolvimento de tecnologia aplicada aos Veículos Aéreos Não Tripulados (VANTs) tem como objetivo melhorar seu desempenho na autonomia de voo e aumentar a precisão na estimação de posição geográfica. À vista disso, manter um baixo custo computacional é fundamental para desenvolver aplicações e executar tarefas tediosas, monótonas e cansativas onde não é necessária tripulação abordo. Tudo isto tem impulsionado o estudo de técnicas para o melhoramento da navegação aérea em VANTs de forma autônoma. Na maioria das vezes, estas técnicas utilizam o sistema americano Global Positioning System (GPS) como ferramenta principal no sistema de navegação autônoma em VANTs e em outros veículos autônomos por ser um sistema de baixo custo.

No entanto, existem sistemas de navegação autônoma de VANTs que realizam estimação de posição geográfica através de imagens capturadas em tempo real ao longo do voo. É uma técnica recente que tem como objetivo reforçar ou, no melhor dos casos, substituir o sinal de GPS que, por alguma razão, pode tornar-se indisponível ou corrompido (Souza et al., 2005). Falhas no sinal de GPS podem ocorrer devido a vários a fatores, tais como alterações na ionosfera, muito comuns na América do Sul (Muella, 2008), vulnerabilidade ao jamming e ataques maliciosos. Além disso, como o GPS é uma tecnologia privada, seu uso em operações militares é restrito (Conte and Doherty, 2008; Shiguemori et al., 2007).

Dentre as técnicas mais utilizadas na literatura para fazer estimação de posição através de imagens, encontram-se o Reconhecimento de Marcos nas imagens capturadas pelo VANT *(landmarks)* (Silva Filho et al., 2014; Luo et al., 2015; Jayatilleke and Nian Zhang, 2013; Cesetti et al., 2009), a Odometria (Reboucas et al., 2013; Chunhui et al., 2014; Romero et al., 2013; Roger-Verdeguer et al., 2012) e o Casamento de Imagens

(Template Matching) (Nazneen et al., 2016; Conte and Doherty, 2008; Silva et al., 2015; Sanjay Kumar Sahani et al., 2011; Baig et al., 2012). Esta última abordagem será o foco e matéria de estudo ao longo deste trabalho, pois é a técnica mais robusta e confiável na estimação de posição, o que faz que tenha sido mais trabalhada e evoluída nos últimos anos. Ao mesmo tempo, esta técnica tem fornecido resultados promissórios no processo de estimação de posição com imagens aéreas, como pode ser observado nos trabalhos de Conte and Doherty (2008), Goltz (2011) e Silva et al. (2015).

A fim de estimar a posição atual da aeronave, as imagens aéreas capturadas por câmera de vídeo são registradas em uma imagem de satélite previamente georeferenciada. Esta ação faz parte da técnica de Casamento de Imagens, tem sido aplicada com sucesso na literatura e proposta como alternativa aos sistemas de navegação autônomos com dependência do GPS (Conte and Doherty, 2008; Goltz and Shiguemori, 2011; Santos et al., 2013; Horta et al., 2012). Como é mostrado em Santos et al. (2013), a posição atual do VANT é estimada através de uniformização da imagem via extração de bordas. Depois, é encontrada a posição da imagem a bordo do VANT sobre a imagem georeferenciada, identificando o índice de correlação espacial mais elevado dentro da matriz de correlação calculada. Além disso, o Casamento de Imagens tem sido aplicado em outras áreas como a localização de objetos em movimento (Pham et al., 2015; Reboucas et al., 2013), o reconhecimento de área (Tahar et al., 2012), mapeamento (Tahar et al., 2012; Tahar and Ahmad, 2011), aplicações militares, na agricultura (Eisenbeiss, 2004; Medeiros, 2007), entre outros.

Na navegação autônoma de VANTs através de imagens, existe uma série de fatores que interferem e dificultam o processo de estimação de posição da aeronave no momento de voo. Alguns deles são: condições do ambiente, condições do tempo, mudanças de luminosidade e perda de estabilidade no VANT devido a ventos fortes. Outros fatores dependem do tipo de câmera utilizada na captura da imagem como: resolução da câmera, distorção de perspectiva devido aos ângulos de captura da imagem, distorção da lente e presença de ruído na captura da imagem. Soma-se a isto que, os mesmos métodos apresentam una série de restrições e condições iniciais que devem-se cumprir para garantir que a técnica seja corretamente executada e assim obter resultados com precisão na estimação de posição do VANT. Em adição, deve ser considerado no processo, o custo computacional avaliado através do tempo de processamento que os métodos exigem para desenvolver sua tarefa. Em Goltz (2011) consideraram-se os tempos de processamento que os operadores estudados gastam para processar a imagem de satélite e as imagens capturadas pelo VANT, além de ser avaliada sua precisão nas estimativas de posição.

Levando a consideração os problemas e restrições anteriormente mencionados, e com a finalidade de melhorar a precisão da técnica de Casamento de Imagens mantendo um

baixo tempo de processamento, neste trabalho foram propostos diferentes métodos para estimar a posição de VANT via Casamento de Imagens.

Como é mencionado em Silva et al. (2015), o método de Casamento de Imagens clássico geralmente pode ser dividido em dois grandes etapas: a extração de características via detecção de bordas, e o cálculo da matriz de correlação espacial entre as imagens a serem analisadas (aérea e georeferenciada). Com isso, a implementação de um extrator de características (algoritmo detector de bordas) robusto junto com um cálculo da matriz de correlação rápido e eficiente entre imagens, garantirá um bom desempenho no método.

Consequentemente, tomando como base o método de Casamento de Imagens clássico, foram propostos dois novos métodos que visam reduzir o tempo de processamento e melhorar a precisão na estimação de posição. O primeiro método é focado na implementação de um algoritmo detector de bordas *canny* adaptativo que melhora consideravelmente o problema de sobrestimação de bordas nas imagens aéreas. Neste método é obtido o grau de homogeneidade presente na imagem analisada, e posteriormente, é ajustada a sensibilidade do detector de bordas. O segundo método propõe fazer o Casamento de Imagens por limiarização ao invés de aplicar detecção de bordas, onde o limiar calculado representa um valor de tolerância ao ruído presente na imagem. Assim, consegue-se extrair e reproduzir um maior número de informações na imagem sem presença de ruído.

Adicionalmente, são propostos outros dois métodos para a correção de distorção projetiva, ajuste de escala e rotação de imagens quando a posição da câmera no momento da captura não é perpendicular à Terra (não possui visada nadir). O primeiro método é paramétrico, ou seja, utiliza o conhecimento prévio dos ângulos de inclinação da aeronave fornecidos pelos sensores inerciais da mesma. Estes ângulos compõem os parâmetros de rotação da matriz de homografia, que é também composta por outras matrizes concatenadas que representam parâmetros intrínsecos da câmera e de translação na imagem. Com estas informações é possível definir a função de transformação homográfica e fazer a correção na imagem analisada. Por sua vez, o segundo método é não-paramétrico, e apresenta uma robusta correção de distorção projetiva e espectral nas imagens aéreas. Este método baseia-se principalmente no casamento de pontos característicos extraídos entre duas imagens (capturada pelo VANT e imagem georeferenciada) utilizando os algoritmos SURF (Bay et al., 2008) e MSAC (Torr and Zisserman, 2000).

1.2 Objetivos e Contribuições do Trabalho

O objetivo principal deste trabalho é propor e desenvolver melhoras no processo de estimação de posição geográfica em VANTs através de imagens aéreas, a fim de reduzir o tempo de processamento e, ao mesmo tempo, melhorar a precisão dos métodos propostos.
Os objetivos específicos são apresentados a seguir.

- Reduzir o tempo de processamento na estimação de posição de VANTs via Casamento de Imagens.
- Melhorar a precisão no cálculo da estimação de posição de VANTs via Casamento de Imagens.
- Propor e implementar métodos que minimizem os fatores que dificultam o processo de estimação de posição de VANTs via Casamento de Imagens.
- Avaliar o desempenho dos métodos propostos em função de: (a) robustez apresentada em diferentes classe de terreno (Vegetação, Urbano e Rodovia) (b) métrica de erro fornecida pela distância média de separação entre a posição calculada pelo método e a posição real (c) tempo de processamento (d) uma análise comparativa entre os métodos propostos.

Como contribuições deste trabalho destaca-se o seguinte:

- Implementação do detector de bordas *canny* adaptativo para estimação de posição via Casamento de Imagens aéreas;
- Cálculo de correlação por limiarização para estimação de posição via Casamento de Imagens aéreas;
- Correção de distorção projetiva em imagens capturadas por VANTs utilizando um método paramétrico;
- Correção de distorção projetiva e espectral em imagens capturadas por VANTs utilizando um método não-paramétrico;
- Caracterização dos algoritmos de extração e casamento de pontos característicos, SURF e MSAC, para encontrar o melhor ponto de operação;
- Avaliação dos métodos propostos neste trabalho em uma simulação de voo real, integrando o método de correção de perspectiva paramétrico e o método de estimação de posição via Casamento de Imagens.

1.3 Organização da Dissertação

Este trabalho está distribuído em outros seis capítulos. No Capítulo 2 é mostrado o estado da arte e conceptos básicos utilizados neste trabalho. No Capítulo 3 são propostos dois novos métodos para melhorar o tempo de processamento e a precisão na estimação de posição em VANTs via Casamento de Imagens. No Capítulo 4 é apresentado um novo método para correção de distorção em imagens aplicada à estimação de posição em VANTs utilizando um método paramétrico. Também, foi desenvolvido outro método que faz correção de distorção projetiva e espectral em imagens capturadas por VANTs utilizando um método não-paramétrico mostrado no Capítulo 5. A discussão e apresentação dos resultados experimentais e análises são mostrados no Capítulo 6. Por fim, as conclusões e trabalhos futuros são abordados no Capítulo 7.

Capítulo 2

Estado da arte

Esse Capítulo aborda o referencial teórico sobre a técnica de estimação de posição através de Casamento de Imagens e apresenta fatores relevantes que influenciam no processo de estimação de posição e sobre as imagens capturadas por VANTs. Inicialmente, na Seção 2.1, é descrita a técnica de estimação de posição geográfica em VANTs através de Casamento de Imagens. Em seguida, a Seção 2.2 contém os diferentes tipos de distorções geométricas que podem apresentar as imagens aéreas capturadas por VANTs, e também, mostra como estas distorções podem ser corrigidas. Por último, a Seção 2.3 descreve o funcionamento do algoritmo SURF, utilizado para extração e casamento de pontos característicos detectados em pares de imagens diferentes com uma mesma cena em comum.

2.1 Estimação de posição em VANTs via Casamento de Imagens

Um dos métodos atualmente mais utilizados para estimação de posição geográfica de VANTs, faz uso do reconhecimento de padrões em imagens aéreas (Khodaskar and Ladhake, 2014; Michalopoulos et al., 2013; Zhang et al., 2013). A técnica consiste em fazer o reconhecimento ou identificação da região atual sobrevoada pelo VANT sobre uma imagem de satélite previamente georeferenciada que contém informações de latitude e longitude. De modo que, encontrar o maior número de correspondências e similaridades entre as imagens vai garantir uma maior precisão e robustez ao método aplicado.

A obtenção de correspondências e similaridades via Casamento de Imagens tem apresentado resultados satisfatórios em diversas aplicações, como é mostrado em (Silva et al., 2015; Goltz, 2011; Conte and Doherty, 2008). Assim, o Casamento de Imagens é aplicado entre a imagem da região sobrevoada capturada por uma câmera abordo

do VANT e a imagem de satélite georeferenciada. Informações adicionais abordo do VANT são também aproveitadas com a finalidade de melhorar o processo de estimação de posição. O uso dos sensores inerciais (INS - *Inertial Navigation System*) como acelerômetros e giroscópios, fornecem uma estimativa de posição com um erro acumulado que aumenta ao longo do tempo (Lima et al., 2005). No entanto, esta informação é aproveitada para criar um recorte aproximado da imagem de satélite e, desta forma, evitar a busca da imagem capturada pelo VANT sobre toda a imagem georeferenciada. Com isto, o custo computacional é reduzido consideravelmente e, por conseguinte, o tempo de processamento requerido no processo de estimação de posição.

Na figura 2.1 é mostrado o procedimento para realizar a estimação de posição via Casamento de Imagens. Neste contexto, existe uma série de condições iniciais que são requeridas para garantir que a técnica forneça bons resultados:

- A imagem capturada pelo VANT deve estar com visada nadir (Silva et al., 2015);
- A imagem capturada pelo VANT deve apresentar fatores de escala e rotação próximos à imagem de satélite georeferenciada;
- Deve-se, reduzir a diferença espectral e realçar as bordas das imagens a serem comparadas (VANT e georeferenciada).

Somam-se a isto outros fatores como a correção de distorção projetiva (perspectiva) e distorção espectral gerada pela diferença de sensores nas imagens a serem testadas que, precisam ser avaliados e resolvidos antes de fazer o processo de estimação de posição para garantir melhores resultados.

Como pode ser apreciado na figura 2.1, o método que normalmente é utilizado toma leitura da região sobrevoada desde uma câmera fixada no VANT e, ao mesmo tempo, faz leitura de uma imagem de satélite georeferenciada que foi previamente embarcada. Depois, com as informações do INS é feito um recorte aproximado da posição onde se espera que esteja voando o VANT. Este recorte resulta ser muito menor que a imagem de satélite georeferenciada e maior que a imagem capturada pelo VANT. Seguidamente, a imagem de satélite é transformada a escala de cinza e é realizada a extração de bordas na imagem. Ao mesmo tempo em que a imagem de satélite é processada, a imagem capturada pelo VANT é sometida a um processo mais complexo. Começando com a conversão de RGB a escala de cinza, para logo ajustar a escala e rotação tomando como referência a leitura dos dados fornecidos pelo INS, a bússola e o altímetro. Posteriormente, o filtro da média e os filtros passa baixa e passa alta são aplicados sobre a imagem para suavizar-la, e reduzir o ruído e as distorções presentes na captura da imagem. Em seguida, a extração de bordas é aplicada com a finalidade de eliminar

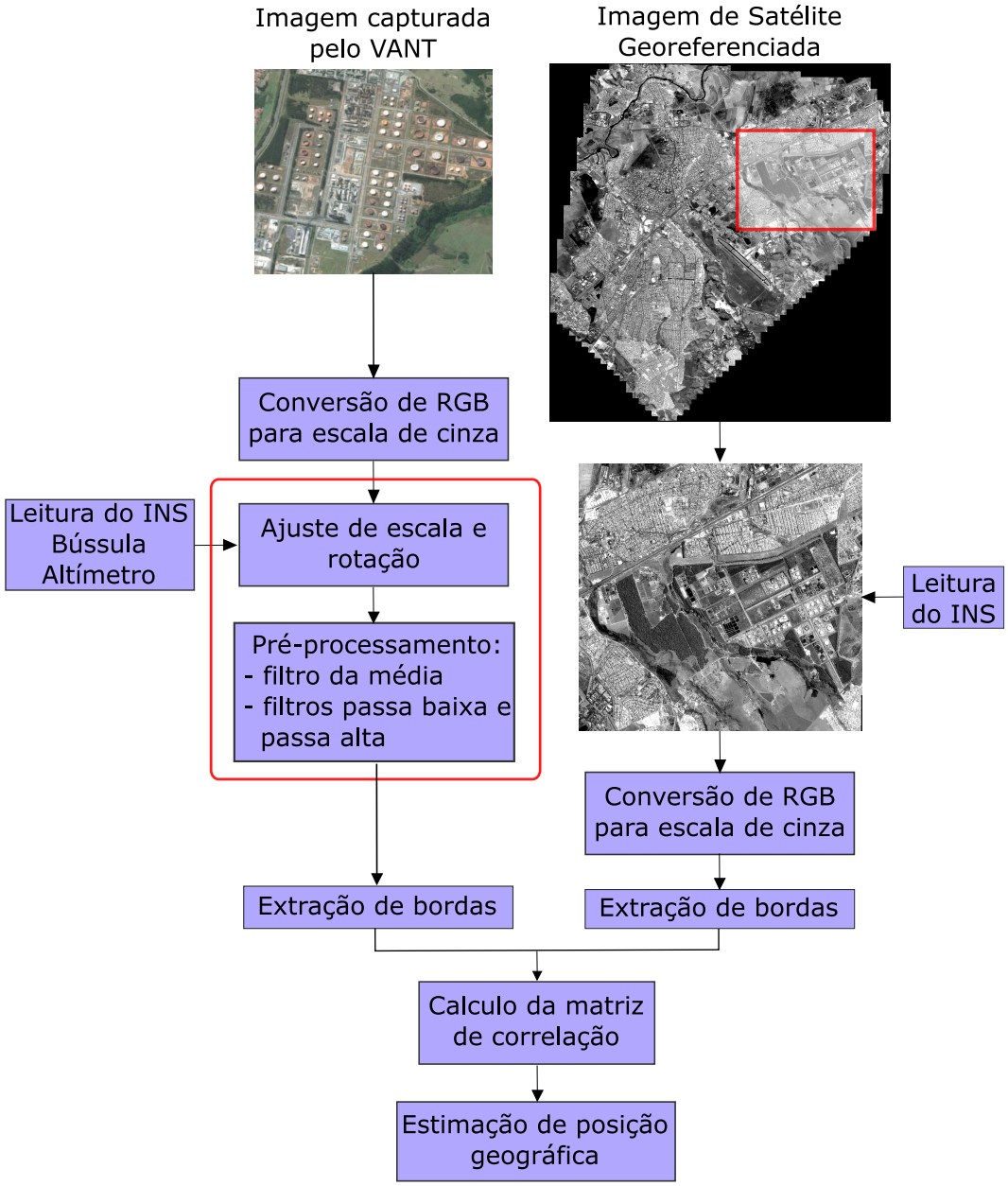


Figura 2.1: Diagrama de estimação de posição geográfica com a técnica Casamento de Imagens.

mudanças grandes de luminosidade entre imagens e caracterizar as imagens baseando-se nas bordas mais representativas que foram detectadas. Finalmente, será estimada a posição geográfica do VANT identificando o maior índice de correlação obtido do cálculo da matriz de correlação entre as duas imagens (VANT e georeferenciada) com as bordas já extraídas.

2.1.1 Casamento de Imagens por Correlação Espacial

Como foi mencionado anteriormente, a detecção de bordas nas imagens representa a etapa anterior ao Casamento de Imagens feito por correlação. Como a imagem de satélite é recortada com base na estimação de posição fornecida pelo INS, nela existe um ponto central p que é atribuído a partir da latitude e longitude informada pelo INS. A partir desse ponto central p, são acrescentados um número constante n de pixeis nas direções Norte, Sul, Leste e Oeste, deixando como resultado final uma imagem quadrada com dimensões de $2n \times 2n$. Esse recorte feito na imagem de satélite georeferenciada será denominado como imagem *target* e a imagem capturada pelo VANT será chamada imagem *template*. A dimensão do *target* (número de pixeis n) será definida dependendo do tamanho da imagem *template*, sempre garantindo que o *target* seja maior que o *template* para poder realizar de forma correta o casamento das imagens.

O processo de casamento das imagens é feito por correlação no domínio espacial. De acordo com (Gonzalez and Woods, 2008), o cálculo da correlação requer o deslocamento da imagem *template* sobre a imagem *target* (ver figura 2.2). De modo que, em cada deslocamento é calculada a suma dos produtos para cada posição (s, t). A equação (2.1) define o cálculo da matriz de correlação no domínio espacial, e é mostrada a seguir

$$c(s,t) = \sum_x \sum_y f(x,y)w(x-s, y-t) \tag{2.1}$$

O termo $c(s,t)$ denota a correlação c na posição (s,t), com $s = 0, 1, ..., M-1$ e $t = 0, 1, ..., N-1$; onde M e N correspondem às dimensões da matriz f, que seria a imagem *target* com as bordas já extraídas. A matriz w com dimensões $J \times K$, e com restrições de $J \leq M$ e $K \leq N$, corresponde à imagem *template* com bordas já extraídas. A soma é realizada sobre a região em que a matriz f e a matriz w se sobrepõem. Como a origem da matriz $f(x,y)$ é definida na parte superior esquerda, o percurso a seguir por parte da matriz $w(x,y)$ será de esquerda a direita percorrendo todas as colunas e descendo uma fila por vez.

Baseando-se na equação (2.1), a estimação de posição do VANT é calculada da seguinte forma: para qualquer valor de posição (s,t) dentro de $f(x,y)$ é calculado um valor em c produto do varrido de $w(x,y)$. De modo que, a obtenção do valor máximo de $c(s,t)$ indica a posição em que $w(x,y)$ obteve um melhor casamento com $f(x,y)$. Assim, o ponto central da imagem *template* corresponderá à localização do VANT sobre a área sobrevoada. Finalmente, esse ponto em comum e de alta correlação $c(s,t)$ entre as duas imagens *(target e template)* contém as informações de latitude e longitude

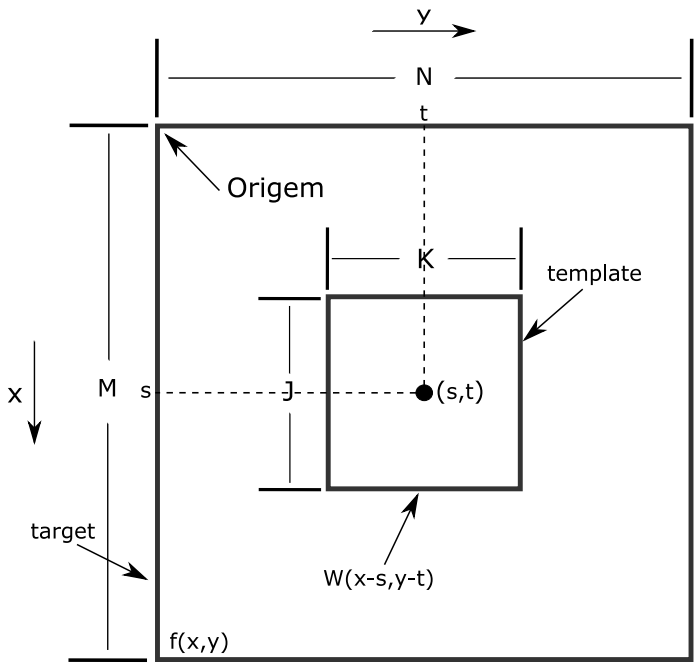


Figura 2.2: Arranjo para a obtenção da correlação de $f(x, y)$ e $w(x, y)$ de (Gonzalez and Woods, 2008).

previamente embarcadas na imagem de satélite georeferenciada. De preferencia, a imagem *template* pode ter dimensões impares para que seja mais fácil atribuir seu pixel central como referência do ponto de máxima correlação (Silva et al., 2015). A figura 2.3 mostra o processo de calculo da matriz de correlação e a obtenção do ponto de máxima correlação entre a imagem capturada pelo VANT *(template)* e a imagem de satélite georeferenciada *(target)*, e como com esse ponto máximo de correlação são obtidos os valores de latitude e longitude.

2.1.2 Métodos de Extração de Bordas

Os algoritmos de detecção de bordas fazem parte da área de segmentação de imagens, e são geralmente baseados em uma das propriedades básicas de valores de níveis de cinza: descontinuidade e similaridade. Tomando como critério de segmentação nas imagens as mudanças abruptas nos níveis de cinza, é possível detectar pontos isolados, detectar linhas e bordas na imagem (Gonzalez and Woods, 2008).

A implementação de algoritmos focados na extração e detecção de bordas com aplicação na estimação de posição geográfica de VANTs, tem como finalidade reduzir ou eliminar as mudanças de luminosidade que podem apresentar as imagens *target* e *template*. A natureza das imagens (capturada pelo VANT e georeferenciada) faz que a diferença de luminosidade entre elas seja considerável, devido a capturas realizadas em tempos diferentes, com distintos sensores e em diferentes condições climáticas. Como foi

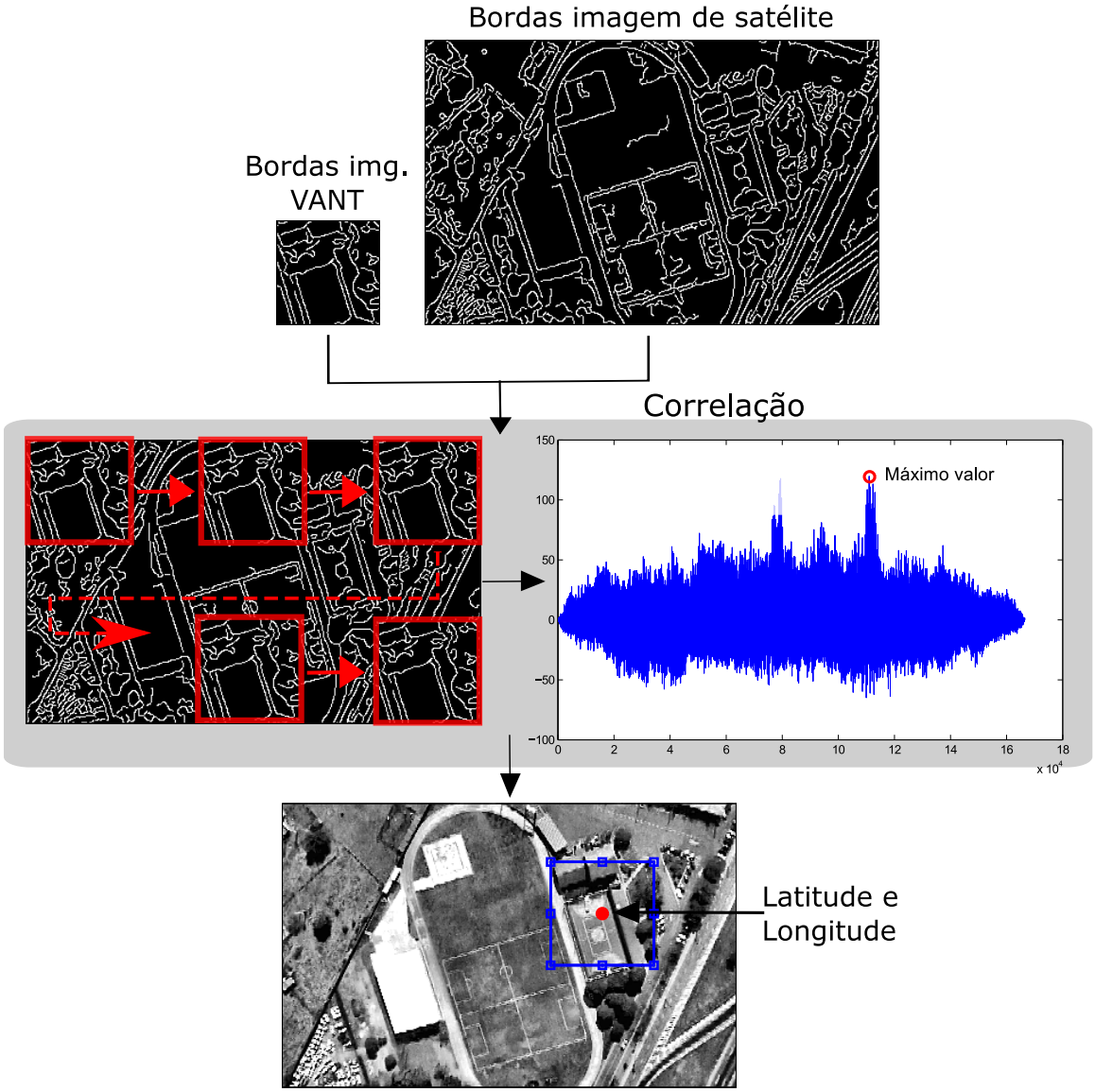


Figura 2.3: Processo de correlação entre as imagens.

observado em trabalhos anteriores (Silva et al., 2015; Goltz, 2011; Conte and Doherty, 2008), a extração de bordas tem demostrado ser robusta a esse tipo de problemas de luminosidade presentes em imagens aéreas.

De acordo com (Gonzalez and Woods, 2008), uma borda é definida como o limite entre duas regiões com propriedades relativamente distintas no nível de cinza. No caso de imagens aéreas, será possível determinar bordas se as regiões analisadas são suficientemente homogêneas entre si, e a transição entre estas regiões é diferenciável. Neste processo, são utilizados operadores de convolução entre matrizes, denotado por $f(x,y)*g(x,y)$. Sendo f a imagem a ser processada e g uma mascara de convolução. A equação (2.2) define a operação de convolução onde a matriz h é a imagem resultante, contendo as bordas detectadas (Goltz, 2011).

$$h(x,y) = f(x,y) * g(x,y) = \sum_{b=-\infty}^{\infty} \sum_{a=-\infty}^{\infty} f(a,b) \times g(x-a, y-b) \tag{2.2}$$

Para a detecção e extração de bordas, o filtro *canny* (Canny, 1986) é o mais implementado por oferecer uma maior robustez e melhor qualidade de bordas sobre as imagens (Goltz, 2011; Silva et al., 2015).

2.1.3 Algoritmo Detector de Bordas Canny

O detector de bordas *canny*, pode também ser chamado como um operador ou filtro de convolução que a diferença de outros detectores de bordas, apresenta uma maior robustez ao ruído presente nas imagens. Segundo o descrito em Canny (1986), para que um filtro seja ótimo na detecção de bordas, ele deve atingir no minimo três critérios de desempenho:

- Uma boa detecção. Isto implica que o detector de bordas deve encontrar somente bordas, ou seja, deve possuir uma baixa probabilidade de falhar na detecção de bordas reais e baixa probabilidade de marcar bordas falsas. Isso faz que o detector apresente uma maximização da razão sinal/ruído.
- Uma boa localização. Os pixeis que o detector marcou como borda devem estar o mais próximo possível do centro da verdadeira borda.
- Apenas uma resposta para uma única borda. Evitar que o detector de bordas identifique múltiplas bordas onde somente deveria existir uma única borda.

Segundo Canny (1986), as equações (2.3) e (2.4) descrevem os dois primeiros critérios definidos anteriormente. O objetivo na detecção e localização de bordas por parte do detector consiste em maximizar o produto dessas equações, Eq.(2.5).

$$\frac{\left|\int_{-w_c}^{+w_c} G(-x) f_c(x) d_x\right|}{n_o \sqrt{\int_{-w_c}^{+w_c} f_c^2(x) d_x}} \tag{2.3}$$

$$\frac{\left|\int_{-w_c}^{+w_c} G'(-x) f_c'(x) d_x\right|}{n_o \sqrt{\int_{-w_c}^{+w_c} f_c'^2(x) d_x}} \tag{2.4}$$

$$\frac{\left|\int_{-w_c}^{+w_c} G(-x) f_c(x) d_x\right|}{n_o \sqrt{\int_{-w_c}^{+w_c} f_c^2(x) d_x}} \cdot \frac{\left|\int_{-w_c}^{+w_c} G'(-x) f_c'(x) d_x\right|}{n_o \sqrt{\int_{-w_c}^{+w_c} f_c'^2(x) d_x}} \tag{2.5}$$

$f_c(x)$ é a resposta de impulso do filtro definido no intervalo $[-w_c, w_c]$. $G(x)$ é uma borda unidimensional e n_o é a amplitude do ruído na imagem. Neste abordagem, é assumido que a borda está centrada em $x = 0$.

A última condição que o detector deve atingir está representada na equação (2.6), e sintetiza matematicamente a ideia de que deve haver uma única resposta para uma única borda verdadeira. De modo que é calculada a distância X_{max} entre máximos adjacentes na resposta do descritor.

$$X_{max}(f_c) = 2\pi \left(\frac{\int_{-\infty}^{+\infty} f_c'^2(x) d_x}{\int_{-\infty}^{+\infty} f_c''^2(x) d_x} \right)^{\frac{1}{2}} \tag{2.6}$$

Assim, ao mesmo tempo em que é maximizada a relação do produto dos dois primeiros critérios (Eq. (2.5)), deve-se também procurar o máximo valor possível de X_{max} (Eq. (2.6)).

De acordo com Silva et al. (2015), o detector de bordas de *canny* pode ser dividido de forma geral em 5 etapas:

- Suavização: é aplicado sobre a imagem analisada um filtro de tipo Gaussiano (Gonzalez and Woods, 2008) que produz o efeito de borramento na imagem com a finalidade de reduzir a existência de ruídos e evitar que sejam detectados como bordas;

- Intensidades de tons de cinza: a aplicação do gradiente sobre a imagem e com a obtenção da magnitude desse gradiente é possível identificar as áreas nas direções x e y separadamente;

- Supressão não máxima: são mantidos os máximos locais da imagem após de aplicar o gradiente com a finalidade de preservar os candidatos de bordas mais fortes;

- Limiar duplo: esta etapa compreende a classificação de bordas em verdadeiras ou falsas. Na maioria dos casos, as bordas detectadas resultam ser verdadeiras, mas outras resultam ser falsas. Isto devido a ruídos grandes que não foram suprimidos na etapa de suavização ou possíveis variações de cores dentro de uma mesma região homogênea. Nesta etapa é definido um limiar, de modo que as bordas maiores a esse limiar são consideradas como verdadeiras. O detector de bordas *canny* utiliza limiar duplo, ou seja, as bordas acima do limiar maior serão bordas fortes e as outras com valores entre os dois limiares serão bordas fracas;

- Rastreamento da borda por histerese: existem as condições de que, se a borda foi forte, pode ser incluída no resultado final da detecção de bordas. Mas se a

borda é fraca, ela será incluída no resultado final, se e somente se, estivera perto de uma borda forte.

Nas figuras 2.4 e 2.5 é mostrada a detecção e extração de bordas feita pelo operador *canny* a uma imagem aérea e de satélite.

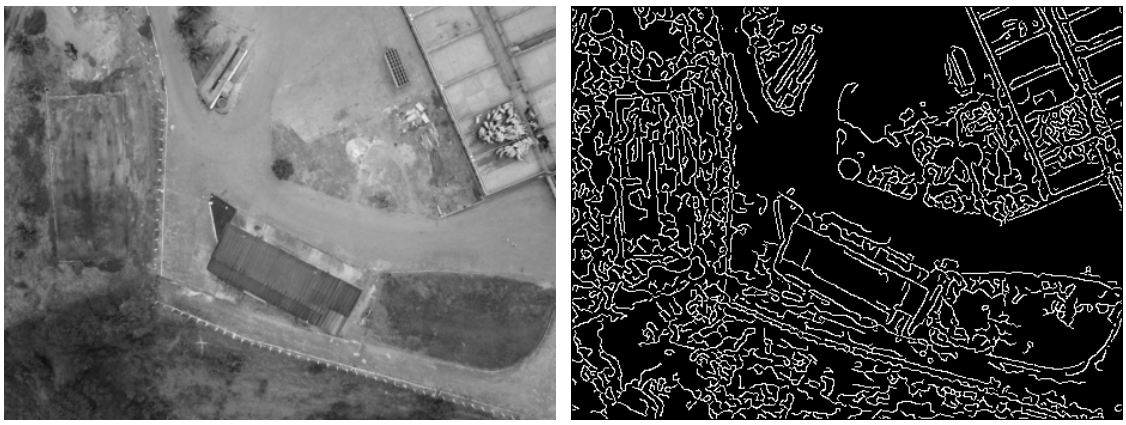


Figura 2.4: Exemplo do detector de bordas *canny* aplicado em uma imagem aérea capturada por VANT.

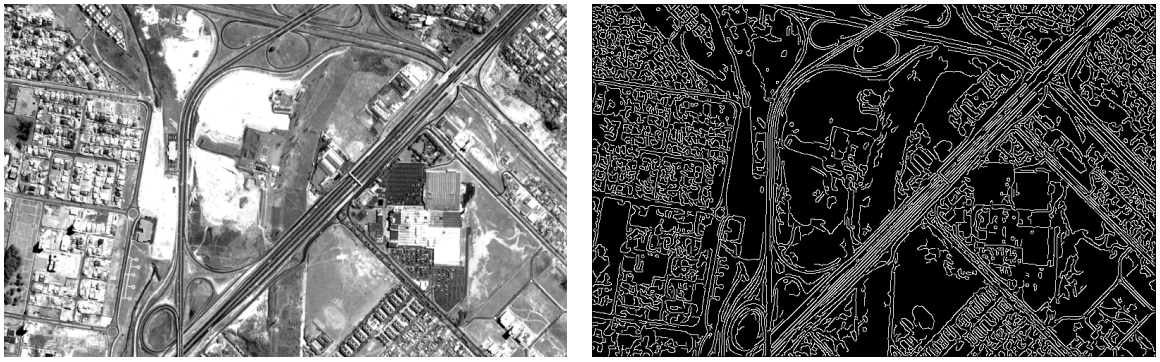


Figura 2.5: Exemplo do detector de bordas *Canny* aplicado em uma imagem de satélite.

2.1.4 Trabalhos Relacionados

As técnicas de processamento de imagens aplicadas na navegação autônoma de VANTs têm sido motivo de pesquisa em diversos trabalhos nos últimos anos. A grande demanda de aplicações onde os VANTs conseguem executar tarefas de forma eficiente e autônoma com tempos baixos de processamento, tem motivado ao perfeicionamento dos métodos para tornar o desenvolvimento dos mesmos ainda mais robustos e eficientes.

Fazer que o VANT tenha a capacidade suficiente de autonomia na navegação, eliminando a dependência do Sistema de Navegação Global por Satélite (GNSS) é um dos maiores desafios atualmente abordados (Merino et al., 2006; Patterson et al., 2011). A

estimação de posição utilizando a técnica de Casamento de Imagens permite a inclusão de novas técnicas de Visão Computacional que podem fazer com que o Casamento de Imagens forneça melhores resultados na navegação autônoma de VANTs.

O estudo apresentado em Conte and Doherty (2008), propõe um sistema integrado de navegação de VANT robusto baseado em Casamento de Imagens aéreas. Este sistema de navegação combina a informação dos sensores inerciais junto com a técnica de Odometria visual e Casamento de Imagens. Com isto, é estimada a posição geográfica do VANT que será utilizada em situações nas que existam falhas no sinal de GPS. Nesse trabalho, é desenvolvida a técnica de Casamento de Imagens, baseada na extração de bordas, obtidas pelo operador *Sobel* e logo, aplicada entre a imagem aérea e a imagem de satélite georeferenciada. De modo que as informações de borda extraídas das imagens analisadas permitem identificar a posição geografia da aeronave no momento de voo.

Em Goltz (2011) foram utilizadas Redes Neurais Artificiais (RNAs) na estimação de posição de VANTs, pois estas apresentam características como baixo custo computacional, tolerância a falhas e robustez ao ruído, tornando-se em uma aplicação com potencial. Por essa razão, no trabalho de Goltz (2011), foram implementadas três tipos de RNAs com diferentes arquiteturas: Rede com Funções de Base Radial (RBF), Rede Perceptron de Múltiplas Camadas (RPMC) e Rede Neural Celular (RNC). Depois, foram comparadas com detectores de bordas tradicionais comummente utilizados no processamento digital de imagens como, o *Roberts*, o *Sobel*, o Laplaciano e o *Canny*. As RNAs aplicaram-se na extração de bordas em imagens aéreas e de satélite para posterior casamento das imagens com o cálculo da correlação no domínio espacial entre as bordas extraídas, a fim de estimar a posição geográfica do VANT.

Por sua parte Silva et al. (2015) apresentou resultados da abordagem de Casamento de Imagens aplicada à estimativa de posição para a navegação autônoma de VANTs. Na técnica de Casamento de Imagens, foram realizadas extrações de bordas por cinco métodos diferentes dos quais, o operador *canny*, a rede neural RPMC treinada sob duas metodologias distintas para representar as bordas em uma imagem aérea e a rede neural RBF foram testadas. A avaliação do método considerou diferentes tipos de terreno (Vegetação, Rodovia e Urbana) dentro de uma rota planejada que será posteriormente sobrevoada pelo VANT.

O estudo de Velho et al. (2015), apresentou um método para realizar estimação de posição em VANTs quando são utilizadas imagens do tipo LIDAR (Laser Imaging Detection and Ranging) que permitem a navegação autônoma sobre o oceano. A pesquisa obtém estimativas de posição do VANT sobre o oceano com base no cálculo dos índices de correlação (Casamento de Imagens) entre as imagens de um sensor LIDAR abordo do VANT. Foram desenvolvidos diferentes tipos de extratores de bordas, aplicando

Rede com Funções de Base Radial (RBF), Rede Perceptron de Múltiplas Camadas (RPMC), o operador *canny* e o operado *sobel*. No final, os resultados das estimativas de posição obtidos com cada extrator foram comparados.

Em sua pesquisa da Silva et al. (2015), motivado pelas aplicações na navegação autônoma de VANTs em voos noturnos, apresentou um estudo para estimar a posição geográfica do VANT com o uso de imagens infravermelhas e a técnica de Casamento de Imagens. A imagem infravermelha obtida pela câmera a bordo da aeronave é procurada sobre uma imagem de satélite georeferenciada na banda do visível. O processo de registro automático entre imagens aéreas e de satélite, é feito com bordas extraídas por aplicação de RNAs configuradas automaticamente com o uso do algoritmo M-PCA (Multiple Particle Collision Algorithm), e pelo cálculo do índice de correlação espacial entre imagens.

O autor Patterson et al. (2011) apresentou um método de estimativa de posição em VANT baseado no registro de imagens aéreas. É fornecido um Sistema de Informações Geográficas (Geographic Information System - GIS) que pode ser utilizado em caso de falhas no sinal de GPS. Esta metodologia não requer de imagem georeferenciada reduzindo o custo de armazenamento. O registro de imagens é realizado com base em características de confiança extraídas das imagens, tais como, estradas, rodovias e água. É utilizado o operador *canny* como extrator de bordas e calculada a matriz de correlação cuzada normalizada para realizar o casamento das imagens.

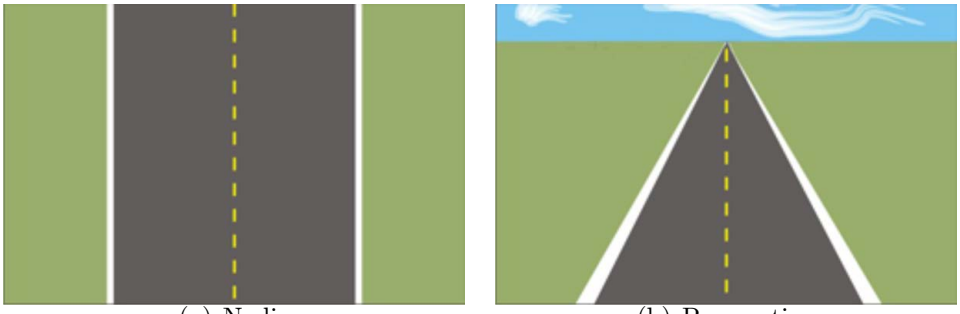
Na pesquisa de Shan et al. (2015) foi proposto um método para a navegação de VANTs assistida por Google Maps quando existem falhas no sinal de GPS. A estimação de posição é efetuada via Casamento de Imagens com filtros de correlação adaptativos. Um fluxo óptico foi utilizado para predizer a posição em frames seguintes. Obteve-se a matriz homográfica, utilizou-se a extração de características HOG (Histograms of Oriented Gradients) para registrar as imagens em Goolge Maps e um filtro de partículas para refinar a localização do VANT.

2.2 Correção de Distorções Geométricas em Imagens Aéreas Capturadas por VANTs

Nas imagens aéreas capturadas por VANTs é muito comum a presença de distorções geométricas sistemáticas que, dependendo do tipo de plataforma e as condições do ambiente, fazem que o grau de distorção sobre a imagem seja maior (Meneses et al., 2012). Desta forma, a correção de distorções geométricas contribuirá à remoção dos erros sistemáticos presentes nessas imagens, que foram ocasionados devido à rotação, inclinação e instabilidade da plataforma que captura a imagem. Ao mesmo tempo, pos-

sibilita a transformação de imagens com perspectiva obtidas pelo VANT, em imagens ortogonais com visada nadir.

Como é mostrado na figura 2.6, a imagem da esquerda 2.6(a) corresponde a uma imagem após a correção de distorção de perspectiva que apresenta a imagem da direita 2.6(b). De modo que, existe uma transformação onde inicialmente as retas paralelas da estrada não são paralelas na fotografia com perspectiva, ou seja a relação de paralelismo não é preservada. Isto faz que o tamanho do pixel varie nas diferentes regiões da imagem, o pixel fica menor na parte imageada que está mais perto do sensor e maior quando existe maior distância com o sensor.



(a) Nadir (b) Perspectiva

Figura 2.6: Exemplo de imagem com visada nadir e com perspectiva.

As distorções geométricas são consideradas desajustes nos pixeis que conformam a imagem aérea, ou seja, o valor do pixel possui uma posição errada dentro da imagem.

As principais distorções geométricas estão relacionadas ao seguintes fatores:

- Rotação da Terra durante o tempo de imageamento (skew): sucede durante o momento de adquisição da imagem, onde a Terra está rotacionando de Oeste para Leste. De modo que é provocado um deslocamento relativo das linhas de imageamento para o Oeste, como é mostrado na figura 2.7. A quantidade de deslocamento e de distorção depende da velocidade relativa da aeronave e da dimensão da área imageada. Embora este tipo de distorção não seja muito comum em VANTs, já que a altura de voo não é muito grande ($<$ 120m) (ANAC, 2015), existem algumas excepções em que as aeronaves podem apresentar este tipo de distorções.

- Variações na altitude, velocidade e atitude do VANT: a variação da altitude na aeronave produz mudanças de escala. Similarmente, a mudança de velocidade conduz à variação de escala ao longo da direção de imageamento, enquanto as variações dos eixos de atitude do VANT (Yaw - guinada, Pitch - arfagem, Roll - rolagem), causam rotações transversais e longitudinais na direção do voo. A

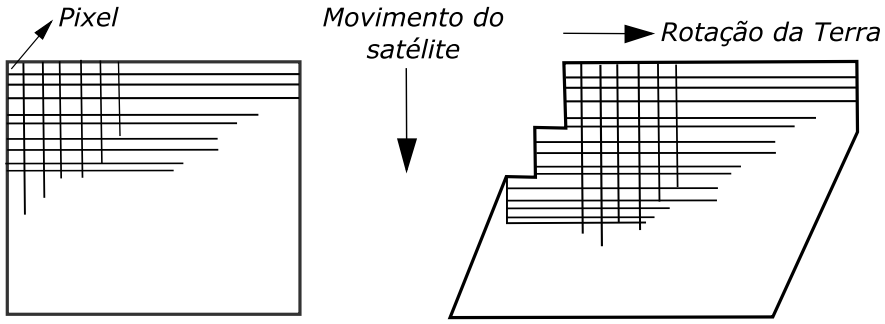


Figura 2.7: Deslocamento na imagem devido ao movimento de rotação da Terra.

figura 2.8 ilustra as diferentes distorções geradas pelas variações anteriormente mencionadas, onde a linha tracejada indica a distorção resultante.

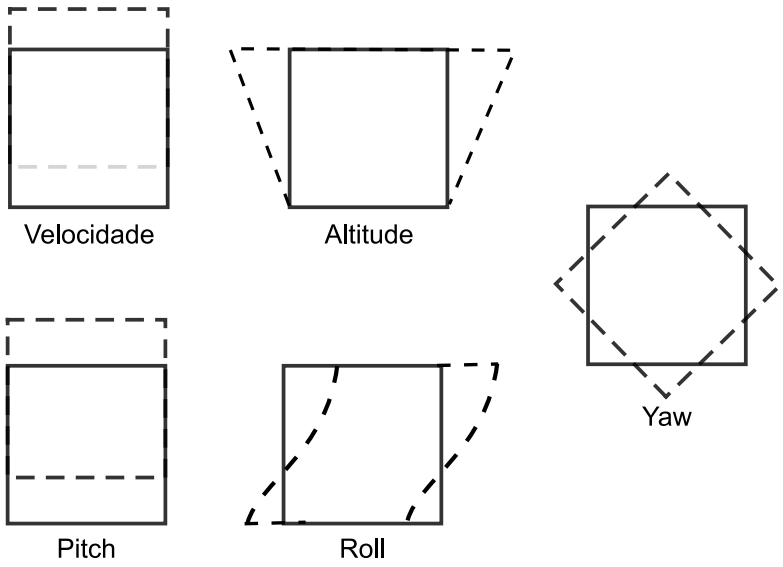


Figura 2.8: Efeitos das distorções geométricas causadas pelo movimento do VANT (instabilidade).

- Distorção da lente: este tipo de distorção depende do tipo de câmera e lente que seja utilizado para fazer o imageamento no VANT. Se o centro da lente da câmera não está alinhado e se existem imprecisões na calibração da câmera antes de ser utilizada, as distorções podem ser de tipo radial ou tangencial (ver fig.2.9). Sendo a primeira mais predominante, pois provoca curvatura nas linhas retas da imagem.

2.2.1 Tipos de Correções Geométricas

No processamento digital de imagens, existem diferentes tipos de correção geométrica. Com o objetivo principal de fazer ortorretificação em imagens aéreas capturadas por VANTS, estas são classificadas a seguir:

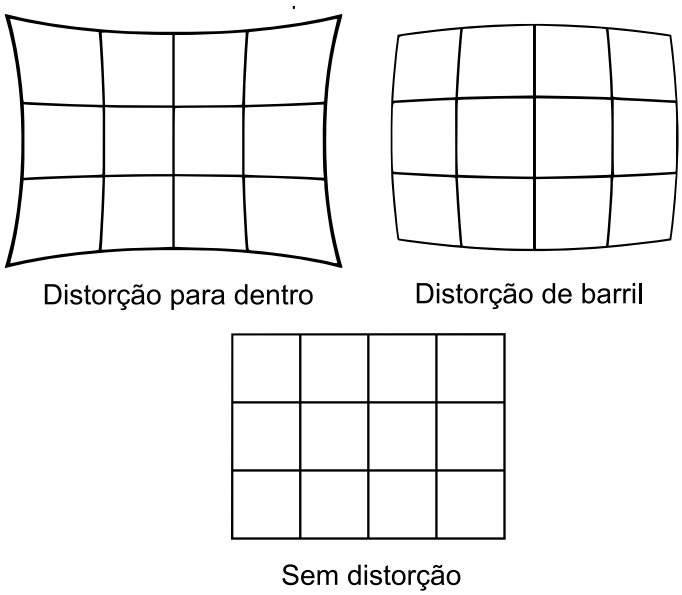


Figura 2.9: Distorções causadas pela lente da câmera abordo do VANT.

- Correção geométrica paramétrica: este tipo de correção baseia-se no conhecimento dos parâmetros intrínsecos da câmera e as informações da sua posição e orientação. Portanto, este tipo de correção, é independente da imagem e não precisa de nenhum processo de identificação de pontos. Um modelo da câmera é requerido. Uma aplicação deste tipo de correção geométrica é proposto no Capítulo 4.

- Correção geométrica não paramétrica: Este tipo de correção geométrica, também conhecida como método baseado em procedimentos de amostragem, utiliza pontos de apoio sobre a imagem. Este método utiliza como base um mapa georeferenciado e compara as diferencias entre as posições de diferentes pontos escolhidos, identificados e registrados em dois imagens (imagem capturada pelo VANT e a imagem de satélite georeferenciada). Desta forma, não é preciso o modelo da câmera, só é necessária uma imagem de satélite georeferenciada tomada como imagem de referência. Uma aplicação de este tipo de correção geométrica é proposto no Capítulo 5.

2.2.2 Modelo de Calibração da Câmera

O modelo da câmera está descrito pelo modelo *pinhole* mostrado na figura 2.10, e é amplamente utilizado na literatura (Hartley and Zisserman, 2000). Este modelo é composto por parâmetros intrínsecos e parâmetros extrínsecos que podem ser formalmente descritos de forma matricial como segue.

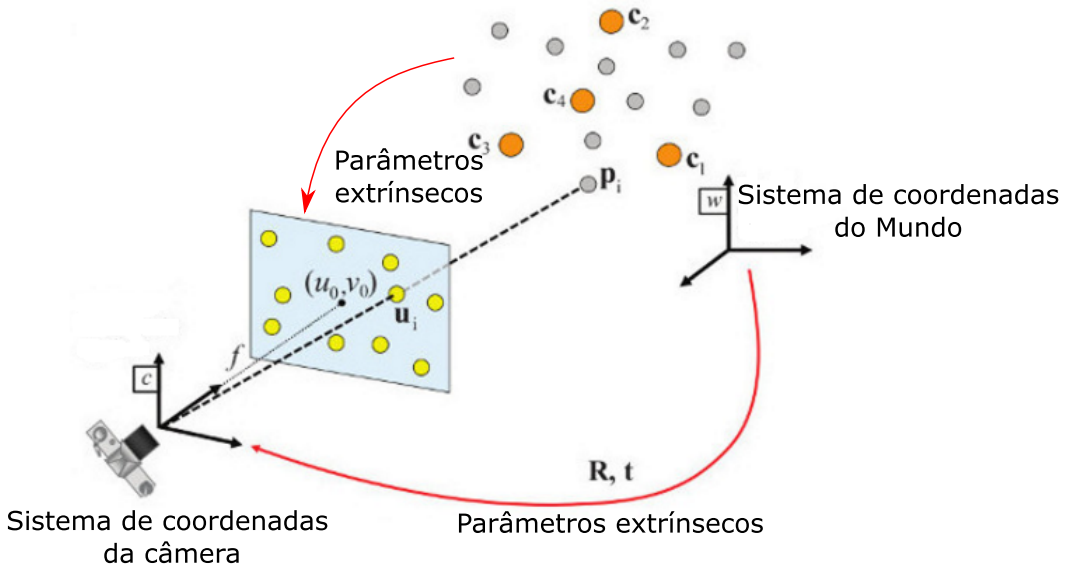


Figura 2.10: Modelo da câmera composto por parâmetros intrínsecos e extrínsecos

- **Parâmetros Intrínsecos**

  Os parâmetros intrínsecos avaliados no modelo da câmera são compostos por especificações próprias do tipo do sensor, fornecidos pelo fabricante. Com isto, tem-se a intenção de obter a relação existente entre um ponto no plano da câmera (coordenadas u, v) e o ponto correspondente no plano do sensor (coordenadas em pixeis). A matriz que representa o modelo dos parâmetros intrínsecos da câmera é mostrada na Eq. (2.7).

$$A = \begin{bmatrix} f_{cam} \cdot k_u & 0 & u_0 \\ 0 & f_{cam} \cdot k_v & v_o \\ 0 & 0 & 1 \end{bmatrix} \tag{2.7}$$

  - f_{cam} é a distância focal da câmera,
  - k_u e k_v são os fatores de escala horizontal e vertical respectivamente, cujas inversas caracterizam o tamanho do pixel na unidade de coordenadas do mundo,
  - u_o e v_o são as coordenadas do ponto central da câmara, isto é, a intersecção entre o eixo ótico e o plano da imagem.

- **Parâmetros Extrínsecos**

  Os parâmetros extrínsecos da câmera estão compostos por uma matriz de rotação que define a orientação do VANT em relação ao eixo inercial. Indica em

que direção a lente da câmera esta apontando no momento de captura da imagem. Para definir a orientação relativa da aeronave, que contém a câmera, são utilizados os ângulos de Tait-Bryan, derivados dos ângulos de Euler, que são amplamente usados em aplicações aeroespaciais (Committee et al., 1995). De modo que, são utilizados três ângulos que descrevem a rotação no espaço Euclidiano tridimensional, por meio de três rotações sucessivas em torno dos eixos do sistema (x, y, z).

A primeira rotação corresponde ao ângulo de rolagem (ϕ) realizado em torno do eixo $\vec{x}$, a segunda rotação corresponde ao ângulo de arfagem (θ) realizado em torno do eixo $\vec{y}$ e a terceira e última rotação corresponde ao ângulo de guinada (ψ) realizado em torno do eixo $\vec{z}$.

$$R_{\phi,\theta,\psi} = R(x,\phi) \cdot R(y,\theta) \cdot R(z,\psi) \tag{2.8}$$

Onde:

$$R_{\phi,\theta,\psi} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix} \cdot \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.9}$$

A matriz de rotação resultante é mostrada em Eq. (2.10).

$$R_{\phi,\theta,\psi} = \begin{bmatrix} \cos\psi\cos\theta & \sin\psi\cos\theta & -\sin\theta \\ \cos\psi\sin\theta\sin\phi - \sin\psi\cos\phi & \sin\psi\sin\theta\sin\phi + \cos\psi\cos\phi & \cos\theta\sin\phi \\ \cos\psi\sin\theta\cos\phi + \sin\psi\sin\phi & \sin\psi\sin\theta\cos\phi - \cos\psi\sin\phi & \cos\theta\cos\phi \end{bmatrix} \tag{2.10}$$

Para simplificar os produtos dos coeficientes das matrizes de rotação, a matriz de rotação resultante será definida como se observa na equação Eq. (2.11). Esta matriz representa o modelo extrínseco e indica a orientação da câmera abordo do VANT.

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \tag{2.11}$$

Para finalizar, através do modelo de calibração da câmera é obtida uma estimação da cena analisada. O modelo inclui os parâmetros intrínsecos e extrínsecos formando uma matriz global, chamada matriz de projeção perspectiva P_p, onde é relacionado um ponto x_π em um sistemas de coordenadas $SM = [x, y, z]^T$ no plano do mundo $(3D)$, e as correspondentes coordenadas $m = [u, v]^T$ do ponto x no plano da imagem $(2D)$, como é mostrado na Eq. (2.12).

$$\begin{bmatrix} u & v & 1 \end{bmatrix}^T = P_p \cdot \begin{bmatrix} x & y & z & 1 \end{bmatrix}^T \tag{2.12}$$

Onde P_p esta definido por:

$$P_p = A \cdot R \tag{2.13}$$

Na Fig. 2.11 é ilustrado como os pontos no plano da imagem $2D$ e o plano da cena $3D$ estão relacionados por uma transformação de plano projetivo.

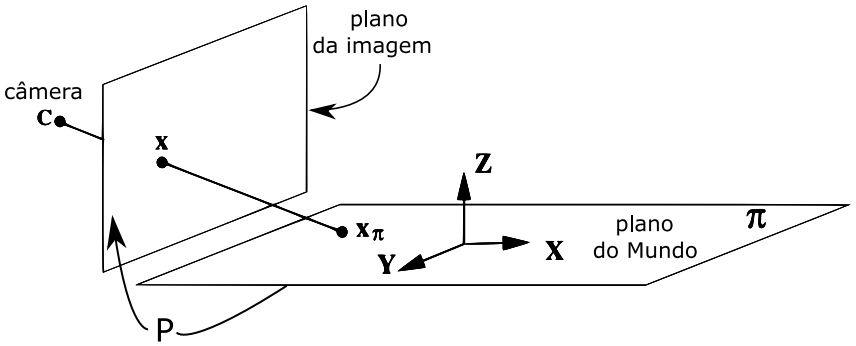


Figura 2.11: Transformação de plano projetivo

2.2.3 Transformadas Geométricas no Plano (2D)

No grupo de transformadas geométricas existe uma classificação hierárquica, onde em cada uma destas transformadas pode-se corrigir um determinado tipo de distorção. De modo que dependendo do grau e complexidade da distorção que apresenta a imagem aérea analisada, pode-se atribuir um tipo de transformada que melhor se ajuste às distorções contidas na imagem (Hartley and Zisserman, 2000).

- Isometrias - Transformação Euclidiana:

  As isometrias *(iso=igual, métrica=medida)*, são transformações do plano $2D$ que preservam a distância Euclidiana e são representadas como:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} \epsilon\cos\theta & -\sin\theta & t_x \\ \epsilon\sin\theta & \cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{2.14}$$

onde o termo ϵ indica a orientação da transformação. Se $\epsilon = 1$ mantem a orientação e se $\epsilon = -1$ é invertida a orientação da transformação. Os termos t_x e t_y correspondem aos fatores de traslação atribuídos em x e y respectivamente. A transformação Euclidiana está composta por uma translação e uma rotação, e pode ser escrita de forma simplificada:

$$X' = H_E X = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} X \tag{2.15}$$

onde R representa a matriz de rotação 2×2 e t é o vetor de traslação. Este tipo de transformação possui três graus de liberdade, dos quais um é para a rotação e dois para a translação. Desta forma, a transformação Euclidiana pode ser estimada a partir de dois pares de pontos correspondentes entre imagens.

- Transformação de Similaridade:

  Uma transformação de similaridade basicamente é uma isometria com adição de um fator de escala s. Isto é representado como:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} s\cos\theta & -s\sin\theta & t_x \\ s\sin\theta & s\cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{2.16}$$

  Os termos t_x e t_y correspondem aos fatores de traslação atribuídos em x e y respectivamente. A equação anterior pode ser escrita de forma mais concisa como:

$$X' = H_s X = \begin{bmatrix} sR & t \\ 0^T & 1 \end{bmatrix} X \tag{2.17}$$

  O termo escalar s, representa o valor de escala atribuído na transformação. Este tipo de transformação possui quatro graus de liberdade, dos quais um é para a rotação, dois graus para a translação e um para a escala. Assim, a transformação de similaridade pode ser calculada a partir de dois pares de pontos correspondentes entre imagens.

- Transformação Affim:

  Uma transformação Affim é uma transformação linear não-singular acompanhada de uma translação t_x e t_y, e está representada matricialmente como:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{2.18}$$

  onde a versão simplificada é escrita como:

$$X' = H_A X = \begin{bmatrix} A & t \\ 0^T & 1 \end{bmatrix} X \tag{2.19}$$

  O termo A é uma matriz não-singular com dimensões 2×2. A transformação Affim tem seis graus de liberdade e pode ser calculada a partir de três pares de pontos correspondentes entre imagens.

- Transformação Projetiva:

  A transformação projetiva e uma transformação linear não-singular de coordenadas homogêneas, e está matricialmente representada como:

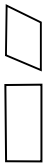
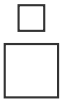
$$X' = H_p X = \begin{bmatrix} A & t \\ V^T & v \end{bmatrix} X \tag{2.20}$$

  onde o termo $V = (v_1, v_2)^T$. A transformação projetiva possui nove graus de liberdade e pode ser calculada a partir de quatro pares de pontos correspondentes entre imagens. Uma transformação projetiva pode ser decomposta em uma cadeia de transformações de uma hierarquia menor, onde a transformação projetiva representa a transformação mais alta e completa dentro das demais transformações anteriormente mencionadas.

$$H = H_s H_A H_p = \begin{bmatrix} sR & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} K & 0 \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} I & 0 \\ V^T & v \end{bmatrix} = \begin{bmatrix} A & t \\ V^T & v \end{bmatrix} \tag{2.21}$$

  Na tabela 2.1, mostra-se um resumo com os diferentes tipos de transformações geométricas que podem ser aplicados para a correção de distorções geométricas em imagens aéreas.

Tabela 2.1: Resumo dos diferentes tipos de transformações geométricas

Transformação	Matriz	Distorção	Propriedades
Projetiva 8° liberdade	$\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$		As linhas paralelas das figuras se cruzam em um ponto no infinito.
Affim 6° liberdade	$\begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$		Paralelismo, as linhas paralelas das figuras não se cruzam no infinito.
Similaridade 4° liberdade	$\begin{bmatrix} sr_{11} & sr_{12} & t_x \\ sr_{21} & sr_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$		Invariantes ao comprimento, ângulos, razão de áreas, linhas paralelas.
Euclidiana 3° liberdade	$\begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$		Invariantes ao comprimento, ângulo, área.

2.2.4 Transformação Linear Direta - DLT

Este algoritmo é utilizado para resolver a matriz homográfica H dado um conjunto suficiente de pontos correspondentes. A equação pode ser expressada em termos do produto de vetores cruzados como $x'_i \times Hx_i = 0$. Esta forma permitirá uma solução simples para H. Esse produto cruzado pode ser definido de forma explicita como (2.22) :

$$x'_i \times Hx_i = \begin{pmatrix} y'_i h^{3T} x_i - w'_i h^{2T} x_i \\ w'_i h^{1T} x_i - x'_i h^{3T} x_i \\ x'_i h^{2T} x_i - y'_i h^{1T} x_i \end{pmatrix} \tag{2.22}$$

Desde $h^{jT} x_i = x_i^T h^j$ para $j = 1, ..., 3$, é definido um conjunto de três equações, que pode ser escrita na forma (2.23):

$$\begin{bmatrix} 0^T & -w_i' X_i^T & y_i' X_i^T \\ w_i' X_i^T & 0^T & -x_i' X_i^T \\ -y_i' X_i T & x_i' X_i^T & 0^T \end{bmatrix} \begin{pmatrix} h^1 \\ h^2 \\ h^3 \end{pmatrix} = 0 \tag{2.23}$$

A equação 2.23 segue o modelo $A_i h = 0$, onde A_i é uma matriz 3×9 e h é um vetor de 9 elementos que constituem as entradas da matriz H. Este modelo representa uma equação linear com o h desconhecido, e os elementos da matriz A são quadráticos nas coordenadas dos pontos conhecidos. Assim, cada correspondência de pontos fornece duas equações nas entradas de H e com isso é possível omitir a terceira equação na busca da solução de H. Portanto, o conjunto de equações pode ser descrito por (2.24):

$$\begin{bmatrix} 0^T & -w_i' X_i^T & y_i' X_i^T \\ w_i' X_i^T & 0^T & -x_i' X_i^T \end{bmatrix} \begin{pmatrix} h^1 \\ h^2 \\ h^3 \end{pmatrix} = 0 \tag{2.24}$$

Onde h^j representa a j_{th} coluna de H, e X_i^T representa um ponto (x_i, y_i, w_i) da imagem georeferenciada, enquanto o ponto (x_i', y_i', w_i') representa um ponto correspondente a imagem capturada pelo VANT. A Eq.(2.24) pode ser expressada da seguinte forma:

$$A_i h = 0 \tag{2.25}$$

Onde A_i representa o i_{th} ponto de correspondências, e é representada por uma nova matriz 2×9 no caso de quatro pontos de controle $A_1 - A_4$ (Hartley and Zisserman, 2000). O Algoritmo 1 representa o funcionamento do DLT.

2.3 SURF

O algoritmo SURF (Speeded-up Robust Transform) (Bay et al., 2008) é um dos mais utilizados na literatura para a extração de pontos característicos invariantes em imagens. A obtenção destes pontos característicos é realizada através da detecção de possíveis pontos de interesse e sua respectiva localização dentro da imagem. Depois, a vizinhança do ponto de interesse é representada como um vetor descritor, geralmente com um tamanho de 64 valores. Este algoritmo além de ser suficientemente robusto e preciso na extração de características, reduz o tempo de processamento utilizando a matriz Hessiana (Neubeck and Van Gool, 2006). O algoritmo SURF está descrito nos seguintes módulos:

Algoritmo 1 Algoritmo DLT

Dados: Dado um número de pontos correspondentes tal que: $n \geq 4$, $(x_i \leftrightarrow x_i')$; determinar a matriz homográfica H tal que, $x_i' \times Hx_i$.

início

(i) Para cada par de pontos correspondentes $(x_i \leftrightarrow x_i')$ calcular a matriz A_i baseado em Eq.(2.24).

(ii) Juntar as n 2×9 matrizes de A_i em uma única matriz A de $2n \times 9$.

(iii) Obter o SVD (Singular Value Decomposition). O vetor singular unitário correspondente ao menor valor singular, é a solução h.

(iii) A matriz H é determinada a partir do resultado de h, onde:

$$\text{h}=\begin{pmatrix} h^1 \\ h^2 \\ h^3 \end{pmatrix} \Longrightarrow H = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \quad (2.26)$$

fim

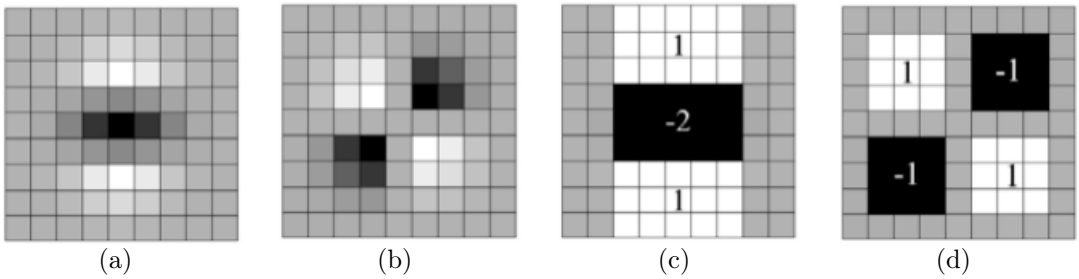
2.3.1 Detecção de Pontos de Interesse

Aqui, a imagem é decomposta em partes integrais (Crow, 1984) para calcular mais rápido os filtros de convolução tipo caixa. O tempo de processamento destas imagens é independente de seu tamanho. É utilizado o determinante da matriz Hessiana para localizar os pontos de interesse e definir a escala. Dado um ponto P em uma imagem I, a matriz Hessiana $H = (x, \sigma_s)$ em P com escala σ_s, é definido como:

$$H(x, \sigma_s) = \begin{bmatrix} L_{xx}(x, \sigma_s) & L_{xy}(x, \sigma_s) \\ L_{xy}(x, \sigma_s) & L_{yy}(x, \sigma_s) \end{bmatrix} \quad (2.27)$$

Onde $L_{xx}(x, \sigma_s)$ é a convolução da derivada da segunda ordem da Gaussiana $\frac{\partial^2}{\partial x^2} g(\sigma_s)$ com a imagem I em um ponto P. De forma similar para $L_{x,y}(x, y, \sigma_s)$ e $L_{y,y}(x, y, \sigma_s)$, as Gaussianas que são ótimas para determinar a escala são discretizadas e recortadas como é mostrado na figura 2.12. De modo que a aproximação da matriz Hessiana utilizando os filtros tipo caixa são mostrados nas figuras 2.12(c), 2.12(d). O determinante Hessiano representa a resposta BLOB (Binary Large OBject - Objeto Grade Binário) das imagens em diferentes escalas (Bay et al., 2008) e é definido como (2.28):

$$det(H_{aprox}) = D_{xx}D_{yy} - (wD_{xy})^2 \quad (2.28)$$



(a) (b) (c) (d)

Figura 2.12: Filtro tipo caixa. Derivada parcial da segunda ordem da função gaussiana em imagens discretas na direção $y-(L_{yy})$ (2.12(a)) e na direção $xy-(L_{xy})$ (2.12(b)). A Figura 2.12(c) e 2.12(d) representam a aproximação com o filtro tipo caixa das Figuras 2.12(a) 2.12(b).

Onde w é um peso atribuído, cuja função é conservar a energia entre o kernel gaussiano e sua aproximação (Bay et al., 2008). Como o filtro tipo caixa faz sua aproximação gaussiana com uma dimensão de 9×9 e $\sigma = 1,2$, isto representa a escala mais baixa ou seja, a resolução mais alta para fazer o cálculo da resposta BLOB. Assim o peso relativo w é determinado como se mostra na equação (2.29).

$$w = \frac{|L_{xy}(1.2)|_F \, |D_{yy}(9)|_F}{|L_{yy}(1.2)|_F \, |D_{xy}(9)|_F} = 0.912... \cong 0.9 \tag{2.29}$$

O espaço-escala implementado em SURF é definido por uma pirâmide de imagens nas quais foram aplicados filtros tipo caixa de diferentes tamanhos e sub-amostrados, para assim obter o maior nível da pirâmide (Ali and Hussain, 2012). Redimensionar os tamanhos dos filtros no lugar da imagem melhora o tempo de processamento e evita o *aliasing* nas imagens. Além disso, predominam os componentes de alta frequência que possuem uma melhor invariância de escala. A subtração entre camadas nas piramides de imagens faz a obtenção de DoG (Diferença de Gaussiana) onde as bordas e os BLOBs podem ser obtidos. Na figura 2.13 é mostrado a diferença entre mudar a escala da imagem (esquerda) e mudar o tamanho do filtro (direita).

2.3.2 Localização de Pontos de Interesse

Os pontos de interesse são localizados através da supressão não-máxima feita nas diferentes escalas da imagem (Neubeck and Van Gool, 2006). Para obter pontos de interesse que resultem ser invariantes à orientação, é calculada a $Haar-Wavelet$ nas direções x e y numa região circular com raio $6s$, onde s é a escala do ponto de interesse. São utilizados dois filtros para calcular a resposta na direção x ou y em qualquer escala. A figura 2.14 mostra o formato dos filtros $Haar-Wavelet$ onde as partes escuras tem um peso de -1 e as partes claras tem um peso $+1$.

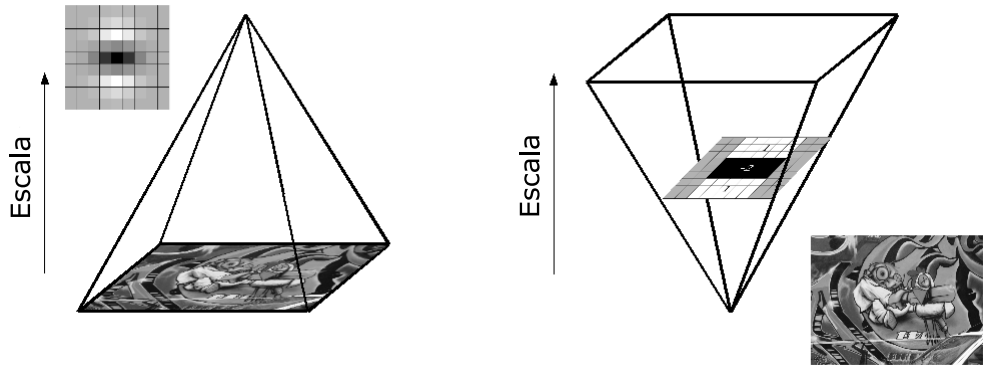


Figura 2.13: Espaço-escala implementado em SURF.

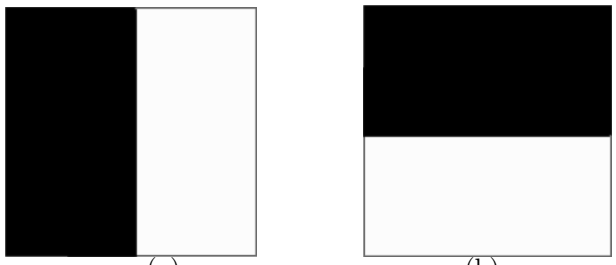


Figura 2.14: Filtros $Haar - wavelet$ para obter as respostas em x 2.14(a) e y 2.14(b).

Após de ser calculado para todos os vizinhos, a orientação dominante é obtida pela soma de todos os resultados contidos dentro de uma janela deslizante com ângulo de $\pi/3$ (Romero Cortijo and Cazorla Quevedo, 2009).

Uma vez definida a região de orientação, esta é dividida em 4 sub-regiões de tamanho $4x4$. Posteriormente, a $Haar-Wavelet$ é calculada para cada sub-região preservando a informação espacial. Por último, são suavizados os resultados através de uma Gaussiana ($\sigma = 3.3s$), obtendo d_x e d_y. Para cada sub-região são adicionados os resultados d_x, d_y e seus respectivos valores absolutos $|d_x|$,$|d_y|$. Dessa forma, cada sub-região fornece um vetor $v = (\sum d_x, \sum d_y, \sum |\mathrm{d}_x|, \sum |\mathrm{d}_y|)$. De modo que, concatenando este vetor para todas as sub-regioes $4x$ resultará um vetor descritor de comprimento igual a 64, assim o descritor SURF é obtido pela união dos vetores v das sub-regiões. A figura 2.15 mostra as propriedades do descritor para três padrões de imagem com intensidades diferentes dentro de uma sub-região. As respostas do descritor são invariantes a iluminação e ao contraste. Quando existe uma região muito homogênea os valores do descritor acostumam ser baixos, se a intensidade aumenta gradualmente na direção x os valores do descritor resultam ser elevados.

O casamento de pontos característicos extraídos com SURF é realizado utilizando a matriz Hessiana quem determina o signo do Laplaciano. Isto oferece uma indexação

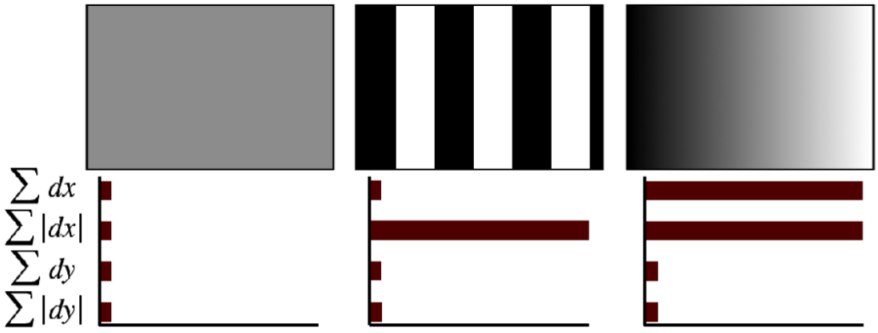


Figura 2.15: Comportamento do descritor para três padrões de imagem com intensidades diferentes dentro de uma sub-região.

mais rápida dos pontos de interesse. O casamento de um par de pontos característicos só é descartado quando existe entre eles uma diferença significativa no contraste e quando os pontos característicos candidatos que possuem o mesmo brilho estão já casados. A figura 2.16 apresenta a forma como é rejeitado um casamento de dois pontos característicos com diferente contraste.

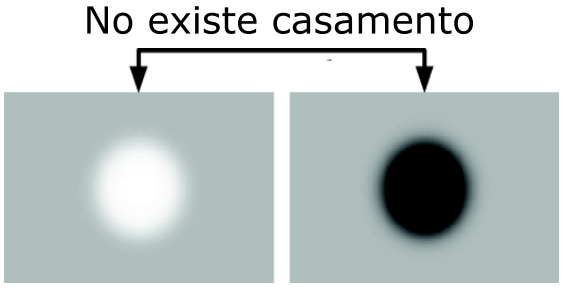


Figura 2.16: Não casamento de pontos característicos devido à diferenças de contraste.

2.3.3 Trabalhos Relacionados

Na literatura, a correção de perspectiva tem sido amplamente abordada em diferentes tipos de aplicações como é descrito em Silva (2006). Esta correção é aplicada em imagens de prédios históricos com a finalidade de conservar e realizar futuras restaurações neles. Outra abordagem é discutida em Wang and Yang (2011) onde é proposto um método de correção geométrica polinomial por partes sobre imagens aéreas com amplo ângulo de visão e uma distorção geométrica grande. Em Aracena Pizarro et al. (2005), os autores apresentam uma comparação das diferentes técnicas ou metodologias de calibração de câmeras, uma delas desenvolvida por Tuceryan et al. (1995); Trucco and Verri (1998) e a outra através de homografia proposta em Zhang et al. (1995). Em Li et al. (2011) os autores propõem um algoritmo que utiliza a informação dos parâmetros de posição de voo (atitude) para fazer uma rápida correção da distorção geométrica nas

imagens capturadas pelo VANT. Portanto, para assegurar corretas estimativas de posição em VANTs via Casamento de Imagens é essencial que ambas imagens (câmera do VANT e imagem georeferenciada) tenham visada nadir (Goltz and Shiguemori, 2008).

Diferentes abordagens e metodologias têm sido desenvolvidas e implementadas para lidar com problemas de diferença espectral e distorção de perspectiva em imagens capturadas por VANTs e imagens de satélite. Por exemplo, Sedaghat et al. (2011) propuseram um algoritmo de casamento totalmente automatizado, chamado Uniform Robust SIFT (UR-SIFT). UR-SIFT é aplicável às diferentes imagens ópticas de sensoreamento remoto com diferente iluminação, rotação e até cinco vezes em diferença da escala. Yu et al. (2008) propõem uma técnica de registro de imagens não-rígida rápida e totalmente automático aplicado em imagens de sensoreamento remoto. Estas imagens possuem variação de iluminação e resolução, diferentes perspectivas e deformações locais. A técnica realiza um processo de pré-registro usando SIFT (Scale Invariant Feature Transform) e um modelo de transformação afim. Então, *Harris corners* e *cross-matching* são usados para fazer o casamento de pares de pontos entre as imagens de referência e de entrada. Também é usada a TINs (Triangulated Irregular Networks) com a finalidade de corrigir as deformações na imagem de entrada.

Em outra abordagem Yasein and Agathoklis (2007), um algoritmo para registrar imagens aéreas tendo apenas uma sobreposição parcial e com presença de diferentes distorções geométricas (rotação, translação e escala) devido à condições de detecção diferentes, ruído e *blurred* é proposto. O algoritmo utiliza *mexican-hat wavelets* para realizar a extração de pontos característicos, *zernike moments* para estabelecer os casamentos entre as duas imagens e o *iterative weighted least squares* para estimar os parâmetros de transformação. Em outra abordagem Fan et al. (2011) descreve um novo método de registro de imagens de sensoriamento remoto em casos de disparidade no espectro. O método usa o algoritmo descritor de características SIFT e correlação local baseada na congruência de fase para se obter um conjunto de pontos característicos com melhor precisão. Em Huang et al. (2012) um sistema de registro de imagens provenientes de um VANT foi desenvolvido. Este sistema utiliza SIFT, RANSAC e também uma variação da normalização de informação mutua para refinar o registro entre imagens capturadas por VANT e imagens satélite do Google Maps. Os autores Hong et al. (2013) apresentam um novo algoritmo para mosaico de imagens baseado no algoritmo de *Harris corner detection*, onde os cantos das imagens aéreas são detectados usando *Harris*, em seguida, o registro de características de correlação é usado para encontrar os pares de registros. Por último, os casamentos são filtrados usando o algoritmo de registro com restrição e o método gradual *fade-out* para a fusão das imagens.

Nam and Phillips (2013) introduzem um novo método para fazer registo de ima-

gens provenientes de vários sensores. O método estabelece correspondências de marcos quando existem muitos valores *outliers* devido à diferentes aparências de imagem em imagens multi-modais. O método de casamento proposto nessa pesquisa, aproveita a ordem espacial de detecção de marcos que é preservada, apesar da deformação não-linear nos dados de sensoreamento remoto, o que faz que o algoritmo de casamento seja robusto a *outliers* quando são analisados pares de imagens de sensoreamento remoto reais. Em Salehpour et al. (2014), um novo método foi proposto para fornecer um algoritmo fiável e preciso para buscar e identificar o centro de uma imagem de referência dentro de um conjunto de quadros de vídeo. Utilizando os algoritmos SIFT e PCA-SIFT, alguns casamentos de pontos são obtidos. Além disso, a proposta utiliza o método de teste recursivo em variança e um modelo projetivo preciso, a fim de detectar e eliminar os casamentos com *outliers*. Zheng et al. (2014a) apresentam um descritor de características afim e invariante a escala SIFT para fazer casamento entre imagens capturadas por VANTs e imagens de satélite. Os algoritmos SIFT e ASIFT foram testados e foi introduzido um novo descritor baseado em regiões com forma de pizza. Este descritor é invariante à rotação, afim, escala e quando a dimensão do vetor de características é reduzido. Portanto, o método que foi proposto satisfaze a robustez desejada e possuí uma baixa complexidade computacional. Finalmente Aouf (2012), realizaram uma análise completa dos algoritmos SIFT e SURF usados para a detecção de pontos característicos. Além disso, foram aplicadas técnicas de casamento em imagens aéreas e imagens infravermelho. Também, o algoritmo RANSAC é utilizado a fim de realizar a rejeição de *outliers* encontrados.

Capítulo 3

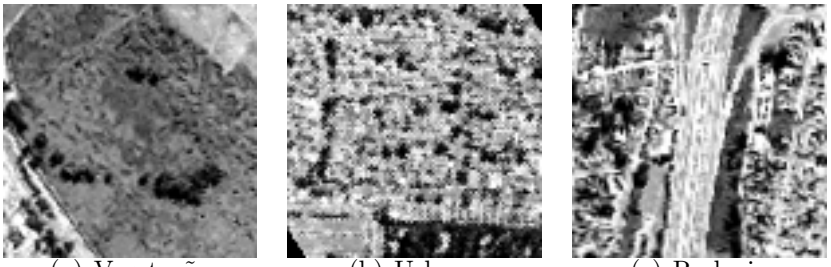
Métodos de Estimação de Posição em VANTs via Casamento de Imagens

Neste Capítulo são mostrados dois novos métodos que contribuem no processo de estimação de posição geográfica em VANTs utilizando Cansamento de imagens. Assim, é possível observar o funcionamento do método Detector de Bordas *canny* Adaptativo descrito na Seção 3.1 e o funcionamento do método de limiarização de imagens descrito na Seção 3.2. Finalmente, na Seção 3.3 é apresentado um resumo do Capítulo.

3.1 Detector de Bordas Canny Adaptativo

A detecção de bordas apresenta uma solução robusta e eficiente a mudanças de luminosidade entre a imagem capturada pelo VANT e a imagem georeferenciada. Porém, resulta difícil encontrar o melhor desempenho para todo tipo de imagens aéreas analisadas. Isso se justifica pelo fato de que a detecção e extração de bordas possui diferentes desempenhos que vão de acordo com cada tipo de imagem analisada. Portanto, generalizar um detector de bordas e fixar seus parâmetros em um só valor vai ter como resultado um alto desempenho apenas para um tipo de imagem específica. Dificilmente vai-se poder abarcar todo tipo de imagem com uma parametrização só do algoritmo detector de bordas. As imagens aéreas e de satélite podem ser classificadas em três classes, dependendo da área sobrevoada que possua maior predominância na imagem: Vegetação, Rodovia e Urbano (Silva et al., 2015). A classe Vegetação é caracterizada por ser imagens com predominância de áreas verdes que resultam ser muito homogêneas, como cultivos, plantações e mato. A classe Urbano é caracterizada por ter grandes variações na escala de cinza e está constituída por áreas mais estruturadas como, casas, edificações, prédios, ruas pequenas, piscinas, etc. Por sua parte, a classe Rodovia é caracterizada por possuir rodovias que se destacam dentro e fora das cida-

des como, grandes rodovias, viadutos, pontes, pista de decolagem no aeroporto, etc. A figura 3.1 mostra exemplos de imagens aéreas capturadas por VANT para cada uma das classes anteriormente mencionadas.



(a) Vegetação (b) Urbano (c) Rodovia

Figura 3.1: Imagen capturadas pelo VANT: (a) Classe Vegetação, (b) Classe Urbano, (c) Classe Rodovia.

De modo que um algoritmo detector de bordas pode extrair uma quantidade razoável de bordas sobre uma classe e ao mesmo tempo gerar sobrestimação de bordas ou um número de bordas insuficientes sobre outra classe. Isto ocasiona uma interpretação deficiente da área sobrevoada e por conseguinte, a estimação de posição via Casamento de Imagens será afetada. A figura 3.2 mostra um exemplo de detecção e extração de bordas com o algoritmo *canny* para cada uma das imagens de exemplo em cada classe (fig. 3.1).



(a) Vegetação (b) Urbano (c) Rodovia

Figura 3.2: Detecção de bordas com *canny*: (a) Classe Vegetação, (b) Classe Urbano, (c) Classe Rodovia.

Como pode ser apreciado na figura 3.2(a), para a classe Vegetação existe uma sobrestimação de bordas que pode causar um cálculo errado da estimação de posição quando se procura a imagem do VANT sobre a imagem de satélite utilizando Casamento de Imagens. A figura 3.3 mostra a detecção de bordas para a imagem da classe Vegetação com dimensão de 95×95 pixeis e a imagem de satélite com dimensão de 400×400 pixeis respectivamente, também, é mostrada a matriz de correlação quando é feito o Casamento de Imagens, a posição calculada (quadrado azul) e a posição verdadeira (quadrado vermelho). É evidenciado que, devido à sobrestimação de bordas

nessa imagem de classe Vegetação, a estimação de posição fornece como resultado uma posição diferente à verdadeira.

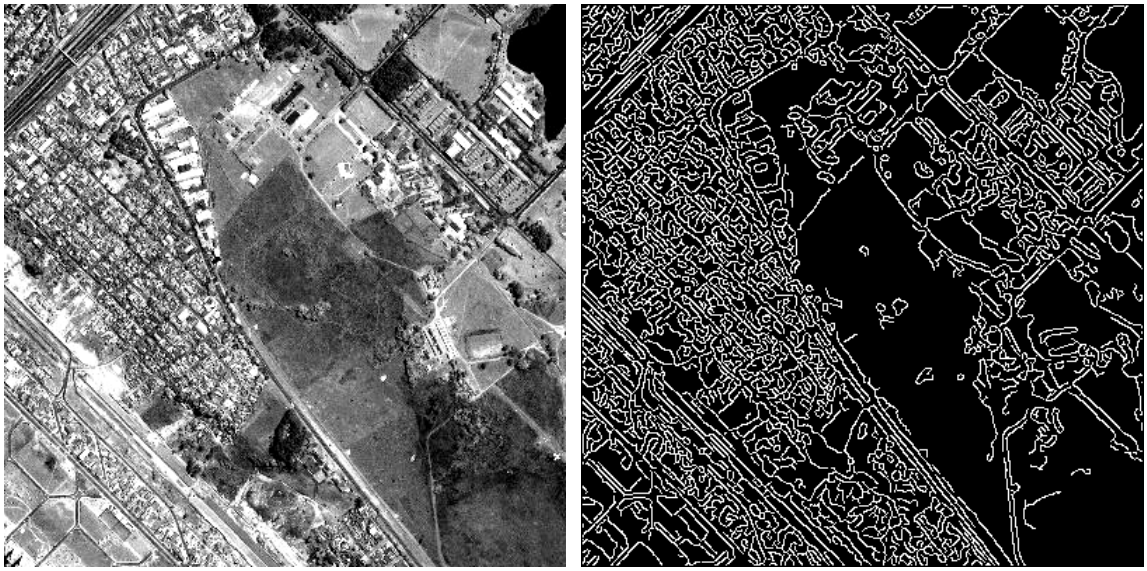


Imagem satélite em tons de cinza (esquerda), detecção de bordas (direita)

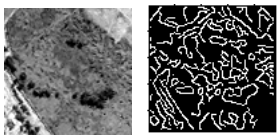
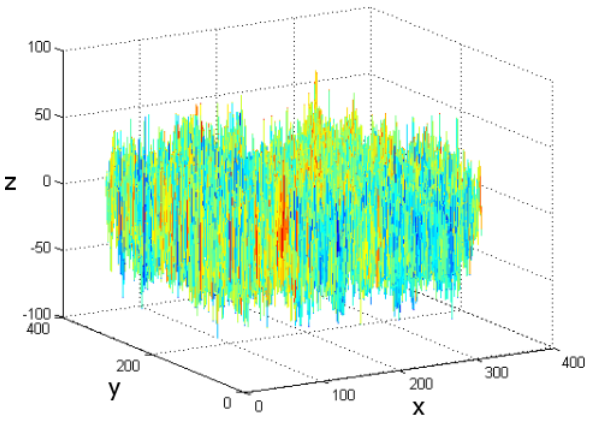
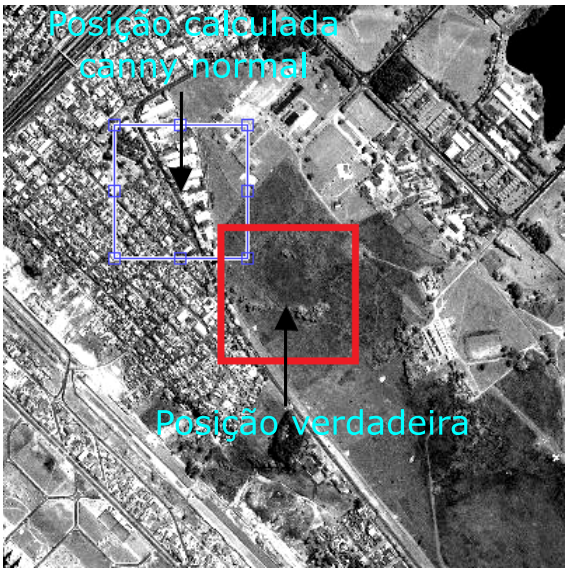


Imagem capturada pelo VANT em tons de cinza (esquerda), detecção de bordas (direita)





Matriz de correlação (esquerda) e resultado da estimação de posição (direita)

Figura 3.3: Resultado da estimação de posição quando existe sobrestimação de bordas em imagem de classe Vegetação (fig 3.2(a)).

Em geral, a classe Vegetação possui imagens que se caracterizam por ser muito

homogêneas e por não possuir grandes variações nos níveis de cinza. Quando ocorre sobrestimação de bordas, são geradas bordas em excesso que não deveriam se apresentar na imagem. Tudo isso é produto de uma alta sensibilidade atribuída ao algoritmo detector de bordas *canny* que pode beneficiar a uma região mais estruturada (classe Urbana), mas acaba afetando à classe Vegetação.

Considerando os problemas anteriormente mencionados, é proposto um novo processo de detecção de bordas adaptativo baseado no algoritmo *canny* que melhora consideravelmente o problema de sobrestimação de bordas nas imagens aéreas e de satélite. Este método faz uma análise estatística para cada imagem capturada pelo VANT a fim de determinar a quantidade de classe Vegetação que a imagem possui. Então, dependendo da quantidade de regiões homogêneas presentes na imagem é calculado um limiar específico. Com essa informação é ajustada a sensibilidade do detector de bordas *canny*, permitindo assim, fazer uma extração de bordas mais ajustável a cada tipo de classe e garantir que não exista sobrestimação de bordas.

Inicialmente sobre a imagem capturada pelo VANT (*template*) é selecionada uma sub-janela com tamanho fixo de 30×30 pixeis. Este tamanho é atribuído como *tamanho template/3*, levando a consideração que o tamanho padrão da imagem *template* utilizado no simulador de voo NAV (Santos et al., 2013; Silva et al., 2015) é de 95×95 pixeis. Desta forma, a sub-janela analisada corresponde a uma terceira parte do tamanho da imagem *template*. Se o tamanho da imagem *template* muda, o tamanho da sub-janela mudará em proporção a ela.

Sobre toda a imagem *template* vai-se deslizar essa sub-janela partindo da parte superior esquerda percorrendo de esquerda a direita e descendo uma fila por vez. Em cada percurso é calculado o desvio padrão e guardado em uma matriz global que armazenara todos os desvios padrão de cada iteração. Este processo tem a intenção de identificar quais regiões contidas dentro da janela avaliada são bem homogêneas indicando presença de Vegetação. Um desvio padrão baixo indica uma região contínua sem mudanças na escala de cinza, comportamento que é somente atribuído a uma região de mato ou Vegetação. Depois, na matriz global resultante, cada coluna é ordenada de menor a maior e é calculada a média. A figura 3.4 mostra o comportamento dos desvios padrão obtidos na imagem Vegetação 3.1(a) (linhas azuis), e a média dos desvios padrão (linha vermelha). Com estas informações, é possível apreciar e identificar se a imagem possui índices de desvio padrão baixo ou alto, o que representa uma maior o menor presença de regiões homogêneas, respectivamente.

Cada um dos resultados dos desvios padrão obtidos representam uma sub-janela de 30×30 pixeis. De modo que são classificados e selecionados os menores desvios padrão por cada coluna formando um vetor resultante de 1×65, isto é devido a que o número de percursos que a sub-janela faz por cada coluna sobre a imagem *template* é

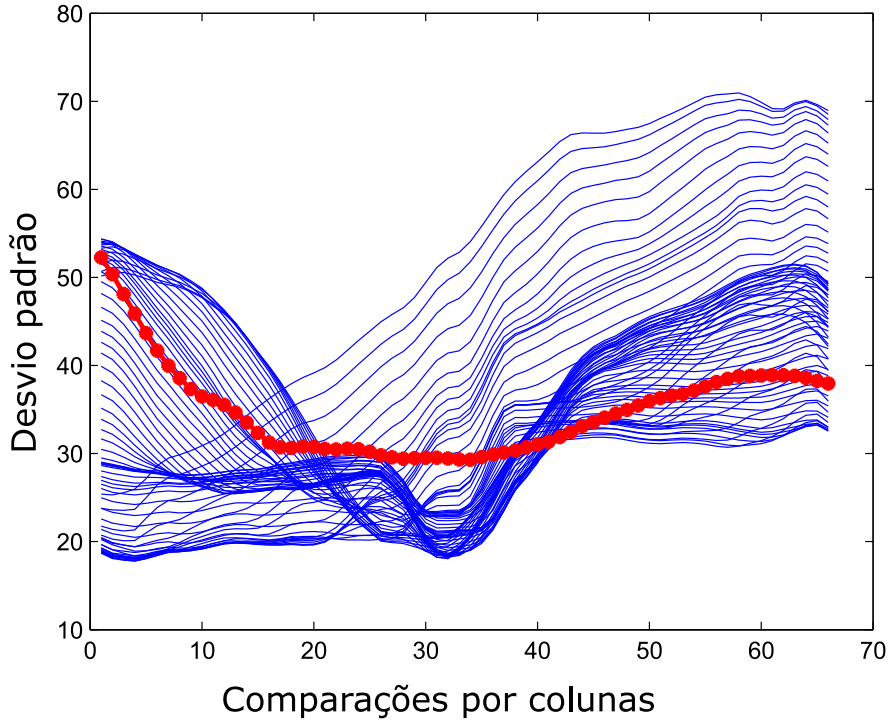


Figura 3.4: Desvio padrão da imagem.

de 65 vezes. Assim, são selecionadas as sub-regiões mais homogêneas dentro da imagem *template*. Por cada janela classificada, é calculada a diferença dos pixeis entre colunas de cada sub-imagem. Desta forma, se o valor de um pixel x_i, y_i na coluna C_i possui um valor muito próximo ao pixel x_i, y_{i+1} na coluna C_{i+1}, a diferença será próxima de zero, indicando uma região homogênea. Passo seguinte, é calcular a média do desvio padrão (para uma distribuição com média em zero) na sub-imagem.

A figura 3.5 representa o histograma de uma sub-imagem analisada onde é observável um maior número de valores próximos a zero, produto da diferença dos pixeis entre colunas feita com anterioridade.

Os resultados geram então um novo vetor de 1×65 que contém os melhores limiares (sub-janelas mais homogêneas que foram encontradas) calculados para a imagem *template*. Como existe uma variação nos limiares, é calculada a média deles, obtendo um limiar final resultante que será o responsável de variar a sensibilidade do algoritmo detector de borda *canny*. A figura 3.6, mostra a variação dos limiares obtidos no vetor resultante (linha azul) e a média calculada para escolher o limiar final (linha vermelha).

O algoritmo detector de bordas *canny* possui dois parâmetros principais que podem ser variados para obter resultados de bordas diferentes, o *threshold* e o *sigma*. Variando o parâmetro *threshold*, serão calculadas e extraídas todas as bordas que estejam acima do valor do *threshold* definido. O parâmetro *sigma* especifica o desvio padrão do filtro Gaussiano que aplica o detector *canny* para suavizar a imagem (Eq. 3.1). Este

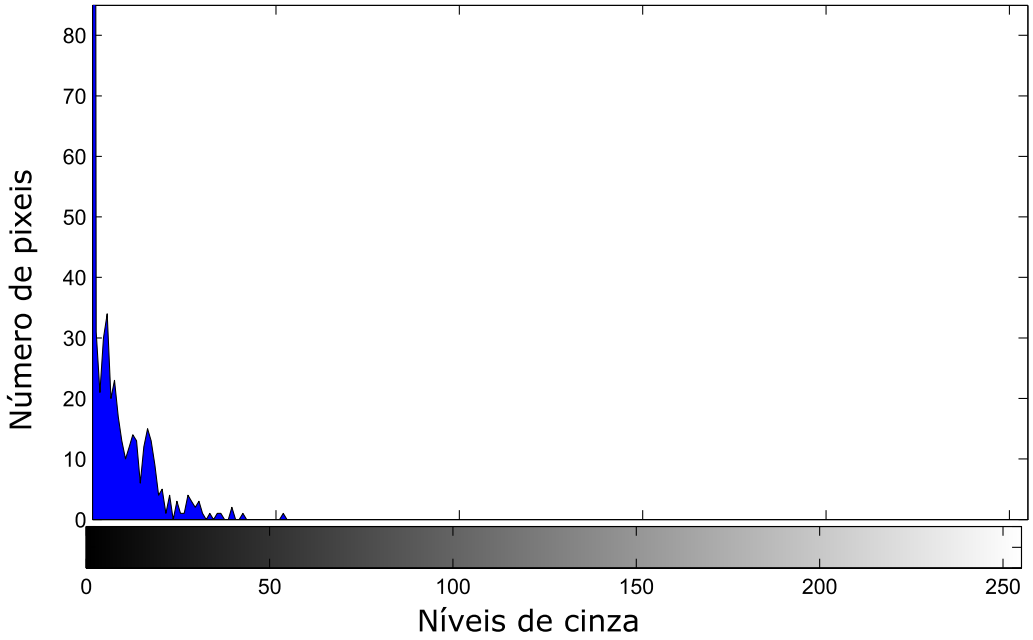


Figura 3.5: Histograma de uma Sub-imagem após de ser calculada a diferença dos pixeis entre colunas.

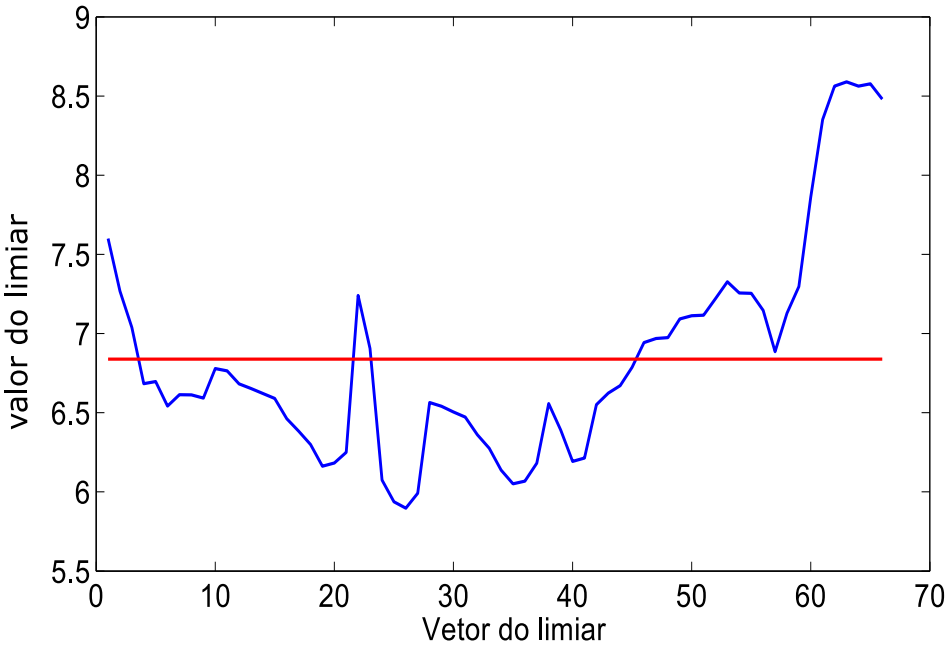
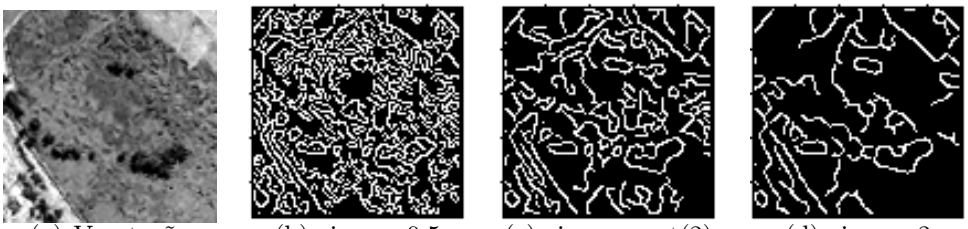


Figura 3.6: Obtenção do limiar final resultante do análise feito sobre a imagem *template*. Variacao do limiar (linha azul), média do limiar (linha vermelha).

parâmetro esta fixado normalmente com um valor de $\sqrt{2}$. Quando o valor atribuído ao *sigma* é grande, menor será a sensibilidade do algoritmo *canny*, e consequentemente a quantidade de bordas extraídas vai ser menor. No caso contrário, quando o valor do *sigma* é menor, a sensibilidade do *canny* será alta e a quantidade de bordas extraídas será bem maior. A figura 3.7 mostra a extração de bordas feita com a algoritmo *canny* sobre a imagem Vegetação (fig. 3.2(a)) para diferentes valores de *sigma* atribuídos.

$$filtro(x,y) = \frac{1}{2\pi\sigma^2} exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \tag{3.1}$$



(a) Vegetação (b) sigma=0.5 (c) sigma=sqrt(2) (d) sigma=2

Figura 3.7: Extração de bordas com *canny* sobre a imagem de classe Vegetação, aplicando diferentes valores de *sigma*.

De modo que, o parâmetro *sigma* é selecionado para trabalhar no método proposto partindo do resultado do limiar final calculado. Assim, é definida mais uma etapa para ajustar o valor do limiar calculado a um valor que seja equivalente ao parâmetro *sigma*. Em outras palavras, um valor grande no limiar indicando uma imagem de classe Urbana ou Rodovia, representará um valor pequeno do parâmetro *sigma* e dessa forma serão melhor detectados os pequenos detalhes presentes na imagem *Template*. Se existe um valor pequeno no limiar indicando uma imagem com classe Vegetação, este representará um valor grande do parâmetro *sigma*, o que permite evitar sobrestimação de bordas que afetem posteriormente o Casamento de Imagens e a estimação de posição.

Portanto, foi definido um conjunto de pontos que obedecem as condições anteriormente mencionadas e com eles foi calculada uma regressão linear. Como resultado final, obtiveram-se todos os valores correspondentes para o *sigma* do filtro Gaussiano quando é variado o valor do limiar. A figura 3.8 mostra a regressão linear, onde a linha azul representa um *sigma* que varia com as mudanças do valor do limiar no intervalo que vai desde $(1, \sqrt{4})$ até $(15, \sqrt{2})$. Este intervalo representa o valor minimo e máximo que pode ser calculado no limiar e seus valores de *sigma* correspondentes. A linha vermelha indica que para limiares maiores de 15 o valor do *sigma* será um valor constante com $\sigma = \sqrt{2}$.

Finalmente, todo o processo de detecção de bordas adaptativo baseado no algoritmo *canny* é aplicado sobre a imagem de classe Vegetação (fig 3.2(a)), onde anteriormente devido à sobrestimação de bordas o cálculo da estimação de posição foi errôneo (ver figura 3.3). A aplicação do método para a imagem de classe Vegetação deu como resultado um limiar com valor de $6,8389$ como pode ser observado na figura 3.6, e corresponde a um $sigma = 1,7557$. Na figura 3.9 é apresentado o resultado do processo de Casamento de Imagens para a imagem de classe Vegetação (fig. 3.2(a)). Os resultados mostram que a detecção e extração de bordas foram diferentes comparados com os

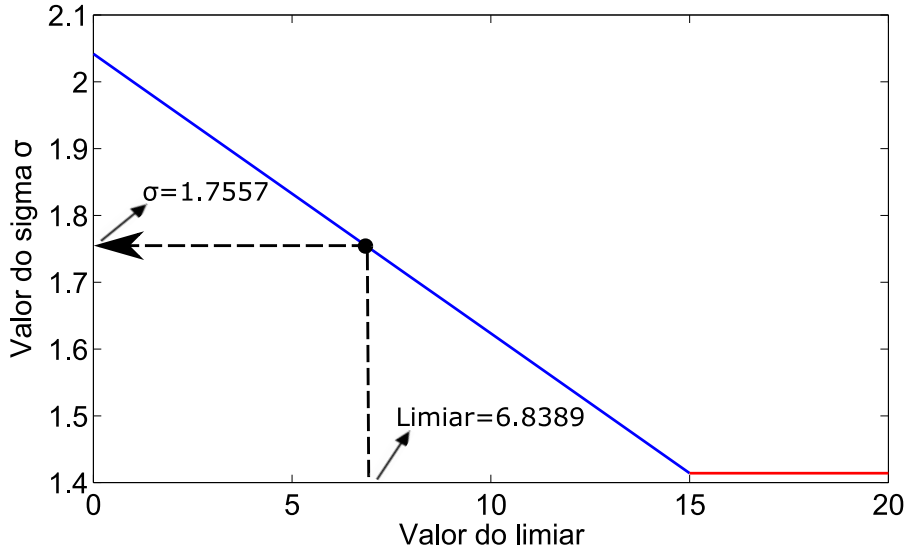


Figura 3.8: Obtenção do *sigma*.

resultados da figura 3.3. A matriz de correlação neste caso apresenta um pico mais dominante e de maior valor. Consequentemente, a estimação de posição para a imagem de teste deu como resultado o valor esperado (quadrado verde), também a maneira de comparação, é mostrado o resultado da estimação de posição obtido anteriormente quando é aplicado o *canny* normal (quadrado azul).

3.2 Cálculo de Correlação por Limiarização de Imagens

Tomando como base os passos feitos no método anterior, é proposto e implementado o cálculo da correlação por limiarização de imagens. Neste caso, toda a análise estatística feita para calcular o limiar que ajusta a borda do algoritmo *canny*, é utilizado neste processo de limiarização pois, o limiar calculado representará um valor de tolerância ao ruído presente na imagem *template*.

Assim, o tamanho definido na sub-janela é de 20×20 pixeis, este tamanho foi reduzido pois pretende-se analisar com maior precisão o ruído presente nas imagens a serem testadas. Esta sub-janela vai percorrer a imagem *template* e vai classificar as regiões mais homogêneas, que correspondem àquelas que apresentem o menor desvio padrão. O vetor resultante dos desvios padrão tem uma dimensão de 1×75, isto porque, a imagem *template* tem dimensões de 95×95 pixeis e o número de percursos que a sub-janela de 20×20 pixeis faz por cada coluna sobre a imagem *template* é de 75 vezes. Cada valor do vetor resultante representa uma sub-janela, e em ela é feita a subtração entre colunas, depois é calculado o desvio padrão da matriz resultante após

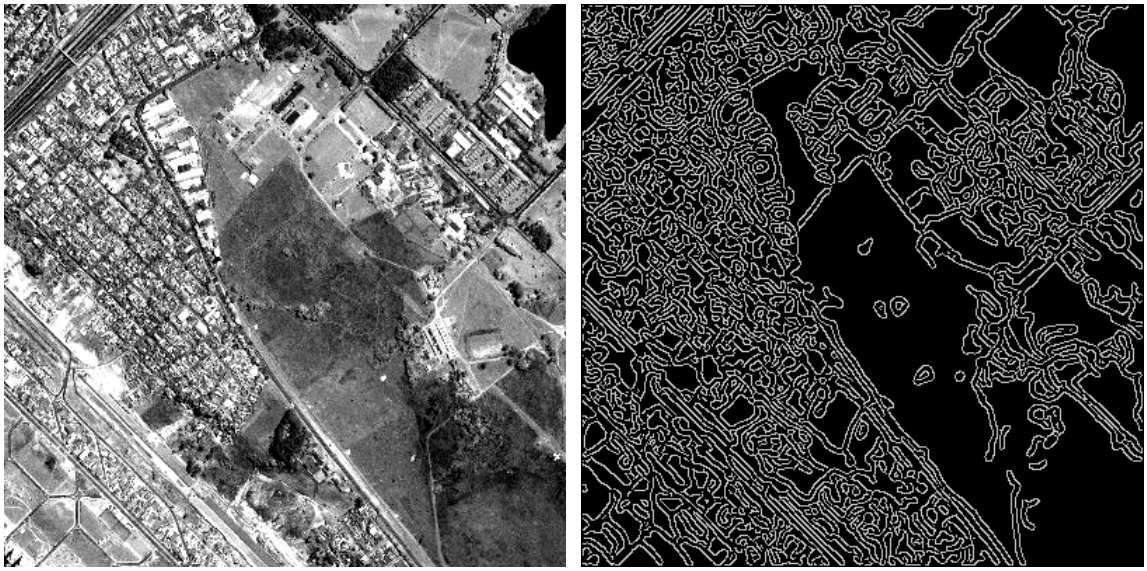


Imagem satélite em tons de cinza (esquerda), detecção de bordas (direita)

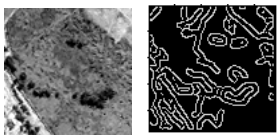
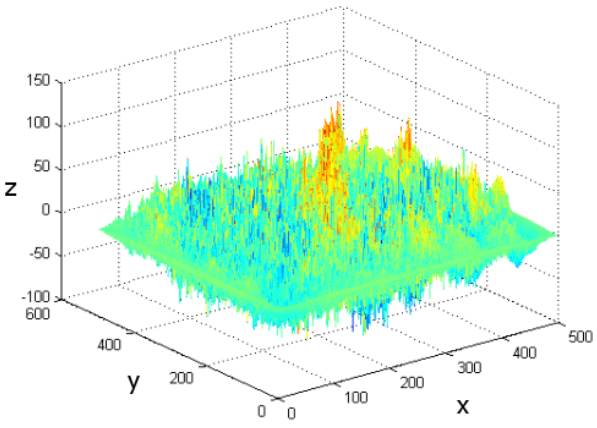
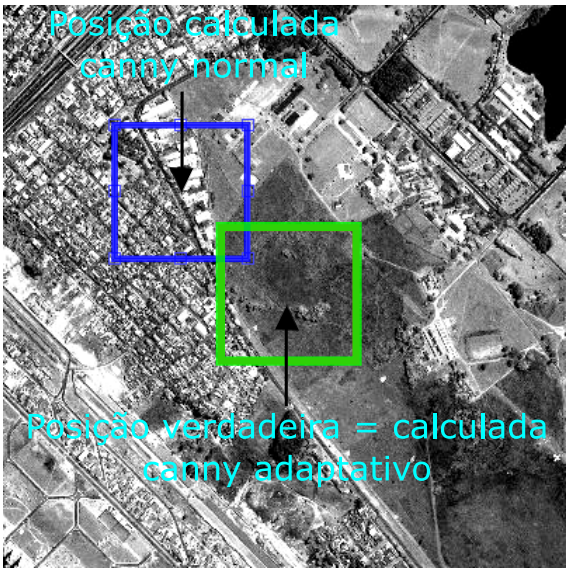


Imagem capturada pelo VANT em tons de cinza (esquerda), detecção de bordas (direita)





Matriz de correlação (esquerda) e resultado da estimação de posição (direita)

Figura 3.9: Resultado da estimação de posição aplicando o método de detecção de bordas com *canny* adaptativo, aplicada em imagem da classe Vegetação (fig 3.2(a)).

a subtração. Finalmente, é estimada a média dos desvios padrão e é multiplicada por 2 para obter a estimativa do limiar α a ser utilizado. Para uma imagem de classe Rodovia, é mostrada na figura 3.10, a variação dos limiares obtidos no vetor resultante

(linha azul) e a média multiplicada por 2 (linha vermelha), que define o valor do limiar final α.

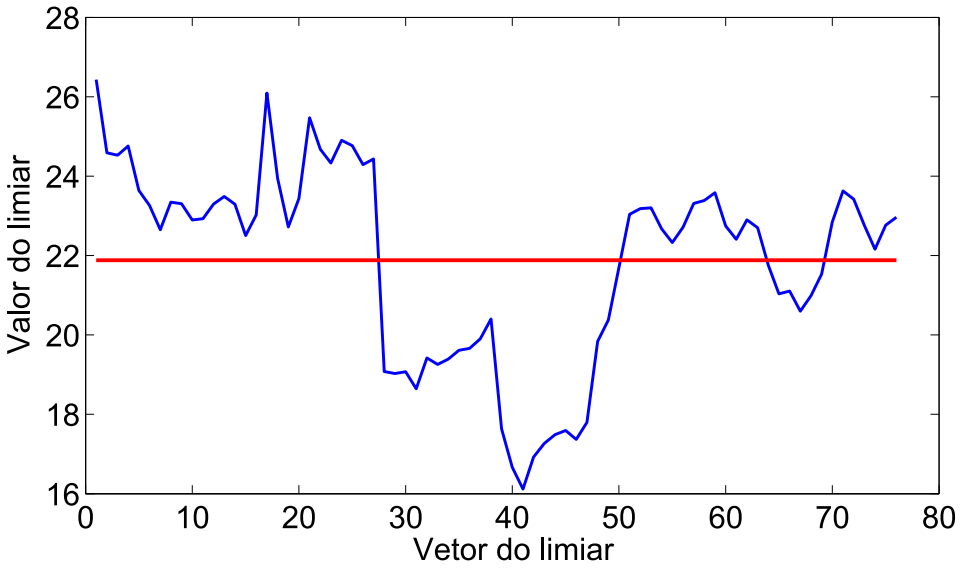


Figura 3.10: Obtenção do limiar α.

Uma vez obtido o valor do limiar α, são definidas umas condições de decisão que definem os valores na limiarização, e são mostradas na equação (3.2).

$$L(x+i, y+i) = \begin{cases} 1, & I(x+i, y+i) > I(x, y) + \alpha \\ 0, & I(x, y) - \alpha \leq I(x+i, y+i) \leq I(x, y) + \alpha \\ -1, & I(x+i, y+i) < I(x, y) - \alpha \end{cases} \tag{3.2}$$

Onde:

- α é o limiar calculado, que define a tolerância ao ruído na imagem.
- $I(x+i, y+i)$ representa o valor em escala de cinza $(0-255)$ do ponto que esta sendo analisado.
- $I(x, y)$ representa o valor em escala de cinza $(0-255)$ do ponto de referência.
- $L(x+i, y+i)$ é a equação de saída da limiarização aplicada nos pixeis da imagem.

A figura 3.11 mostra o processo de cálculo da correlação e estimação de posição por limiarização de imagens, aplicado à imagem de classe Vegetação (fig. 3.2(a)), onde são limiarizadas as imagens de satélite georeferenciada e a capturada pelo VANT. Também é mostrada a matriz de correlação obtida no processo de Casamento de Imagens e, finalmente, é observado a estimação de posição com o valor esperado (quadrado laranja),

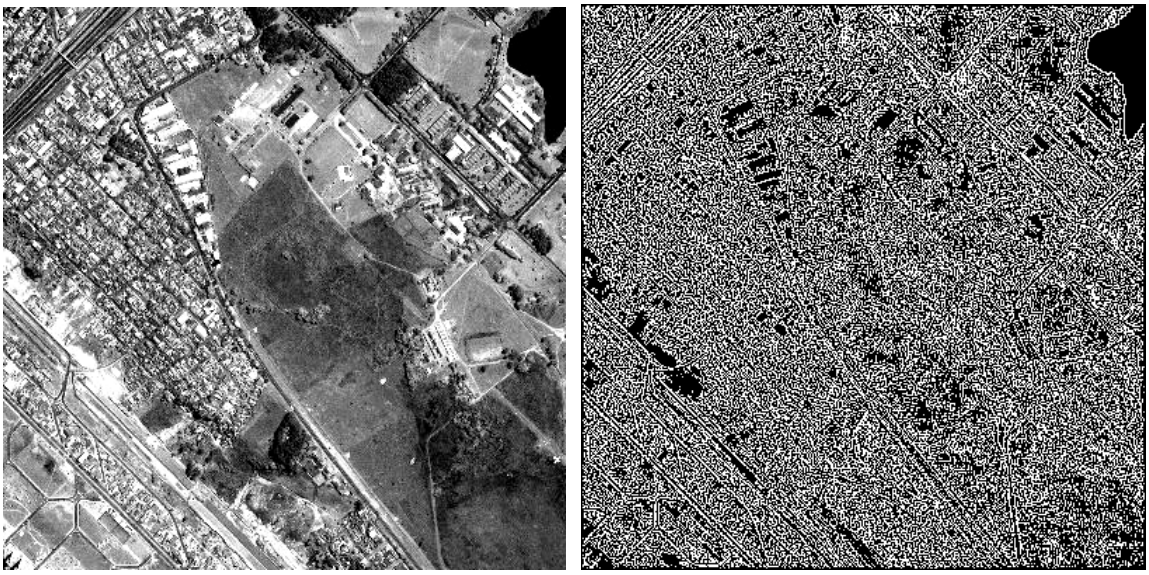


Imagem satélite em tons de cinza (esquerda), limiarização (direita)

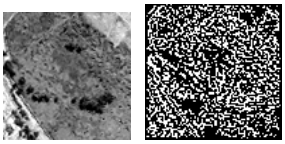
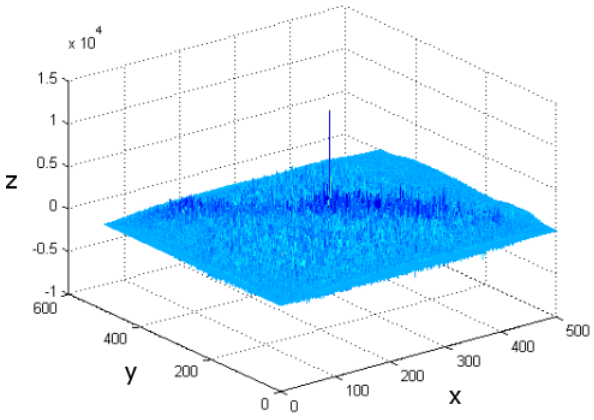
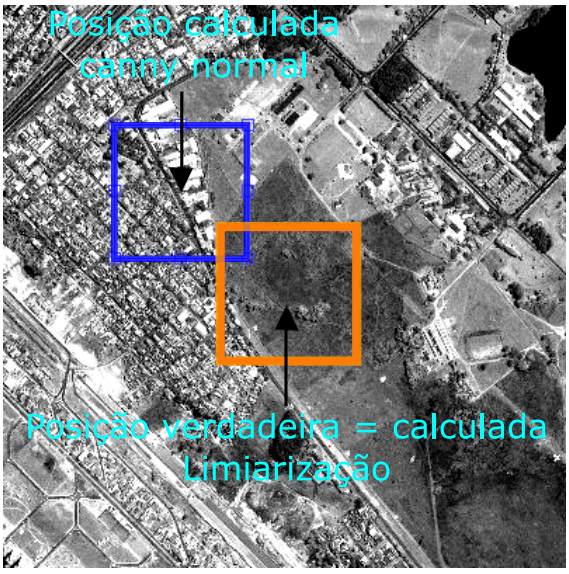


Imagem capturada pelo VANT em tons de cinza (esquerda), limiarização (direita)





Matriz de correlação (esquerda) e resultado da estimação de posição (direita)

Figura 3.11: Resultado da estimação de posição utilizando o método de correlação por limiarização de imagens, aplicada em imagem da classe Vegetação (fig 3.2(a)).

e a maneira de comparação, é mostrado o resultado da estimação de posição obtido anteriormente quando é aplicado o *canny* normal (quadrado azul).

Como foi observado nos resultados anteriores, o cálculo da correlação por limiarização da imagem fornece uma matriz de correlação bem definida, com um pico máximo

bem dominante e diferenciável, que corresponde à máxima correlação encontrada entre as duas imagens. O processo de limiarização consegue extrair e reproduzir um maior número informação, comparado com os algoritmos de detecção de bordas clássicos. Por essa razão, resulta ser factível que o ruído inserido nas imagens aéreas também seja reproduzido, mas este problema é reduzido com as regras de decisão. De modo que se o ponto de referência possui ruído, este é descartado e as comparações feitas em relação a ele, vão ter um valor neutro "0" que não contribuirá, nem atrapalhará o cálculo da matriz de correlação. Além disso, pode-se apreciar que a matriz de correlação apresentada, demostra ser uma matriz bem comportada onde existe só um máximo global como é esperado, e não tem presença de ruído que poderiam ocasionar vários máximos locais.

3.3 Resumo do Capítulo

Neste Capítulo, foram propostos dois métodos de detecção e extração de bordas para a estimação de posição geográfica em VANTs via Casamento de Imagens. A primeira proposta é baseada no algoritmo detector de bordas *canny*, onde foi gerado um novo detector de bordas *canny* adaptativo que melhora o problema de sobrestimação de bordas nas imagens capturadas por VANTs. A segunda proposta mostra um método que realiza limiarização na imagem, baseando-se no grau de homogeneidade que apresenta a área avaliada. Dessa forma, é calculado um limiar com robustez ao ruido presente nas imagens analisadas.

Na figura 3.12, é mostrada (quadrado em cor laranja) a parte onde os métodos propostos neste Capítulo se encaixam (são incorporados- se ajustariam - funcionariam) para funcionar como parte do processo geral de estimação de posição geográfica em VANTs via Casamento de Imagens.

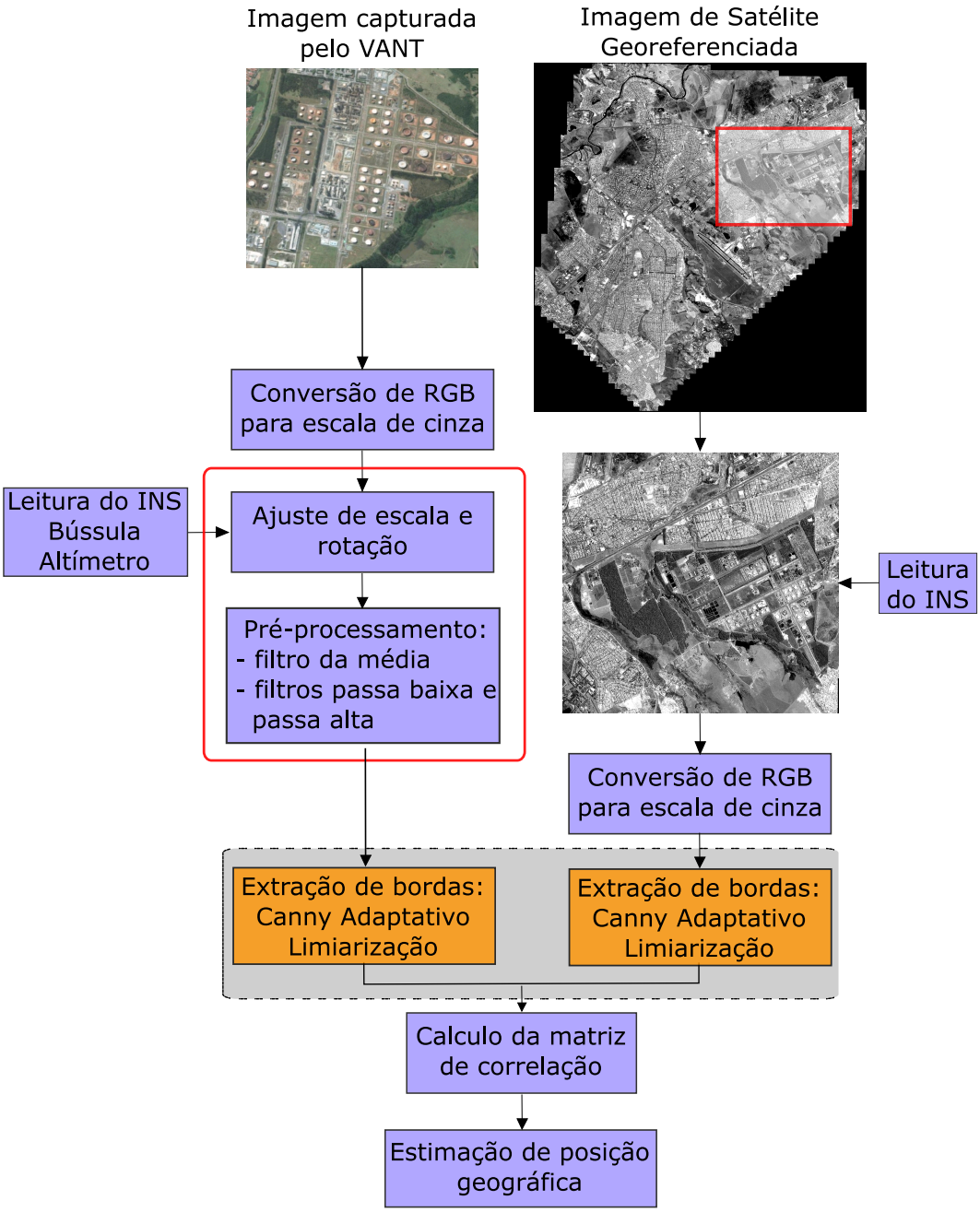


Figura 3.12: Diagrama de estimação de posição geográfica com a técnica Casamento de Imagens incluído os métodos propostos neste capítulo.

Capítulo 4

Correção de Distorção Projetiva em Imagens Capturadas por VANTs Utilizando um Método Paramétrico

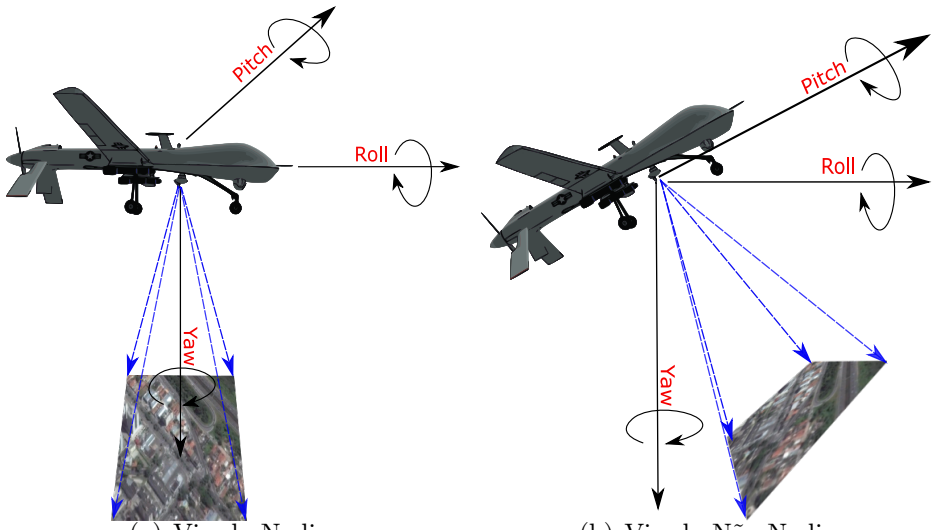
O Casamento de Imagens visa encontrar uma boa precisão na correspondência entre imagens (câmera do VANT e imagem georeferenciada) com a mesma cena, capturadas em tempos diferentes (Sedaghat et al., 2011). Além disso, a posição e a resolução da câmera faz que a imagem aérea apresente distorção espectral e efeitos de perspectiva (Birchfield, 1998) que afetam diretamente a estimação da posição do VANT conforme é ressaltado em (Tu and Dong, 2013; González Jiménez, 1999) e (Pai, 2010). O autor (Toutin, 2004) afirma que as características de imagens aéreas com distorção não conseguem representar com precisão a posição que esta sendo imageada. Adicionalmente, no caso particular de estimação de posição de VANTs via Casamento de Imagens, é essencial que ambas as imagens (aérea e georeferenciada) tenham visada nadir.

Considerando as limitações anteriormente mencionados, este Capítulo é focado na correção de distorção projetiva em imagens capturadas por VANTs quando a câmera não tem visada nadir, ou seja, quando a câmera não possui um ângulo perpendicular à Terra devido a manobras de viragem ou perturbações no ambiente. Como solução, é proposto um método paramétrico, que utiliza o conhecimento prévio dos ângulos de inclinação da aeronave, fornecidos pelos sensores inerciais da mesma. Estes ângulos conformam os parâmetros de rotação na matriz homográfica H, que também é formada por parâmetros intrínsecos da câmera. A maior contribuição deste Capítulo é que o método toma como referência a informação dos ângulos de inclinação do VANT para definir a homografia sem a necessidade de definir pontos de controle (Ground Control Points) (Wang and Yang, 2011) ou utilizar outra imagem como referência. Assim, a metodologia é independente da imagem porque não requer nenhum processo

de identificação de descritores ou extração de características nela.

4.1 Imagens Aéreas com Visada não Nadir

Imagens com visada não nadir possuem um nível considerável de distorção geométrica. A medida que aumenta o grau (ângulo) de distorção o tamanho dos pixeis na imagem é também alterado não uniformemente (Galbraith et al., 2005). Outros fatores, como o tipo de terreno, a inclinação, elevação e curvatura da terra são fatores secundários que podem induzir uma leve distorção na imagem. A perspectiva gerada em imagens com visada não nadir devido a manobras de viragem ou perturbações no ambiente é mostrada na Fig. 4.1, onde os ângulos Yaw, Pitch e Roll equivalem aos ângulos de guinada, arfagem e rolagem no VANT.



(a) Visada Nadir (b) Visada Não Nadir

Figura 4.1: Tipos de Visadas em VANTs.

Para o caso de estudo neste trabalho, as distorções causadas pela curvatura da Terra e tipo do terreno são ignorados pois as imagens testadas são de baixa altitude (altura de voo $<$ 120 m) (ANAC, 2015). Além disso, ignora-se a distorção da lente da câmera. Por conseguinte, apenas a distorção geométrica causada pela instabilidade do VANT será corrigida.

4.2 Visão Geral Técnica

O processo geral de correção da imagem capturada pelo VANT é mostrado na Fig.4.2. Este está composto inicialmente pela leitura da imagem distorcida junto com os ângulos de inclinação do VANT e os parâmetros da câmera. Depois, é calculada a matriz de

rotação e a matriz que define os parâmetros intrínsecos da câmera. Com isso, é obtida a matriz homográfica e por fim é aplicada a transformação projetiva sobre a imagem para corrigir a distorção geométrica nela.

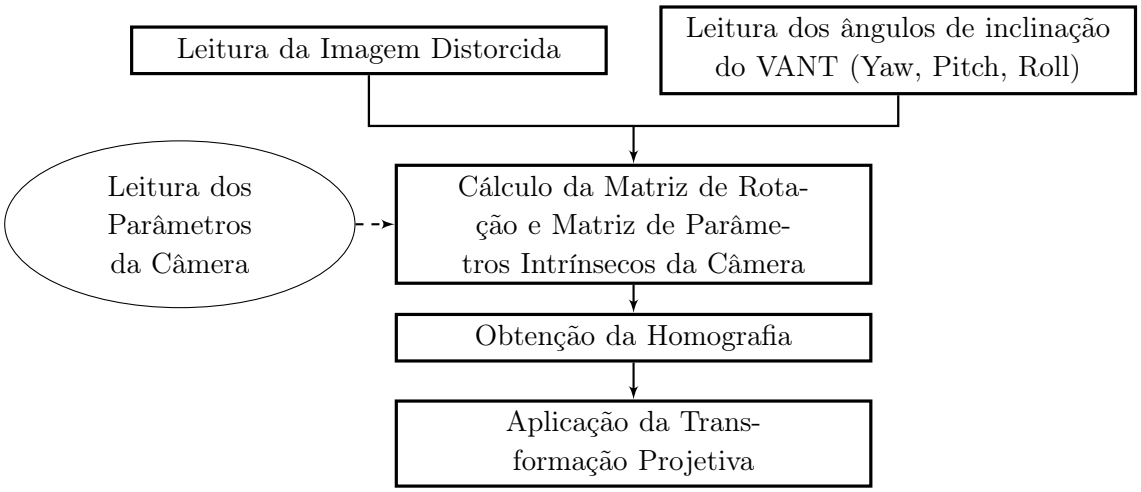


Figura 4.2: Processo Geral de Correção da Imagem

4.3 Cálculo da Matriz de Rotação e Matriz de Parâmetros Intrínsecos da Câmera

Para fazer correção de perspectiva em imagens capturadas por VANTs é necessário definir um modelo analítico da câmera, onde seja possível estimar as magnitudes e características da cena imageada que esta sendo analisada. Desta forma, pontos específicos da imagem poderão ser projetados em um plano partindo do modelo.

O modelo geralmente é obtido através da parametrização da câmera (Zhang et al., 1995), onde são estimados os valores dos parâmetros intrínsecos e extrínsecos do sensor conforme foi descrito anteriormente na Seção 2.2.2. Estes parâmetros definem a forma de como um objeto $3D$ é projetado no plano da imagem e assim, informações podem ser extraídas para posterior análise.

Em nosso caso, tem-se conhecimento das características da câmera e a leitura dos ângulos de inclinação do VANT, e pode-se definir com tais informações, o modelo da câmera, estimar a homografia e posteriormente corrigir a distorção causada pela visada não nadir da câmera abordo do VANT devido à variação dos ângulos (Yaw, Pitch, Roll), através de um método paramétrico de transformação projetiva.

Nesta etapa, a leitura dos parâmetros intrínsecos (Eq. 2.7) que definem parte do modelo é feita uma vez, pois a câmera que faz a captura das imagens é a mesma ao longo do voo, e possui as mesmas características após de sua respetiva calibração que

é feita antes do voo. Portanto, a matriz A vai ser constante no processo de correção de distorção projetiva nas imagens capturadas por VANTs.

Por sua parte, a leitura dos parâmetros extrínsecos (Eq. 2.11) varia com cada imagem capturada pelo VANT, pois em cada instante de captura de imagem a aeronave pode variar seus ângulos de inclinação (Yaw, Pitch, Roll), e com isto, os valores da matriz R. Dessa forma, os ângulos de inclinação do VANT são lidos com cada captura da imagem, para assim atualizar os parâmetros extrínsecos e garantir que a correção de distorção projetiva nas imagens seja correta.

4.4 Análise de Homografia e Transformação Projetiva

A correção de imagens pode ser considerada como o processo de transformar a projeção de uma imagem com distorção (inclinação na lente da câmera) a uma imagem livre de distorção com visada ortogonal do terreno. Existe então, uma matriz de homografia $H(x, y)$ que associa duas imagens. Nesta matriz, cada pixel j da imagem de referência com coordenadas (x, y) é mapeado a uma nova posição de uma imagem destino com coordenadas (u, v) como é indicado na seguinte equação:

$$\begin{aligned} u_j &= H_x(x, y) \\ v_j &= H_y(x, y) \end{aligned} \tag{4.1}$$

A correção de imagens baseadas no modelo de calibração do sensor gera esta homografia entre planos, a qual por definição estabelece que, cada ponto é correspondido por outro ponto e cada reta é correspondida por outra reta. Em consequência, a matriz de homografia consegue atribuir o valor de cada pixel na imagem original para uma nova posição do pixel na imagem corrigida. Na Eq. (4.2) é observada a matriz homográfica (Hartley and Zisserman, 2000).

$$H = A' \cdot (R - t_p n_p^T / d) \cdot A^{-1} \tag{4.2}$$

- A é a matriz de calibração da câmera (parâmetros intrínsecos),
- R é a matriz de rotação,
- t_p é a posição da visão da cena,
- n_p é um vetor normal ao plano da cena,

- d é a distância do ponto no plano da imagem até o ponto no plano do mundo.

A matriz homográfica é caracterizada por ser uma matriz de dimensões 3×3, invertível e não singular. H pode ser representada como se observa na equação Eq. (4.3). Onde cada um dos coeficientes $h_{11} - h_{32}$ já são conhecidos utilizando a Eq. (4.2), com $h_{33} = 1$. Todos os nove coeficientes podem ter valores diferentes.

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \tag{4.3}$$

A matriz H (Eq. (4.3)), é equivalente a estas duas expressões, Eq. (4.4).

$$\begin{aligned} u &= \frac{h_{11}x + h_{21}y + h_{31}}{h_{13}x + h_{23}y + 1} \\ v &= \frac{h_{12}x + h_{22}y + h_{32}}{h_{13}x + h_{23}y + 1} \end{aligned} \tag{4.4}$$

Concluindo, a matriz homográfica H define a relação entre os planos através de uma transformação projetiva. Esta transformação projetiva possui 8 graus de liberdade e define um nível mais completo de homografia em comparação com a transformação afim ou euclidiana.

4.5 Resumo do Capítulo

Neste Capítulo, foi proposto um método paramétrico que realiza correção de distorção projetiva em imagens capturadas por VANTs. Este método toma leitura os ângulos de inclinação do VANT e as características da câmera para realizar as correções sobre a imagem. Isto faz com que o método apresente vantagens e robustez na retificação de imagens, pois este não é afetado pelas mudanças de luminosidade ou ruido inserido nas imagens capturadas pelo VANT.

Dessa forma, na figura 4.3, é mostrado (quadrado em cor laranja) a parte onde o método paramétrico proposto neste Capítulo se incorpora para trabalhar como parte do processo geral de estimação de posição geográfica em VANTs via Casamento de Imagens.

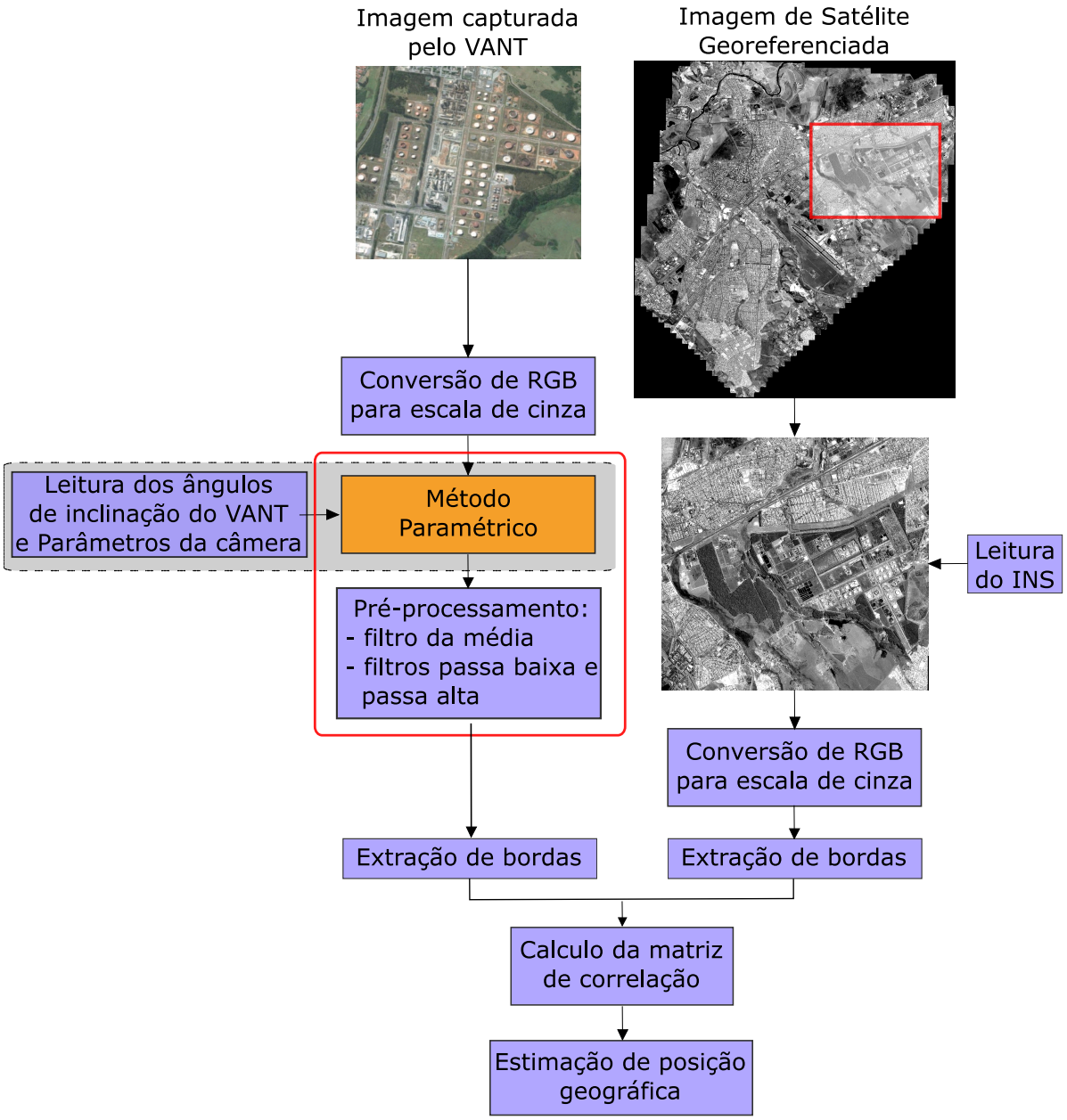


Figura 4.3: Diagrama de estimação de posição geográfica com a técnica Casamento de Imagens incluído o método proposto neste capítulo.

Capítulo 5

Correção de Distorção Projetiva e Espectral em Imagens Capturadas por VANTs Utilizando um Método não Paramétrico

Este Capítulo apresenta outro método diferente para corrigir as distorções tanto projetivas e espectrais nas imagens capturadas por veículos aéreos não tripulados. As distorções projetivas que são ocasionadas quando a câmera não está com visada nadir (ângulo perpendicular ao solo) devido às manobras de viragem ou perturbações no ambiente. As distorções espectrais, por sua vez, podem ocorrer quando as imagens analisadas são adquiridos a partir de sensores diferentes. De fato, as imagens georeferenciadas geralmente apresentam maior resolução que as imagens aéreas e muitas vezes são obtidas de algum banco de dados conhecidos, como o *Google Earth*.

De um modo geral, obter uma adequada precisão no processo de Casamento de Imagens não é simples. As distorções apresentadas no registro da imagem existem principalmente devido ao processo de aquisição da imagem. Na maioria dos casos, as imagens são capturados em tempos distintos e condições climáticas totalmente diferentes, além da utilização de distintos tipos de sensores e diferente resposta espectral (Sedaghat et al., 2011). Tudo isto tende a degradar o desempenho do Casamento de Imagens, já que as regiões em comuns entre as imagens (aérea e georeferenciada), podem gerar distintos descritores. Além disso, a posição da câmara de vídeo e sua resolução pode afetar imagens por adição de ruído, aumento na distorção e efeitos de perspectiva (Birchfield, 1998).

Portanto, este Capítulo é focado na correção de distorção projetiva em imagens capturadas por VANTs. Como solução é proposto um método não paramétrico, que

como seu nome indica, é caraterizado por não utilizar parâmetros da câmera. Requer apenas definir pontos correspondentes entre a imagem capturada pelo VANT e a imagem georeferenciada.

5.1 Visão Geral Técnica

Para a correção de distorção projetiva e espectral em imagens capturadas por VANTs é feito em cada par de imagens (aérea e georeferenciada), os passos descritos na Fig.5.1. Onde, inicialmente as imagens são submetidas a uma etapa de pré-processamento, em seguida pontos característicos são obtidos e casados entre as imagens aplicando o algoritmo SURF (Speeded-Up Robust Transform) (Bay et al., 2008) e com o algoritmo MSAC (M-estimator Sample Concensus) (Torr and Zisserman, 2000) são eliminados os casamentos de pontos característicos atípicos (*outliers*). Finalmente, os parâmetros que compõem a matriz homográfica H são estimados e a transformação projetiva é aplicada na imagem capturada pelo VANT. Este método será melhor detalhado e discutido nas seguintes seções.

5.2 Pré-processamento

Esta etapa está centrada no pré-processamento das imagens antes de fazer a extração de características. A intenção desta etapa é aumentar o número de casamentos de pontos característicos que são extraídos por SURF quando às imagens obtidas são a partir de diferentes sensores e em distintos tempos de captura. Neste tipo de situações, o brilho e a diferença de contraste complica o casamento de pontos característicos, independentemente do tipo do algoritmo descritor que se esteja utilizando. Isto é devido à falta de uniformidade e a diferença na distribuição nos níveis de cinza entre as imagens (câmera do VANT e georeferenciada), o que dá como resultado pontos característicos com diferenças significativas no contraste e no nível de brilho. Dessa forma, os casamentos de pontos característicos não poderão ser calculados.

5.2.1 Equalização de Histograma

A equalização de histograma (Tu and Dong, 2013) é uma técnica muito eficaz utilizada para obter consideráveis melhorias na imagem para um posterior análise e extração de características. Assim, o princípio de equalização de histograma se fundamenta na ideia de que, através de transformações feitas na imagem em tons de cinza e a partir do conhecimento sobre a distribuição de probabilidade dos pixeis na imagem original, será gerada uma nova imagem cuja distribuição de probabilidade vai garantir uniformidade

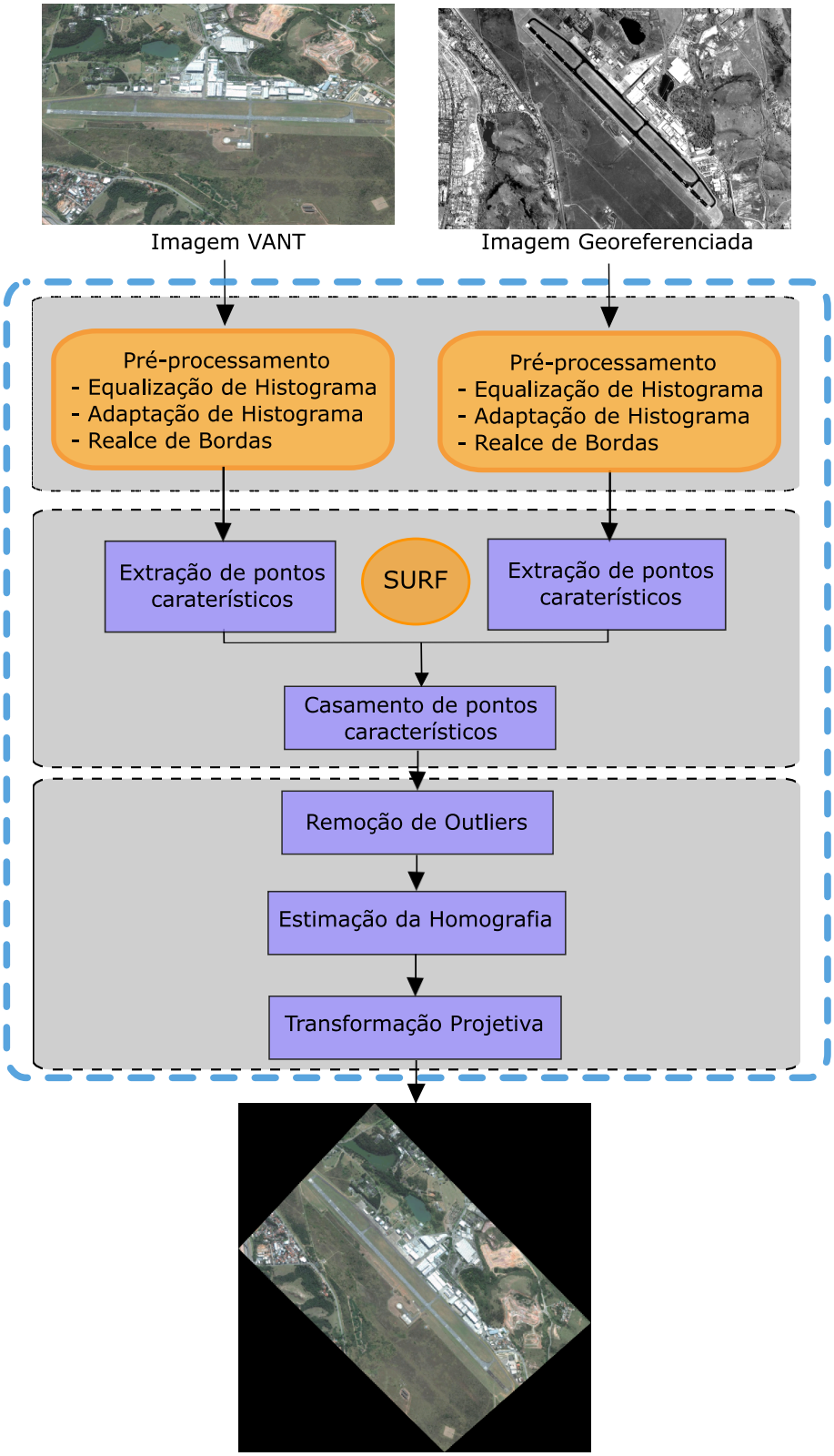


Figura 5.1: Arquitetura para correção de distorção projetiva e espectral em imagens aéreas.

nos níveis de cinza da imagem. Isto é conseguido utilizando a função de distribuição cumulativa, como função de mapeamento (Sasi and Jayasree, 2013). Assim, a nova imagem apresentará um melhor contraste com uma ampla faixa dinâmica ao longo do histograma. A figura 5.2 representa o diagrama esquemático de equalização de histograma. Nele, o eixo x mostra os níveis de cinza da imagem, enquanto o eixo y indica as probabilidades que cada nível de cinza possui.

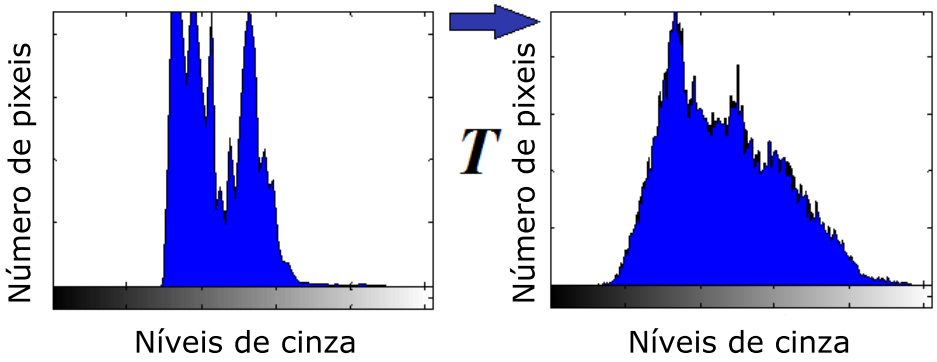


Figura 5.2: Diagrama esquemático de equalização de histograma.

Se uma imagem possui n números de pixeis, que estão distribuídos em níveis de intensidade discretos L_p (0-255), onde n_k é o número de pixeis com i_k nível de intensidade. A função de densidade de probabilidade (f_i) e a função de densidade cumulativa (F_k) são definidas pela equação Eq.(5.1) e Eq.(5.2) respectivamente.

$$f_i(i_k) = \frac{n_k}{n} \tag{5.1}$$

$$F_k(i_k) = \sum_{j=0}^{k} f_i(i_k) \tag{5.2}$$

Além disso, existem outras técnicas que são comumente implementadas no pré-processamento, dependendo da aplicação. Ao longo desta etapa também foram aplicados ajuste de intensidade, adaptação de histograma e realce de bordas nas imagens.

5.2.2 Realce de Bordas e Ajuste de Intensidade

O realce de imagens utilizado para conseguir uma maior nitidez sobre uma cena imageada é uma das etapas mais importantes no pre-processamento de imagens (Alsam et al., 2015). A imagem original é processada de tal forma que a imagem resultante fornecerá melhores condições para aplicações posteriores específicas do problema abordado. Nesta abordagem a utilização de filtros espaciais são bem comuns, e tem como

principal objetivo a de identificar e ressaltar as variações de intensidade presentes ao longo da imagem.

O processo para melhorar a nitidez é baseado na diferenciação espacial. O filtro espacial é aplicado sobre a imagem e posteriormente a imagem original é subtraída da imagem filtrada. Uma máscara de enfoque é utilizada para perfilar a imagem, ela não pode criar detalhes adicionais, mas pode melhorar consideravelmente os detalhes que na imagem original não eram possível apreciar. O filtro utilizado nesta aplicação é o filtro Laplaciano (Gonzalez and Woods, 2008), este filtro utiliza a derivada de segunda ordem para a detecção das bordas mas não a direção delas. Como sua resposta esta independente da direção das bordas o filtro Laplaciano só responde as transições de intensidade na imagem. O filtro é definido por uma função bidimensional $f_L(x, y)$ onde a derivada de segunda ordem é mostrada na Eq.(5.3)

$$\nabla^2 f_L = \frac{\partial^2 f_L}{\partial x^2} + \frac{\partial^2 f_L}{\partial y^2} \tag{5.3}$$

Como a derivada da segunda ordem é um operador linear, a equação anterior pode ser expressado da seguinte forma (5.4):

$$\begin{aligned} \frac{\partial^2 f_L}{\partial x^2} &= f_L(x+1, y) + f_L(x-1, y) - 2f_L(x, y) \\ \frac{\partial^2 f_L}{\partial y^2} &= f_L(x, y+1) + f_L(x, y-1) - 2f_L(x, y) \end{aligned} \tag{5.4}$$

Portanto, o Laplaciano de forma discreta é dada pela seguinte equação (5.5):

$$\nabla^2 f_L(x, y) = f_L(x+1, y) + f_L(x-1, y) + f_L(x, y+1) + f_L(x, y-1) - 4f_L(x, y) \tag{5.5}$$

A anterior equação é implementada usando a máscara de filtro mostrada em (5.6), com um coeficiente central de -4. De modo que, o pixel central é subtraído com os pixeis vizinhos para depois ser calculada a média. Assim, quando o pixel central passasse por uma borda esta diferença será grande e nos outros casos a diferença será pequena.

$$\frac{1}{k}\begin{bmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \tag{5.6}$$

Destacando as descontinuidades de intensidade, realçando as variações de intensidade e conservando a nitidez na imagem ao mesmo tempo, quando a imagem original é subtraída da imagem com o filtro Laplaciano. Essa operação esta dada pela seguinte

equação (5.7):

$$g_L(x, y) = f_L(x, y) - [\nabla^2 f_L(x, y)] \tag{5.7}$$

A Figura 5.3 mostra os resultados obtidos após do realce das imagens utilizando o filtro Laplaciano. Nas figuras 5.3(a) e 5.3(b) são mostradas as imagens (georeferenciada e capturada pelo VANT) resultantes da equalização e adaptação de histograma que serão as entradas do filtro Laplaciano. As figuras 5.3(c) e 5.3(d) é evidenciado as imagens resultantes da aplicação do filtro Laplaciano. Por último, é mostrado as imagens finais 5.3(e) e 5.3(f), uma vez que a imagem inicial foi subtraída da imagem filtrada.

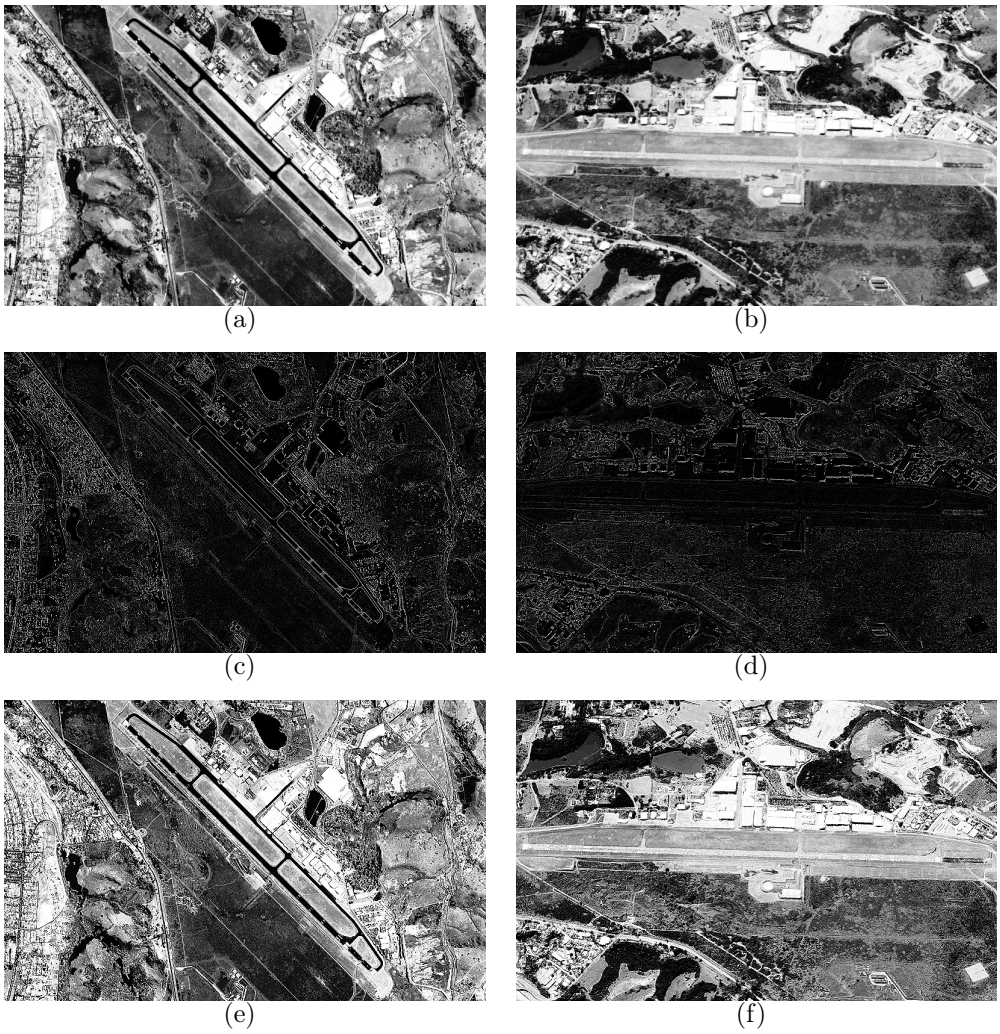


Figura 5.3: (a) Imagem georeferenciada em tons de cinza (b) Imagem capturada pelo VANT em tons de cinza; (c) Imagem georeferenciada depois de ser aplicado o filtro Laplaciano (d) Imagem capturada pelo VANT depois de ser aplicado o filtro Laplaciano; (e) Imagem georeferenciada depois da subtração com a imagem filtrada (f) Imagem capturada pelo VANT depois da subtração com a imagem filtrada.

Finalmente, a imagem 5.4 mostra o resultado final de todo o pré-processamento aplicado em cada uma das imagens para garantir casamento de pontos característicos entre elas, quando o algoritmo SURF seja aplicado. A Figura 5.4(a) mostra a imagem original em tons de cinza capturada pelo VANT e a 5.4(c) mostra seu histograma correspondente. O resultado obtido após o pré-processamento é mostrado na Figura 5.4(b). É evidenciado um aumento na escala dinâmica dos níveis de cinza e, consequentemente, é produzido um aumento no contraste da imagem. Além disso, apresenta um maior realce de bordas e uma distribuição mais uniforme nos níveis de cinza, conforme ilustrado na Figura 5.4(d), que corresponde ao histograma da imagem resultante pré-processada. Por sua vez, a imagem georeferenciada também é analisada, a imagem original em tons de cinza e seu histograma são mostrados nas Figuras 5.4(e) e 5.4(g), respectivamente. As Figuras 5.4(f) e 5.4(h) ilustram os resultados obtidos do pré-processamento e seu histograma respectivamente. Desta etapa, é importante ressaltar a grande similaridade que existe nos histogramas das imagens capturada pelo VANT e georeferenciada depois de ser pré-processadas. Isto proporciona imagens com maior qualidade para ser utilizadas na extração e casamento de pontos características com SURF.

5.3 Extração de Características e Casamento de Pontos Característicos com SURF

O algoritmo SURF (Speeded-up Robust Transform) (Bay et al., 2008) é aplicado sobre as imagens (aérea e georeferenciada) já pré-processadas. Inicialmente pontos característicos são extraídos em cada imagem por separado. Nesse processo, SURF faz detecção de pontos de interesse em diferentes variações de escala na imagem, onde a imagem é suavizada e avaliada com o uso de DoG (Diferença de Gaussiana) e o determinante da matriz Hessiana para obter os BLOBs que contém a informação dos pontos característicos. Em seguida é calcula a respectiva localização dos pontos de interesse sobre a imagem e com ela é definido o vetor descritor com tamanho de 64 valores de características. No processo de casamento de pontos característicos, são descartados os pontos característicos com grande diferença no contraste ou aqueles que possuem o mesmo brilho mas que estão já casados. Uma das virtudes deste algoritmo SURF é a maneira como ele varia o tamanho do filtro tipo caixa ao invés de variar o tamanho da imagem (mudar a escala), fazendo que seja reduzido o tempo de processamento.

A figura 5.5(a) mostra as imagens de teste já pré-processadas. A figura 5.5(b) ilustra os resultados da detecção de pontos característicos obtidos a partir das imagens de teste. A figura 5.5(c) apresenta as localizações dos pontos de interesse correspondentes

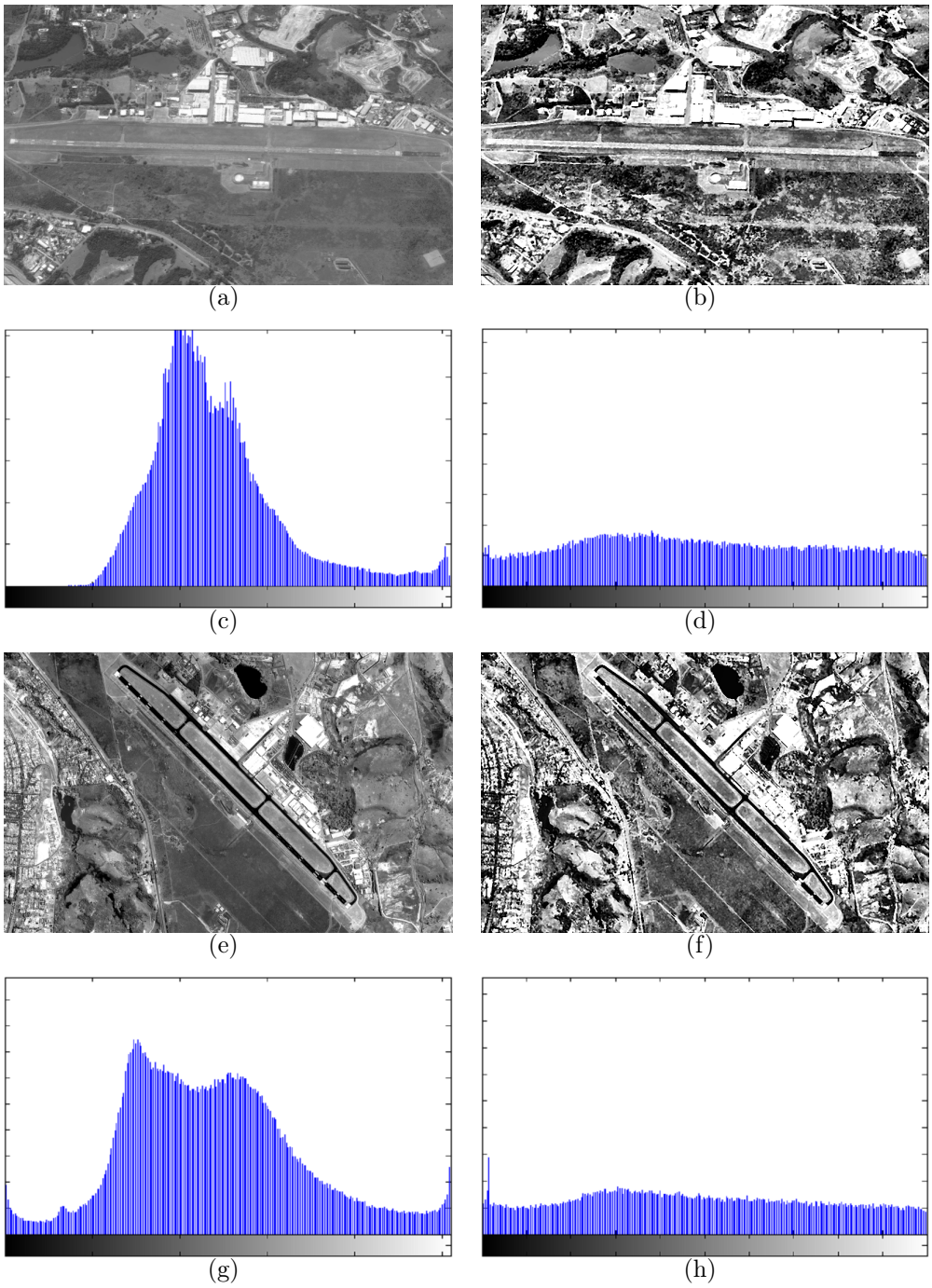


Figura 5.4: Pré-processamento: (a) Imagem capturada pelo VANT original em tons de cinza e (c) Seu histograma correspondente; (b) Imagem capturada pelo VANT depois do Pré-processamento e (d) Seu histograma correspondente; (e) Imagem georeferenciada original em tons de cinza e (g) Seu histograma correspondente; (f) Imagem georeferenciada depois do pré-processamento e (h) Seu histograma correspondente.

a cada detecção de ponto característico encontrado nas imagens. A figura 5.5(d) ilustra os resultados de casamento obtidos da aplicação do algoritmo SURF a um par de imagens, onde a imagem da esquerda representa a imagem georeferenciada 5.4(f) e a

imagem da direita representa a imagem capturada pelo VANT com visada não nadir, rotação, traslação e diferença de escala 5.4(b).

5.3.1 Remoção de Outliers com MSAC

Após do casamento de pontos característicos nas imagens, é implementado o algoritmo MSAC (M-estimator Sample Concensus) (Torr and Zisserman, 2000) que constitui uma melhoria do algoritmo RANSAC (Fischler and Bolles, 1981; Dubrofsky, 2009). O MSAC permite a remoção de *outliers* e encontra a melhor correspondência de pontos característicos entre duas imagens. Neste algoritmo, é utilizada uma técnica de amostragem aleatória que define um número K de hipóteses que representam os K possíveis modelos matemáticos que podem ser adotados. Cada hipótese é gerada escolhendo aleatoriamente um conjunto de amostras iniciais de todos os casamentos de pontos característicos obtidos na etapa anterior. É classificado como um valor atípico (outlier), o ponto característico cuja distância geométrica seja maior que o limiar predefinido (threshold) t_o. De todas as K hipóteses, é selecionada a que minimiza o número de valores atípicos (*outliers*) e que, por conseguinte maximiza o número de *inliers*. Portanto, escolhendo a hipótese com o valor mais baixo na função de custo f_{custo}, Eq.(5.8), é minimizada a soma dos resíduos ao quadrado (r^2) para os *inliers* (Torr and Zisserman, 2000).

$$f_{custo} = \sum_j \rho(r_j^2) \tag{5.8}$$

Onde o termo residual ρ é:

$$\rho(r^2) = \begin{cases} r^2 & r^2 < t_o^2 \\ t_o^2 & r^2 \geq t_o^2 \end{cases} \tag{5.9}$$

MSAC é melhor que RANSAC porque, na avaliação da função de custo, os *inliers* são ponderados com base no valor da distância geométrica, enquanto RANSAC pondera estes *inliers* com 0, de modo que o algoritmo não é sensível ao limiar utilizado para determinar se a correspondência entre pontos é *inlier* ou não.

A Figura 5.5(d) ilustra os resultados obtidos depois do filtragem de *outliers* com MSAC, as imagens testadas foram obtidas com distintos sensores e em diferentes tempos de captura.

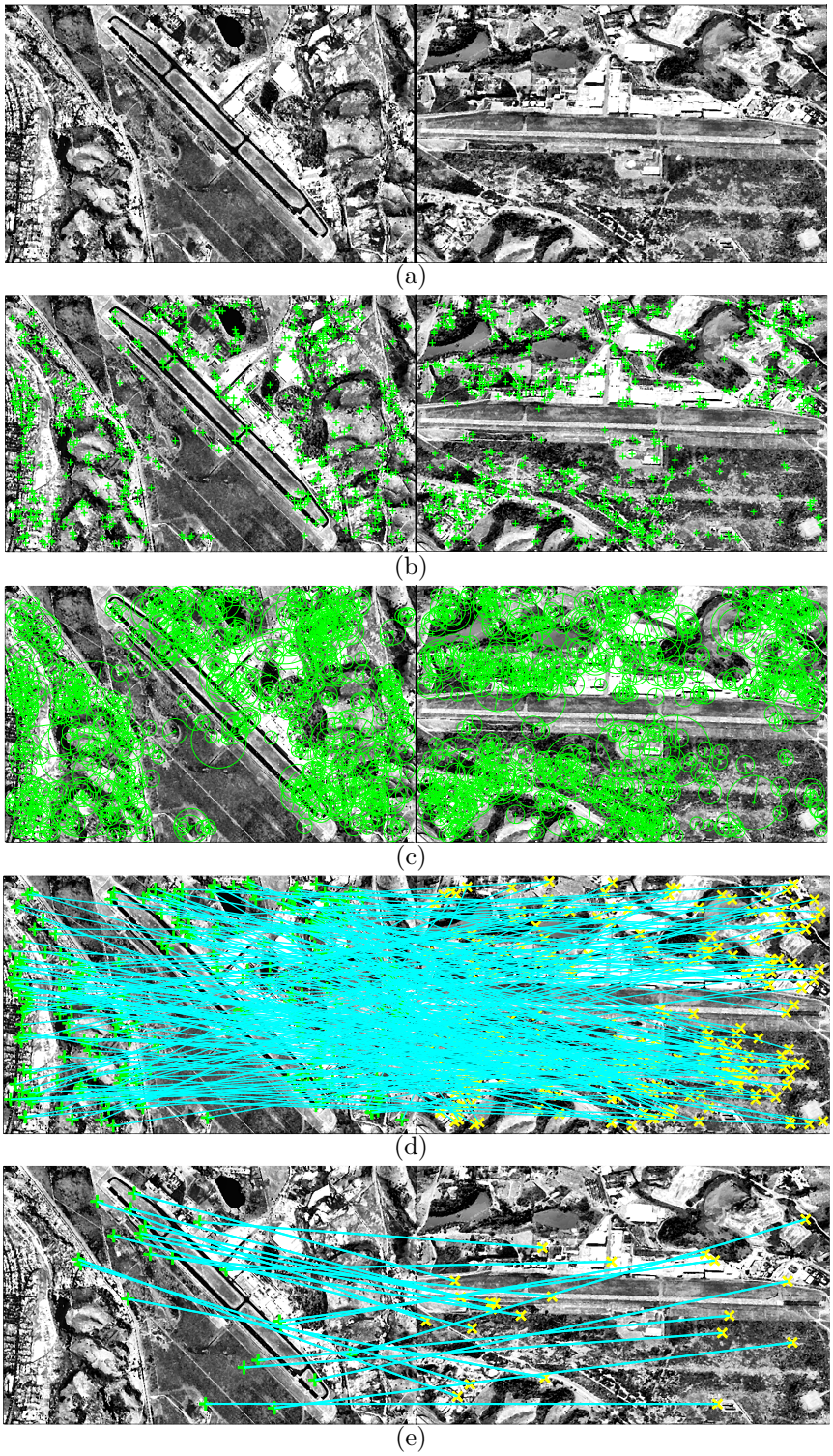


Figura 5.5: Detecção e casamento de pontos característicos com SURF e MSAC: (a) Imagens de teste já pré-processadas; (b) Detecção de pontos característicos na imagem georeferenciada (esquerda) e na imagem do VANT (direita); (c) Localização dos pontos característicos na imagem georeferenciada (esquerda) e na imagem do VANT (direita); (d) Casamento de pontos característicos com SURF entre a imagem georeferenciada (esquerda) e a imagem do VANT (direita); (e) Casamento de pontos característicos entre a imagem georeferenciada (esquerda) e a imagem do VANT (direita) usando SURF+MSAC.

5.4 Estimação de Homografia a Partir do Casamento de Pontos Característicos

As correspondências de pontos característicos obtidos da etapa anterior, são utilizadas para a obtenção da matriz homográfica H que, é representada por uma matriz 3×3 que relaciona os pontos característicos entre a imagem georeferenciada e a imagem capturada pelo VANT, da seguinte forma:

$$X = HX' \tag{5.10}$$

Onde X e X' representam os pontos característicos de referência e de destino, respectivamente. Uma matriz de 3×3 que seja invertível e não singular pode ser considerada como homografia. H esta definida da seguinte forma:

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \tag{5.11}$$

Considerar o número de pontos correspondentes necessários para calcular a transformação projetiva H, é a primeira questão a ser avaliada. Existe um limite inferior de pontos correspondentes a ser considerado, que depende do número de graus de liberdade e o número de restrições que existem no modelo de transformação homográfica. Cada correspondência de pontos representam duas restrições e possuem dois graus de liberdade que correspondem aos componentes X e Y. Portanto, é necessário definir pelo menos quatro correspondências de pontos conhecidos para calcular H totalmente.

Os parâmetros $h_{11} - h_{32}$ podem ser determinados através da solução das equações definidas em (5.12) que, representam um problema não linear de mínimos quadrados. Existem muitos métodos de solução, neste trabalho é implementado o algoritmo DLT (Direct Linear Transform) para estimar os parâmetros e obter a matriz de transformação H ideal.

$$\begin{bmatrix} x^{1'} \\ y^{1'} \\ x^{2'} \\ y^{2'} \\ x^{3'} \\ y^{3'} \\ x^{4'} \\ y^{4'} \end{bmatrix} = \begin{bmatrix} x^1 & y^1 & 1 & 0 & 0 & 0 & -x^1x^{1'} & -y^1x^{1'} \\ 0 & 0 & 0 & x^1 & y^1 & 1 & -x^1y^{1'} & -y^1y^{1'} \\ x^2 & y^2 & 1 & 0 & 0 & 0 & -x^2x^{2'} & -y^2x^{2'} \\ 0 & 0 & 0 & x^2 & y^2 & 1 & -x^2y^{2'} & -y^2y^{2'} \\ x^3 & y^3 & 1 & 0 & 0 & 0 & -x^3x^{3'} & -y^3x^{3'} \\ 0 & 0 & 0 & x^3 & y^3 & 1 & -x^3y^{3'} & -y^3y^{3'} \\ x^4 & y^4 & 1 & 0 & 0 & 0 & -x^4x^{4'} & -y^4y^{4'} \\ 0 & 0 & 0 & x^4 & y^4 & 1 & -x^4y^{4'} & -y^4y^{4'} \end{bmatrix} \cdot \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \end{bmatrix} \tag{5.12}$$

5.5 Resumo do Capítulo

Ao longo deste Capítulo, foi apresentado um método não paramétrico para corrigir a distorção projetiva e espectral em imagens capturadas por VANTs. Este método utiliza a imagem georeferenciada como referência para encontrar casamentos de pontos característicos sobre a imagem capturada pelo VANT. O diferencial deste método é que resulta ser robusto à diferencia de sensores entre as imagens (capturada pelo VANT e georeferenciada). Além disso, não precisa de leitura de ângulos de inclinação do VANT, nem parâmetros da câmera para corrigir as imagens. Na figura 5.6, é mostrado onde o método proposto é encaixado para funcionar como parte do processo geral de estimação de posição geográfica em VANTs via Casamento de Imagens.

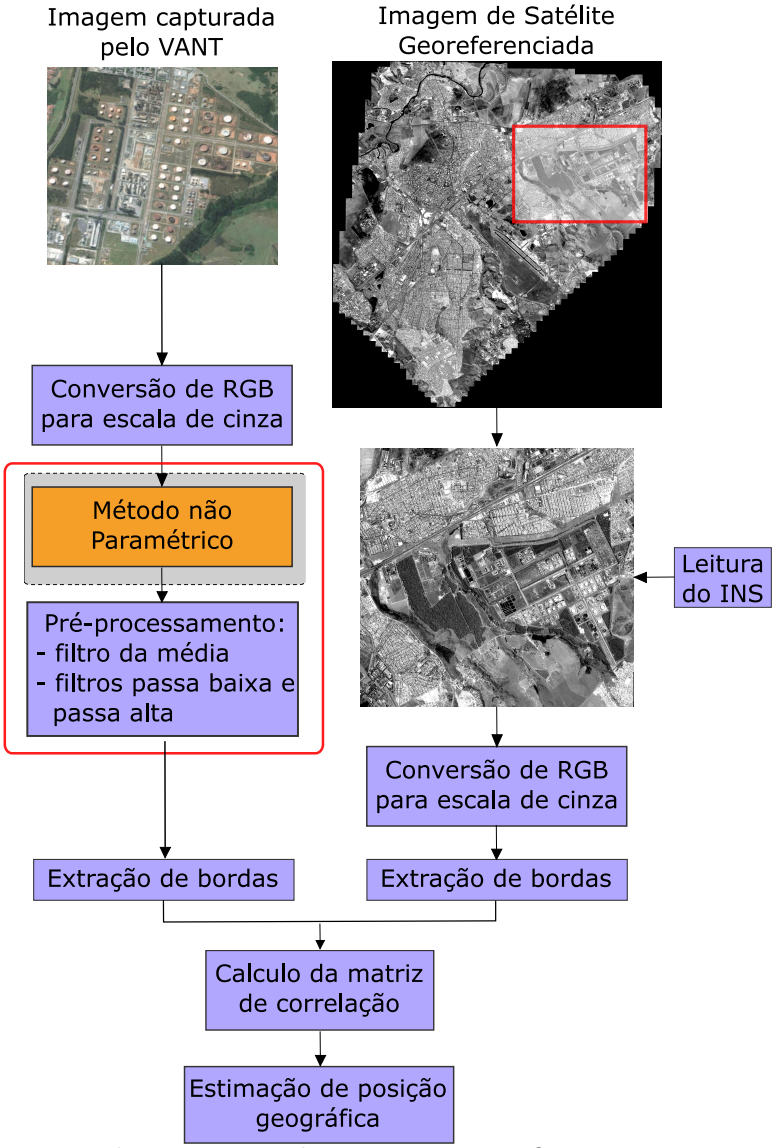


Figura 5.6: Diagrama de estimação de posição geográfica com a técnica Casamento de Imagens incluído o método proposto neste Capítulo.

Capítulo 6

Resultados Experimentais e Análises

Neste Capítulo são apresentados os resultados dos experimentos realizados na avaliação dos diferentes métodos propostos ao longo deste trabalho. Inicialmente, na Seção 6.1, é definida a Área de trabalho e o Software onde foram desenvolvidos os métodos propostos. Em seguida, a Seção 6.2, mostram-se os experimentos realizados para avaliar o desempenho dos métodos desenvolvidos de Casamento de Imagens descritos no Capítulo 3. Na Seção 6.3, avalia-se a correção de imagens aéreas com distorção projetiva utilizando o método paramétrico proposto no Capítulo 4. Por sua parte, na Seção 6.4, o método de correção de imagens aéreas com distorção projetiva e espectral proposto no Capítulo 5 é avaliado em imagens de igual sensor assim como em imagens de diferente sensor. Finalmente, na Seção 6.5, é avaliado em conjunto o método paramétrico de correção de perspectiva e o método de estimação de posição via Casamento de Imagens sobre uma simulação de voo feito de VANT.

6.1 Área de Trabalho e Software

Os métodos desenvolvidos ao longo deste trabalho foram implementados em MATLAB 2014a, utilizando as bibliotecas de Processamento Digital de Imagens e Visão Computacional que o Software MATLAB oferece. A área de trabalho opera sobre um processador Intel® Core $i5-4200U$ CPU $1.60GHz$ $2.30GHz$ com 8 GB de RAM, rodando com sistema operacional Windows 8.1 Pro de 64 bits. Alguns métodos também foram desenvolvidos e implementados em $C++$ baseado na biblioteca de *Processamento Digital de Imagens - OpenCV* 3.0, rodando com sistema operacional LINUX, Ubuntu 14.04.5 LTS (secção 6.3).

6.2 Casamento de Imagens

Dois experimentos foram realizados com o objetivo de avaliar o desempenho dos métodos desenvolvidos de Casamento de Imagens descritos no Capítulo 3. Os métodos foram avaliados sobre os diferentes tipos de classes definidas anteriormente (Vegetação, Urbano e Rodovia). Desta forma, foi definida uma rota no simulador de voo NAV (Silva et al., 2015; Santos et al., 2013), e as imagens extraídas ao longo do percurso do VANT foram testadas e comparadas junto com o mapa georeferenciado previamente embarcado no VANT (Silva et al., 2015). Estas imagens foram obtidas com o satélite *Quickbird*, com uma resolução espacial de $0,6\ m$ e pertencem à região de São José dos Campos no Estado de São Paulo - Brasil (anexo A1.1).

O primeiro experimento teve como objetivo testar a acurácia dos métodos propostos com imagens de diferente sensor (sensor de câmera acoplado no VANT e sensor do satélite) ao longo dos diferentes tipos de classes definidas: Vegetação, Rodovia e Urbano. A imagem georeferenciada e a imagem capturada pelo VANT apresentam diferenças de luminosidade na faixa do visível e a imagem aérea apresenta em algumas regiões, um alto grau de degradação e ruído constante. Algumas dessas imagens são mostradas no anexo A1.2. Na sequência, foi planejada uma rota específica no simulador de voo na qual foram analisadas 132 pares de imagens (georeferenciada e capturada pelo VANT) para este experimento, dos quais o 42% das imagens capturadas pelo VANT pertencem à classe Vegetação, o 39% pertencem à classe Urbana e o 19% pertencem à classe Rodovia. A rota planejada utilizada nos testes é mostrada na figura 6.1. O simulador de voo descrito em (Silva et al., 2015; Santos et al., 2013) possui um protocolo de comunicação TCP-IP entre o software *FligthGear 2015* (Sorton and Hammaker, 2005), que modela o comportamento da aeronave e os módulos de programação de Visão Computacional desenvolvidos para fazer estimação de posição através de imagens, também descritos em (Silva et al., 2015; Santos et al., 2013).

O segundo experimento foi feito com imagens do mesmo sensor (sensor de câmera acoplado no VANT e sensor do satélite), com a finalidade de testar o comportamento e acurácia dos métodos propostos para os diferentes tipos de terreno imageados pelo VANT. A mesma rota planejada no teste anterior foi utilizada neste experimento. A figura do anexo A1.3, mostra algumas das imagens utilizadas na simulação de voo. Obtiveram-se 158 pares de imagens (georeferenciada e capturada pelo VANT) a serem analisadas, dos quais o 48% das imagens capturadas pelo VANT pertencem à classe Vegetação, o 37% pertencem à classe Urbana e o 15% pertencem à classe Rodovia.

As dimensões da imagem georeferenciada para os dois experimentos é de 401×401 pixeis e a imagem capturada pelo VANT apresenta dimensões de 95×95 pixeis. O número de pares de imagens (georeferenciada e capturada pelo VANT) analisadas no



Figura 6.1: Rota planejada para avaliação dos testes.

primeiro e segundo experimento estão acima de 100 pares de imagens, isto com o objetivo de aumentar a frequência de captura das imagens e garantir um cobertura completa das áreas sobrevoadas pelo VANT na rota planejada. Em trabalhos anteriores como o de Silva et al. (2015), o máximo números de estimativas em uma rota planejada foi de 87.

A medida de acurácia na estimação de posição é definida pela distância Euclidiana entre a posição estimada pelos diferentes métodos aplicados via Casamento de Imagens e a posição real. Assim, o pixel de maior correlação após o casamento, define a *latitude, longitude* da posição estimada. O cálculo dessa distância D está definido pela seguinte equação:

$$D = R_T \; arcos(cos(lat_1)cos(lat_2)cos(lon_1 - lon_2) + sin(lat_1)sin(lat_2)) \tag{6.1}$$

Onde

- R_T é o raio médio da circunferência da Terra,
- lat_1 e lon_1 são a latitude e longitude da posição real,
- lat_2 e lon_2 correspondem à latitude e longitude da posição estimada pelos métodos aplicados,

- D representa a distância de separação em metros entre a posição estimada e a posição real.

Após a aplicação dos métodos *canny* adaptativo e limiarização, mostram-se os resultados comparativos correspondentes ao experimento 1 na tabela 6.1. De forma que, são mostrados os resultados do método de estimação de posição com extrator de bordas *canny* utilizado em (Silva et al., 2015), e comparados com o extrator de bordas *canny* que utiliza a função do cálculo da matriz de correlação otimizada. Também, são comparados com o método do detector de bordas *canny* adaptativo e o método por limiarização que foram descritos ao longo do Capítulo 3.

Uma métrica de avaliação para os métodos neste experimento é o valor da distância média em metros como métrica de erro, e indica quão separada está a posição que foi estimada em cada um dos métodos com relação à posição real onde o VANT deveria estar. Também, é calculada a variância (Eq. 6.2) e o desvio padrão (Eq. 6.3) em função da distância média. O número de falsos positivos (FP) estimado representa o número de estimativas de posição na qual sua distância média esteve por cima de 100 metros, esta distância foi utilizada em trabalhos anteriores como um índice de distância máxima permissivo na estimação de posição do VANT (Silva et al., 2015). Por último, é obtido o tempo de processamento que cada método demora em fazer a estimação de posição, esta métrica resulta ser importante pois indica quais dos métodos têm potencial para ser implementada em tempo real.

$$Var(distância) = \sum \frac{(distância_i - \overline{distância})^2}{\#\ imagens} \tag{6.2}$$

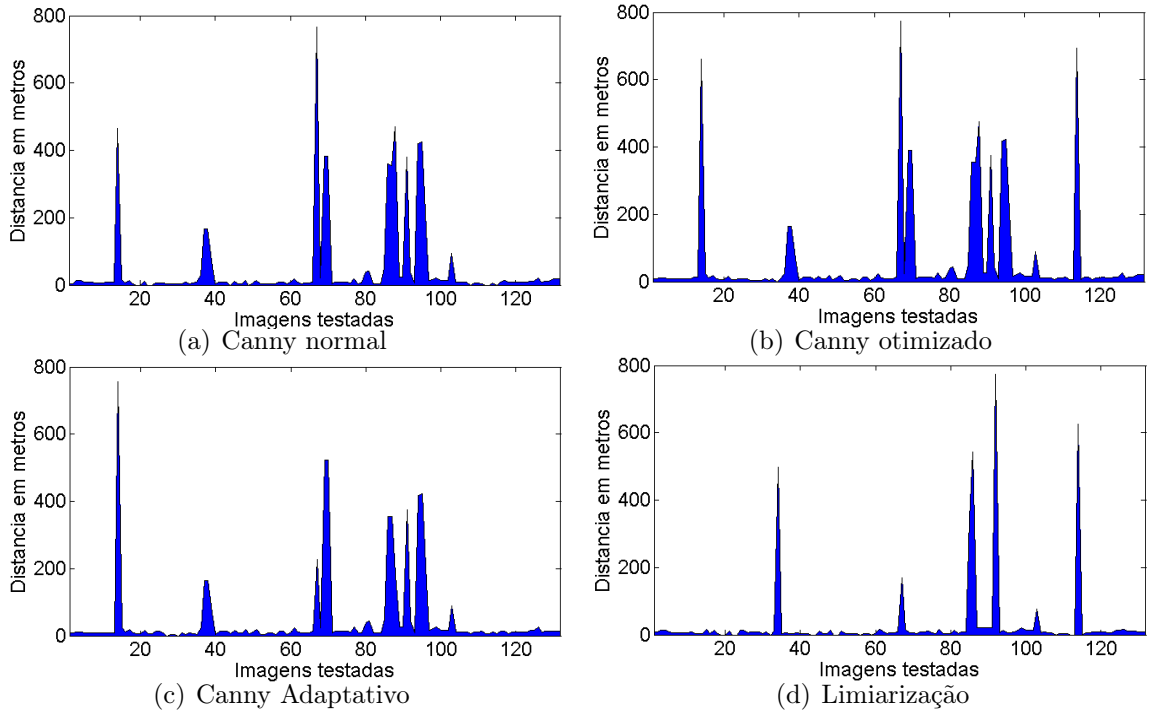
$$DP(distância) = \sqrt{\sum \frac{(distância_i - \overline{distância})^2}{\#\ imagens}} \tag{6.3}$$

Segundo os resultados consignados na tabela 6.1, ao longo de toda a rota planejada, o método com *canny* normal teve uma distância média baixa de $47,2$ metros, mas seu tempo de processamento foi de $8,499$ segundos, considerado como um tempo de processamento grande devido a um alto custo computacional. Em contraste, o método de *canny* com a função do cálculo da matriz de correlação otimizada oferece o tempo de processamento mais baixo de todos os métodos comparados, $0,52$ segundos, com uma distância média de $56,67$ metros. Portanto, neste método existe um ganho em tempo de processamento mas tem-se uma perda em precisão. Por sua parte, os resultados mostrados no método com *canny* adaptativo, fornecem uma distância média baixa de $47,44$ metros e um tempo de processamento baixo de $0,75$ segundos, sendo assim a melhor relação entre precisão e tempo de processamento de todos os métodos testados para este banco de imagens de diferente sensor. Por último, o método de estimação

de posição por limiarização com 30, 17 metros, apresenta a menor distância média de todos os métodos e um tempo de processamento de 1, 61 segundos. Com este método, obtêm-se uma maior precisão mas o tempo de processamento aumenta em relação a essa precisão. A figura 6.2 ilustra as distâncias obtidas pelos métodos aplicados para todas as 132 imagens de diferente sensor. O eixo y representa a distância em metros e o eixo x indica o número de imagens testadas.

Tabela 6.1: Resultados comparativos de desempenho nos diferentes métodos aplicados para a estimação de posição em imagens de diferente sensor.

Método	*# Imagens*	*Dist. média (m)*	*variância*	*Desv. padrão*	*FP(%) >100 m*	*VP(%) <100 m*	*Tempo Processamento (s)*
Canny normal	132	47,2005	1,46e+04	120,92	9,84(13)	90,15	8,499
Canny otimizado	132	56,6775	1,91e+04	138,20	10,60(14)	89,40	**0,5267**
Canny Adaptativo	132	47,7783	1,36e+04	116,76	9,84(13)	90,15	**0,7581**
Limiarização	132	**30,1748**	*1,20e+04*	*109,68*	**4,54(6)**	95,45	**1,6107**



(a) Canny normal (b) Canny otimizado

(c) Canny Adaptativo (d) Limiarização

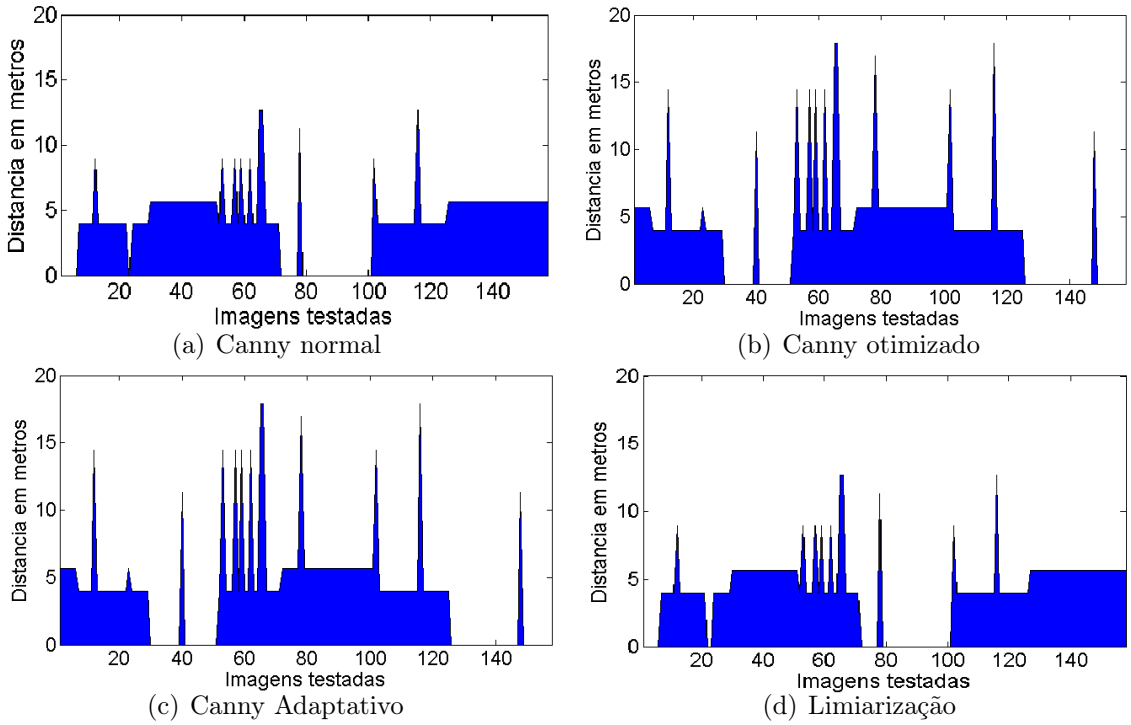
Figura 6.2: Imagens capturadas por VANT de diferente sensor (experimento 1).

A tabela 6.2 mostra os resultados comparativos após a aplicação dos métodos propostos ao experimento 2. É evidente que o método com *canny* normal apresenta o maior tempo de processamento, com 8, 394 segundos e uma distância média de 4, 0646 metros, definindo-se portanto, como o método com maior custo computacional. Os métodos de *canny* com a função do cálculo da matriz de correlação otimizada e *canny* adaptativo tiveram o mesmo desempenho para este experimento, com distância média de 3, 871 e tempo de processamento baixo de 0, 5269 e 0, 7556 segundos, respectivamente.

O método de estimação de posição por limiarização apresenta um bom desempenho, com uma distância média de $4,0288$ metros e um tempo de processamento de $1,6097$ segundos. Em comparação com os métodos anteriores o tempo de processamento é um pouco maior mas segue sendo bem baixo comparado com o método de *canny* normal, que é o método utilizado atualmente. A figura 6.3 ilustra as distâncias obtidos pelos métodos aplicados para todas as 158 imagens de igual sensor. Onde o eixo y representa a distância em metros e o eixo x indica o número de imagens testadas.

Tabela 6.2: Resultados comparativos de desempenho nos diferentes métodos aplicados para a estimação de posição em imagens de igual sensor.

Método	***# Imagens***	***Dist. média (m)***	***variância***	***Desv. padrão***	***FP(%) >100 m***	***VP(%) <100 m***	***Tempo Processamento (s)***
Canny normal	158	4,0646	7,3355	2,7084	0	100	8,394
Canny otimizado	158	3,871	15,6134	3,9514	0	100	*0,5269*
Canny Adaptativo	158	3,871	15,6134	3,9514	0	100	*0,7556*
Limiarização	158	4,0288	7,4232	2,7246	0	100	*1,6097*



(a) Canny normal (b) Canny otimizado

(c) Canny Adaptativo (d) Limiarização

Figura 6.3: Imagens capturadas por VANT de igual sensor (experimento 2).

De forma geral, os métodos de *canny* adaptativo e limiarização obtiveram um bom desempenho nos dois experimentos realizados. Com estes métodos propostos conseguiu-se reduzir consideravelmente o tempo de processamento que é exigido para realizar a estimação de posição geográfica em um par de imagens (capturada pelo VANT e geo-referenciada). Além disso, os valores de distância média entre a posição estimada pelos métodos propostos e a posição verdadeira foram baixos, distâncias médias menores a

50 metros. Particularmente, o método de *canny* adaptativo resultou ser $11,21$ vezes mais rápido que o método de *canny* normal sem ter perda de precisão na estimação de posição. Por sua parte, o método por limiarização resultou ser $5,27$ vezes mais rápido que o método de *canny* normal e além disso, teve melhoras na precisão de estimação de posição pois a distância média foi reduzida em 17 metros. Com isso, a relação de precisão versus tempo de processamento permitirá em futuros experimentos realizar testes em tempo real.

6.3 Correção de Distorção Projetiva com Método Paramétrico

Para a avaliação do desempenho deste método foram utilizadas imagens aéreas sub-amostradas com dimensão de 800×533 pixeis. As imagens capturadas pelo VANT foram obtidas de uma câmera SONY ILCE-7R com 24,3 Megapixels, distância focal $f_{cam} = 45\ mm$ e com dimensão de 7360×4912. A taxa de sub-amostragem aplicada foi $1/9,2$ da dimensão original, este sub-amostragem é realizado pois caso contrário, trabalhar com imagens de alta resolução aumentará consideravelmente o tempo de processamento na retificação da imagem. As imagens capturadas pelo VANT possuem variações de rotação nos ângulos (Yaw, Picth, Roll) e perspectiva.

A fim de testar o método proposto, foram pré-selecionadas 55 de 1096 imagens aéreas capturadas pelo VANT em um voo real sobre a região de São Carlos no Estado de São Paulo - Brasil (Anexo A2.3). Estas imagens possuem distorção geométrica devido a variações nos ângulos (Yaw, Pitch, Roll) da aeronave no momento de captura da imagem. Seis das 55 imagens com maiores ângulos de distorção foram escolhidas para ser mostradas nas figuras: 6.4 - 6.9, parte esquerda. Na parte direita, apresentam-se os resultados depois de realizar a correção da distorção projetiva nas imagens aplicando o método paramétrico proposto no Capítulo 4. Todo o conjunto de imagens foi fornecido pelo Instituto de Estudos Avançados da Aeronáutica IEAV, situado na cidade de São José dos Campos.



Figura 6.4: Imagem com distorção (esquerda) - Imagem corrigida (direita).

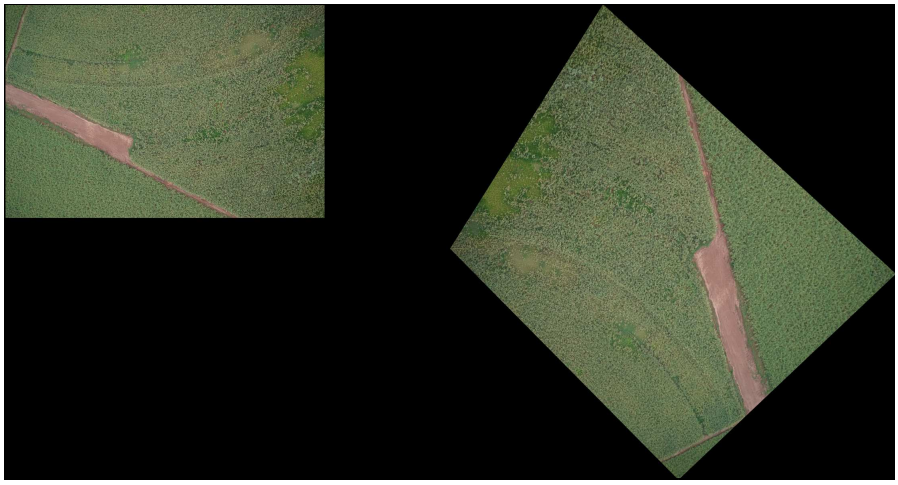


Figura 6.5: Imagem com distorção (esquerda) - Imagem corrigida (direita).



Figura 6.6: Imagem com distorção (esquerda) - Imagem corrigida (direita).

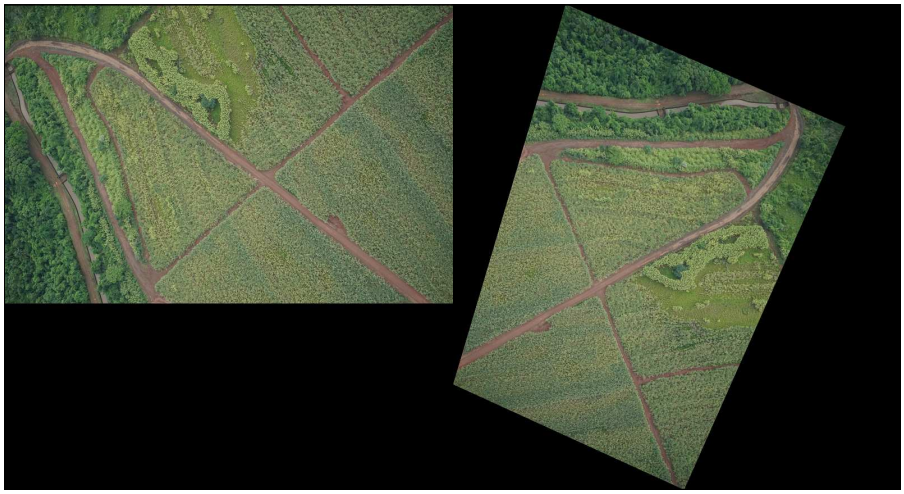


Figura 6.7: Imagem com distorção (esquerda) - Imagem corrigida (direita).

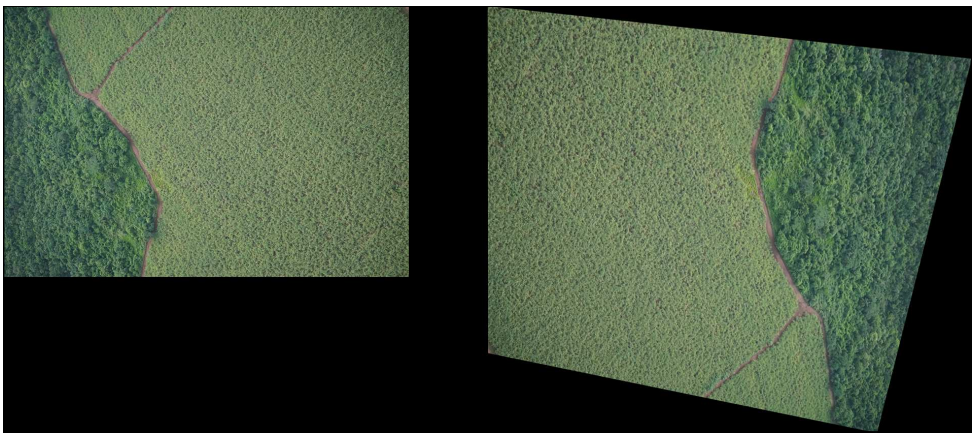


Figura 6.8: Imagem com distorção (esquerda) - Imagem corrigida (direita).

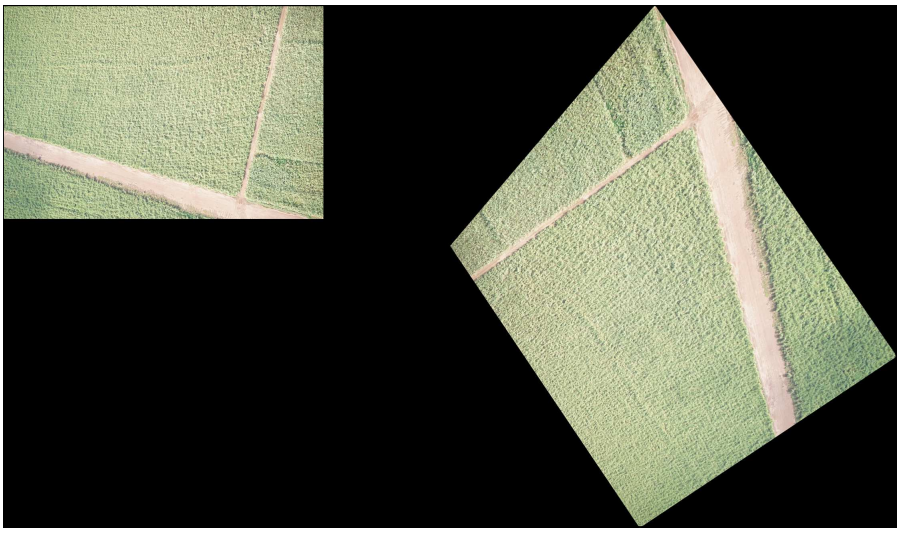


Figura 6.9: Imagem com distorção (esquerda) - Imagem corrigida (direita).

A Tabela 6.3 apresenta a latitude, longitude e os ângulos de inclinação (Yaw, Pitch, Roll) de cada imagem com distorção. Também, são mostrados os respectivos tempos de processamento que o método exige para corrigir cada imagem de teste.

Tabela 6.3: Ângulos (Yaw, Pitch, Roll) de cada imagem com distorção e seus respectivos tempos de processamento na correção. Latitude e Longitude mostra a posição da imagem aérea que foi capturada para comparar com Google Maps.

Imagem (Lat,Long)	***Ângulos de inclinação do VANT***			***Tempo de processamento [s]***
	Yaw	***Pitch***	***Roll***	**Matlab - OpenCV**
fig.6.4 (-22.07636831,-48.05341833)	112,6	4,9	7,7	0,4052 - **0,0390**
fig.6.5 (-22.07919877,-48.05222637)	224,3	16,7	11,4	0,4724 - **0,0400**
fig.6.6 (-22.08255284,-48.07370971)	250,9	8,5	-7,6	0,3982 - **0,0330**
fig.6.7 (-22.07686876,-48.05280350)	113	6,9	12,6	0,3792 - **0,0340**
fig.6.8 (-22.07975411,-48.05124583)	191,1	11,5	18,2	0,4184 - **0,0290**
fig.6.9 (-22.08146956,-48.05155315)	234,4	17,9	11,4	0,4678 - **0,0410**

Os resultados mostram que em todos os casos de teste avaliados, o método proposto consegue fazer correção de perspectiva nas imagens em um tempo de processamento baixo, na faixa de $0,37 - 0,47$ segundos. Este tempo de processamento é um fator importante para a estimação da posição de VANTs através da técnica de Casamento de Imagens, pois com um menor tempo de processamento na correção de perspectiva, vai-se garantir um maior número de estimativas de posição ao longo do voo. Paralelamente, o método também foi implementada em $C++$ baseado na biblioteca de *Processamento Digital de Imagens - OpenCV* 3.0, onde os resultados mostram que o tempo de processamento foi bem menor, ou seja foi 10 vezes mais rápido a resposta deste método em *OpenCV* que comparado com *Matlab*. Os tempos estiveram na faixa de $29 - 41$ milissegundos, fazendo ainda mais fatível a implementação em tempo real devido a seu baixo custo computacional.

6.4 Correção de Distorção Projetiva com Método não Paramétrico

Para este método não paramétrico proposto no Capítulo 5, os resultados foram classificados em duas partes. A primeira contém a análise do método não paramétrico aplicado a imagens aéreas obtidas do mesmo sensor e que apresentam distorção de perspectiva, diferenças de escala, rotação e translação. Na segunda parte, a análise foi focada nas imagens aéreas de diferente sensor que apresentam diferença espectral, distorção de perspectiva, diferença de escala, rotação e translação. O nível de complexidade nas imagens testadas na segunda parte é maior, pois a diferença de sensores e a distorção espectral, fazem que os algoritmos de detecção e casamento de pontos característicos

reduzam consideravelmente sua qualidade ou, em alguns casos, não consigam encontrar casamento dos pontos extraídos entre as imagens analisadas (georeferenciada e VANT). Para a avaliação do desempenho do método, foram utilizadas imagens com dimensão de 800×500 pixeis (imagens georeferenciadas e imagens capturadas pelo VANT). As imagens georeferenciadas possuem visada nadir e as imagens capturadas pelo VANT tem variações de escala, rotação, translação e perspectiva.

6.4.1 Imagens de Igual Sensor

Para cada conjunto de imagens analisado, com o propósito de observar os benefícios que o pré-processamento oferece, foi implementado o método não paramétrico sem a etapa de pré-processamento e, logo, comparado com todo o método, incluindo o pre-processamento. Tudo isto tem a finalidade de observar e ressaltar as vantagens que este pré-processamento oferece, e como ele melhora a qualidade no casamento de pontos característicos.

Foram testadas imagens em diferentes ambientes e diferentes tipos de áreas sobrevoadas, principalmente áreas da classe Urbana, áreas com presença de Vegetação e áreas com presença de Rodovias em setores semi-urbanos. Estas imagens apresentam perspectiva, diferenças de escala e rotação. As imagens mostradas nas figuras 6.10 - 6.16 apresentam: (a) Imagem com visada nadir (esquerda), Imagem com perspectiva (direita); (b) Casamento de pontos característicos sem pré-processamento; (c) Casamento de pontos característicos com pré-processamento; (d) Resultados da correção de distorção sem pré-processamento; (e) Resultados da correção de distorção com pré-processamento.

Como análise dos resultados, apresenta-se a tabela 6.4. Nela mostram-se para cada par de imagens, o número de pontos característicos detectados em cada imagem, o número de casamentos estabelecidos entre pares de imagens e o tempo de processamento exigido para cada conjunto de imagens. Além disso, também calculam-se a *eficiência* (Eq.(6.4)) e 1-*precisão* (Eq.(6.5)) para avaliar a robustez do algoritmo SURF.

$$eficiência = \frac{\#pontos\ característicos\ extraídos\ que\ foram\ casados}{\#pontos\ característicos\ extraídos} \tag{6.4}$$

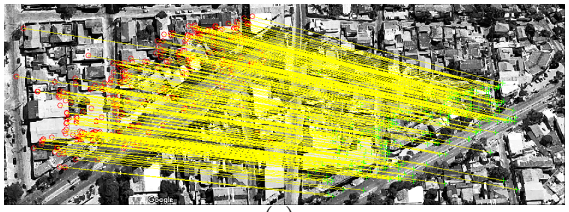
$$1 - precisão = \frac{\#\ casamentos\ rejeitados}{\#\ (casamentos\ aceptados + \ casamentos\ rejeitados)} \tag{6.5}$$



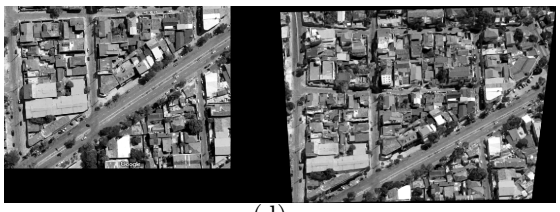
(a)



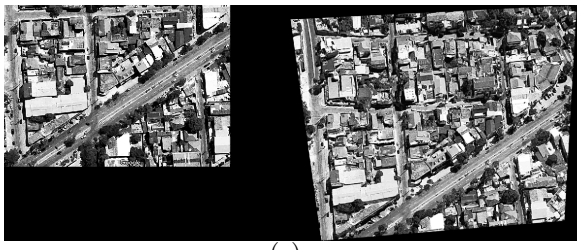
(b)



(c)



(d)

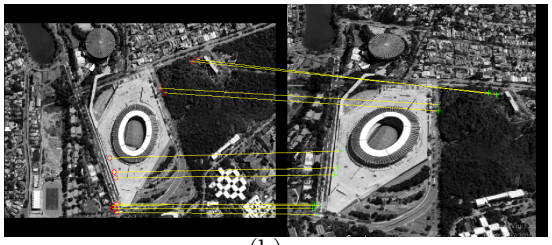


(e)

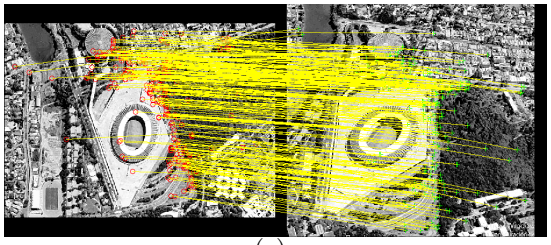
Figura 6.10: Imagem urbana.



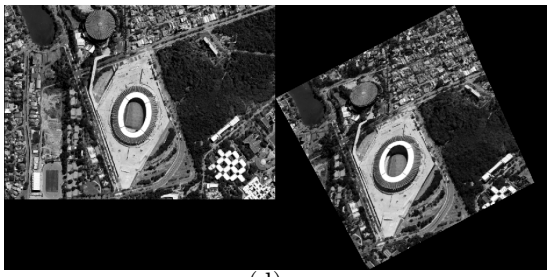
(a)



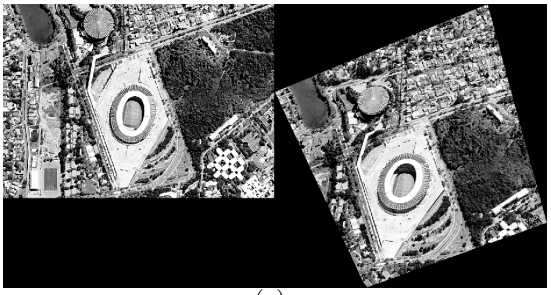
(b)



(c)

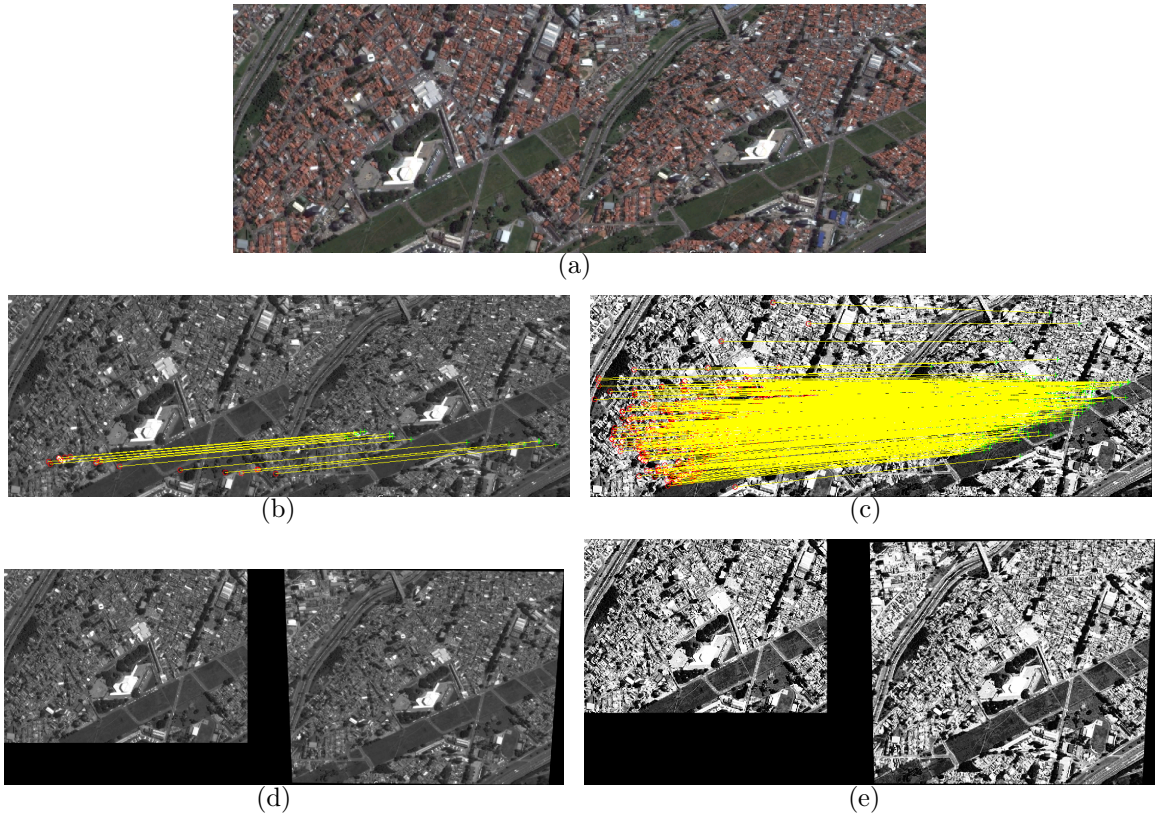


(d)



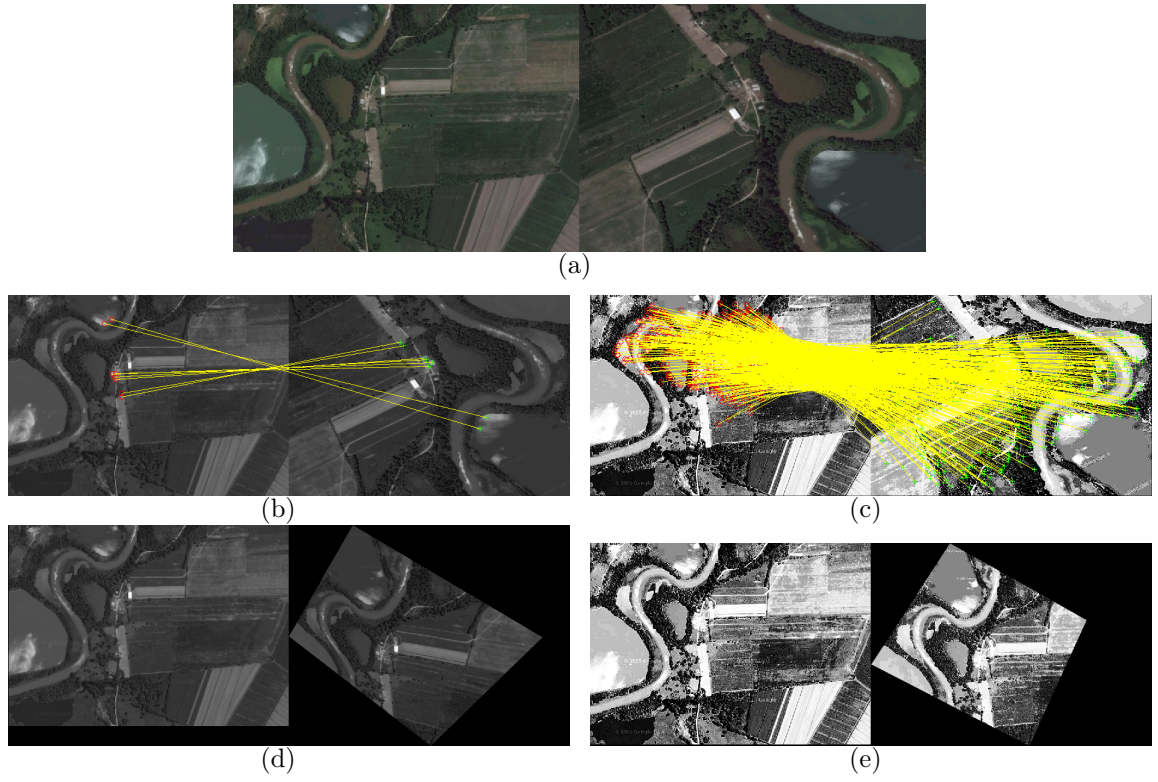
(e)

Figura 6.11: Imagem urbana (Mineirão).



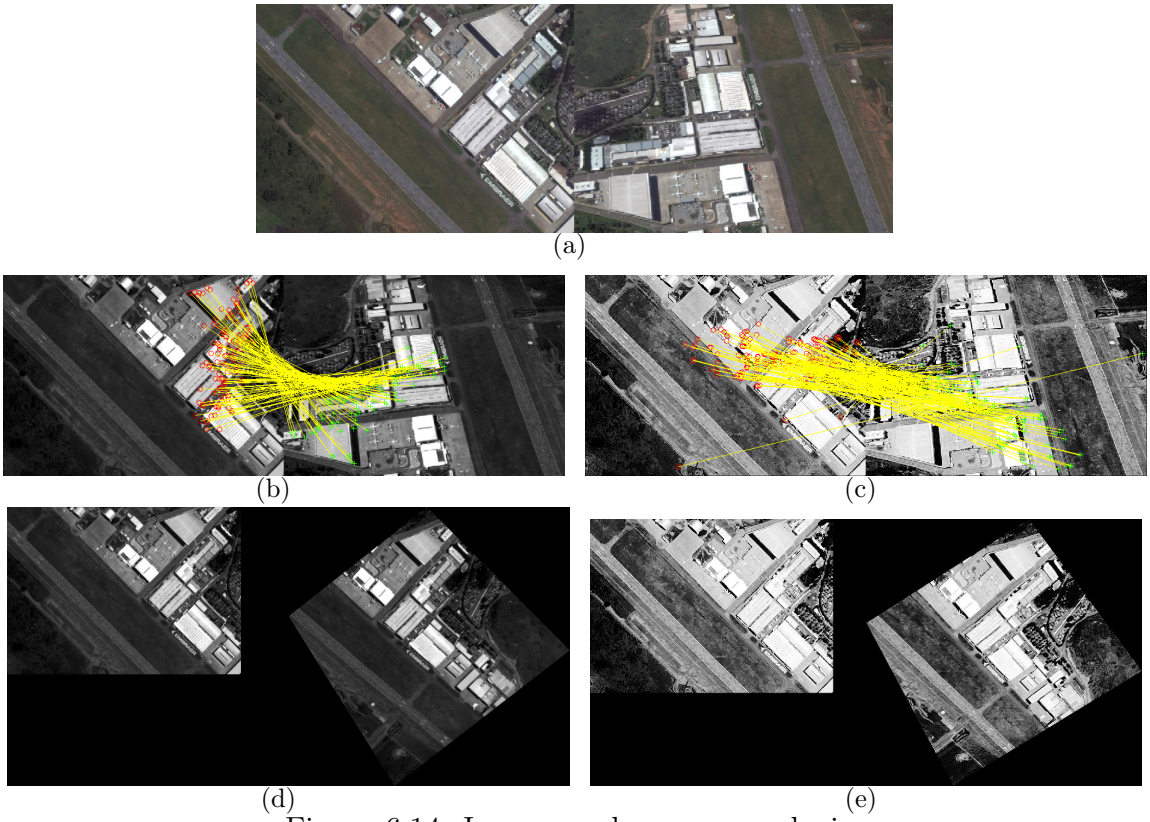
(a) (b) (c) (d) (e)

Figura 6.12: Imagem urbana com presença de gramado.



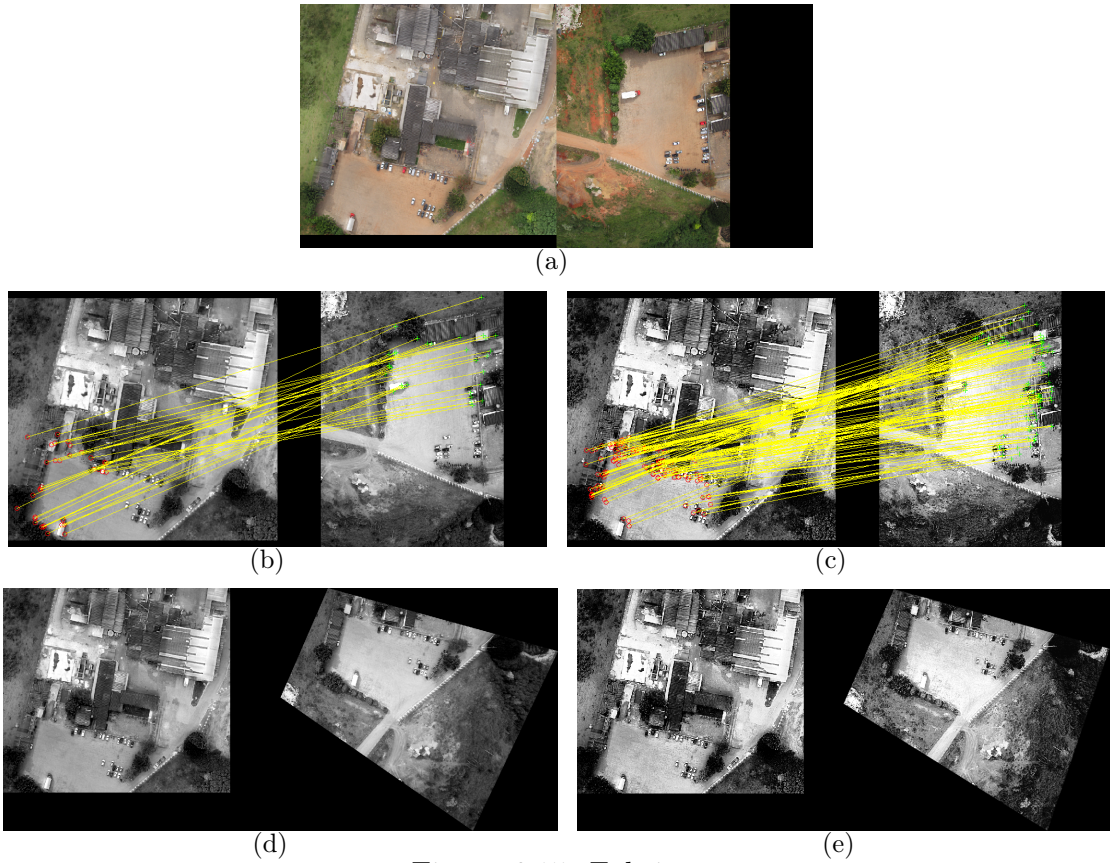
(a) (b) (c) (d) (e)

Figura 6.13: Imagem Vegetação.



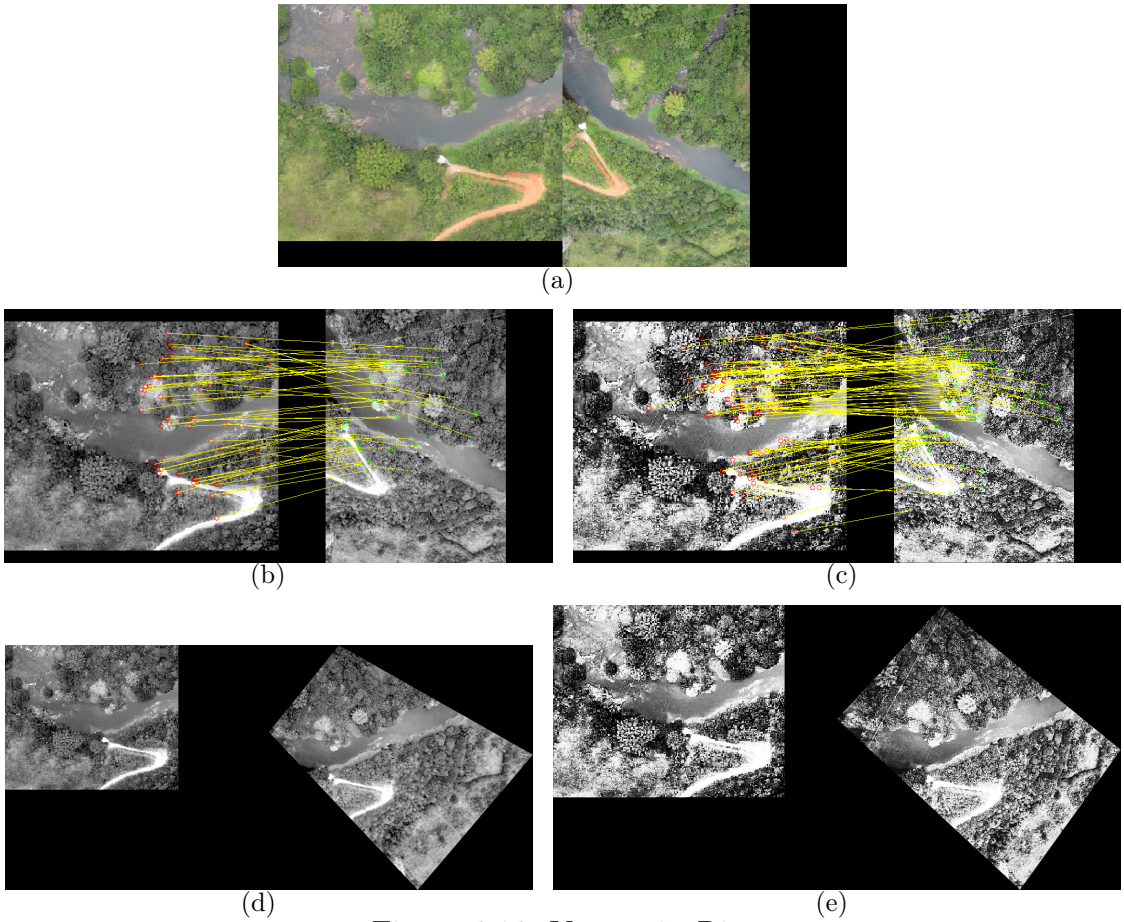
(a) (b) (c) (d) (e)

Figura 6.14: Imagem urbana com rodovia.



(a) (b) (c) (d) (e)

Figura 6.15: Fábrica.



(a) (b) (c) (d) (e)

Figura 6.16: Vegetação Rio.

Por sua vez, os resultados de *eficiência* e *1-precisão* obtidos no experimento, mostraram ser sempre superiores quando a etapa de pré-processamento é aplicada. A métrica de *eficiência* indica a relação de quanto foi o número de pares de pontos característicos extraídos em cada imagem que se tornaram casamentos, sobre o total de pontos característicos extraídos. A métrica *1-precisão* indica a capacidade que teve o algoritmo para rejeitar *outliers*, ou seja, a capacidade de excluir casamentos incorretos para assim, garantir os melhores casamentos que definiram o modelo da matriz homográfica H.

Conforme indicado na tabela 6.4, é evidente o aumento no número de casamentos quando o método proposto inclui a etapa de pré-processamento e com ela, o aumento no tempo de processamento também é evidenciado. No entanto, a relação do incremento no número de casamentos, mostrada na figura 6.17 (a) é bem maior comparado com o aumento no tempo de processamento, mostrado na figura 6.17 (b). Ou seja, a relação de casamento versus tempo de processamento quando a etapa de pré-processamento é aplicada, foi maior comparado com o método sem pré-processamento na maioria dos experimentos. Isto pode ser apreciado na figura 6.17 (c), onde essa relação foi estimada a partir da divisão entre o número de casamentos e o tempo de processamento gasto

Tabela 6.4: Resultados de extração e casamento de pontos característicos com e sem pré-processamento em imagens com igual sensor. Tempo de processamento para cada caso avaliado.

Imagem	*Método*	*Número de pontos característicos*		*Número de casamentos*	*Eficiência*	*1-precisão*	*Tempo de Processamento*
		Georeferenciada	*VANT*				*/s/*
Urbana	Sem pré-processamento	2045	2360	8	0,6842	0,9190	1,4237
	Com pré-processamento	4658	5208	**409**	0,7853	0,9251	1,9727
Mineirão	Sem pré-processamento	2045	2117	11	0,8217	0,8877	1,3136
	Com pré-processamento	4647	4883	**270**	0,9104	0,9285	1,8554
Urbana com presença de gramado	Sem pré-processamento	1240	1260	17	0,8236	0,8471	1,0960
	Com pré-processamento	4647	5085	**558**	0,8989	0,8829	1,8795
Urbana com rodovia	Sem pré-processamento	672	868	122	0,6247	0,8067	0,9934
	Com pré-processamento	2131	3147	**177**	0,7059	0,9071	1,0084
Vegetação	Sem pré-processamento	85	114	8	0,7298	0,7833	0,5921
	Com pré-processamento	3822	4031	**547**	0,9028	0,8669	1,4724
Vegetação e Rio	Sem pré-processamento	1852	1673	49	0,9706	0,9717	1,2868
	Com pré-processamento	6793	5075	**100**	0,9861	0,9847	3,4550
Fábrica	Sem pré-processamento	1195	787	36	0,9474	0,9424	0,8168
	Com pré-processamento	4064	2674	**148**	0,9599	0,9540	1,4289

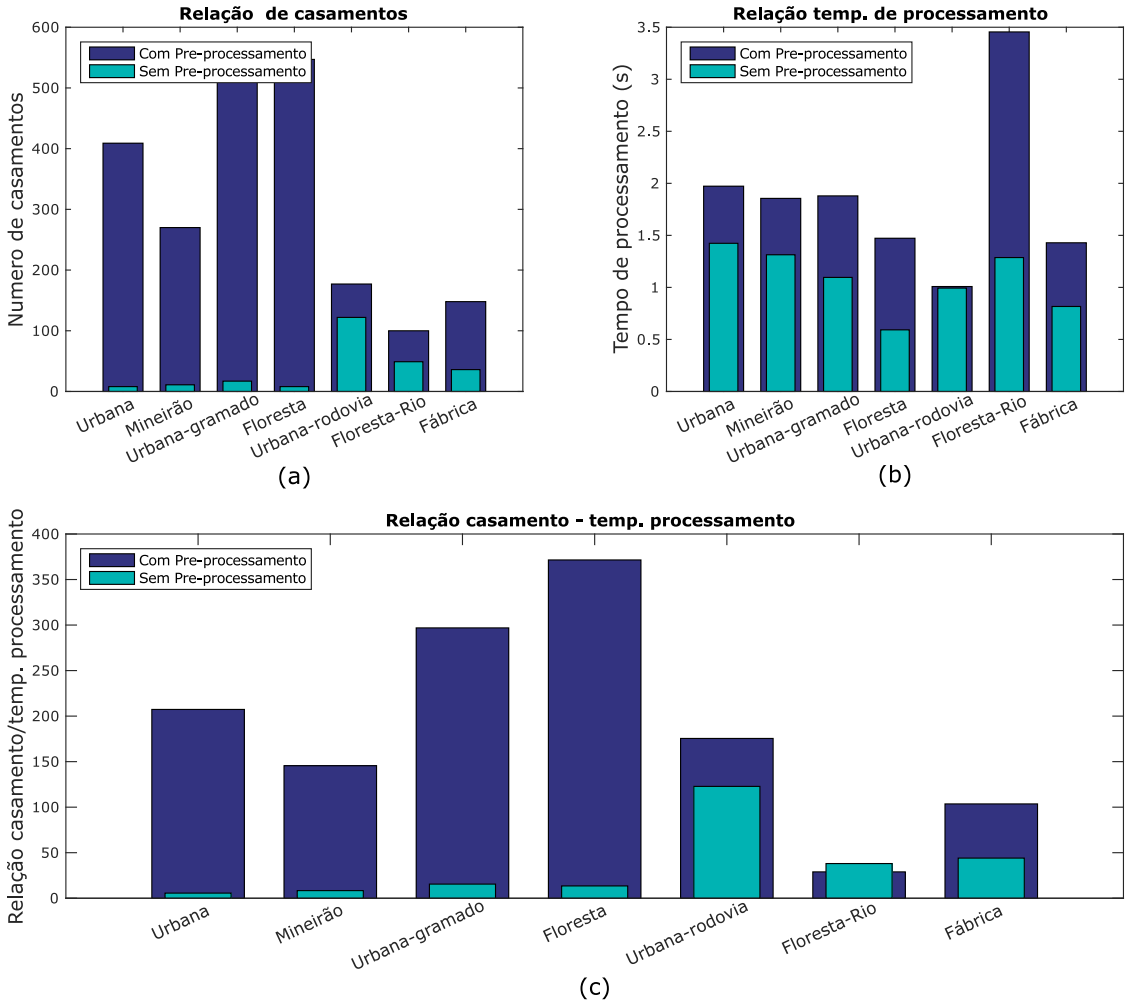


Figura 6.17: (a) Relação de número de casamentos com e sem pré-processamento; (b) Relação de tempos de processamento com e sem pré-processamento; (c) Relação de casamento vs. tempo de processamento.

na simulação. Portanto, se existe um valor grande nessa relação, fornecerá-se uma maior robustez no cálculo da matriz homográfica devido ao incremento de número de casamentos estimados em um tempo baixo de processamento. Dessa forma, se o objetivo da aplicação final é obter um grande volume de casamentos entre pontos característicos, a etapa de pré-processamento vai garantir esse objetivo. Por outro lado, o método aplicado sem pré-processamento obtêm um número baixo de casamentos que em algumas situações, devido ao tipo de imagem, por exemplo a imagem tipo Urbana, figura 6.10(b) ou Vegetação, figura 6.13(b) onde o número de casamentos foi menor, reduzindo a probabilidade de garantir uma boa correção de perspectiva na imagem capturada pelo VANT.

6.4.2 Imagens de Diferente Sensor

Nesta seção, implementou-se o método proposto na figura 5.1 para cada par de imagens mostrado nas figuras 6.18 - 6.22 (a) e (b). Estas imagens obtidas de diferentes sensores e ambientes, compõem os casos de teste avaliados na correção de distorção projetiva nas imagens capturadas pelo VANT. A fim de abarcar todas as classes existentes em um voo real, as imagens escolhidas contém presença de áreas urbanas combinadas com gramados, e também existem imagens com presença de rodovias predominantes e edificações ao redor. As imagens georeferenciadas em sua maioria, foram obtidas do satélite *Quickbird* e possui a característica de visada nadir. Por sua parte, as imagens capturadas pelo VANT são obtidas de um sensor diferente gerando uma diferença de resolução, brilho e luminosidade considerável. Também com variação de escala, rotação e com distorção projetiva.

Os resultados do método proposto foram comparados com os do método ASIFT (Yu and Morel, 2011), por ser uns dos algoritmos mais robustos e utilizados na literatura (Ye et al., 2015; Morel and Yu, 2009; Zheng et al., 2014b). O método ASIFT é uma versão melhorada do método SIFT que apresenta invariância à transformada afim, fazendo simulações de dois parâmetros (θ, ϕ) que são os que definem os valores de orientação do eixo da câmera que captura a imagem (Morel and Yu, 2009).

Pode-se detalhar que ambas imagens (figuras 6.18 - 6.22 (a) e (b)) correspondem a cenas diferentes, o que implica que existem diferenças em tempos de capturas e de sensores utilizados. A análise destes fatores permitiram avaliar a robustez do método. Por sua parte, as figuras 6.18 - 6.22 (c), mostram os resultados obtidos da aplicação do método proposto, uma vez que foi corrigida a distorção projetiva nas imagens capturadas pelo VANT. Assim, a imagem capturada pelo VANT é alinhada com a imagem georeferenciada, mostrando que, as duas imagens apresentam a mesma escala, orientação e que a perspectiva é retificada.



(a) (b)



(c)

Figura 6.18: O resultado do método proposto para o caso de teste Airport: (a) Imagem georeferenciada; (b) Imagem capturada pelo VANT; (c) A saída do método proposto.

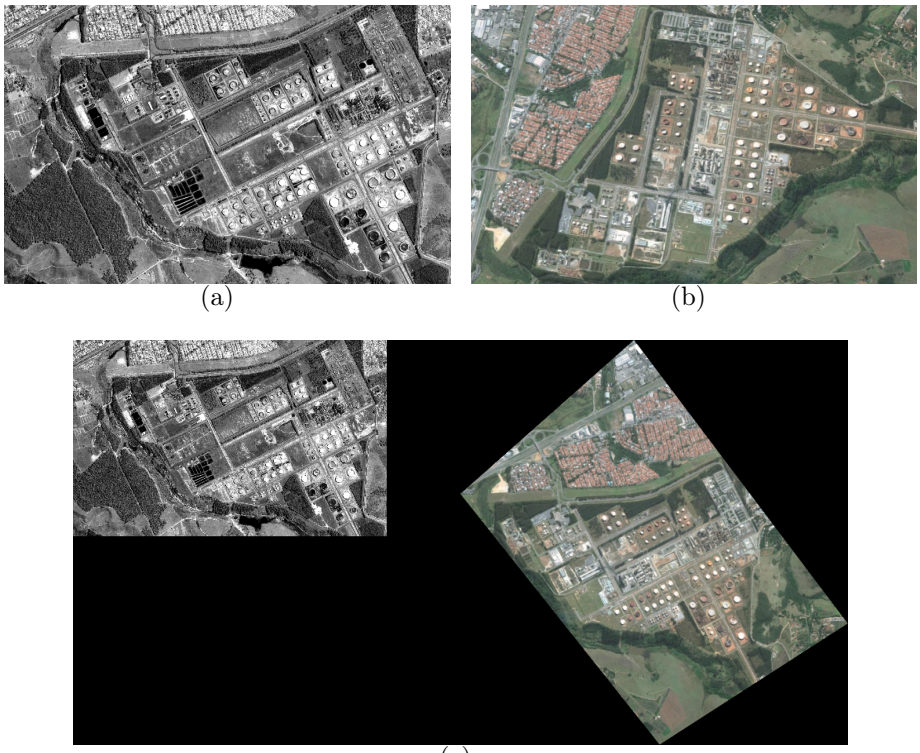


(a) (b)



(c)

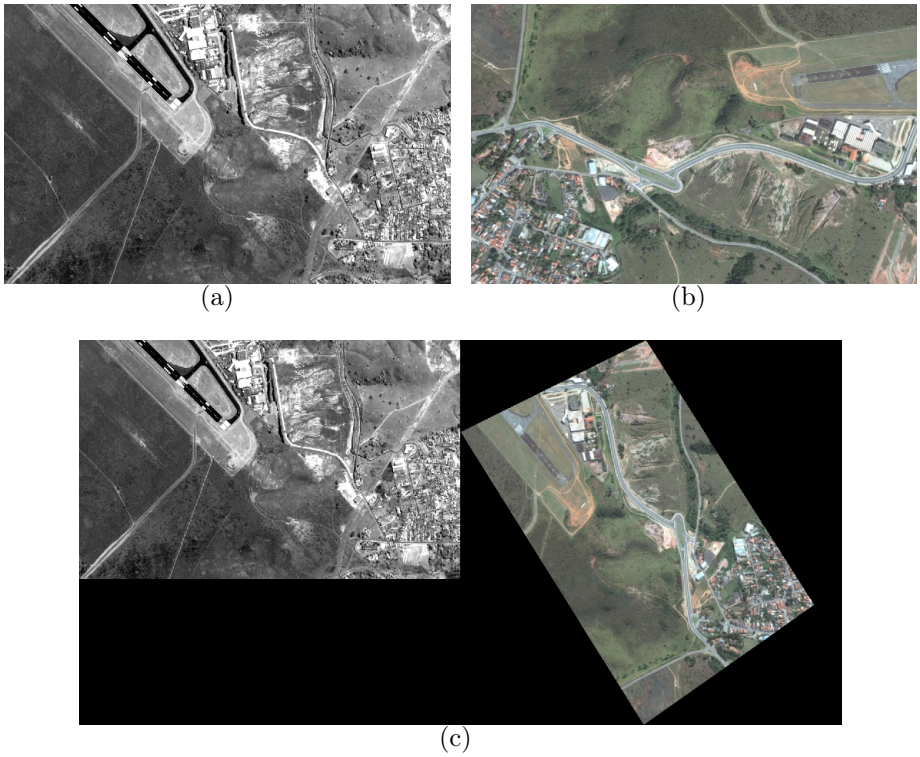
Figura 6.19: O resultado do método proposto para o caso de teste Stadium: (a) Imagem georeferenciada; (b) Imagem capturada pelo VANT; (c) A saída do método proposto.



(a) (b)

(c)

Figura 6.20: O resultado do método proposto para o caso de teste Tanks: (a) Imagem georeferenciada; (b) Imagem capturada pelo VANT; (c) A saída do método proposto.



(a) (b)

(c)

Figura 6.21: O resultado do método proposto para o caso de teste Bush: (a) Imagem georeferenciada; (b) Imagem capturada pelo VANT; (c) A saída do método proposto.

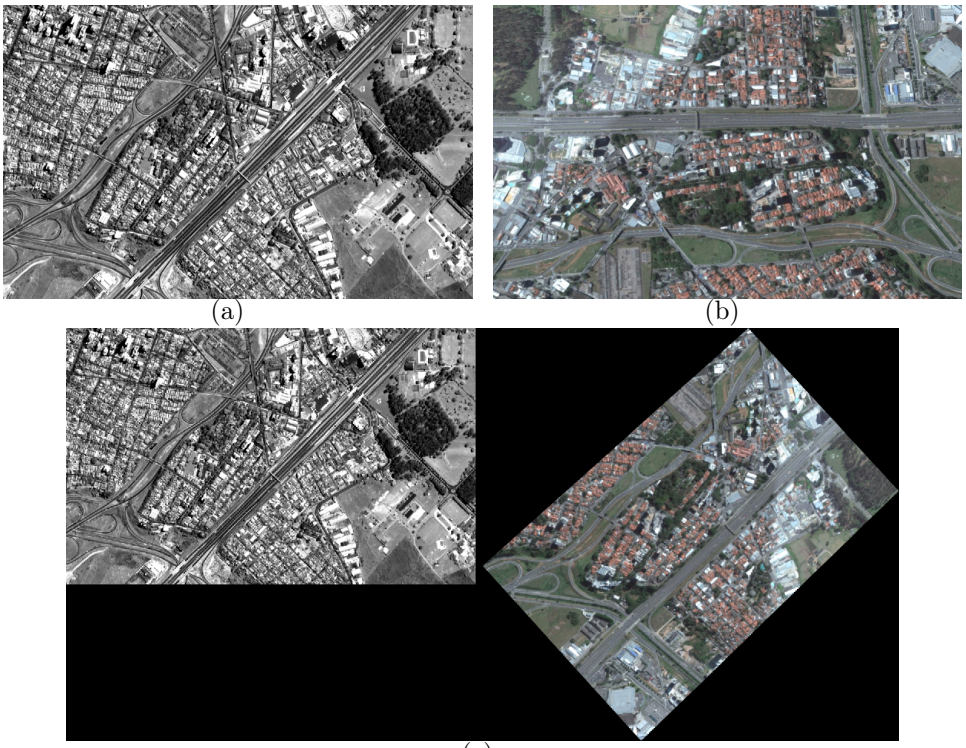


Figura 6.22: O resultado do método proposto para o caso de teste Highway: (a) Imagem georeferenciada; (b) Imagem capturada pelo VANT; (c) A saída do método proposto.

Na tabela 6.5, são comparados os resultados do algoritmo ASIFT e do método proposto para todas as imagens de teste. Primeiramente, é mostrado o número de pontos característicos detectados e o número de casamentos obtidos depois de 30 simulações para cada caso de teste. Em seguida, encontram-se os respectivos tempos de processamento.

Segundo os resultados, em todos os casos de teste avaliados, o algoritmo ASIFT apresentou um maior número de pontos característicos detectados comparado com o número de pontos característicos obtidos com o método proposto. Isto é porque o algoritmo SURF implementado neste trabalho não faz variação de escala nas imagens. Dessa forma, SURF não calcula pontos característicos de uma mesma posição (x, y) com diferente escala ou orientação, enquanto o algoritmo ASIFT sim, o que é refletido em um maior número de pontos característicos. Consequentemente, o tempo de processamento que cada algoritmo exige para realizar a extração de pontos característicos e também, o tempo de casamento das imagens está diretamente associado com o número de pontos característicos detectados.

Ainda na tabela 6.5 é observado que, o método proposto é consideravelmente mais rápido que ASIFT. A relação de tempos mostra que, em média o algoritmo ASIFT demora 17 vezes mais tempo que o algoritmo proposto. É importante ressaltar que

Tabela 6.5: Média de pontos característicos extraídos das imagens. Média, valor mínimo e máximo de casamento de pontos característicos nas imagens. Tempos de processamento obtidos de cada método em cada caso de teste depois de 30 execuções.

		Número de pts característicos		***Número de casamentos***			***Média exec.***
Imagem	***Método***	*Média Geo.*	*Média VANT*	*Média*	*Min*	*Max*	*tempo [s]*
Airport	ASIFT	37508	24293	3	0[a]	26	50,66
	Proposto	3799	3765	35	25	44	3,52
Stadium	ASIFT	17565	24335	0	0[b]	0	29,9
	Proposto	2168	2271	29	22	45	2,48
Tanks	ASIFT	45770	32936	20	19	22	74,83
	Proposto	4537	3912	42	30	61	3,88
Bush	ASIFT	29425	26251	10	0[c]	25	42,53
	Proposto	3493	3390	35	27	48	3,30
Highway	ASIFT	48267	41012	0	0[d]	0	92,43
	Proposto	4448	4090	40	30	52	3,91

[a] 27 dos 30 testes tiveram zero casamentos.
[b] 30 dos 30 testes tiveram zero casamentos.
[c] 15 dos 30 testes tiveram zero casamentos.
[d] 30 dos 30 testes tiveram zero casamentos.

para o cálculo do tempo de processamento foi incluída a etapa de pré-processamento no método proposto.

Finalmente, uma das métricas que melhor avalia a robustez do método proposto é o número de casamentos de pontos característicos obtidos para cada par de imagens de teste. A tabela 6.5 mostra que para todos os casos de teste, o número de casamentos obtidos com o algoritmo ASIFT foi sempre menor e, em alguns casos, não conseguiu realizar o casamento de pontos característicos entre a imagem georeferenciada e a imagem capturada pelo VANT. Em contraste, o método proposto obteve melhores resultados, o número de casamentos obtidos entre as imagens avaliadas variam em um intervalo de 22 a 61 casamentos de pontos de característicos. Este fato reafirma a robustez do método proposto para ser aplicado em situações onde a transformação projetiva deve ser feita com imagens obtidas a partir de sensores diferentes.

6.4.3 Caracterização dos Algoritmos SURF e MSAC

Nesta seção, tem como objetivo caracterizar e avaliar o desempenho do método proposto, a fim de encontrar o melhor ponto de operação dos algoritmos SURF e MSAC para todas as imagens de teste avaliadas. Métricas como eficiência, *recall* e *precisão* são utilizadas neste experimento. Os resultados mostraram que é possível reduzir consideravelmente o tempo de processamento que o método exige para corrigir as imagens. Dessa forma, tornando-se factível que o método proposto seja implementado em condições de voo real. Detalhes sobre o desenvolvimento deste experimento são apresentados

nos anexos A4.1, A4.2.

Os resultados observados na tabela comparativa 6.6, mostram que existe uma melhora considerável no desempenho e no tempo de processamento do método proposto após a caracterização dos algoritmos SURF e MSAC. Com o método trabalhando no melhor ponto de operação, foi possível aumentar o número de casamentos calculados entre os pares de imagens de teste, e ao mesmo tempo reduzir consideravelmente o tempo de processamento em cada execução. Assim, o método proposto apresenta um menor custo computacional, com grande potencial de ser embarcada e implementada em simulações de voo em tempo real.

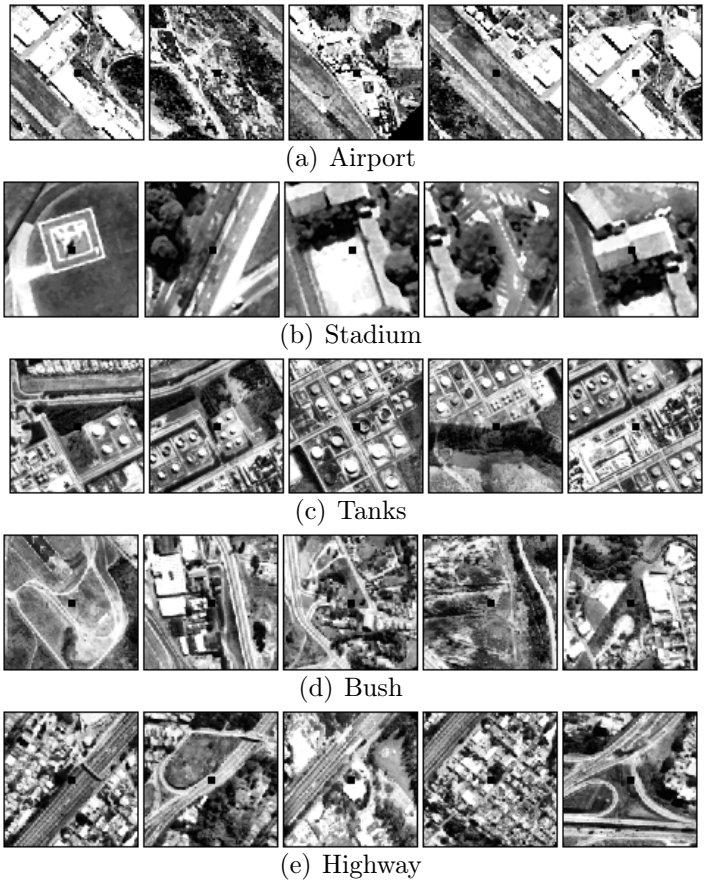
Tabela 6.6: Média de pontos característicos extraídos das imagens. Média, valor mínimo e máximo de casamento de pontos característicos nas imagens. Tempos de processamento obtidos de cada método em cada caso de teste depois de 30 execuções. Caracterizados os algoritmos SURF e MSAC e encontrar o melhor ponto de operação.

		Número de Keypoints		***Número de casamentos***			***Média exec.***
Imagem	***Método***	*Média Geo.*	*Média VANT*	*Média*	*Min*	*Max*	*tempo [s]*
Airport	ASIFT	37508	24293	3	0	26	50,66
	Proposto	3799	3765	35	25	44	3,52
	Prop. caracterizado	**1637**	**1541**	**41**	**33**	**51**	**0,9993**
Stadium	ASIFT	17565	24335	0	0	0	29,9
	Proposto	2168	2271	29	22	45	2,48
	Prop. caracterizado	**1224**	**1265**	**58**	**35**	**68**	**0,6616**
Tanks	ASIFT	45770	32936	20	19	22	74,83
	Proposto	4537	3912	42	30	61	3,88
	Prop. caracterizado	**1792**	**1694**	**47**	**32**	**55**	**1,0987**
Bush	ASIFT	29425	26251	10	0	25	42,53
	Proposto	3493	3390	35	27	48	3,30
	Prop. caracterizado	**1277**	**1424**	**43**	**32**	**50**	**0,8160**
Highway	ASIFT	48267	41012	0	0	0	92,43
	Proposto	4448	4090	40	30	52	3,91
	Prop. caracterizado	**1949**	**1893**	**43**	**32**	**54**	**1,2058**

6.4.4 Estimação de Posição para Imagens de Diferente Sensor

A fim de avaliar a precisão na estimação de posição com a técnica de Casamento de Imagens sobre as imagens que foram o resultado de aplicar nosso método de correção de perspectiva em imagens com diferença espectral e com visada não nadir (Capítulo 5), realizou-se o seguinte: (1) selecionaram-se 5 diferentes sub-regiões de forma aleatória para cada imagem capturada pelo VANT que foi corrigida por nosso método (figura 6.23). Cada sub-região contém sub-regiões de terreno diferentes com a finalidade de testar a precisão da estimação de posição após de ser aplicado o método de correção nas imagens. Cada sub-região representa uma janela de 90×90 pixeis, e seu ponto central corresponde à posição em termos de latitude e longitude onde o VANT fez a

captura da imagem. (2) identificaram-se as sub-regiões sobre a imagem georeferenciada utilizando a técnica de Casamento de Imagens. (3) calculou-se a diferença em metros entre a posição conhecida e a identificada em termos de latitude e longitude.



(a) Airport

(b) Stadium

(c) Tanks

(d) Bush

(e) Highway

Figura 6.23: Seleção de sub-regiões para cada imagem de teste.

Depois de selecionar as sub-regiões, a extração de características em cada imagem é feita através da extração de bordas, para este propósito, foi utilizado o detector de bordas *canny* adaptativo. Seguidamente, é calculada a matriz de correlação espacial entre pixeis, onde é identificado o ponto máximo da matriz de correlação. Este ponto máximo é definido como o pixel central da janela de 90×90 pixeis que compõe o novo padrão identificado que da a posição geográfica do VANT e a informação de latitude e longitude estimada na imagem georeferenciada. Na figura 6.24 é mostrado um exemplo do procedimento de reconhecimento de padrões e estimação de posição feita para cada uma das sub-regiões selecionadas em imagens de diferente sensor.

A tabela 6.7, mostra os resultados do reconhecimento de padrões e a estimação de posição (latitude e longitude) entre diferentes sub-regiões selecionadas na imagem do VANT e a imagem georeferenciada. Para isto, estimou-se a posição do padrão de cada sub-região identificada na imagem georeferenciada (latitude e longitude) e calculou-se a distância em metros entre essa posição e a posição da sub-região analisada.

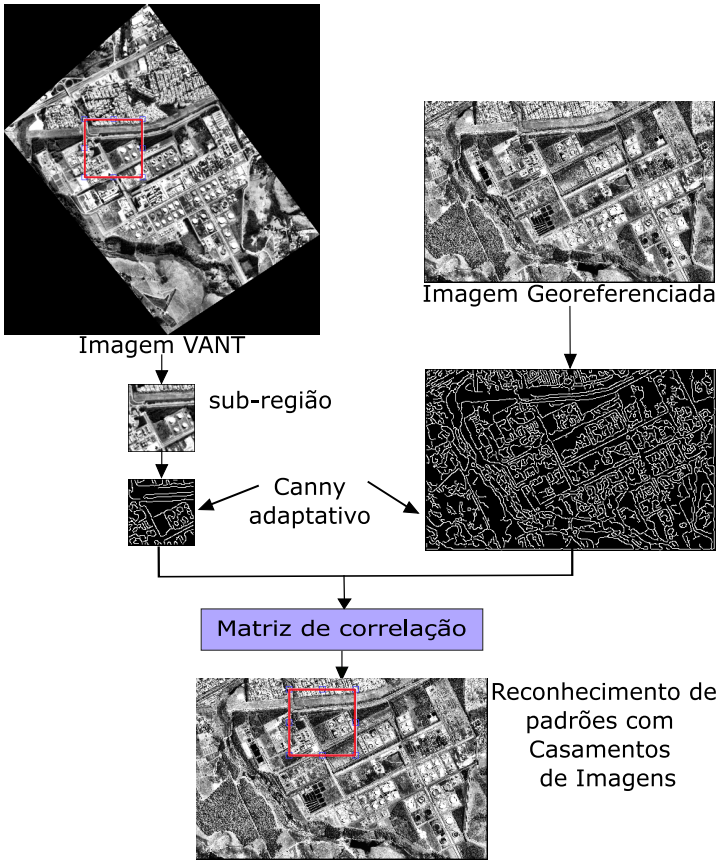


Figura 6.24: Fluxograma de Reconhecimento de Padrões e estimação de posição com a técnica Casamento de Imagens de diferente sensor.

Os resultados mostram que, a posição dos padrões identificados sobre a imagem georeferenciada foram muito próximos, e que as imagens identificadas correspondem à mesma área que a das sub-regiões selecionadas como pode ser apreciado na figura 6.25. As distâncias foram menores a 70 metros no caso da imagem Airport e na média de 50,81 metros. No caso da imagem Stadium a precisão foi bem maior, as distâncias foram em média de 2,19 metros. Para o caso da imagem Bush, a distância em média também foi baixa mantendo um valor de 10,81 metros. No caso das imagens Tanks e Highway os resultados nas distâncias tiveram um comportamento similar, os valores em média foram de 22,45 e 17,40 metros, respectivamente. Conforme observado, foi obtida uma boa estimação de posição e acerto no reconhecimento do mesmo padrão em outra imagem com o método proposto neste trabalho, mesmo utilizando informação de sensores totalmente diferentes. Assim, uma boa precisão na estimação de posição através de imagens é em grande parte garantida se existe uma boa correção de perspectiva nas imagens capturadas pelo VANT.

Tabela 6.7: Estimativa de posição entre as sub-regiões da imagem capturada pelo VANT e a imagem georeferenciada utilizando a técnica de Casamento de Imagens - *template matching*.

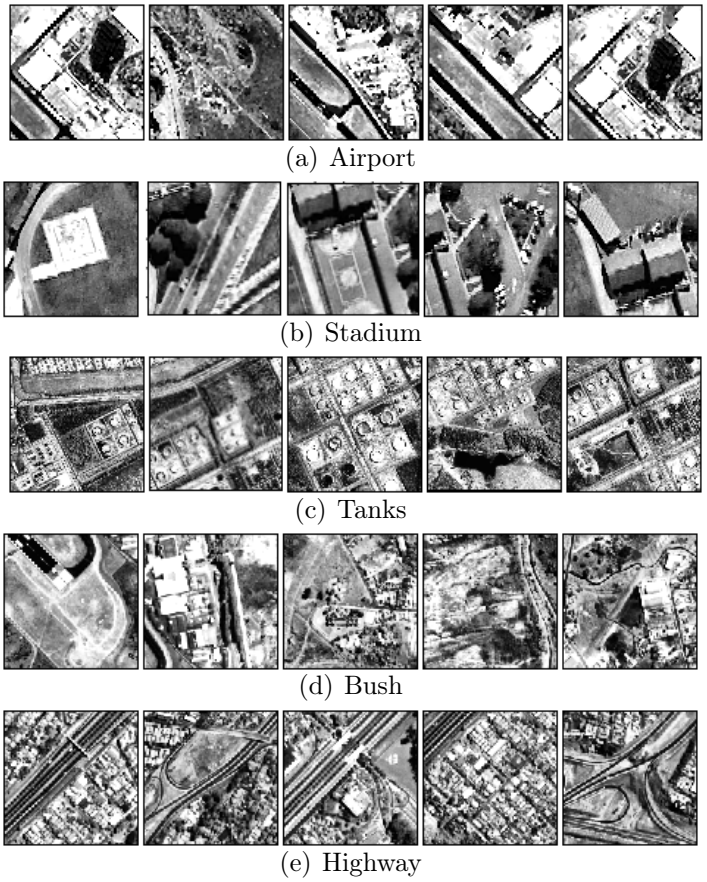
Imagem	***Posição das sub-regiões nas Imagens capturadas pelo VANT [lat,lon]***	***Posição estimada na Imagem Georeferenciada [lat,lon]***	***Distância em metros [m]***	***Média [m]***
Airport	-23.227648°,-45.858111°	-23.227663°, -45.857970°	15,7152	50,8126
	-23.228700°,-45.868022°	-23.229368°,-45.867770°	58,9480	
	-23.234497°, -45.851103°	-23.234700°, -45.852377°	70,2527	
	-23.225497°, -45.862714°	-23.225867°, -45.863248°	65,9421	
	-23.225979°,-45.858263°	-23.226266°,-45.857930°	43,2054	
Stadium	-23.194998°, -45.874642°	-23.195012°, -45.874640°	1,1094	2,1929
	-23.196000°, -45.874073°	-23.196001°, -45.874056°	1,8911	
	-23.195538°, -45.874168°	-23.195529°, -45.874150°	2,1192	
	-23.195588°, -45.873902°	-23.195600°, -45.873901°	0,9383	
	-23.195365°,-45.874109°	-23.195315°, -45.874136°	4,9067	
Tanks	-23.187477°, -45.833300°	-23.187447°,-45.833446°	16,3924	22,4549
	-23.187222°, -45.822812°	-23.187140°,-45.822953°	16,9178	
	-23.196229°, -45.818409°	-23.196028°,-45.818264°	22,4445	
	-23.200992°,-45.826438°	-23.200650°,-45.826312°	30,0395	
	-23.190936°,-45.822653°	-23.191007°,-45.822420°	26,4787	
Bush	-23.239361°, -45.851278°	-23.239234°,-45.851217°	11,9696	10,8138
	-23.236565°, -45.849537°	-23.236661°,-45.849563°	7,9970	
	-23.243996°, -45.844198°	-23.243984°,-45.844252°	2,9233	
	-23.237857°, -45.846221°	-23.237690°,-45.846301°	15,7271	
	-23.241028°, -45.843403°	-23.240975°,-45.843537°	15,4524	
Highway	-23.208125°, -45.887900°	-23.208378°,-45.887562°	42,3914	17,4010
	-23.205420°, -45.890706°	-23.205411°,-45.890672°	3,8431	
	-23.205416°, -45.883582°	-23.205291°,-45.883595°	9,8102	
	-23.210890°, -45.888042°	-23.210725°,-45.888098°	14,2396	
	-23.209692°, -45.894621°	-23.209603°,-45.894758°	16,7208	

6.5 Análise do Vídeo

O objetivo deste experimento é avaliar os métodos propostos neste trabalho em uma simulação de voo real, onde seja integrado o método de correção de perspectiva e o método de estimação de posição via Casamento de Imagens. Com isto, quer-se melhorar a estimação de posição, dar uma maior robustez e precisão no ajuste e correção da imagem capturada pelo VANT antes de ser procurada sobre a imagem georeferenciada.

Foi disponibilizado um banco de imagens adquirido partir de um voo de um VANT sobre uma região de São Carlos no estado de São Paulo - Brasil. A XMobots, empresa brasileira especializada em VANTs, forneceu ao projeto PITER (Processamento de Imagens em Tempo Real) um vídeo de uma área com predominância de classe Vegetação. Um exemplo das imagens obtidas do vídeo é mostrada na figura 6.26. A imagem georeferenciada utilizada para este experimento encontra-se no anexo A3.1.

O processo geral de retificação da imagem capturada pelo VANT foi feito baseado no método proposto no Capítulo 4, onde é feita a correção de perspectiva e o ajuste de escala e rotação na imagem do VANT de forma paramétrica. Em seguida, a imagem



(a) Airport

(b) Stadium

(c) Tanks

(d) Bush

(e) Highway

Figura 6.25: sub-regiões encontrados utilizando *template matching* em cada imagem.



Figura 6.26: Imagem capturada pelo VANT na região de São Carlos-SP.

capturada pelo VANT já corrigida é procurada sobre a imagem georeferenciada e, posteriormente, foi calculada a diferença em metros com respeito à posição (latitude, longitude) que foi identificada e estimada sobre a imagem georeferenciada. Na figura 4.3 é mostrado o procedimento de correção e ajuste da imagem capturada pelo VANT utilizando o método paramétrico e o processo de estimação de posição feita para cada

uma das imagens que foram capturadas pelo VANT.

Dessa forma, foram analisados dois trechos de rota sobre a mesma região de voo. Foi feita uma análise comparativa entre os resultados da estimação de posição obtida com nosso método onde é incluída a correção de perspectiva, ajuste de escala e rotação na imagem do VANT (fig.4.3). Seguidamente, isto foi comparado com a estimação de posição utilizando só o ajuste de rotação e escala (fig.2.1). Nas figuras 6.27 e 6.29 são mostrados os resultados da estimação de posição para cada um dos trechos de rota analisados, onde a latitude e longitude da posição real é mostrada em vermelho, da posição estimada com nosso método mostrado em cor amarelo e, finalmente, em cor verde, a estimação de posição sem correção de perspectiva, ou seja, com apenas o ajuste de rotação.

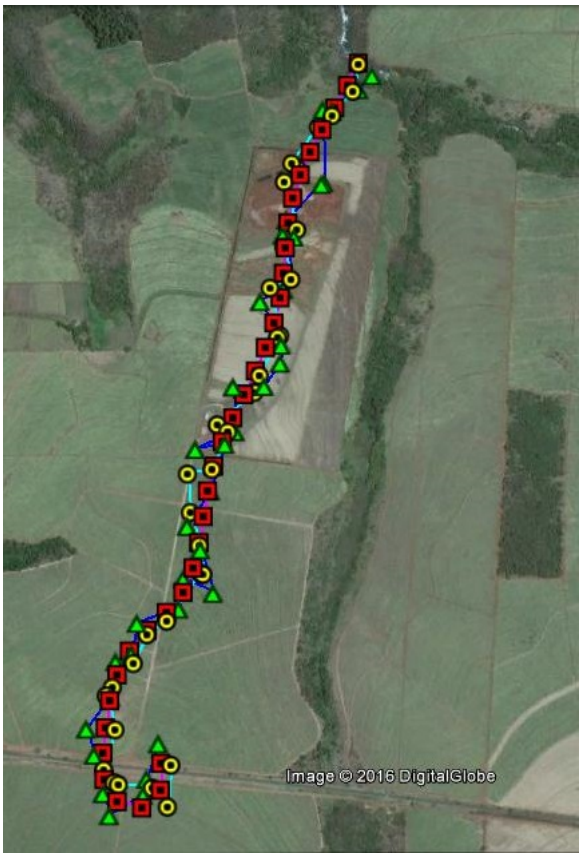


Figura 6.27: Resultado da estimação de posição para o trecho 1.

As figuras 6.28 e 6.30 mostram o comportamento das distâncias de separação em metros entre a posição real e as posições estimadas incluindo a correção de perspectiva e só com rotação para todas as imagens pertencentes a cada trecho de rota simulada. Também, foram geradas umas tabelas comparativas 6.8 e 6.9 que, mostram o desempenho da aplicação de correção de perspectiva versus a aplicação de rotação no processo de estimação de posição de VANTs através de imagens. Foram calculados e consignados nas tabelas a distância média, o desvio padrão, a variância, a porcentagem de falsos

positivos (FP) (número de estimativas de posição na qual sua distância média esteve acima de 100 metros) e verdadeiros positivos (VP) para cada trecho de rota simulado.

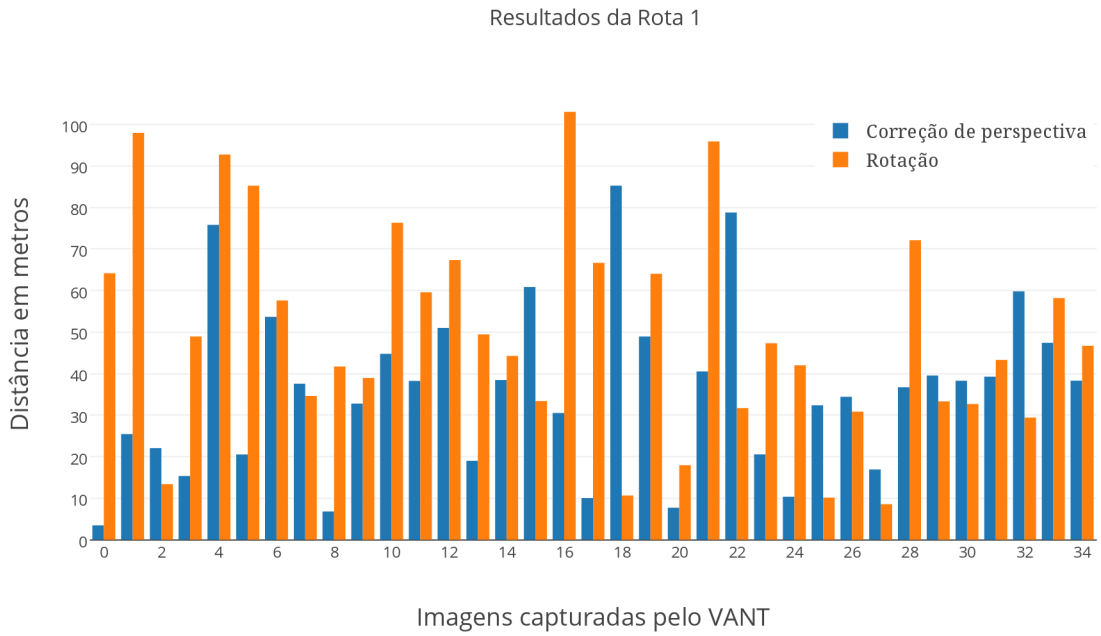


Figura 6.28: Distância média calculada para o trecho 1.

Tabela 6.8: Desempenho da aplicação de Correção de perspectiva vs. rotação para estimação de posição na rota 1

Método	Distância média	Desvio padrão	Variância	FP(%)	VP(%)
Correção de perspectiva	36,0447	20,2698	410,8649	0	100 %
Rotação	50,0113	25,6991	660,4414	2,8571 %	97,1428 %

Dos resultados da tabela 6.8 pode-se concluir que para o trecho de rota 1, a estimação de posição aplicando o método de correção de perspectiva, ajuste de escala e rotação na imagem do VANT antes de calcular a matriz de correlação, obteve melhor desempenho que aplicando só o ajuste de rotação. Com a correção de perspectiva, a distância média de separação entre a posição estimada e a posição real foi 14 metros menor em comparação com a distância média obtida com rotação. Os valores de desvio padrão e variância pertencentes à correção de perspectiva foram também menores que os correspondentes à rotação. Além disso, com a correção de perspectiva não houve presença de FP, enquanto com a rotação foi detectado um FP.



Figura 6.29: Resultado da estimação de posição para o trecho 2.

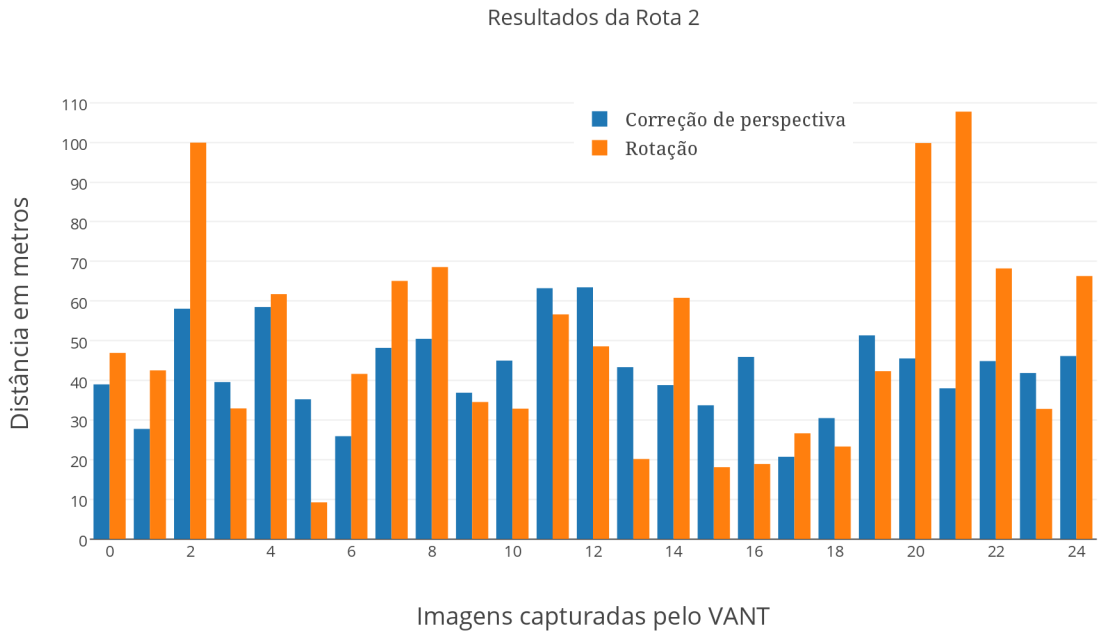


Figura 6.30: Distância média calculada para o trecho 2.

Tabela 6.9: Desempenho da aplicação de Correção de perspectiva vs. rotação para estimação de posição na rota 2

Método	**Distância média**	**Desvio padrão**	**Variância**	**FP(%)**	**VP(%)**
Correção de perspectiva	42,8718	11,0206	121,4536	0	100 %
Rotação	49,0536	26,5257	703,6140	2,8571 %	97,1428 %

Os resultados na tabela 6.9 mostram que para o trecho de rota 2, a estimação de posição aplicando o método de correção de perspectiva, ajuste de escala e rotação na imagem do VANT antes de calcular a matriz de correlação, obteve melhor desempenho que aplicando só o ajuste de rotação. Com a correção de perspectiva, a distância média de separação entre a posição estimada e a posição real foi $6,2$ metros menor em comparação com a distância média obtida com rotação. Os valores de desvio padrão e variância pertencentes à correção de perspectiva foram também menores que os correspondentes à rotação. Além disso, com a correção de perspectiva não houve presença de FP, enquanto com a rotação foi detectado um FP.

Capítulo 7

Conclusões e Trabalhos Futuros

7.1 Conclusões

Apresentados os métodos e seus respectivos resultados, esta secção busca ressaltar algumas considerações importantes deste trabalho.

7.1.1 Quanto à Estimação de Posição via Casamento de Imagens

Dois novos métodos foram apresentados para fazer estimação de posição via Casamento de Imagens capturadas por VANTs e de satélite previamente georeferenciada. Com esses métodos foi possível dar solução aos problemas de sobrestimação de bordas e inclusão de ruído que afetam à comparação das imagens no cálculo da matriz de correlação e, consequentemente, à estimação de posição. Neste aspecto, concluiu-se que os métodos de *canny* adaptativo e limiarização apresentaram um bom desempenho. Através da aplicação destes métodos, foi possível reduzir consideravelmente o tempo de processamento e obter valores baixos de distância média em comparação com os demais métodos testados.

7.1.2 Quanto à Correção de Distorção Projetiva

- **Método Paramétrico**

  Este novo método foi proposto para fazer correção de distorção projetiva em imagens capturadas por VANTs quando a câmera não tem visada nadir. A novidade deste método radica principalmente no fato de que é utilizada como referência a informação dos ângulos de inclinação do VANT e os parâmetros intrínsecos da câmera para definir a matriz de homografia e fazer a transformação projetiva nas imagens. Assim, o método demostra ser independente das características da

imagem e não requer de descritores, nem extração de características nela. Além disso, é um método que não se vê afetado por diferenças em tempos de captura ou mudanças espectrais ou físicas que possam ocorrer na região sobrevoada.

Os resultados dos experimentos mostram que todo o método exige um curto tempo de processamento na correção. Portanto, é factível de ser implementado em condições reais de voo devido a seu baixo custo computacional. Em adição, o método foi desenvolvido em dois entornos de programação: Matlab e OpenCV, permitindo que seja testado com maior facilidade.

- **Método não Paramétrico**

  Apresentou-se um novo método não paramétrico para fazer correção de distorção projetiva em imagens capturadas por VANTs quando a câmera não tem visada nadir. A novidade deste método está principalmente na correção de distorção projetiva em imagens com diferentes sensores (aérea e georeferenciada). Além das duas imagens serem capturadas em tempos diferentes, a imagem capturada pelo VANT apresenta diferença de escala, rotação, translação e perspectiva. O método proposto, obteve um melhor desempenho que o algoritmo ASIFT na estimação de casamento de pontos característicos e no tempo de processamento. Conclui-se que a correção da distorção projetiva nas imagens de teste foi feita adequadamente conforme foi visto na secção de resultados.

  Também, demostrou-se que a inclusão de uma etapa de pré-processamento no método proposto, melhora os resultados significativamente com um aumento no casamento de pontos característicos e, consequentemente, na correção da distorção projetiva. De modo que os histogramas representativos nas imagens tornaram-se equivalentes e equalizados entre eles, reduzindo as distorções espectrais causadas pela diferença de sensores. Portanto, uma etapa de pré-processamento antes de utilizar qualquer algoritmo extrator de pontos característicos e casamento em imagens (ASIFT, SIFT, SURF), permitirá a obtenção de um maior número de resultados. Isto porque a implementação de diferencia de Gaussiana (DoG) convoluida com a imagem nos algoritmos empregados para fazer o casamento, fornece imagens suavizadas em diferentes escalas. Por isso, se as propriedades entre as imagens (georeferenciada e capturada pelo VANT) não apresentam similaridade na distribuição dos níveis de cinza, os pontos característicos detectados em cada uma delas não vão ser casados porque a magnitude do gradiente e a orientação de cada ponto característico será diferente entre imagens.

  Neste trabalho foi utilizado o algoritmo SURF para a extração e casamento de pontos característicos e foi implementado o algoritmo MSAC para a eliminação

de *outliers*. Embora este método não seja o mais robusto, consegue obter bons resultados com baixo tempo de processamento. Esta ação se justifica porque, a natureza do problema precisa de um método de baixo tempo de processamento e eficiência na correção de perspectiva em imagens capturadas pelo VANT. Além disso, não é necessária a busca de um volume grande de casamentos entre imagens porque a correção de perspectiva pode ser feita no mínimo com quatro casamentos de pontos característicos. Em adição, o método proposto não precisa de parâmetros da câmera, nem precisa do conhecimento dos ângulos de inclinação do VANT.

- **Quanto aos Testes com Igual Sensor**

  Nos testes do método não paramétrico com imagens de igual sensor, demostrou-se que a inclusão da etapa de pré-processamento na análise das imagens capturadas por VANTs, melhora os resultados do algoritmo SURF significativamente. Houve um aumento no número de casamento de pontos característicos, uma maior robustez no cálculo da matriz homográfica e um tempo baixo de processamento. É concluído que, a relação de casamento versus tempo de processamento quando a etapa de pré-processamento é aplicada, é maior comparado com o método sem pré-processamento. Dessa forma, se o objetivo na análise das imagens fosse obter um grande volume de casamentos entre pontos característicos, a etapa de pré-processamento vai garantir esse objetivo.

- **Quanto à Eficiência, Recall e Precision**

  Com o cálculo da eficiência, *recall* e *precision* no método não paramétrico, conseguiu-se caracterizar os algoritmos SURF e MSAC. De modo que foi possível avaliar o melhor ponto de operação dos algoritmos e do método quando se trabalha com imagens de diferentes sensores. Com isto, foi determinado o valor do *threshold* que fornece a melhor eficiência para todas as imagens que foram avaliadas. Ao mesmo tempo, foi evidenciada uma melhora no desempenho do método na etapa de rejeição de outliers. Estimou-se a faixa de valores do parâmetro *td* que fornece a melhor rejeição de outliers e, ao mesmo tempo, proporciona o maior número de casamentos de pontos característicos corretos.

  Com os resultados do *threshold* e do *td* foram obtidas melhoras consideráveis no desempenho e no custo computacional do método proposto. Desta forma, com o método trabalhando no melhor ponto de operação, o número de casamentos aumentou e o tempo de processamento em cada execução foi reduzido consideravelmente, apresentando grande potencial para ser embarcado e implementado em simulações de voo em tempo real.

- **Quanto aos Testes com Diferente Sensor**

  Uma boa precisão na estimação de posição através de imagens aéreas obtidas de diferentes sensores, é em grande parte garantida se existe uma boa correção de perspectiva, rotação e escala nas imagens capturadas pelo VANT. Nos resultados, foi avaliada a robustez na estimação de posição e precisão no reconhecimento de padrões entre uma imagem corrigida com nosso método (Cap. 5) e a imagem georeferenciada, onde como resultado, a estimação de posição das sub-regiões sobre a imagem georeferenciada resultou ser muito próxima.

7.1.3 Quanto ao Teste do Vídeo

Nos testes de estimação de posição sobre o vídeo, o método implementado na correção de perspectiva demostrou ser independente das características da imagem. Portanto, o processo não é afetado por mudanças espectrais ou físicas presentes nas imagens. Com isto, tem-se a intenção de melhorar a estimação de posição, dar uma maior robustez e precisão no ajuste da imagem capturada pelo VANT antes de ser procurada sobre a imagem georeferenciada. Os resultados mostraram que quando foi aplicada a correção de perspectiva na estimação de posição, esta melhorou, pois fornece posições mais precisas e com menor distância de separação comparado com a posição real.

Em conclusão, no processo de navegação autônoma de VANTs através de imagens, a correção de perspectiva aplicada às imagens capturadas pela aeronave obtiveram melhores resultados comparado com a aplicação de apenas a rotação de imagens. Isto porque a correção de perspectiva melhora a percepção da imagem e o ajuste de rotação ao longo dos três ângulos de inclinação do VANT (Yaw, Pitch, Roll). Em contrapartida, o método que foi aplicado só a rotação é mais limitado pois o ajuste é feito apenas no ângulo Yaw.

7.2 Publicações em Eventos Científicos

- Publicação do artigo "CORREÇÃO DE PERSPECTIVA EM IMAGENS APLICADA À NAVEGAÇÃO AUTÔNOMA DE VANTs" no XII Simpósio de Mecânica computacional 2016 (Jaimes et al., 2016).

- Participação na modalidade pôster, "Correção de Perspectiva aplicado à Navegação autônoma de VANTs através de imagens" no Geoscience and Remote Sensing Society (GRSS)-YP & International Society for Photogrammetry and Remote Sensing (ISPRS) Summer School 2016.

- Participação na modalidade poster, "Correção de distorção projetiva em imagens aéreas capturadas por VANTs utilizando uma metodologia não-paramétrica (imagens de igual sensor)" no Geoscience and Remote Sensing Society (GRSS)-YP & International Society for Photogrammetry and Remote Sensing (ISPRS) Summer School 2016.

- Publicação do artigo "Implementação de uma Rede RBF como Classificador Semi-supervisionado Baseado em COP-Kmeans" no XXI Congresso brasileiro de automática - CBA2016.

- Participação na modalidade poster, "Finding the best operating point to perform the robust projective and spectral distortion correction in UAV images"no 4th Conference of Computacional Interdisciplinary Sciences - CCIS2016.

7.3 Trabalhos Futuros

Como proposta de continuidade deste trabalho, algumas melhoras e sugestões para trabalhos futuros são definidas.

- Melhorar a precisão na correção de distorção projetiva.

- Implementar paralelamente os dois métodos que fazem correção de distorção projetiva, o método paramétrico e o método não paramétrico para obter a homografia e aplicar a transformação projetiva na imagem de forma mais robusta, aproveitando as vantagens que cada um destes métodos oferece.

- Criar uma métrica de decisão no uso dos métodos de correção de distorção projetiva. Dessa forma, dependendo o grau de distorção projetiva apresentado na imagem a ser avaliada será utilizado o método que melhor forneça resultados para esse caso específico.

- Avaliar a distribuição dos pontos característicos extraídos que são utilizados no casamento das imagens.

- Implementar o método para a sobrestimação de bordas nos diferentes extratores de bordas desenvolvidos com Redes Neurais Artificiais.

- Estudar técnicas que permitam fazer fusão de métodos na estimação de posição.

Referências Bibliográficas

Ali, S. and Hussain, M. (2012). Panoramic image construction using feature based registration methods. In *2012 15th International Multitopic Conference (INMIC)*, pages 209–214. IEEE. (Citado em pagina 28.)

Alsam, A., Farup, I., and Rivertz, H. J. (2015). Iterative sharpening for image contrast enhancement. In *Colour and Visual Computing Symposium (CVCS), 2015*, pages 1–4. IEEE. (Citado em pagina 55.)

ANAC (2015). ANAC propõe regras para RPA e aeromodelos â ANAC, https://www.anac.gov.br/noticias/2015/anac-propoe-regras-para-rpa-e-aeromodelos. (Citado em paginas 17 and 47.)

Aouf, N. (2012). SIFT and SURF feature analysis in visible and infrared imaging for UAVs. *2012 IEEE 11th International Conference on Cybernetic Intelligent Systems (CIS)*, pages 46–51. (Citado em pagina 32.)

Aracena Pizarro, D., Campos, P., and Tozzi, C. L. (2005). Comparación de Técnicas de Calibración de Cámaras Digitales. *Revista Facultad de Ingeniería - Universidad de Tarapacá*, 13:57–67. (Citado em pagina 30.)

Baig, A., Chaudhry, M. A., and Mahmood, A. (2012). Local normalized cross correlation for geo-registration. In *Proceedings of 2012 9th International Bhurban Conference on Applied Sciences & Technology (IBCAST)*, pages 70–74. IEEE. (Citado em pagina 2.)

Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L. (2008). Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*, 110:346–359. (Citado em paginas 3, 26, 27, 28, 53, and 58.)

Birchfield, S. (1998). An introduction to projective geometry (for computer vision). *Unpublished note, Stanford university.* (Citado em paginas 46 and 52.)

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, (6):679–698. (Citado em pagina 12.)

Castro, C. L. D. (2011). *Novos critérios para seleção de modelos neurais em problemas de classificação com dados desbalanceados*. PhD thesis, Universidade Federal de Minas Gerais. (Citado em paginas 118 and 119.)

Cesetti, A., Frontoni, E., Mancini, A., Zingaretti, P., and Longhi, S. (2009). Vision-based autonomous navigation and landing of an unmanned aerial vehicle using natural landmarks. In *2009 17th Mediterranean Conference on Control and Automation*, pages 910–915. IEEE. (Citado em pagina 1.)

Chunhui, Z., Rongzhi, W., Tianwu, Z., and Quan, P. (2014). Visual odometry and scene matching integrated navigation system in uav. In *Information Fusion (FUSION), 2014 17th International Conference on*, pages 1–6. IEEE. (Citado em pagina 1.)

Committee, D. S. et al. (1995). Ieee standard for distributed interactive simulation-application protocols. *IEEE Computer Society*, 1278. (Citado em pagina 21.)

Conte, G. and Doherty, P. (2008). An integrated UAV navigation system based on aerial image matching. *Proceedings of the IEEE Aerospace Conference*, pages 1–10. (Citado em paginas 1, 2, 6, 11, and 15.)

Crow, F. C. (1984). Summed-area tables for texture mapping. In *Proceedings of the 11th annual conference on Computer graphics and interactive techniques - SIGGRAPH '84*, volume 18, pages 207–212, New York, New York, USA. ACM Press. (Citado em pagina 27.)

da Silva, W., Shiguemori, E. H., Vijaykumar, N. L., and de Campos Velho, H. F. (2015). Estimation of uav position with use of thermal infrared images. In *2015 9th International Conference on Sensing Technology (ICST)*, pages 828–833. IEEE. (Citado em pagina 16.)

Dubrofsky, E. (2009). *Homography estimation*. PhD thesis, UNIVERSITY OF BRITISH COLUMBIA (Vancouver). (Citado em pagina 60.)

Eisenbeiss, H. (2004). A mini unmanned aerial vehicle (uav): system overview and image acquisition. *International Archives of Photogrammetry. Remote Sensing and Spatial Information Sciences*, 36. (Citado em pagina 2.)

Fan, D., Ye, Y., Pan, L., and Yan, S. (2011). A remote sensing adapted image registration method based on sift and phase congruency. In *2011 International Conference on Image Analysis and Signal Processing*, pages 326–331. IEEE. (Citado em pagina 31.)

Fawcett, T. (2006). An introduction to roc analysis. *Pattern recognition letters*, 27(8):861–874. (Citado em pagina 118.)

Fischler, M. A. and Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24:381–395. (Citado em pagina 60.)

Galbraith, A., Theiler, J., Thome, K., and Ziolkowski, R. (2005). Resolution enhancement of multilook imagery for the multispectral thermal imager. *IEEE Transactions on Geoscience and Remote Sensing*, 43:1964–1977. (Citado em pagina 47.)

Goltz, G. A. M. (2011). Redes neurais artificiais em imagens para estimação da posição de um vant. *Sao José dos Campos*, page 117. (Citado em paginas 2, 6, 11, 12, and 15.)

Goltz, G. A. M. and Shiguemori, E. H. (2008). Aplicação do algoritmo sift em imagens de navegação autônoma. In *Workshop Anual de Pesquisa e Desenvolvimento do IEAv*, pages 35–35. (Citado em pagina 31.)

Goltz, G. A. M. and Shiguemori, E. H. (2011). Position estimation of uav by image processing with neural networks. *X Congresso Brasileiro de Inteligência computacional*, pages 9–17. (Citado em pagina 2.)

Gonzalez, R. C. and Woods, R. E. (2008). Digital image processing. *Nueva Jersey*. (Citado em paginas x, 9, 10, 11, 13, and 56.)

González Jiménez, A. J. (1999). Visión por computador. (Citado em pagina 46.)

Haibo He, H. and Garcia, E. (2009). Learning from Imbalanced Data. *IEEE Transactions on Knowledge and Data Engineering*, 21(9):1263–1284. (Citado em pagina 119.)

Hartley, R. I. and Zisserman, A. (2000). *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521623049. (Citado em paginas 19, 22, 26, and 49.)

Hong, Y. X., Jie, Z. Q., Dan Dan, Z., Xin, S. X., and Jing, X. (2013). UAV image automatic mosaic method based on matching of feature points. In *2013 Chinese Automation Congress*, pages 33–36. IEEE. (Citado em pagina 31.)

Horta, E., Shiguemori, E. H., Velho, H. F. C., and Braga, A. P. (2012). Extração de características e casamento de padrões aplicados à estimação de posição de um VANT. *Congresso Brasileiro de Automática*, 1:5045–5050. (Citado em pagina 2.)

Huang, S.-M., Huang, C.-C., and Chou, C.-C. (2012). Image registration among UAV image sequence and Google satellite image under quality mismatch. In *2012 12th International Conference on ITS Telecommunications*, pages 311–315. IEEE. (Citado em pagina 31.)

Jaimes, B. A., de Castro, C. L., and Torres, F. S. (2016). CORREÇÃO DE PERSPECTIVA EM IMAGENS APLICADA À NAVEGAÇÃO AUTÔNOMA DE VANTs. In *XII Simposio de Mecânica Computacional*, volume V.1, pages 103–110, Diamantina, 2016. (Citado em pagina 97.)

Jayatilleke, L. and Nian Zhang (2013). Landmark-based localization for Unmanned Aerial Vehicles. In *2013 IEEE International Systems Conference (SysCon)*, pages 448–451. IEEE. (Citado em pagina 1.)

Khodaskar, A. and Ladhake, S. (2014). Pattern recognition: Advanced development, techniques and application for image retrieval. In *Communication and Network Technologies (ICCNT), 2014 International Conference on*, pages 74–78. IEEE. (Citado em pagina 6.)

Li, Y., Li, Z., Tong, L., Yan, Y., and Guo, D. (2011). A geometrical rectification algorithm of UAV remote sensing images based on flight attitude parameters. In *2011 IEEE International Geoscience and Remote Sensing Symposium*, pages 487–490. IEEE. (Citado em pagina 30.)

Lima, S. R. S., Krueger, C. P., and Seeber, G. (2005). Integração gps/ins utilizando sensores inerciais baseados em sistemas microeletromecânicos (mems). *Boletim de Ciências Geodésicas*, 11(2):291–292. (Citado em pagina 7.)

Luo, K., Miao, H., and Wang, L. (2015). Recognition algorithm of landmark for quadrotors aircraft based on image feature of corner points. In *2015 IEEE International Conference on Information and Automation*, pages 1437–1440. IEEE. (Citado em pagina 1.)

Medeiros, F. A. (2007). Desenvolvimento de um veículo aéreo não tripulado para aplicação em agricultura de precisão. *Universidade Federal de Santa Maria, Santa Maria*. (Citado em pagina 2.)

Meneses, P. R., Almeida, T. d., et al. (2012). Introdução ao processamento de imagens de sensoriamento remoto. *Brasília: UNB/CNPq*. (Citado em pagina 16.)

Merino, L., Wiklund, J., Caballero, F., Moe, A., De Dios, J., Forssen, P.-E., Nordberg, K., and Ollero, A. (2006). Vision-based multi-UAV position estimation. *IEEE Robotics & Automation Magazine*, 13(3):53–62. (Citado em pagina 14.)

Michalopoulos, K., Tsakalakis, M., and Bourbakis, N. (2013). Detecting texture paths and patterns in aerial images. In *Information, Intelligence, Systems and Applications (IISA), 2013 Fourth International Conference on*, pages 1–5. IEEE. (Citado em pagina 6.)

Morel, J.-M. and Yu, G. (2009). Asift: A new framework for fully affine invariant image comparison. *SIAM Journal on Imaging Sciences*, 2(2):438–469. (Citado em pagina 80.)

Muella, M. T. d. A. H. (2008). *Morfologia Dinâmica Das Irregularidades Ionosfericas Pequena Escala Imageamento Ionosferico Gps*. PhD thesis. (Citado em pagina 1.)

Nam, M. and Phillips, R. (2013). Multi-modal aerial image registration using spatial structure. In *2013 Asilomar Conference on Signals, Systems and Computers*, pages 569–573. IEEE. (Citado em pagina 31.)

Nazneen, N., Shafiq, M., and Hameed, A. (2016). Template matching of aerial images using gpu. In *2016 13th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*, pages 206–212. IEEE. (Citado em pagina 2.)

Neubeck, A. and Van Gool, L. (2006). Efficient Non-Maximum Suppression. In *18th International Conference on Pattern Recognition (ICPR'06)*, volume 3, pages 850–855. IEEE. (Citado em paginas 26 and 28.)

Pai, D. T. (2010). *Auto rectification for robotic helicopter aerial imaging*. PhD thesis, San Diego State University. (Citado em pagina 46.)

Pang-Ning, T., Steinbach, M., Kumar, V., et al. (2006). Introduction to data mining. In *Library of congress*, volume 74. (Citado em pagina 119.)

Patterson, T., McClean, S., Morrow, P., and Parr, G. (2011). Utilizing Geographic Information System Data for Unmanned Aerial Vehicle Position Estimation. In *2011 Canadian Conference on Computer and Robot Vision*, pages 8–15. IEEE. (Citado em paginas 14 and 16.)

Pham, I., Jalovecky, R., and Polasek, M. (2015). Using template matching for object recognition in infrared video sequences. In *2015 IEEE/AIAA 34th Digital Avionics Systems Conference (DASC)*, pages 8C5–1–8C5–9. (Citado em pagina 2.)

Reboucas, R. A., Eller, Q. d. C., Habermann, M., and Shiguemori, E. H. (2013). Visual Odometry and Moving Objects Localization Using ORB and RANSAC in Aerial Images Acquired by Unmanned Aerial Vehicles. In *2013 BRICS Congress on Computational Intelligence and 11th Brazilian Congress on Computational Intelligence*, pages 441–446. IEEE. (Citado em paginas 1 and 2.)

Roger-Verdeguer, J. F., Mannberg, M., and Savvaris, A. (2012). Visual odometry with failure detection for the aegis uav. In *2012 IEEE International Conference*

on Imaging Systems and Techniques Proceedings, pages 291–296. IEEE. (Citado em pagina 1.)

Romero, H., Salazar, S., Santos, O., and Lozano, R. (2013). Visual odometry for autonomous outdoor flight of a quadrotor uav. In *Unmanned Aircraft Systems (ICUAS), 2013 International Conference on*, pages 678–684. IEEE. (Citado em pagina 1.)

Romero Cortijo, A. M. and Cazorla Quevedo, M. Á. (2009). Comparativa de detectores de características visuales y su aplicación al SLAM. (Citado em pagina 29.)

Salehpour, M., Behrad, A., Sadjedi, H., and Hafezi, M. (2014). Precise positioning of reference image in video frames. In *2014 22nd Iranian Conference on Electrical Engineering (ICEE)*, pages 903–907. IEEE. (Citado em pagina 32.)

Sanjay Kumar Sahani, S. K., Adhikari, G., and Das, B. (2011). A fast template matching algorithm for aerial object tracking. In *2011 International Conference on Image Information Processing*, pages 1–6. IEEE. (Citado em pagina 2.)

Santos, A., Torres, L., Fraga, H., Velho, D. C., Shiguemori, E. H., and Pádua Braga, A. (2013). Controle e navegação de VANTs através de utilização de template matching. *Computer on the Beach*, pages 70–78. (Citado em paginas 2, 36, and 65.)

Sasi, N. and Jayasree, V. (2013). Contrast Limited Adaptive Histogram Equalization for Qualitative Enhancement of Myocardial Perfusion Images. *Engineering*, 5:326. (Citado em pagina 55.)

Sedaghat, A., Mokhtarzade, M., and Ebadi, H. (2011). Uniform Robust Scale-Invariant Feature Matching for Optical Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, 49(11):4516–4527. (Citado em paginas 31, 46, and 52.)

Shan, M., Wang, F., Lin, F., Gao, Z., Tang, Y. Z., and Chen, B. M. (2015). Google map aided visual navigation for UAVs in GPS-denied environment. In *2015 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pages 114–119. IEEE. (Citado em pagina 16.)

Shiguemori, E. H., Martins, M. P., and Monteiro, M. V. T. (2007). Landmarks recognition for autonomous aerial navigation by neural networks and Gabor transform. *Proc. SPIE*, 6497:64970R–64970R–9. (Citado em pagina 1.)

Silva, C. A. O., Goltz, G. A. M., Shiguemori, E. H., de Castro, C. L., Velho, H. F. d. C., and de Braga, A. P. (2015). Avaliação da Técnica de Casamento de Imagens para Navegação Autônoma de VANTs. In *Anais do 12. Congresso Brasileiro de*

Inteligência Computacional, pages 1–6. ABRICOM. (Citado em paginas 2, 3, 6, 7, 10, 11, 12, 13, 15, 33, 36, 65, 66, and 67.)

Silva, D. C. D. (2006). Retificação de Imagens Digitais de Fachadas com Pré-correção das Distorções das Lentes. (Citado em pagina 30.)

Silva Filho, P., Rodrigues, M., Saotome, O., and Shiguemori, E. H. (2014). Fuzzy-based automatic landmark recognition in aerial images using orb for aerial auto-localization. In *International Symposium on Visual Computing*, pages 467–476. Springer. (Citado em pagina 1.)

Sorton, E. F. and Hammaker, S. (2005). Simulated flight testing of an autonomous unmanned aerial vehicle using flightgear. *American Institute of Aeronautics and Astronautics, AIAA 2005*, 7083. (Citado em pagina 65.)

Souza, C., Dias, M., and Santos, J. (2005). Análise da Vulnerabilidade de Receptores GPS Comerciais sob Ação de Interferência Intencional. *VIII Simpósio de Guerra Eletrônica (VIII SIGE)*. (Citado em pagina 1.)

Sun, Y., Kamel, M. S., Wong, A. K., and Wang, Y. (2007). Cost-sensitive boosting for classification of imbalanced data. *Pattern Recognition*, 40(12):3358–3378. (Citado em pagina 119.)

Tahar, K. N. and Ahmad, A. (2011). UAV-based stereo vision for photogrammetric survey in aerial terrain mapping. In *2011 IEEE International Conference on Computer Applications and Industrial Electronics (ICCAIE)*, pages 443–447. IEEE. (Citado em pagina 2.)

Tahar, K. N., Ahmad, A., Akib, W. A. A. W. M., and Mohd, W. M. N. W. (2012). Aerial mapping using autonomous fixed-wing unmanned aerial vehicle. In *2012 IEEE 8th International Colloquium on Signal Processing and its Applications*, pages 164–168. IEEE. (Citado em pagina 2.)

Torr, P. and Zisserman, A. (2000). MLESAC: A new robust estimator with application to estimating image geometry. *Computer Vision and Image Understanding*, 78:138–156. (Citado em paginas 3, 53, and 60.)

Toutin, T. (2004). Geometric processing of remote sensing images: Models. *Algorithms and*. (Citado em pagina 46.)

Trucco, E. and Verri, A. (1998). *Introductory techniques for 3-D computer vision*, volume 201. Prentice Hall Englewood Cliffs. (Citado em pagina 30.)

Tu, L. and Dong, C. (2013). Histogram equalization and image feature matching. In *2013 6th International Congress on Image and Signal Processing (CISP)*, volume 01, pages 443–447. IEEE. (Citado em paginas 46 and 53.)

Tuceryan, M., Greer, D., Whitaker, R., Breen, D., Crampton, C., Rose, E., and Ahlers, K. (1995). Calibration requirements and procedures for a monitor-based augmented reality system. *IEEE Transactions on Visualization and Computer Graphics*, 1:255–273. (Citado em pagina 30.)

Velho, H. F. d. C. et al. (2015). Estimation of uav position using lidar images for autonomous navigation over the ocean. In *2015 9th International Conference on Sensing Technology (ICST)*, pages 811–816. IEEE. (Citado em pagina 15.)

Wang, L. and Yang, R. (2011). Global stereo matching leveraged by sparse ground control points. In *CVPR 2011*, pages 3033–3040. IEEE. (Citado em paginas 30 and 46.)

Yasein, M. S. and Agathoklis, P. (2007). A Robust, Feature-based Algorithm for Aerial Image Registration. In *2007 IEEE International Symposium on Industrial Electronics*, pages 1731–1736. IEEE. (Citado em pagina 31.)

Ye, X., Li, P., and Zhang, J. (2015). Fully affine invariant matching algorithm based on nonlinear scale space for side scan sonar image. In *2015 IEEE International Conference on Mechatronics and Automation (ICMA)*, pages 2387–2391. IEEE. (Citado em pagina 80.)

Yu, G. and Morel, J.-M. (2011). ASIFT: An Algorithm for Fully Affine Invariant Comparison. *Image Processing On Line*, 1. (Citado em pagina 80.)

Yu, L., Zhang, D., and Holden, E.-J. (2008). A fast and fully automatic registration approach based on point features for multi-source remote-sensing images. *Computers & Geosciences*, 34(7):838–848. (Citado em pagina 31.)

Zhang, L., Han, Y., Yang, Y., Song, M., Yan, S., and Tian, Q. (2013). Discovering discriminative graphlets for aerial image categories recognition. *IEEE Transactions on Image Processing*, 22(12):5071–5084. (Citado em pagina 6.)

Zhang, Z., Deriche, R., Faugeras, O., and Luong, Q.-T. (1995). A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial intelligence*, 78:87–119. (Citado em paginas 30 and 48.)

Zheng, M., Wu, C., Chen, D., and Meng, Z. (2014a). Rotation and affine-invariant SIFT descriptor for matching UAV images with satellite images. In *Proceedings of*

2014 IEEE Chinese Guidance, Navigation and Control Conference, pages 2624–2628. IEEE. (Citado em pagina 32.)

Zheng, M., Wu, C., Chen, D., and Meng, Z. (2014b). Rotation and affine-invariant sift descriptor for matching uav images with satellite images. In *Guidance, Navigation and Control Conference (CGNCC), 2014 IEEE Chinese*, pages 2624–2628. IEEE. (Citado em pagina 80.)

Anexo 1

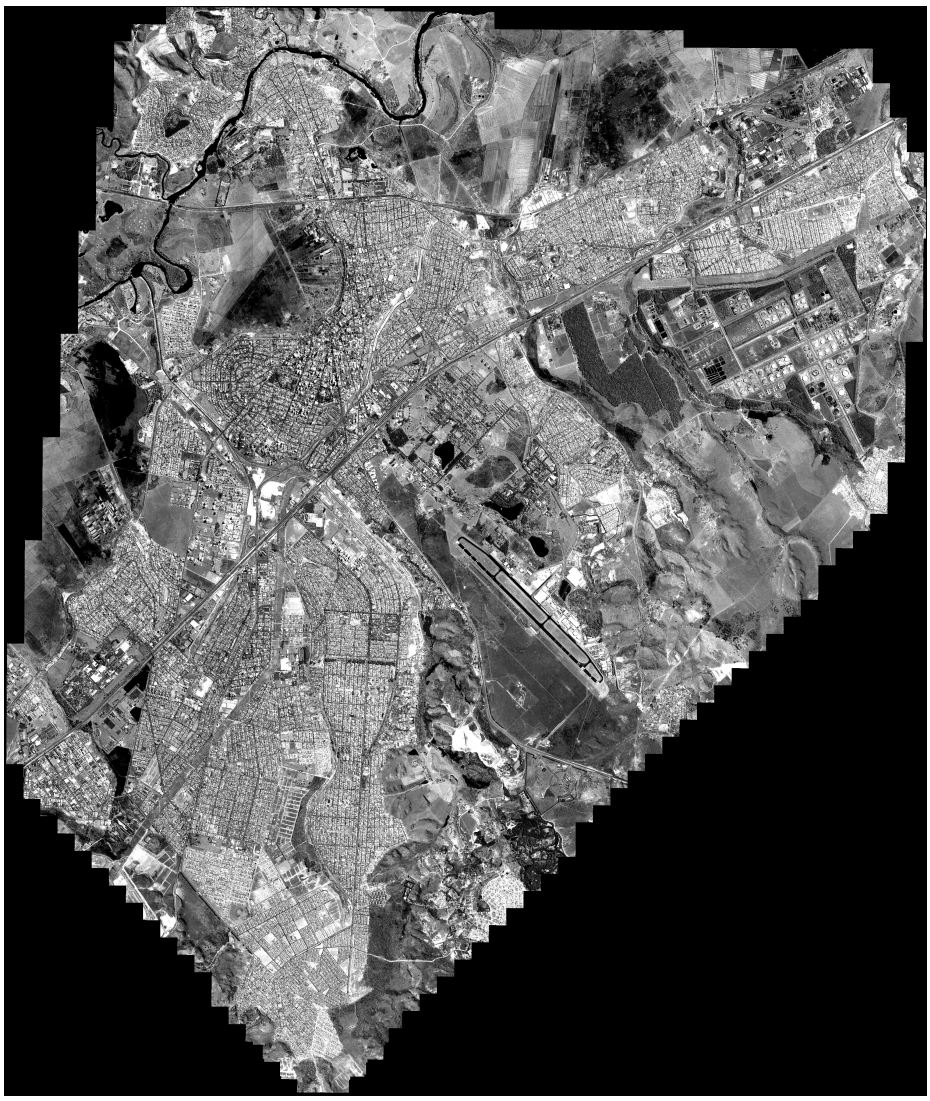


Figura A1.1: Mapa georreferenciado obtido do satélite *Quickbird*, com uma resolução espacial de 0.6 m sobre a região de São José dos Campos no estado de São Paulo - Brasil.

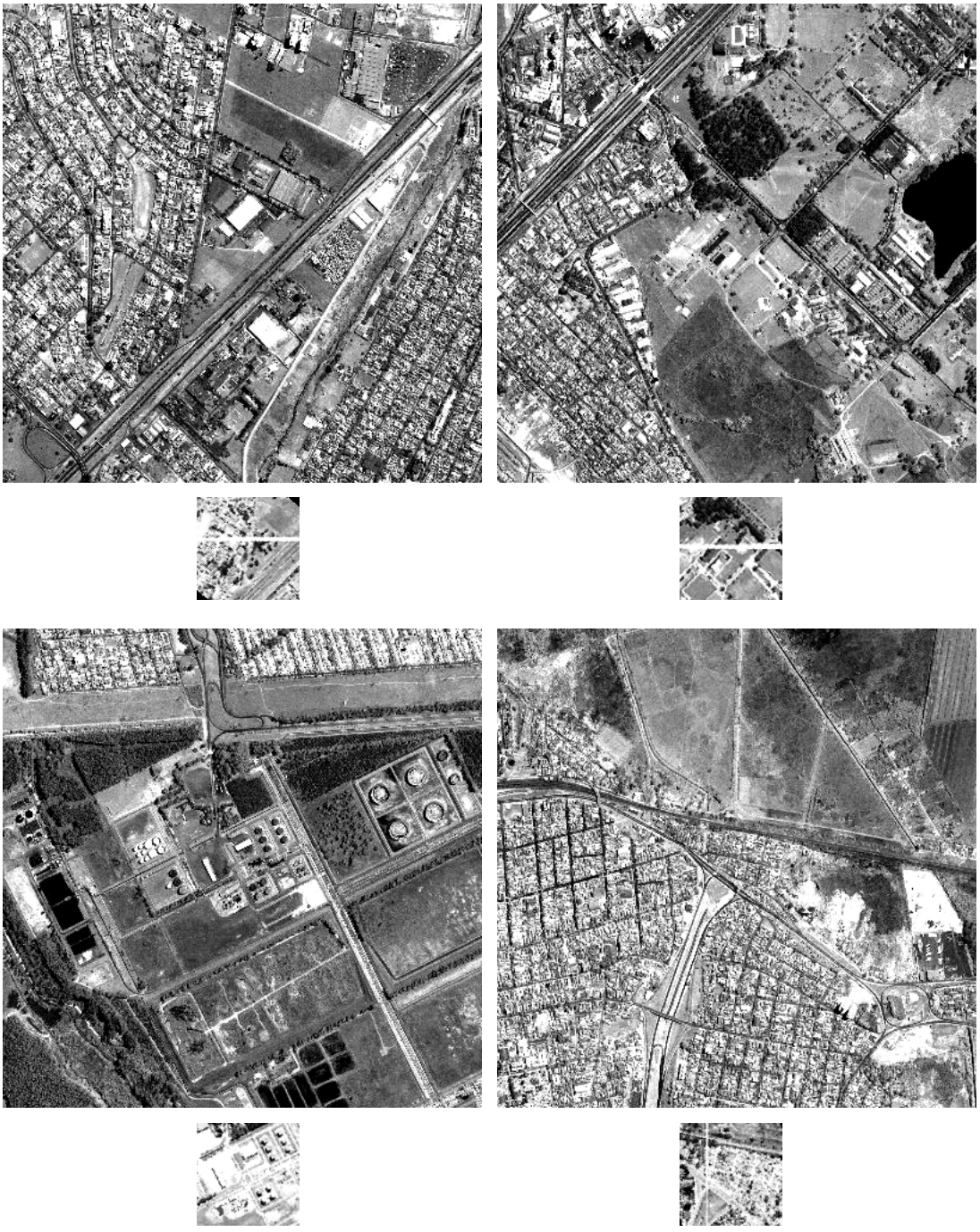


Figura A1.2: Imagens georeferenciadas e imagens capturadas por VANT com diferença de luminosidade e presença de ruido (experimento 1).

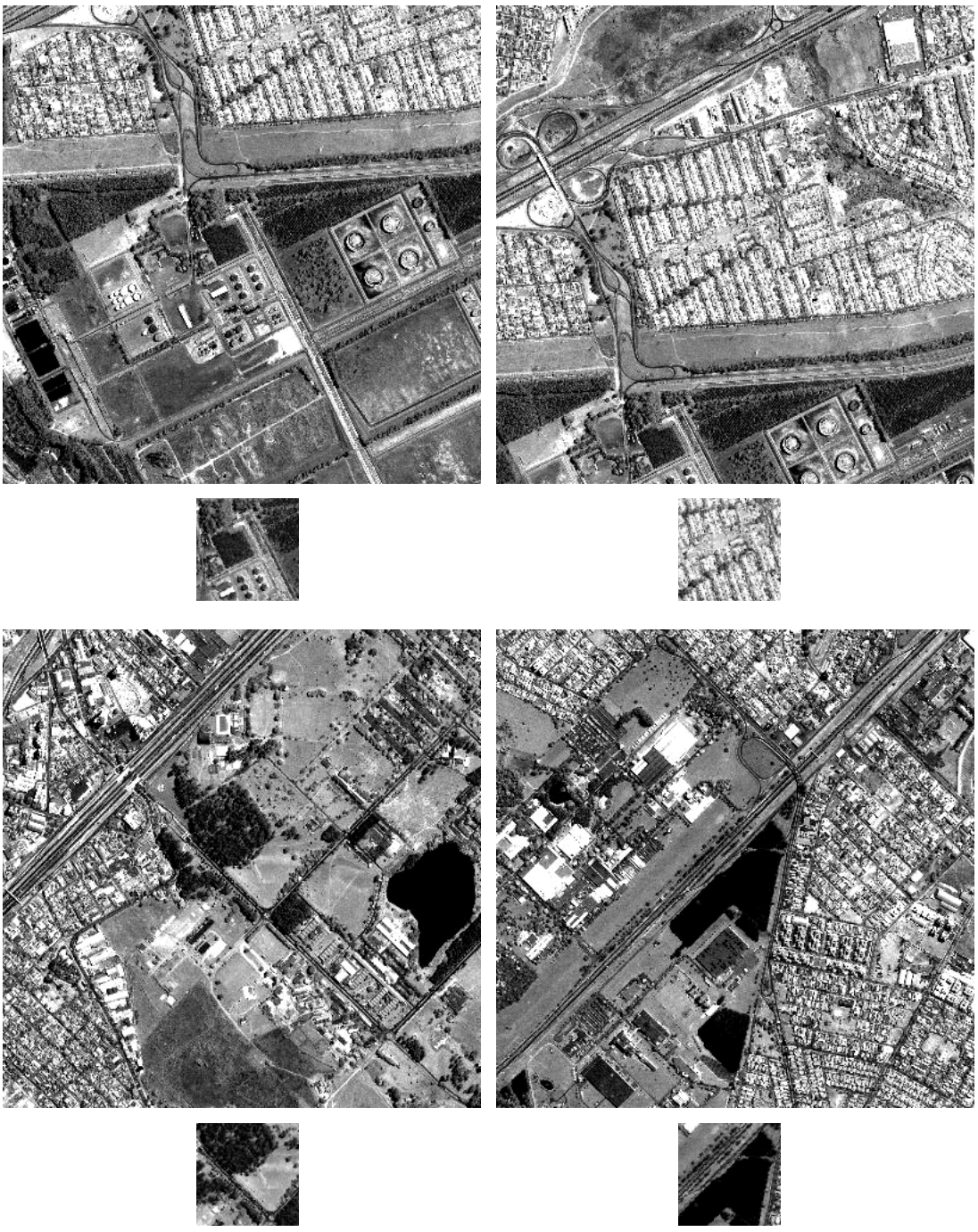


Figura A1.3: Imagens georeferenciadas e Imagens capturadas por VANT de igual sensor (experimento 2).

Anexo 2



Figura A2.1: Imagens capturadas por VANT sobre a região de São Carlos, SP - Brasil.



Figura A2.2: Imagens capturadas por VANT sobre a região de São Carlos, SP - Brasil.

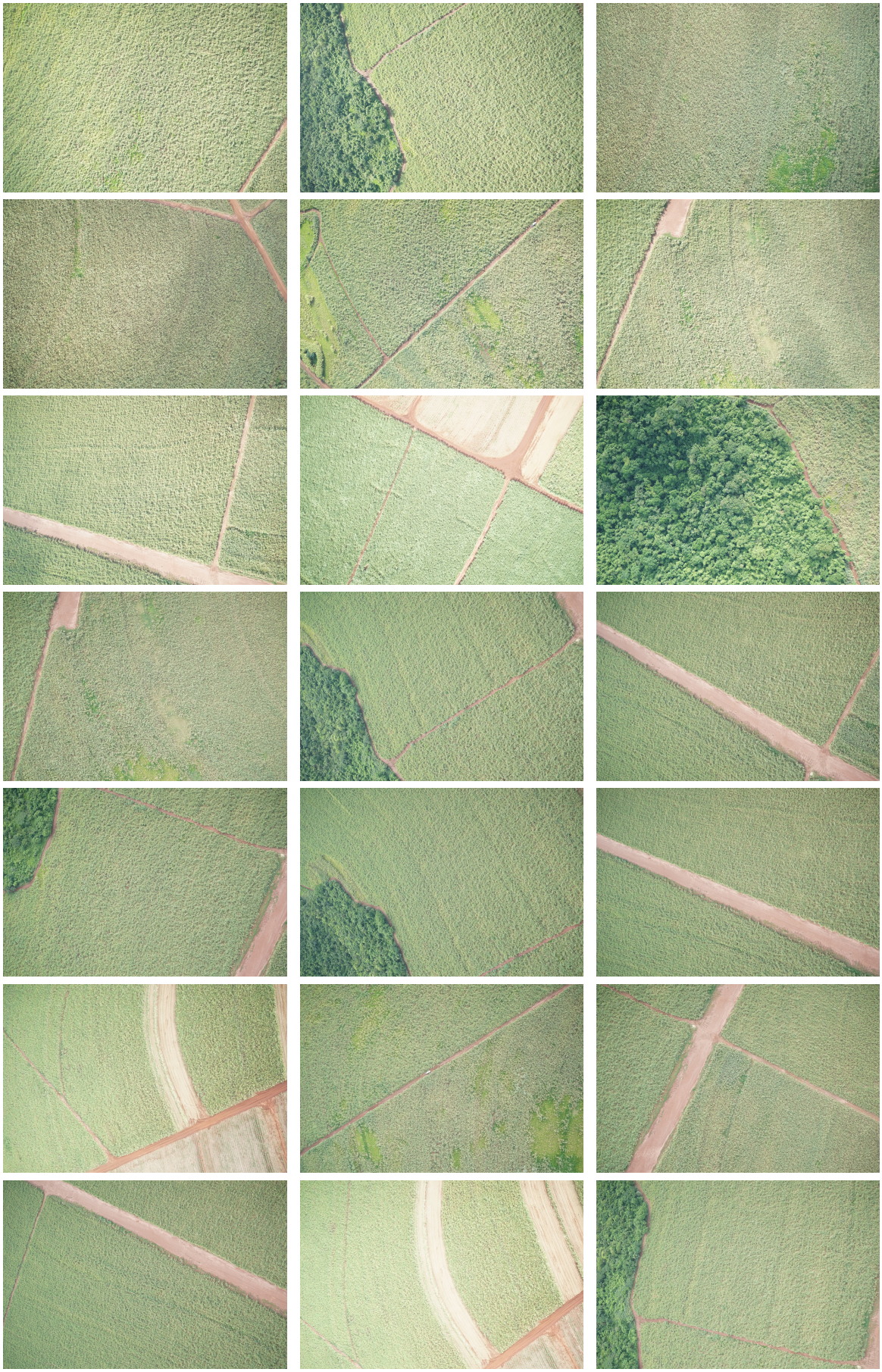


Figura A2.3: Imagens capturadas por VANT sobre a região de São Carlos, SP - Brasil.

Anexo 3



Figura A3.1: Mapa georreferenciado sobre a região de São José dos Campos no estado de São Paulo - Brasil.

Anexo 4

A4.1 Teste de Eficiência em Imagens de Diferente Sensor

Nesta seção, tem-se como objetivo caracterizar os parâmetros utilizados nos algoritmos SURF e MSAC. Também, avaliar o melhor ponto de operação do algoritmo e da metodologia quando se trabalha com imagens de diferentes sensores. De modo que para cada par de imagens analisadas, foi variado o valor do *threshold* que o algoritmo SURF define para fazer detecção e extração de pontos característicos sobre as imagens. Mantendo fixos os valores do número de oitavas igual a 4 e o número de níveis de escala igual a 6, foi feita a variação do *threshold* na faixa de 1 até 10000. Assim, em cada iteração foram calculados o número de pontos característicos extraídos nas duas imagens e o número de casamentos estabelecidos por cada iteração. Finalmente, com essas informações foi estimada a eficiência (Eq.(6.4)) para cada variação do *threshold*. Com isto, tem-se a intenção de observar o comportamento do descritor e determinar o *threshold* que forneça a melhor eficiência para todas as imagens de diferentes sensores que foram avaliadas.

A figura A4.1 e a figura A4.2 mostram os resultados do número de casamentos e o número de pontos característicos extraídos nas duas imagens respectivamente. Observa-se que a medida que o valor do *threshold* aumenta, o número de casamentos e pontos característicos seguem um comportamento exponencial decrescente, indicando uma perda na capacidade do algoritmo SURF para a geração de pontos característicos e, por conseguinte, o número de casamentos também é diminuído. Na figura A4.3 é observado o comportamento do algoritmo SURF conforme a sua eficiência quando é variado o *threshold*. Como conclusão, pode-se afirmar que os valores do *threshold* que fornecem um melhor desempenho para as diferentes imagens analisadas está no intervalo de 1 até 1000. Dessa forma, para nosso trabalho o valor do *threshold* é fixado em 800 para garantir uma eficiência alta e obter o maior numero de casamentos possível em qualquer uma das imagens testadas.

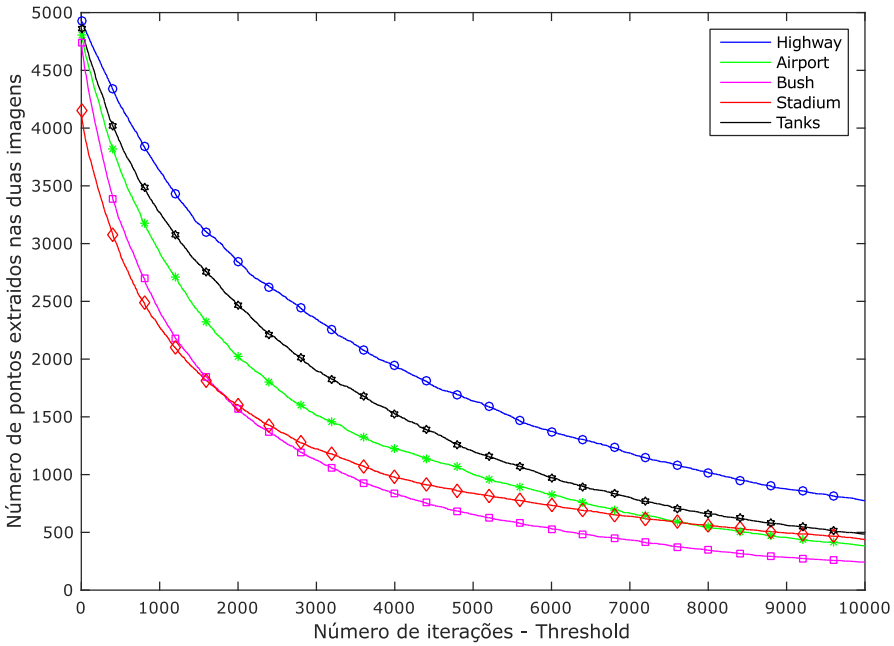


Figura A4.1: Número de pontos vs. variação do *threshold*.

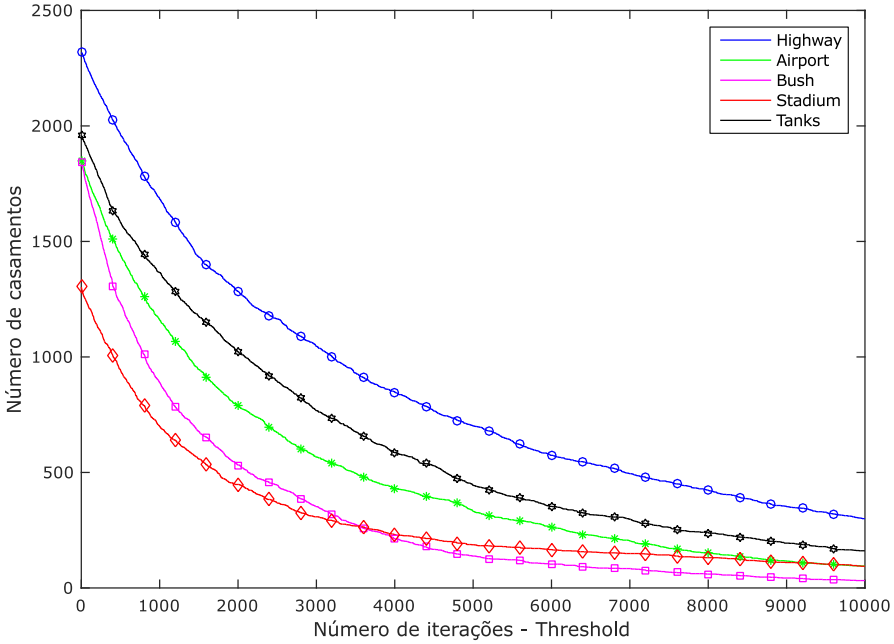


Figura A4.2: Número de casamentos vs. variação do *threshold*.

A4.2 Teste de *Recall* e *Precision* em Imagens de Diferente Sensor

Nesta secção é mostrado outro experimento feito para avaliar o desempenho do nosso algoritmo descritor SURF, e a etapa de rejeição de casamentos de pontos característicos

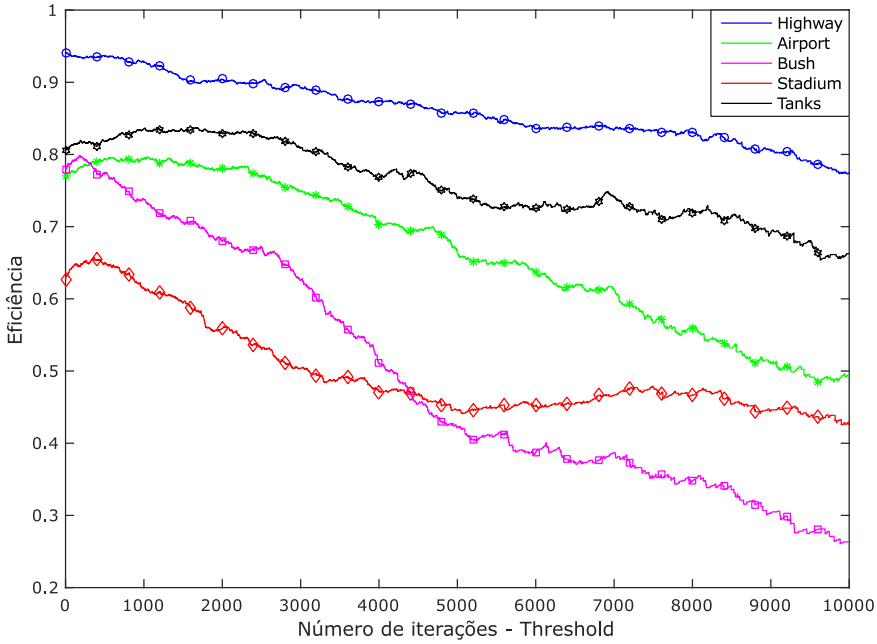


Figura A4.3: Eficiência Vs. variação do *threshold*.

(*outliers*), que foi realizada utilizando o algoritmo MSAC aplicado nas imagens de diferentes sensores. Para este experimento, foi variado o parâmetro t_d, que faz referência à distância máxima em pixeis permissível para que um ponto característico extraído ($P2'$) seja diferente da posição projetada (real $P2$) do seu ponto característico associado ($P1$) no cassamento. Tudo isto pode ser apreciado na figura A4.4. E tem a intenção de encontrar a faixa de valores do parâmetro t_d que forneça a melhor rejeição de *outliers* e ao mesmo tempo garanta o maior número de casamentos de pontos característicos corretos para estimar o modelo da matriz homográfica H.



Figura A4.4: Distância máxima em pixeis permissível para o casamento, t_d.

A rejeição de *outliers* feita pero algoritmo MSAC nas imagens de diferentes sensores, torna-se em outras palavras um problema de classificação de casamentos de pontos característicos. Assim, podem-se atribuir dois tipos de classes: uma classe 1 que corresponderia aos casamentos de pontos característicos que estejam corretos, ou seja,

os casamentos cuja distância máxima em pixeis em algum dos pontos característicos esteja por baixo do parâmetro t_d. E outra classe 0 que corresponderia aos casamentos de pontos característicos errados, ou seja, os casamentos cuja distância máxima em pixeis em algum dos pontos característicos esteja por cima do parâmetro t_d. Pode-se observar que o parâmetro t_d funciona como limiar de decisão ou *threshold*, e é quem define a probabilidade de erro/acerto para cada classe.

Assim, a maneira mais eficaz de avaliar a rejeição de *outliers* é através da identificação dos acertos (ou erros) cometidos para cada classe (Castro, 2011). Neste caso, todo o processo de classificação pode ser descrito por meio de uma matriz de confusão (Fawcett, 2006), como é ilustrado na tabela A4.1. Cada elemento dessa matriz fornece todos os possíveis acertos/erros que podem ocorrer em um classificador, onde a diagonal principal representa as classificações corretas e os elementos fora dessa diagonal representam as classificações erradas cometidas, que são mostradas a seguir:

- TN (verdadeiros negativos); são todos os casamentos que ficaram como *outliers* após da execução do algoritmo MSAC, classe 0.
- TP (verdadeiros positivos); são todos os casamentos que conseguiram passar como bom casamento após da execução do algoritmo MSAC, classe 1.
- FP (falsos positivos); são todos os casamentos que estão dentro da classe 1, classificados como bons casamentos mas em realidade são *outliers*.
- FN (falsos negativos); são todos os casamentos que estão na classe 0, que são bons casamentos mas não passaram na classe 1.

Tabela A4.1: Matriz de confusão para um classificador binário.

	classe = 0	***classe = 1***
real negativo	**TN**	**FP**
real positivo	**FN**	**TP**

A faixa de valores na qual foi variado o parâmetro t_d, vai desde 1 até 30 pixeis. Para cada iteração realizada foram obtidos os valores de TN, TP, FP e FN a fim de estabelecer métricas que avaliem de forma direta e independente o desempenho do descritor SURF e o classificador MSAC a partir da matriz de confusão (tabela A4.1). Levando em conta a natureza do nosso problema de classificação e sendo este considerado como um problema de classes desbalanceadas, onde existe um maior número de *outliers* (classe 0) e um número menor de corretos casamentos (classe 1), o critério de avaliação na classificação deve estar focada na análise e detecção da classe minoritária

e ao mesmo tempo, deve ser considerada com a mesma relevância, a discriminação de ambas classes (Castro, 2011).

Algumas das métricas mais utilizadas em problemas de classes desbalanceadas são *Recall* e *Precision* (Haibo He and Garcia, 2009; Sun et al., 2007; Pang-Ning et al., 2006). *Recall*, definido pela Eq.(A4.1), é a taxa de verdadeiros positivos e indica a razão entre o número de verdadeiros positivos (TP) e o número total de real positivo, que contém os verdadeiros positivos (TP) e os falsos negativos (FN).

$$Recall = \frac{TP}{TP + FN} \tag{A4.1}$$

Precision, é definido pela Eq.(A4.2) e representa a razão entre o número de verdadeiros positivos (TP) e o número total de positivos, que contém os verdadeiros positivos (TP) e os falsos positivos (FP).

$$Precision = \frac{TP}{TP + FP} \tag{A4.2}$$

As figuras A4.5 e A4.7 representam os resultados de *Recall* e *Precision* respectivamente. Estes resultados foram obtidos das imagens de diferentes sensores quando foi variado o parâmetro t_d (figuras 6.18 - 6.22 (a) e (b)). Para nosso caso, a medida de *Recall* reflexa a capacidade que tem o algoritmo MSAC para encontrar e classificar o maior número de corretos casamentos (TP) dentro de um conjunto de pontos característicos fornecidos pelo algoritmo SURF. E também que, após da sua classificação exista o menor número de corretos casamentos que foram rejeitados (FN).

Uma pós-análise dos resultados de *Recall* é mostrada na figura A4.6, onde o valor da média para cada variação de t_d representa a qualidade do algoritmo MSAC em função do *Recall* sobre as imagens de teste. O desvio padrão representa a faixa de erro ou a dispersão que podem apresentar os resultados. Pode-se concluir, na figura A4.5 que existe uma grande variação nos valores de *Recall* quando o valor de t_d é menor a 7 pixeis, o que significa que em $t_d \leq 7$, o desvio padrão resultante entorno da média é grande (ver figura A4.6). Em contraste, na faixa de valores onde $8 \leq t_d \leq 17$ pixeis, a variação foi bem menor, obtendo um desvio padrão entorno da média bem próximo de 0, como pode-se evidenciar na figura A4.6. Por sua vez, valores de *Recall* para quando $18 \leq t_d \leq 30$ pixeis apresentaram um desvio padrão irregular, onde em alguns casos este valor foi baixo e em outros foi alto em torno da média (ver figura A4.6). Em suma, o melhor valor esperado do parâmetro t_d em função do *Recall* será aquele que apresente um valor médio perto de 1 (máximo valor de *Recall*) e um desvio padrão baixo entorno dessa média. Em consequência, os valores na faixa de $8 \leq t_d \leq 17$ pixeis, são os que melhor se ajustam a este critério de valor esperado.

De forma similar, a medida de *Precision* (figura A4.7) representa a capacidade que

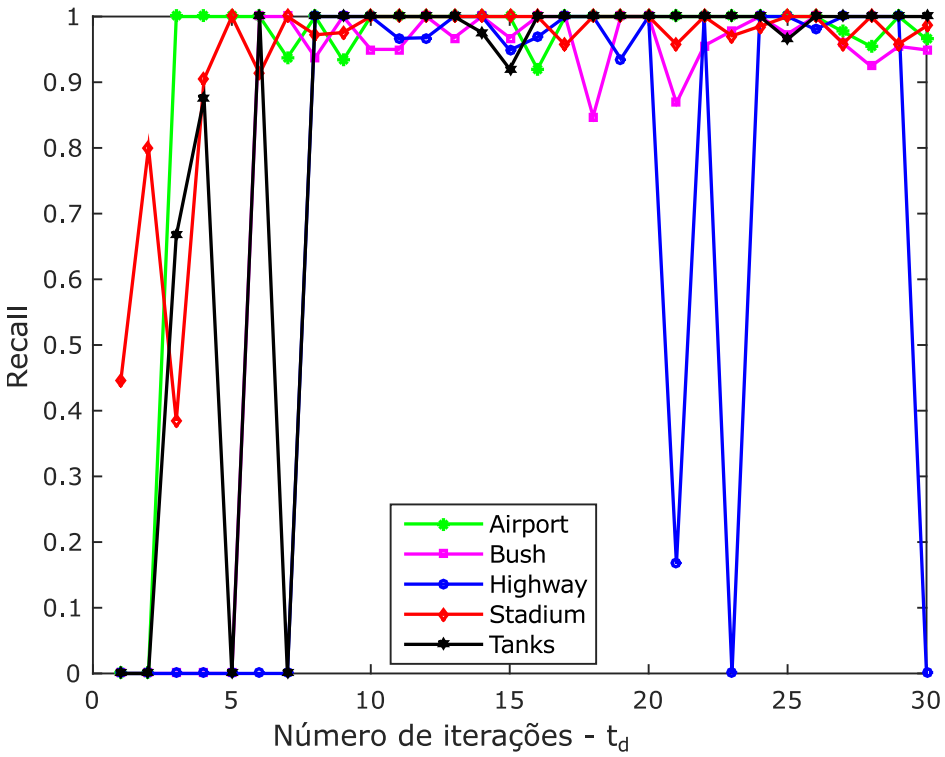


Figura A4.5: Resultado de *Recall* nas imagens de teste.

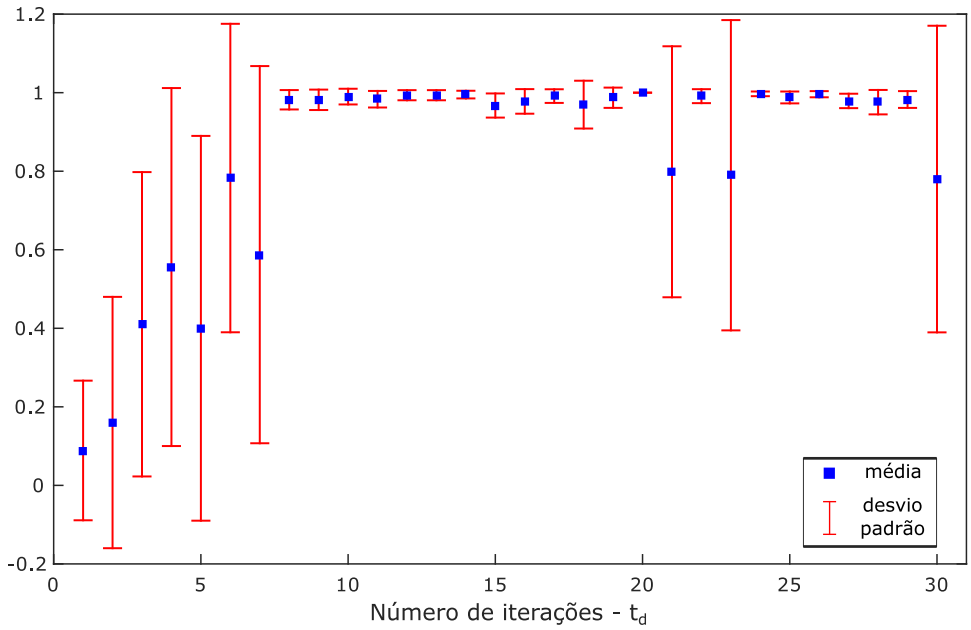


Figura A4.6: Resultado da média e desvio padrão para *Recall*.

tem o algoritmo MSAC na classificação de boa qualidade (precisão), ou seja, que dentro do conjunto de corretos casamentos (TP) já classificados, exista o menor número de *outliers* (FP). A pós-análise dos resultados de *Precision* podem ser apreciados na figura A4.8, onde são mostrados os valores da média e desvio padrão para cada valor

de t_d. Na figura A4.7, quando $t_d \leq 7$ pixeis, é evidenciada uma grande variação na medida de *Precision* e, por conseguinte, o desvio padrão em torno da média também é grande (ver figura A4.8). No entanto, na faixa de valores onde $8 \leq t_d \leq 20$ pixeis, a variação de *Precision* foi menor e o desvio padrão em torno da média é relativamente baixo como é mostrado na figura A4.8. Por sua vez, quando $21 \leq t_d \leq 30$ pixeis, os valores de *Precision* mostram um comportamento oscilante, isto é comprovado nos valores da média e em seus respectivos valores de desvio padrão (ver figura A4.8). Em conclusão, o melhor valor esperado para o parâmetro t_d em função de *Precision* será o t_d com desvio padrão baixo e uma média perto do máximo valor ($Precision = 1$). Em consequência, os valores na faixa de $8 \leq t_d \leq 20$ pixeis, resultaram ser os que melhor se ajustam a este critério de valor esperado.

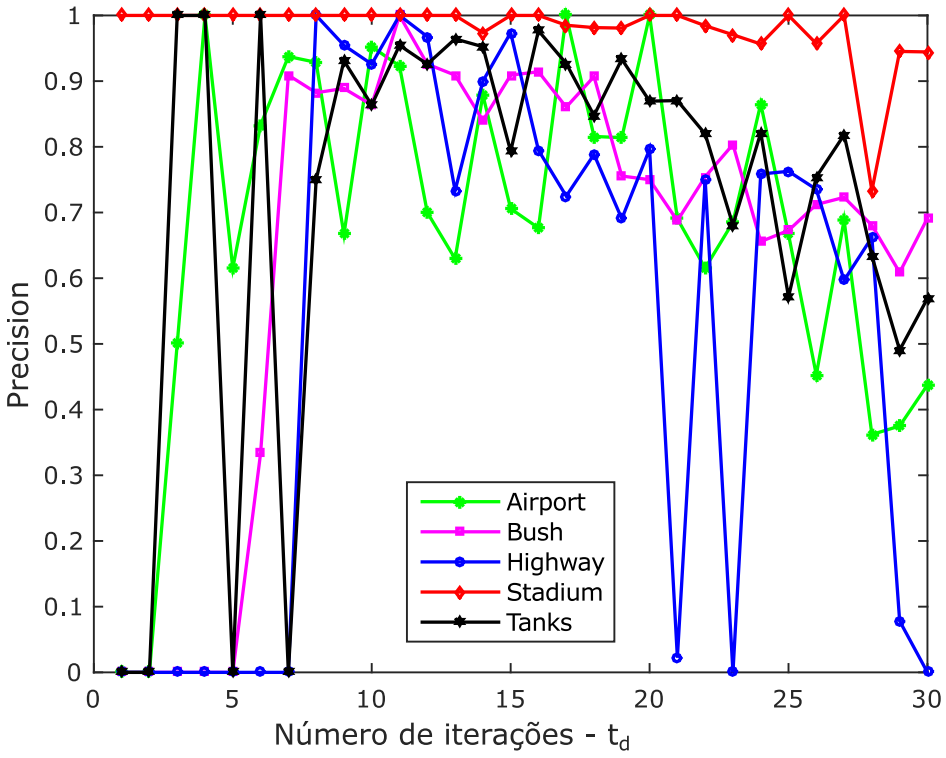


Figura A4.7: Resultado de *Precision* nas imagens de teste.

Na figura A4.9 é mostrado o gráfico de *Recall* vs $1 - Precision$ (Eq.(A4.3)) para cada par de imagens de diferentes sensores que foram testadas. Neste gráfico, é possível evidenciar com quais imagens de teste o algoritmo SURF e o algoritmo MSAC fornecem um melhor desempenho. Dessa forma, mostra-se que para as imagens de teste *stadium* (figura 6.19), os algoritmos SURF e MSAC apresentaram o melhor resultado em comparação com as demais imagens de teste. Em contraste, é observado que o desempenho mais baixo foi atribuído às imagens de teste *Highway* (figura 6.19). Pode-se dizer então que a diferença entre estes resultados é influenciada diretamente pelo tipo

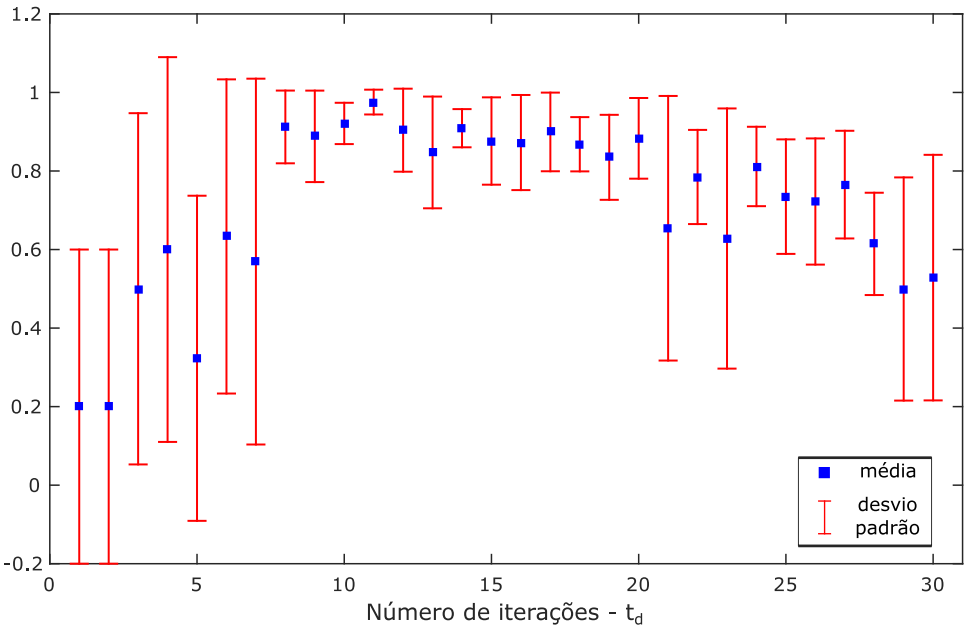


Figura A4.8: Resultado da média e desvio padrão para *Precision*.

de características e o nível complexidade que cada imagem de teste apresenta. No caso das imagens *Highway*, são imagens aéreas capturadas sobre uma região urbana bem estruturada, que apresentam um grau de rotação e perspectiva maior em comparação com as outras imagens de teste. As imagens *stadium* por sua parte, são imagens aéreas capturadas sobre uma região semi-urbana com maior presença de gramado e o grau de perspectiva e rotação é de nível meio em comparação com as outras imagens de teste.

$$1 - Precision = \frac{FP}{TP + FP} \tag{A4.3}$$

Baseando-se na aplicação dos experimentos anteriormente feitos, foram definidos e caracterizados os melhores valores do parâmetro *treshold* e do parâmetro t_d que influenciam nos resultados dos algoritmos SURF e MSAC. Posteriormente, com estes valores foi simulado novamente o experimento do cálculo de pontos característicos e casamentos sobre as imagens de teste após 30 execuções, e os resultados deste experimento são mostrados na tabela A4.2.

Os resultados observados na tabela comparativa A4.2, mostram que existe uma melhora considerável no desempenho e no custo computacional do método proposto após a caracterização dos algoritmos SURF e MSAC. Com o método trabalhando no melhor ponto de operação, foi possível aumentar o número de casamentos calculados entre os pares de imagens de teste, e ao mesmo tempo reduzir consideravelmente o tempo de processamento em cada execução. Assim, o método proposto apresenta um menor custo computacional, com grande potencial de ser embarcada e implementada em simulações de voo em tempo real.

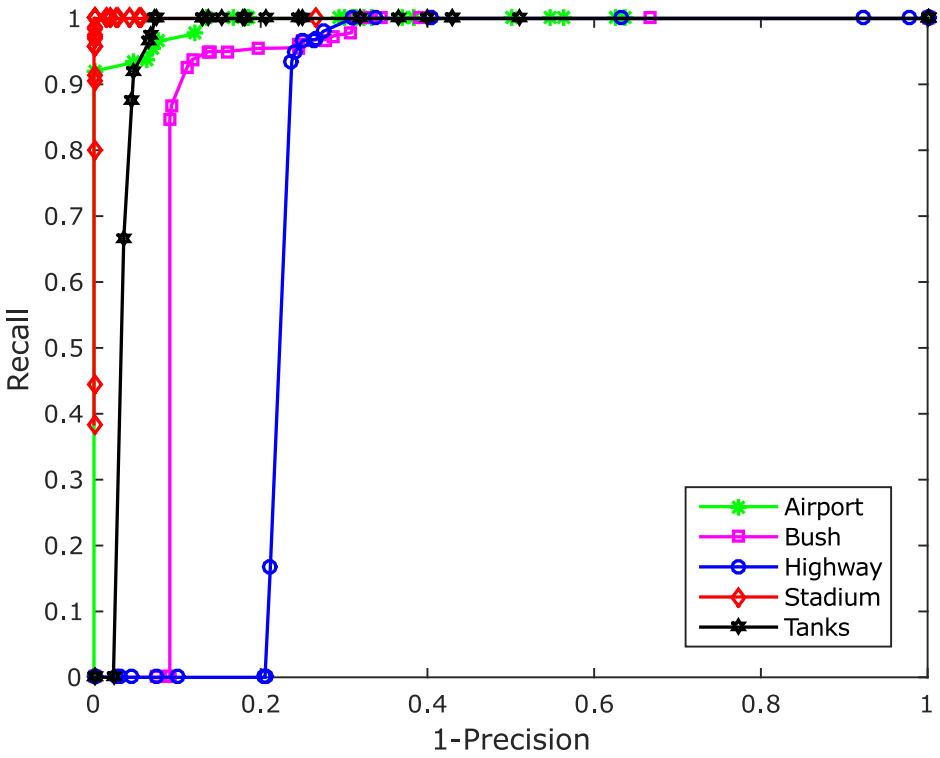


Figura A4.9: Gráfico de *Recall* vs. $1 - Precision$.

Tabela A4.2: Média de pontos característicos extraídos das imagens. Média, valor mínimo e máximo de casamento de pontos característicos nas imagens. Tempos de processamento obtidos de cada método em cada caso de teste depois de 30 execuções. Caracterizados os algoritmos SURF e MSAC e encontrar o melhor ponto de operação.

		Número de Keypoints		***Número de casamentos***			***Média exec.***
Imagem	***Método***	*Média Geo.*	*Média VANT*	*Média*	*Min*	*Max*	*tempo [s]*
Airport	ASIFT	37508	24293	3	0	26	50,66
	Proposto	3799	3765	35	25	44	3,52
	Prop. caracterizado	**1637**	**1541**	**41**	**33**	**51**	**0,9993**
Stadium	ASIFT	17565	24335	0	0	0	29,9
	Proposto	2168	2271	29	22	45	2,48
	Prop. caracterizado	**1224**	**1265**	**58**	**35**	**68**	**0,6616**
Tanks	ASIFT	45770	32936	20	19	22	74,83
	Proposto	4537	3912	42	30	61	3,88
	Prop. caracterizado	**1792**	**1694**	**47**	**32**	**55**	**1,0987**
Bush	ASIFT	29425	26251	10	0	25	42,53
	Proposto	3493	3390	35	27	48	3,30
	Prop. caracterizado	**1277**	**1424**	**43**	**32**	**50**	**0,8160**
Highway	ASIFT	48267	41012	0	0	0	92,43
	Proposto	4448	4090	40	30	52	3,91
	Prop. caracterizado	**1949**	**1893**	**43**	**32**	**54**	**1,2058**