



CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL

DESCRITOR LOCAL BASEADO EM FILTROS LOG-GABOR PARA CORRESPONDÊNCIA DE PONTOS-CHAVE ENTRE IMAGENS MULTIESPECTRAIS

CRISTIANO FRAGA GUIMARÃES NUNES

Orientador: Prof. Dr. Flávio Luis Cardeal Pádua
Centro Federal de Educação Tecnológica de Minas Gerais

BELO HORIZONTE
FEVEREIRO DE 2017

CRISTIANO FRAGA GUIMARÃES NUNES

**DESCRITOR LOCAL BASEADO EM FILTROS
LOG-GABOR PARA CORRESPONDÊNCIA DE
PONTOS-CHAVE ENTRE IMAGENS
MULTIESPECTRAIS**

Dissertação apresentada ao Programa de Pós-graduação em Modelagem Matemática e Computacional do Centro Federal de Educação Tecnológica de Minas Gerais, como requisito parcial para a obtenção do título de Mestre em Modelagem Matemática e Computacional.

Área de concentração: Modelagem Matemática e Computacional.

Linha de pesquisa: Métodos Matemáticos Aplicados.

Orientador: Prof. Dr. Flávio Luis Cardeal Pádua
Centro Federal de Educação Tecnológica de
Minas Gerais

CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL
BELO HORIZONTE
FEVEREIRO DE 2017

N972d Nunes, Cristiano Fraga Guimarães
Descritor local baseado em filtros Log-Gabor para correspondência de pontos-chave entre imagens multiespectrais / Cristiano Fraga Guimarães Nunes. – 2017.
xiv,71 f. : il, tabs., grafs., fotos.

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Modelagem Matemática e Computacional.
Orientador: Flávio Luis Cardeal Pádua.
Bibliografia: f. 64-71.
Dissertação (mestrado) – Centro Federal de Educação Tecnológica de Minas Gerais.

1. Imagem Multiespectral – Teses. 2. Filtros digitais (Matemática) – Teses. 3. Processamento de imagens – Técnicas digitais – Teses. 4. Algoritmos de computador – Teses. I. Pádua, Flávio Luis Cardeal.
II. Centro Federal de Educação Tecnológica de Minas Gerais.
III. Título.

CDD 006.37

Elaboração da ficha catalográfica pela Biblioteca-Campus II / CEFET-MG

Esta folha deverá ser substituída pela cópia digitalizada da folha de aprovação fornecida.

Dedico esse trabalho aos meus familiares e amigos.

Agradecimentos

Com muito esforço, concluo mais uma fase em minha vida e, é imprescindível agradecer a todos àqueles que se fizeram presentes e ajudaram de forma direta ou indireta a concluir mais essa etapa.

Agradeço primeiramente a Deus que me deu forças, saúde e sabedoria para caminhar essa jornada.

Aos meus pais e irmãs, pessoas especiais na minha vida que contribuíram na formação do meu caráter, apoiaram continuamente e permitiram eu seguir meu sonho, mesmo eu estando ausente em vários momentos.

Ao Artur, pelas várias horas de estudos e pelo apoio constante por meio de conselhos e sugestões pertinentes para a execução deste trabalho.

Ao meu orientador Prof. Flávio Cardeal, que sempre acreditou em mim. Obrigado por me ajudar neste percurso, pelo seu apoio, paciência e suporte. Sua orientação transcendeu a contribuição acadêmica e profissional. Muito obrigado.

Agradeço aos colegas de trabalho do Escritório de Projetos do CEFET-MG, pela compreensão e pela oportunidade de realizar o mestrado, e que tornam o meu local de trabalho um lugar mais divertido.

A todos os colegas do Grupo de Pesquisas Interdisciplinares em Informação Multimídia do CEFET-MG e amigos do mestrado pela convivência e pelo apoio.

“O fator decisivo para vencer o maior obstáculo é, invariavelmente, ultrapassar o obstáculo anterior.” (Henry Ford)

Resumo

Este trabalho apresenta um novo método descritor de características locais em imagens multiespectrais para ser aplicado em abordagens de correspondência de pontos-chave. O problema de correspondência de pontos-chave em imagens multiespectrais está presente em um amplo conjunto de aplicações em Visão Computacional e Sensoriamento Remoto, tais como: monitoramento de desastres naturais, sistemas de vigilância por vídeo, navegação de aeronaves guiadas por imagem, dentre outras. Diferentemente da grande maioria das abordagens anteriores, baseadas em métodos descritores projetados exclusivamente para imagens obtidas no espectro de luz visível, este trabalho propõe a utilização de um novo descritor local desenvolvido especialmente para trabalhar com imagens multiespectrais. O descritor proposto, denominado *Multispectral Feature Descriptor* (MFD), baseia-se na extração de informações das imagens nos domínios do espaço e da frequência por meio da utilização de filtros Log-Gabor. Para avaliar a eficácia e a eficiência do método proposto, utilizou-se duas bases de dados compostas por imagens obtidas nos espectros de luz visível e infravermelho, amplamente utilizadas na literatura, e comparou-se seu desempenho com algoritmos do estado da arte, tais como, EOH e LGHD. Os resultados experimentais obtidos demonstram que o descritor MFD possui maior eficiência computacional que os algoritmos descritores EOH e LGHD, bem como uma eficácia (precisão e revocação) estatisticamente comparável às de tais métodos.

Palavras-chave: Descritor de características locais. Imagens multiespectrais. Filtros Log-Gabor. Correspondência de pontos-chave.

Abstract

This work presents a new local feature descriptor for multispectral images to be applied in keypoint matching approaches. The problem of matching keypoints in multispectral images may be found in a wide range of applications in Computer Vision and Remote Sensing, such as natural disaster monitoring, video surveillance systems, image-guided aircraft navigation, among others. Unlike the most of previous approaches based on descriptors designed to work with images taken in the visible light spectrum, this work proposes a new local descriptor especially developed for multispectral images. The proposed descriptor, called Multispectral Feature Descriptor (MFD), is based on the extraction of image information in the space and frequency domains through the use of Log-Gabor filters. To evaluate the performance of the proposed method, we used two databases composed of images obtained in visible light and infrared spectra, widely used in the literature, and compared their performance with state-of-the-art algorithms, such as EOH and LGHD. The experimental results showed that the MFD descriptor has better computational efficiency than the EOH and LGHD descriptor algorithms, as well as precision and recall values statistically comparable to the corresponding ones of those algorithms.

Palavras-chave: Local feature descriptor. Multispectral images. Log-Gabor filters. Keypoints matching.

Lista de Figuras

Figura 1 – Exemplo de variação em uma das componentes do vetor gradiente. . . .	2
Figura 2 – Etapas do processo de correspondência de pontos, entre duas imagens, que utilizam técnicas baseadas em características locais.	4
Figura 3 – Cenário no qual o método descritor para correspondência de pontos-chave em imagens multiespectrais é aplicável.	5
Figura 4 – Problema de pesquisa: Descrição em imagens multiespectrais para correspondência de pontos-chave.	6
Figura 5 – Exemplos de aplicações onde é fundamental a correspondência de pontos em imagens multiespectrais.	8
Figura 6 – Visão geral do processo SIFT-EOH para correspondência de pontos-chave em imagens multiespectrais.	14
Figura 7 – Espectro eletromagnético.	20
Figura 8 – Exemplo de imagem multiespectral.	20
Figura 9 – Exemplo de imagens de resposta aos filtros Log-Gabor.	22
Figura 10 – Diferenças na extração de bordas entre uma imagem do espectro visível e uma imagem do espectro infravermelho de uma mesma cena.	23
Figura 11 – Pirâmide de escala montada pelo algoritmo SIFT.	24
Figura 12 – Detecção de máximos e mínimos locais do algoritmo SIFT.	25
Figura 13 – Construção do descritor SIFT.	27
Figura 14 – Aproximações de núcleo Gaussiano do algoritmo SURF.	28
Figura 15 – Construção do descritor SURF.	28
Figura 16 – Detector de cantos FAST.	29
Figura 17 – Padrão de 128 pares de pontos de amostragem do algoritmo BRIEF. . .	30
Figura 18 – Pirâmide de escala montada pelo algoritmo BRISK.	32
Figura 19 – Padrão de amostragem do algoritmo BRISK.	33
Figura 20 – Construção do descritor EOH.	34
Figura 21 – Exemplos de erro de sobreposição entre duas regiões.	37
Figura 22 – Exemplo de aplicação das métricas de repetibilidade, precisão e revocação. .	39
Figura 23 – Sistema de correspondência de pontos em imagens multiespectrais. . .	41
Figura 24 – Extração da sub-região e cálculo do histograma do MFD.	44
Figura 25 – Montagem final do descritor MFD.	45
Figura 26 – Metodologia utilizada para avaliação dos algoritmos detectores.	47
Figura 27 – Metodologia utilizada para avaliação dos algoritmos descritores.	47
Figura 28 – Exemplos de imagens da base de dados RGB-NIR.	49
Figura 29 – Exemplos de pares de imagens da base de dados RGB-LWIR.	50

Figura 30 – Quantidade média de pontos detectados em imagens de espectro visível e imagens de espectro infravermelho.	52
Figura 31 – Valor médio de repetibilidade para os métodos detectores.	52
Figura 32 – Tempo médio de processamento para os métodos detectores.	53
Figura 33 – Valores médios de precisão, revocação e medida F1 para a variação na quantidade de filtros de orientação do método descritor MFD.	54
Figura 34 – Tempo médio de processamento para a variação na quantidade de filtros de orientação do método MFD.	55
Figura 35 – Valores médios de precisão, revocação e medida F1 para a variação na quantidade de filtros de escala do método descritor MFD.	56
Figura 36 – Tempo médio de processamento para a variação na quantidade de filtros de escala do método MFD.	57
Figura 37 – Valores médios de precisão e revocação para os métodos descritores (para a base de dados RGB-NIR).	58
Figura 38 – Resultado de precisão para os métodos descritores (para a base de dados RGB-NIR).	59
Figura 39 – Valores médios de precisão e revocação para os métodos descritores (para a base de dados RGB-LWIR).	60
Figura 40 – Valores médios de precisão para os métodos descritores (para a base de dados RGB-LWIR).	60
Figura 41 – Valores médios do tempo de processamento dos métodos descritores. .	61

Lista de Tabelas

Tabela 1 – Resultados dos métodos detectores.	51
Tabela 2 – Valores médios de precisão, revocação e medida F1 para a variação na quantidade de filtros de orientação do método MFD.	54
Tabela 3 – Valores médios de precisão, revocação e medida F1 para a variação na quantidade de filtros de escala do método MFD.	56
Tabela 4 – Resultados dos métodos descritores (base de dados RGB-NIR).	58
Tabela 5 – Resultados dos métodos descritores (base de dados RGB-LWIR).	59

Lista de Quadros

Quadro 1 – Principais métodos descritores apresentados na revisão bibliográfica. .	18
Quadro 2 – Principais características dos algoritmos apresentados.	36
Quadro 3 – Tamanhos dos descritores dos métodos de descrição para imagens multiespectrais.	57
Quadro 4 – Parâmetros do algoritmo SIFT.	73
Quadro 5 – Parâmetros do algoritmo SURF.	73
Quadro 6 – Parâmetros do algoritmo ORB.	74
Quadro 7 – Parâmetros do algoritmo FAST.	74
Quadro 8 – Parâmetros do algoritmo BRISK.	74
Quadro 9 – Parâmetros do algoritmo EOH.	74
Quadro 10 – Parâmetros do algoritmo LGHD.	74

Lista de Abreviaturas e Siglas

ANOVA	<i>Analysis of Variance</i>
BRIEF	<i>Binary Robust Independent Elementary Features</i>
BRISK	<i>Binary Robust Invariant Scalable Keypoints</i>
EHD	<i>Edge Histogram Descriptor</i>
EOH	<i>Edge Oriented Histogram</i>
FAST	<i>Features from Accelerated Segment Test</i>
FREAK	<i>Fast Retina Keypoint</i>
LGHD	<i>Log-Gabor Histogram Descriptor</i>
LWIR	<i>Long-Wave Infrared</i>
MFD	<i>Multispectral Feature Descriptor</i>
MPEG	<i>Moving Picture Experts Group</i>
MWIR	<i>Mid-Wave Infrared</i>
NIR	<i>Near-Infrared</i>
ORB	<i>Oriented FAST and Rotated BRIEF</i>
RANSAC	<i>RANdom SAmple Consensus</i>
SAR	<i>Synthetic Aperture Radar</i>
SIFT	<i>Scale-Invariant Feature Transform</i>
SURF	<i>Speeded Up Robust Features</i>
SWIR	<i>Short-Wave Infrared</i>
VLWIR	<i>Very Long-Wave Infrared</i>

Sumário

1 – Introdução	1
1.1 Caracterização do problema	5
1.2 Motivação	7
1.3 Objetivos: Geral e Específicos	8
1.4 Contribuições	9
1.5 Organização do texto	9
2 – Trabalhos Relacionados	11
2.1 Métodos detectores e descritores para correspondência de imagens em geral	11
2.2 Métodos descritores para imagens multiespectrais	13
2.3 Discussão	17
3 – Fundamentação Teórica	19
3.1 Imagens multiespectrais	19
3.2 Filtros Log-Gabor	20
3.3 Detecção e descrição de pontos-chave	24
3.3.1 Algoritmo SIFT	24
3.3.2 Algoritmo SURF	26
3.3.3 Algoritmo FAST	29
3.3.4 Algoritmo ORB	29
3.3.5 Algoritmo BRISK	32
3.3.6 Algoritmo EOH	34
3.3.7 Algoritmo LGHD	35
3.4 Correspondência de descritores	36
3.5 Métricas de avaliação	37
3.5.1 Repetibilidade	38
3.5.2 Precisão	38
3.5.3 Revocação	38
3.5.4 Medida F1	39
4 – Metodologia	41
4.1 <i>Multispectral Feature Descriptor</i> (MFD)	42
4.2 Avaliação dos algoritmos detectores	46
4.3 Avaliação dos algoritmos descritores	47
5 – Resultados	49

5.1	Avaliação dos Algoritmos Detectores	50
5.2	Análise de Parâmetros do Descritor MFD	53
5.3	Avaliação de Desempenho do Descritor MFD	57
6	– Conclusão	63
6.1	Trabalhos futuros	64
	Referências	65
	 Apêndices	 71
	 APÊNDICE A – Parâmetros de simulação	 73

1 Introdução

O problema de correspondência em Visão Computacional visa encontrar entidades equivalentes (por exemplo, pontos ou algum outro objeto de interesse) entre duas ou mais imagens de uma mesma cena (YU et al., 2012). Por meio da obtenção das correspondências entre essas entidades, diversas aplicações em Visão Computacional e Sensoriamento Remoto podem ser realizadas, tais como monitoramento de desmatamentos, calibração de câmeras, reconstrução de cenas tridimensionais, localização e mapeamento simultâneo, registro de imagens e formação de imagens panorâmicas (RICAURTE et al., 2014; YU et al., 2012; LOWE, 2004).

No contexto do problema de correspondência, uma das etapas mais críticas consiste em descrever pontos ou objetos de interesse da forma mais representativa possível, de modo a discriminá-los das demais regiões da imagem que representam a cena monitorada (AGUILERA; SAPP; TOLEDO, 2015). A descrição dessas entidades se torna uma tarefa desafiadora devido às ambiguidades encontradas no mundo e a presença de ruídos no processo de aquisição de imagens, ainda mais quando há mudanças no ângulo de visão, mudanças de escala e oclusão (DEDEEPA; SANDHYA, 2016; LI et al., 2015a). Por exemplo, diferentes pontos no mundo podem possuir aparência idêntica nas imagens obtidas quando observados sob pontos de vista diferentes dificultando-se assim a discriminação e correspondência de suas projeções em imagens distintas.

Neste cenário, deve-se destacar os esforços recentes de pesquisadores no que diz respeito ao estudo da correspondência de pontos em imagens multiespectrais¹, sobretudo, no âmbito do problema de registro de imagens² (CAMPO; RUIZ; SAPP, 2012). Para Li et al. (2015b), o registro de imagens multiespectrais tem sido um problema desafiador devido à falta de relações explícitas ou implícitas entre pontos correspondentes dessas imagens.

Um dos principais problemas encontrados durante as descrições de pontos em imagens multiespectrais se baseia no fato de que as direções dos gradientes de intensidade associados a tais pontos podem variar significativamente entre as imagens obtidas em diferentes faixas do espectro eletromagnético (JIN et al., 2016; AGUILERA; SAPP; TOLEDO, 2015). A Figura 1 ilustra as taxas de variação (em azul) de uma das componentes do vetor gradiente ao longo de um segmento de reta (em vermelho). O segmento de reta indicado em cada par de imagens contém pontos correspondentes (em verde). Essas variações, comumente de natureza não-linear, podem afetar a imagem de diferentes modos.

¹O conceito de imagens multiespectrais será melhor explicado na Seção 3.1 desse trabalho.

²O registro de imagens é o processo de alinhamento espacial de duas ou mais imagens de uma mesma cena, registrada de diferentes pontos de vista, eventualmente em diferentes instantes de tempo e por diferentes sensores (SAXENA; SINGH, 2014).

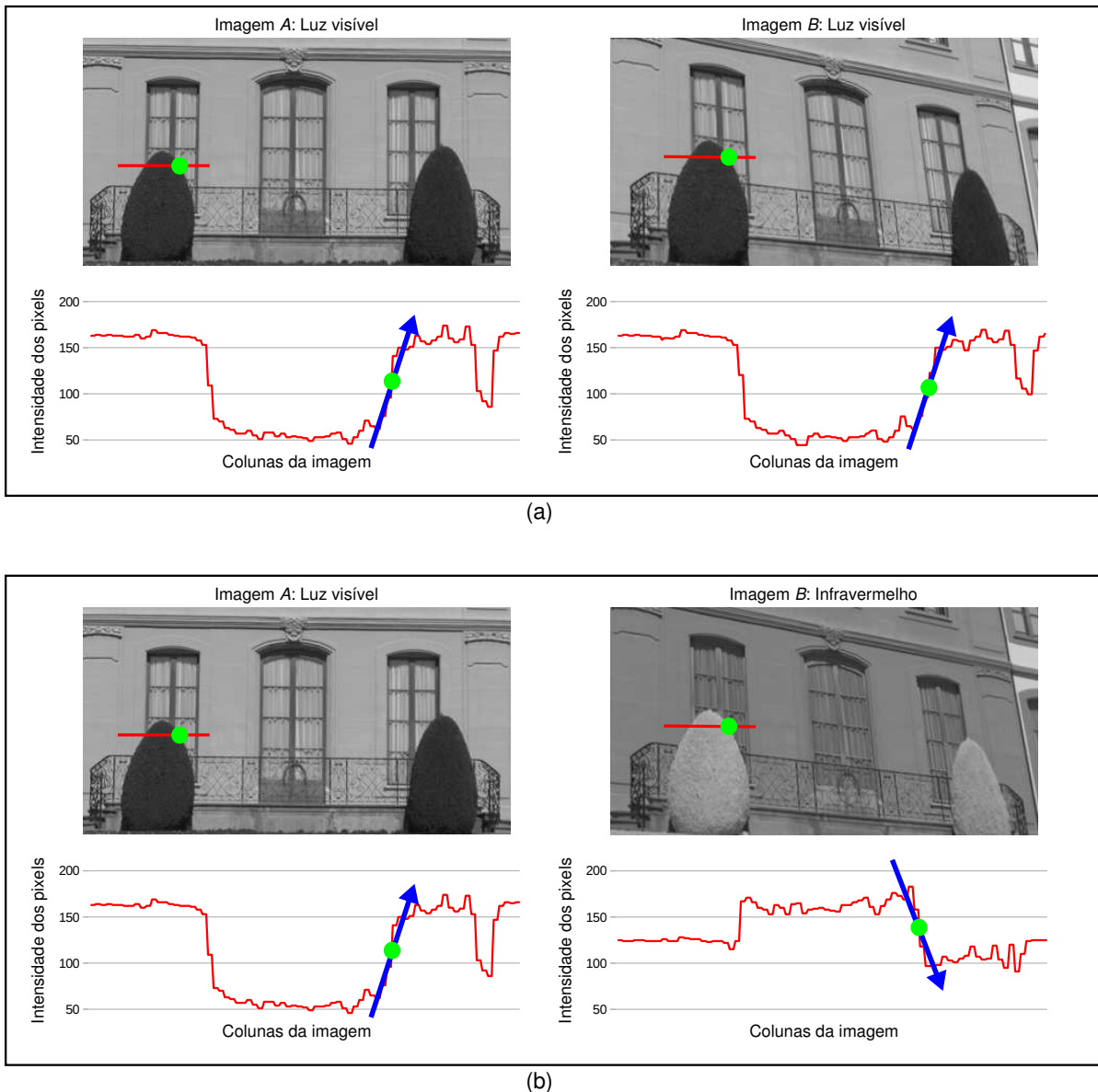


Figura 1 – Taxas de variação de uma das componentes do vetor gradiente em pares de imagens ao longo de um segmento de reta (em vermelho). Em (a) é exibido um exemplo de pares de imagens de mesmo espectro, e (b) um exemplo de pares de imagens de espectro diferentes. Observa-se que as taxas de variação (setas em azul) são semelhantes para imagens de mesmo espectro, mas diferentes em imagens de espectro diferentes, em um determinado ponto (em verde). A imagem *A* possui um ponto de vista diferente em relação à imagem *B* em ambos os casos.

No entanto, mesmo com essas variações nos gradientes de intensidade, as aparências globais das formas dos objetos, as quais podem, por exemplo, ser representadas pelas suas bordas ou texturas, tendem a permanecer razoavelmente invariantes (AGUILERA; SAPP; TOLEDO, 2015).

Uma vez que as bordas de um mesmo objeto em diferentes imagens definem formas com um alto grau de similaridade, mesmo com gradientes diferentes, é possível extrair

características que as represente de maneira robusta e estável, nas diferentes imagens, por meio de técnicas invariantes às diferenças de gradiente como, por exemplo, filtros de Log-Gabor (KOVESI, 1999), os quais são comumente utilizados em Visão Computacional para detecção de bordas (MOUATS et al., 2015; AGUILERA; SAPPA; TOLEDO, 2015).

Técnicas para descrição de pontos em imagens multiespectrais têm sido utilizadas em várias aplicações, incluindo vigilância de pedestres (TORABI; MASSÉ; BILODEAU, 2012), navegação teleguiada (OSBORNE et al., 2011) e reconhecimento de objetos (JANNEY; BOOTH, 2015). Essas imagens são, na maioria das vezes, capturadas em diferentes momentos, em diferentes pontos de visão e em faixas de frequência distintas do espectro eletromagnético, tais como, espectro visível e infravermelho. Elas fornecem uma informação complementar ao espectro visível e torna a análise de imagens mais confiável. Para utilizar essas imagens, é necessário realizar o registro das mesmas, de modo a estarem alinhadas em um sistema de coordenadas comum. Em seguida uma fusão pode ser realizada entre essas imagens, permitindo-se a obtenção de informações da cena que não seriam facilmente obtidas por imagens individuais (LI et al., 2015c).

Durante a década passada, métodos foram propostos para o registro de imagens multiespectrais, sendo tais métodos predominantemente baseados na extração de características locais, as quais eram utilizadas em uma etapa de correspondências entre pontos-chave³ (LI et al., 2015c). Nestes métodos, os elementos a serem comparados consistem em conjuntos numéricos denominados descritores de características, e entre os critérios de similaridade utilizados, pode-se citar a distância geométrica entre tais descritores (RICAURTE et al., 2014; JIANG; XU; LIU, 2013).

O presente trabalho aborda o desenvolvimento de um novo método descritor para correspondências de pontos-chave em imagens multiespectrais. Sendo assim, este trabalho está intrinsecamente ligado às pesquisas de registro de imagens que utilizam técnicas baseadas em características, uma vez que a correspondência dos pontos-chave é realizada, principalmente, por meio da extração de características locais fornecidas por um método descritor de imagem.

A tarefa de correspondência de pontos, que utiliza a abordagem baseada em características, pode ser dividida em três etapas principais, quais sejam: (i) detecção, (ii) descrição e (iii) correspondência (JIN et al., 2016; BRITO et al., 2016; LI et al., 2015c; AGUILERA et al., 2012a). Alguns autores ainda consideram uma quarta etapa denominada refinamento (JIN et al., 2016; LI et al., 2015c). A Figura 2 ilustra as principais etapas do processo de correspondência de pontos em um par de imagens. Na etapa de detecção, alguns operadores são aplicados às imagens com a finalidade de encontrar um conjunto de

³Pontos-chave, também conhecidos como “pontos de interesse”, são pontos estáveis e distintivos em uma imagem, de tal forma que estes pontos possam ser unicamente identificados por meio de um critério de detecção bem definido, mesmo sob condições adversas (LINDEBERG, 2015).

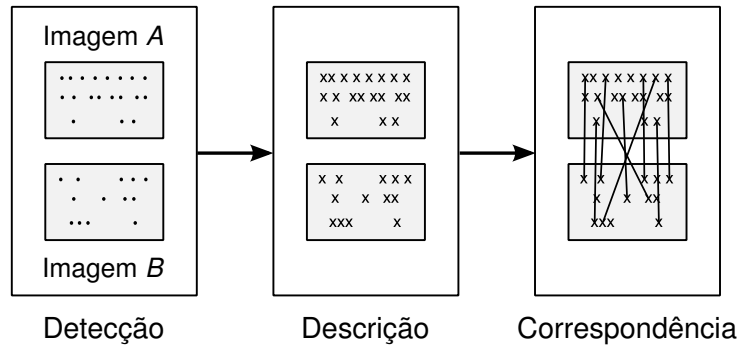


Figura 2 – Etapas do processo de correspondência de pontos, entre duas imagens, que utilizam técnicas baseadas em características locais.

pontos-chave. A etapa de descrição tem o objetivo de extrair informações referentes aos pontos detectados, de tal forma que, estes pontos possam ser discriminados, fornecendo uma espécie de assinatura a cada ponto. Então, um ponto-chave é representado por meio de um descritor (vetor numérico de dimensão n). Esse descritor é utilizado para computar a similaridade entre os pontos-chave. Por fim, na etapa de correspondência, cada ponto em uma imagem A é correspondido ao ponto mais similar em uma imagem B , dado um critério de similaridade (por exemplo, distância euclidiana ou distância de Hamming) baseado em seus descritores.

Há uma vasta literatura que propõe detectores e descritores de pontos-chave. Técnicas típicas para realizar a detecção e a descrição de pontos-chave incluem: SIFT (*Scale-Invariant Feature Transform*) (LOWE, 2004), SURF (*Speeded Up Robust Features*) (BAY; TUYTELAARS; GOOL, 2006), ORB (*Oriented FAST and Rotated BRIEF*) (RUBLEE et al., 2011), BRISK (*Binary Robust Invariant Scalable Keypoints*) (LEUTENEGGER; CHLI; SIEGWART, 2011), FREAK (*Fast Retina Keypoint*) (ALAHÍ; ORTIZ; VANDERGHEYNST, 2012), dentre outras. No entanto, tais técnicas foram fundamentalmente projetadas considerando-se as características de imagens obtidas no espectro visível. Portanto, embora essas técnicas funcionem bem em imagens deste espectro, elas frequentemente não possuem resultados satisfatórios quando utilizadas em imagens multiespectrais (LI et al., 2016; KUPFER; NETANYAHU; SHIMSHONI, 2015).

Com base nesse cenário e na necessidade de técnicas específicas para a solução desse problema, este trabalho propõe um novo método descritor para realizar correspondências de pontos-chave em imagens multiespectrais. A Figura 3 ilustra o cenário considerado neste trabalho em que o método descritor para correspondência de pontos-chave em imagens multiespectrais é aplicável.

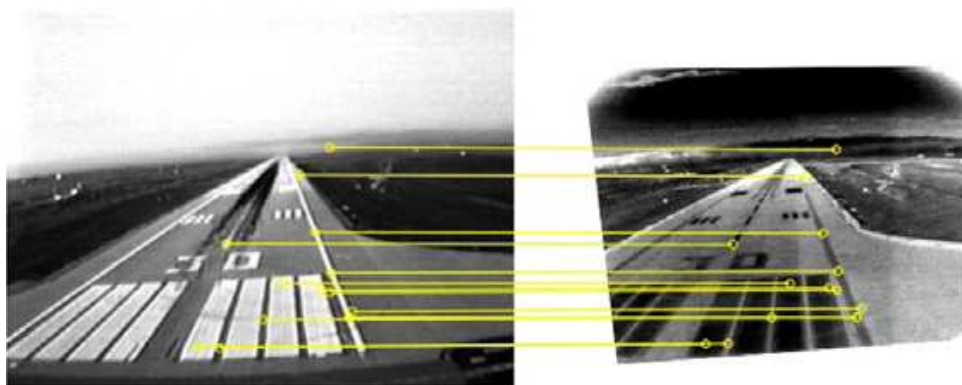


Figura 3 – Exemplo de cenário no qual o método descritor para correspondência de pontos-chave em imagens multiespectrais é aplicável. A figura é composta por um par de imagens: a primeira imagem foi capturada por uma câmera que opera no espectro de luz visível enquanto a segunda imagem foi capturada por uma câmera que opera no espectro infravermelho. Nesse exemplo são exibidos os pontos correspondentes entre o par de imagens.

Fonte: Adaptado de [Li et al. \(2016\)](#).

1.1 Caracterização do problema

Este trabalho apresenta um novo método descritor de características locais em imagens multiespectrais para ser aplicado em abordagens de correspondência de pontos-chave. O descritor proposto, denominado *Multispectral Feature Descriptor* (MFD), baseia-se na extração de informações das imagens nos domínios do espaço e da frequência por meio da utilização de filtros Log-Gabor.

A hipótese de pesquisa deste trabalho é de que ao utilizar filtros de Log-Gabor para extração de características locais em imagens multiespectrais para compor um descritor de pontos-chave obtenha-se correspondências com valores de revocação e precisão comparáveis aos dos métodos do estado da arte, porém com um custo computacional significativamente inferior, o qual amplia a utilização do método em aplicações onde este requisito seja crucial.

O cenário básico estudado neste trabalho, objetivando-se avaliar o método desenvolvido, é composto por um par de imagens obtidas em faixas distintas do espectro eletromagnético, mais especificamente, o espectro de luz visível e o espectro do infravermelho. Embora, o método proposto é potencialmente aplicável em imagens obtidas em outras faixas do espectro, escolheu-se avaliar seu desempenho com imagens obtidas nas faixas da luz visível e infravermelho, em função de suas amplas utilizações em aplicações científico-tecnológicas.

Ambas as imagens são obtidas de uma mesma cena, a partir de pontos de visão e/ou instantes de tempo diferentes. O descritor proposto neste trabalho possui uma aplicação

direta em sistemas de registro de imagens multiespectrais.

A Figura 4 exibe um diagrama que ilustra as principais etapas associadas ao problema de pesquisa abordado neste trabalho. Nessa figura é exibido um processo de correspondência de pontos-chave em imagens multiespectrais, composto por três etapas: detecção, descrição e correspondência. Uma câmera de luz visível captura uma imagem, denominada Imagem A, e uma câmera de infravermelho captura outra imagem, denominada Imagem B. Em cada uma das imagens são detectados pontos que serão correspondidos de forma a representar o mesmo ponto no mundo real. Em amarelo na Figura 4, destaca-se a etapa fundamentalmente estudada neste trabalho, qual seja, a etapa de descrição de cada ponto detectado (Etapa 2 da Figura 4).

O problema de descrição em imagens multiespectrais possui características peculiares, e para que o descritor proposto neste trabalho tenha sucesso, é necessário que

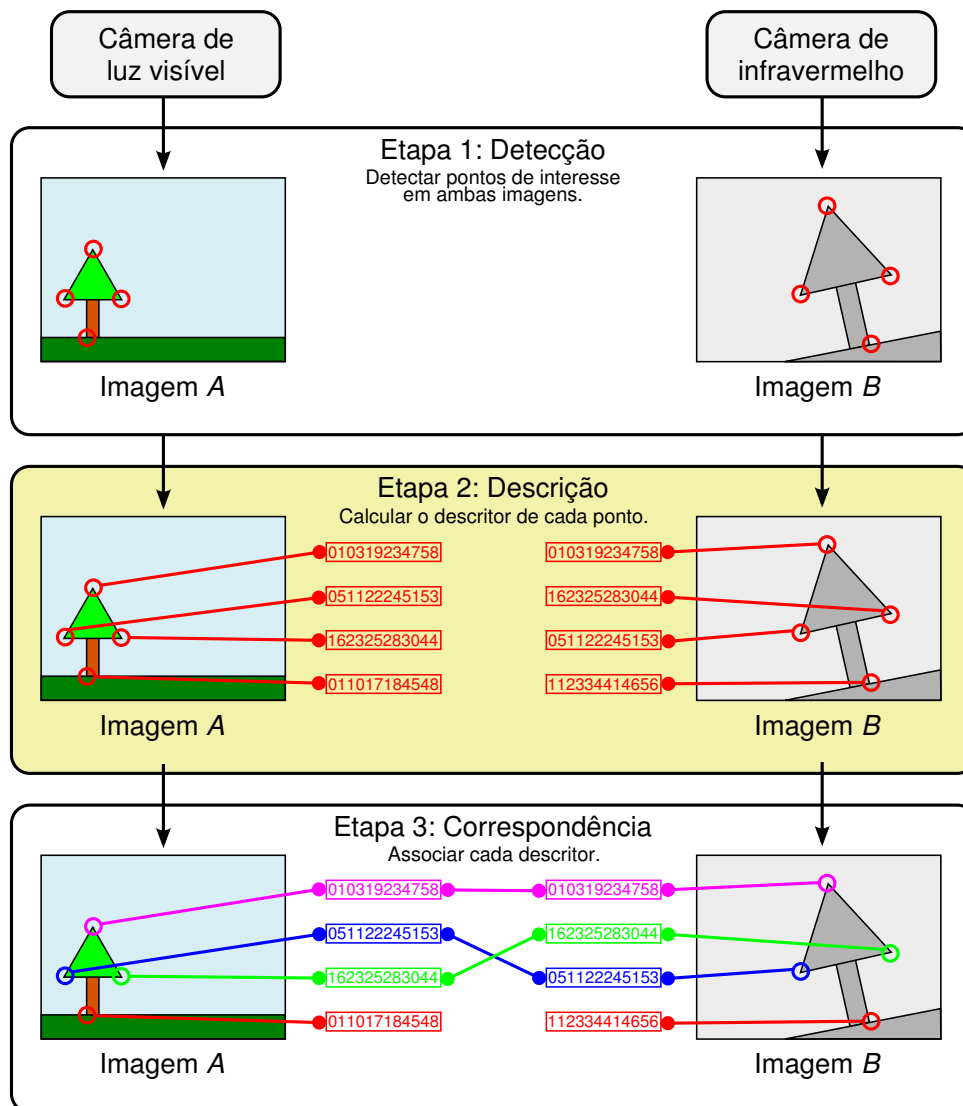


Figura 4 – Problema de pesquisa: Descrição em imagens multiespectrais para correspondência de pontos-chave.

sejam atendidos alguns requisitos, tais como: tolerância às variações dos gradientes de intensidade entre as imagens de diferentes espectros; tolerância às faixas de frequências de espectro, tais como: infravermelho próximo e infravermelho de ondas longas; baixo custo computacional, que poderá ser alcançado por meio da utilização de descritores de dimensões menores e de abordagens eficientes.

Para construir um sistema de correspondência de pontos desse trabalho é necessário que se tenha um método detector para realizar a detecção de pontos-chave em ambas as imagens. Para tanto, foi avaliado e selecionado um método detector de pontos, dentre os métodos do estado da arte, mais adequado para imagens multiespectrais. O desenvolvimento de um novo método para as etapas de detecção e correspondência (Etapa 1 e Etapa 3 da [Figura 4](#)) não é o escopo deste trabalho.

1.2 Motivação

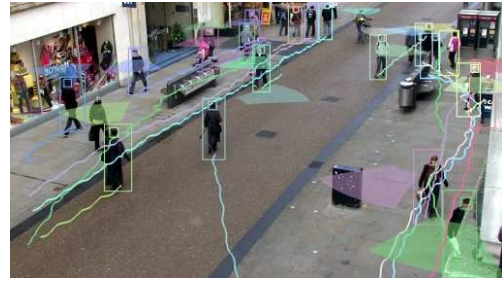
Encontrar correspondências em imagens multiespectrais consiste em uma parte fundamental de diversos problemas da Visão Computacional, como monitoramento de desastres naturais ([BALKAYA et al., 2015](#)), sistemas de vigilância por vídeo ([TORABI; MASSÉ; BILODEAU, 2012](#)), navegação de aeronaves guiadas por imagem ([OSBORNE et al., 2011](#)), melhoria de resolução de imagens ([SUN et al., 2015](#)), processamento de imagens multiespectrais em drones ([BOKOLONGA et al., 2016](#)), dentre outras, conforme os exemplos ilustrados na [Figura 5](#). No entanto, esta é uma tarefa muito difícil para as comunidades de Visão Computacional e Sensoriamento Remoto, pelo fato de que os pontos nos quais se desejam mapear entre imagens multiespectrais são, na maioria das vezes, localmente diferentes ([QIN et al., 2014](#)).

Com a finalidade de construir descritores confiáveis, pesquisadores têm adaptado descritores desenvolvidos para imagens do espectro visível para imagens multiespectrais ([DELLINGER et al., 2015](#); [SALEEM; SABLATNIG, 2014](#); [SALEEM; SABLATNIG, 2013](#)). Tais descritores adaptados ainda não fornecem uma alta taxa de correspondências corretas em certas imagens multiespectrais, especialmente quando a diferença entre a banda de duas imagens multiespectrais é muito grande. Além desses descritores não fornecerem um bom desempenho, é necessário muitas das vezes realizar uma etapa de refinamento para descartar as correspondências incorretas ([LI et al., 2016](#); [DELLINGER et al., 2015](#)).

Entre as técnicas de refinamento para remoção de correspondências incorretas, se destaca o método RANSAC (*RANdom SAmple Consensus*) ([FISCHLER; BOLLES, 1981](#)) e seus variantes ([WU et al., 2015](#); [TAN et al., 2014](#)). O problema com o método RANSAC e seus variantes é que seu desempenho é afetado pela alta taxa de correspondências incorretas, uma vez que esse algoritmo utiliza apenas correspondências como conjunto de entrada, desconsiderando outras informações da imagem ([LI et al., 2016](#); [LI et al., 2015c](#)).



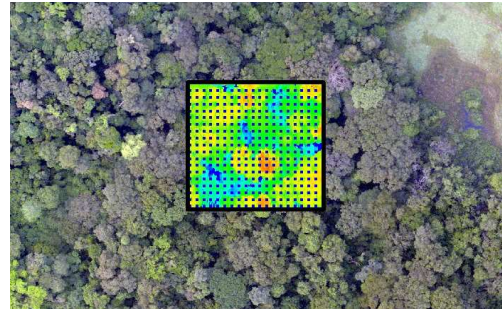
(a) Monitoramento de desastres naturais



(b) Sistemas de vigilância por vídeo



(c) Navegação de aeronaves guiadas por imagem



(d) Melhoria de resolução de imagens

Figura 5 – Exemplos de aplicações onde é fundamental a correspondência de pontos em imagens multiespectrais.

Nesse sentido, este trabalho busca melhorar a taxa de correspondências corretas por meio da utilização de um novo método descritor projetado especialmente para trabalhar com imagens multiespectrais. Além da obtenção de uma melhor taxa de correspondência correta, um outro fator que motiva a realização desse trabalho é a construção de um método descritor para correspondência que seja rápido, uma vez que grande parte dos trabalhos na literatura (KUPFER; NETANYAHU; SHIMSHONI, 2015; LI et al., 2015a) baseiam-se no método SIFT como base para detecção e descrição de pontos-chave, e já foi mostrado em alguns trabalhos (RUBLEE et al., 2011; LEUTENEGGER; CHLI; SIEGWART, 2011) que o SIFT consiste em uma técnica que possui alto custo computacional devido a sua complexidade e ineficiência para o tratamento de imagens multiespectrais (KUPFER; NETANYAHU; SHIMSHONI, 2015).

1.3 Objetivos: Geral e Específicos

O objetivo geral deste trabalho consiste em apresentar um novo método descritor de características locais em imagens multiespectrais para ser aplicado em abordagens de correspondência de pontos-chave. A principal contribuição desse trabalho consiste em explorar técnicas do estado da arte e construir um novo método descritor com uma eficácia maior e mais rápido que os métodos atualmente existentes. Por meio deste trabalho, busca-se também contribuir com pesquisas na área de Visão Computacional baseadas

na utilização ou mesmo no desenvolvimento de novas técnicas para correspondência de pontos.

Para que tal objetivo geral fosse atingido, alguns objetivos específicos foram considerados. A seguir se encontra uma lista desses objetivos:

- Pesquisar e implementar algoritmos do estado da arte para detecção automática de pontos-chave em imagens multiespectrais;
- Pesquisar e implementar algoritmos do estado da arte para discriminar cada ponto detectado, para que estes possam ser unicamente identificados;
- Realizar experimentos com diferentes bases de dados de imagens multiespectrais. Por meio destes experimentos, avaliou-se a eficácia e a eficiência do método descritor proposto, comparando-o com métodos descritores do estado da arte.

1.4 Contribuições

A principal contribuição desse trabalho consiste em apresentar um novo método descritor de características locais em imagens multiespectrais que possa ser aplicado em abordagens de correspondência de pontos-chave. Assim, pode contribuir para o desenvolvimento em pesquisas na área de Visão Computacional e Sensoriamento Remoto, baseadas na utilização ou mesmo no desenvolvimento de métodos descritores.

Diferentemente de alguns trabalhos recentemente propostos (como, por exemplo, os trabalhos de [Saleem e Sablatnig \(2014\)](#), [Dellinger et al. \(2015\)](#), dentre outros detalhados no [Capítulo 2](#)), o método descritor proposto não se baseia no método SIFT. Além disso, não utiliza o cálculo de gradientes para a descrição de pontos-chave, que em imagens de espectro diferentes não é desejável, e propõe uma nova forma mais eficiente para a utilização de filtros Log-Gabor, com descritores de dimensões menores, em comparação com o descritor LGHD proposto no trabalho de [Aguilera, Sappa e Toledo \(2015\)](#).

1.5 Organização do texto

Este trabalho está organizado em seis capítulos, incluindo o presente. O [Capítulo 2](#), apresenta os principais estudos relacionados ao tema deste trabalho, descrevendo seus resultados e suas contribuições. A fundamentação teórica utilizada para o desenvolvimento do trabalho encontra-se no [Capítulo 3](#). No [Capítulo 4](#) é apresentado o descritor proposto para correspondência de pontos em imagens multiespectrais, com a descrição dos principais conceitos e teorias que fundamentam o desenvolvimento desse método computacional. No

Capítulo 5 são apresentados e discutidos os resultados experimentais obtidos. Por fim, no Capítulo 6 são apresentadas as conclusões e perspectivas de trabalhos futuros.

2 Trabalhos Relacionados

Este capítulo apresenta alguns dos principais estudos relacionados ao problema de pesquisa abordado, comparando-os sempre que possível com a metodologia proposta neste trabalho. Com o objetivo de organizar os assuntos aqui apresentados, este capítulo está dividido em três seções.

A primeira seção descreve trabalhos que propõe detectores e descritores para imagens obtidas no espectro de luz visível. Vale ressaltar a existência de uma diversidade de pesquisas nesta área e que, portanto, são apresentadas e discutidas apenas um subconjunto das mesmas. Este subconjunto é composto por trabalhos que foram identificados como mais diretamente relacionados com o presente trabalho e cujas metodologias se assemelham à metodologia aqui apresentada. A segunda seção aborda trabalhos relacionados a correspondência de pontos-chave bem como trabalhos que tratam da construção de descritores para imagens multiespectrais. Por fim, na última seção é feita uma discussão final sobre o estado da arte no que diz respeito a descrição de pontos-chave em imagens multiespectrais, os principais desafios atuais neste tópico e as relações entre a metodologia aqui proposta e as estratégias utilizadas em trabalhos anteriores.

2.1 Métodos detectores e descritores para correspondência de imagens em geral

No âmbito da solução do problema de correspondência de pontos em imagens que utiliza a abordagem baseada em características, os algoritmos detectores e descritores de pontos-chave são de grande importância. O algoritmo SIFT (*Scale-Invariant Feature Transform*), proposto por [Lowe \(2004\)](#), é um dos mais populares algoritmos utilizados em correspondência de pontos em imagens, sendo composto por um método detector e um método descritor de pontos. Este algoritmo se mostra robusto às variações de translação, rotação e escala no processo de extração de características.

O algoritmo SURF (*Speeded Up Robust Features*), proposto por [Bay, Tuytelaars e Gool \(2006\)](#), é baseado no algoritmo SIFT. Este algoritmo apresenta um desempenho de tempo de processamento melhor, em comparação com o algoritmo SIFT. Mais recentemente, algoritmos mais rápidos foram propostos, e ao contrário dos algoritmos SIFT e SURF, eles utilizam um vetor binário para a representação do descritor em vez de vetores que utilizam uma representação numérica em pontos flutuantes.

No trabalho de [Mikolajczyk et al. \(2005\)](#) foi realizado um estudo sobre algoritmos detectores de pontos-chave, e os autores definem a métrica de repetibilidade para avaliar o

desempenho de detecção desses algoritmos. Em um outro trabalho do mesmo grupo de pesquisa, em Mikolajczyk e Schmid (2005), foi realizado um levantamento sobre algoritmos descritores de pontos-chave. Neste trabalho são utilizadas duas métricas para avaliar os algoritmos descritores: precisão e revocação. Os autores mostram em seu trabalho que os algoritmos baseados no algoritmo SIFT podem localizar com boa exatidão pontos-chave correspondentes mesmo sob condições extremas. A metodologia de avaliação e as métricas de repetibilidade, precisão e revocação, definidas nesses trabalhos, serviram como base para avaliar a eficácia do método descritor proposto nesse trabalho.

Algumas técnicas sacrificam a qualidade do processo de descrição para alcançar a vantagem do baixo tempo de processamento. Por exemplo, o algoritmo BRIEF (*Binary Robust Independent Elementary Features*) - proposto por Calonder et al. (2010) segue esta estratégia, mas não é capaz de encontrar correspondências corretas quando a escala da imagem é alterada. Baseados neste algoritmo, os algoritmos ORB (*Oriented FAST and Rotated BRIEF*) (RUBLEE et al., 2011), BRISK (*Binary Robust Invariant Scalable Keypoints*) (LEUTENEGGER; CHLI; SIEGWART, 2011) e FREAK (*Fast Retina Keypoint*) (ALAH!; ORTIZ; VANDERGHEYNST, 2012) foram desenvolvidos para manter um balanço entre o custo computacional e qualidade do processo de descrição, visando ser uma alternativa eficiente comparada ao SIFT e ao SURF. Tais métodos descritores foram estudados e avaliados em alguns trabalhos como, por exemplo, no estudo comparativo apresentado por Heinly, Dunn e Frahm (2012), onde são avaliados os algoritmos detectores e descritores binários (SIFT, SURF, BRIEF, ORB e BRISK) para o processo de correspondência de pontos em imagens.

No estudo apresentado por Jiang, Xu e Liu (2013) é feita uma avaliação sobre quais algoritmos detectores e descritores de características são mais apropriados para a solução do problema de odometria visual¹. Neste trabalho são analisados algoritmos detectores e descritores (SIFT, SURF, FAST, BRIEF, BRISK) e avaliados sua robustez, precisão e custo computacional. A metodologia utilizada para a avaliação dos algoritmos contém um processo de correspondência de pontos em imagens, de forma similar ao processo utilizado nesse trabalho. Primeiro é necessário detectar os pontos-chave nas duas imagens utilizando um detector de pontos-chave. Depois, utiliza-se um descritor de características para representar um ponto-chave detectado em uma forma distintiva. A correspondência dos pontos é obtida por meio da comparação dos descritores de características baseadas em um critério de similaridade, no qual foi utilizado a distância de Hamming para os descritores binários e a distância euclidiana (norma L_2) para os demais descritores. Nos resultados os autores concluem que os detectores baseados em cantos (por exemplo, FAST e BRISK) são superiores para a solução em projetos de odometria visual estéreo e outras aplicações gerais, devido ao seu baixo custo computacional e bom desempenho.

¹Odometria visual, em robótica, consiste em uma técnica para estimar a posição e o movimento de robôs por meio de uma única câmera ou múltiplas câmeras.

Como o presente trabalho não aborda a construção de um método detector, se faz necessário definir um detector específico de pontos-chave para avaliar o desempenho do método descritor proposto. Para tanto, os trabalhos de Jiang, Xu e Liu (2013) e Heinly, Dunn e Frahm (2012) serviram de base metodológica para avaliar e selecionar um método detector que seja mais adequado para o processo de correspondência de pontos-chave desse trabalho. Sendo assim, para selecionar um detector mais apropriado para imagens multiespectrais, foram avaliados (ver Seção 4.3) os métodos detectores SIFT, SURF, FAST, BRISK e ORB por serem métodos comumente utilizados na literatura, em especial o método detector SIFT que é utilizado em grande parte dos trabalhos aqui apresentados e o método detector FAST por ter sido empregado no trabalho de Aguilera, Sappa e Toledo (2015) como detector de pontos em imagens que contém variações não lineares.

Uma característica importante que é exigida nos métodos detectores é que seu funcionamento não deve se basear no cálculo de gradientes para encontrar os pontos-chave na imagem, visto que o gradiente entre imagens de espectros diferentes podem diferir. Na Seção 3.3 deste trabalho será descrito o funcionamento de cada detector aqui avaliado.

2.2 Métodos descritores para imagens multiespectrais

No trabalho de Brown e Susstrunk (2011) são apresentadas modificações no algoritmo SIFT para tratar da descrição de pontos-chave em imagens multiespectrais para categorização de cenas. O método descritor resultante foi denominado de MSIFT. As modificações consistem basicamente em calcular a diferença de gaussianas no componente de luminância da imagem e calcular, em seguida, histogramas de gradientes orientados em cada canal RGB da imagem. Os autores ainda incluem a técnica de análise de componentes principais (PCA) para reduzir a dimensão do descritor resultante (dimensão 512). Outra contribuição desse trabalho consiste na criação de uma base de dados² formadas por centenas de pares de imagens do espectro visível e do espectro infravermelho. Os autores mostram em seu trabalho que as imagens de espectro visível possuem baixa correlação com imagens do espectro infravermelho, e que ainda assim, o algoritmo descritor proposto, pode explorar essas diferenças entre os pares de imagens para serem aplicadas em categorização de cenas. Apesar da melhora na eficácia do algoritmo SIFT, os autores perdem na eficiência do algoritmo, ao processar todos os canais RGB da imagem de entrada e incluir uma etapa de redução de dimensionalidade nessa abordagem.

Com o objetivo de criar um método descritor mais eficiente e diferente das técnicas utilizadas no algoritmo SIFT, nos trabalhos de Aguilera et al. (2012a) e Aguilera et al. (2012b) é apresentado um novo algoritmo descritor que extrai características por meio do cálculo de bordas orientadas, denominado EOH (*Edge Oriented Histogram*). Os autores

²Disponível em: <http://ivrl.epfl.ch/supplementary_material/cvpr11/index.html>.

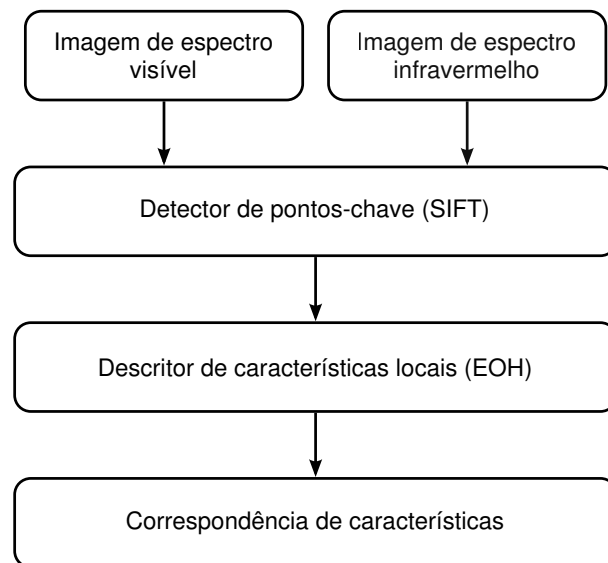


Figura 6 – Visão geral do processo SIFT-EOH para correspondência de pontos-chave em imagens multiespectrais.

Fonte: Adaptado de [Aguilera et al. \(2012a\)](#).

avaliam o método descritor em um processo de correspondência de pontos em imagens multiespectrais, denominado de SIFT-EOH, que é formado pelo método detector de pontos SIFT e o método descritor EOH. A [Figura 6](#) ilustra o processo de correspondência desenvolvido nesse trabalho. O método descritor EOH é baseado no método descritor EHD (*Edge Histogram Descriptor*) definido no padrão MPEG-7 ([SIKORA, 2001](#)) e incorpora informações espaciais de contornos da vizinhança de cada ponto detectado. Para extrair as informações espaciais de contorno, os autores utilizam o detector de bordas Canny ([CANNY, 1986](#)). Segundo [Qin et al. \(2014\)](#), o desempenho do descritor EOH em relação à precisão, no entanto, não é satisfatória principalmente pelo fato de que o detector Canny utilizado no EOH necessita de limiares definidos para um determinado conjunto específico de imagens, e a extração de bordas em imagens multiespectrais podem variar, sendo difícil a seleção de limiares adequados para a extração de bordas em imagens multiespectrais. Apesar da limitação do descritor EOH, ele é um método que tem sido empregado em pesquisas recentes que abordam o processamento de imagens multiespectrais em geral, como em pesquisas de registro de imagens multiespectrais ([JIN et al., 2016](#)), identificação aérea de veículos ([JANNEY; BOOTH, 2015](#)), reconhecimento de faces ([MISHRA et al., 2016](#)), dentre outras.

Uma outra abordagem baseada no método descritor SIFT, denominada de NG-SIFT, é apresentada nos trabalhos de [Saleem e Sablatnig \(2013\)](#) e [Saleem e Sablatnig \(2014\)](#). Esse método descritor é construído sob a hipótese de que as bordas em imagens multiespectrais são razoavelmente invariantes. Como principal modificação no algoritmo SIFT original, é utilizado o cálculo de contraste local no lugar de magnitude de gradientes

para a construção do descritor. Nestes trabalhos os autores mostram que o descritor proposto possui uma precisão melhor, em comparação com o algoritmo SIFT comum quando aplicado às imagens multiespectrais. Para a avaliação da eficácia, os autores utilizam as métricas de precisão e revocação, e apesar desse algoritmo possuir uma precisão superior em comparação com o algoritmo SIFT original, os autores não aperfeiçoaram a eficiência em relação ao tempo de processamento do algoritmo SIFT. Sendo assim, o processo de descrição de pontos em imagens multiespectrais apresentado nestes trabalhos possui uma complexidade similar ao descritor SIFT original.

Os algoritmos descritores construídos originalmente para trabalhar com imagens do espectro visível foram avaliados em relação a sua eficácia em imagens multiespectrais. No estudo empírico apresentado por [Ricaurte et al. \(2014\)](#) são avaliados o comportamento de métodos descritores clássicos quando utilizados em imagens do espectro infravermelho de ondas longas e comparados com resultados obtidos com imagens no espectro visível. Nesse estudo, são avaliadas a robustez dos descritores quanto a alguns fatores como: mudanças de rotação, escala, borramento (do inglês *blur*) e ruídos adicionais. Foram avaliados a precisão e revocação dos algoritmos SIFT, SURF, ORB, BRISK, BRIEF e FREAK. Os autores concluem que o algoritmo SIFT obteve uma eficácia (precisão e revocação) melhor, em grande parte dos experimentos.

Em [Qin et al. \(2014\)](#) é proposto um algoritmo descritor de pontos-chave para imagens multiespectrais baseado em histograma de gradientes inspirado no algoritmo EOH, denominado de HCGEC. Diferentemente do algoritmo EOH que utiliza o detector Canny para extrair informações de bordas, o algoritmo codifica as estruturas de bordas da imagem apenas por meio de gradientes e suprime informações de textura, que de acordo com os autores, é desejável em imagens multiespectrais. Os autores utilizam duas bases de dados multiespectrais para demonstrar que o descritor proposto pode produzir resultados comparáveis ao descritor EOH e melhor que os descritores SIFT e SURF. Para a avaliação do descritor, os autores utilizam a métrica de precisão e revocação de pontos-chave. Apesar de o algoritmo proposto apresentar bons valores de precisão e revocação, o algoritmo possui uma complexidade computacional alta.

No trabalho de [Dellinger et al. \(2015\)](#) são apresentadas modificações no algoritmo SIFT para a utilização em imagens SAR (*Synthetic Aperture Radar*). O novo algoritmo proposto recebe o nome de SAR-SIFT. Esse algoritmo inclui tanto a detecção de pontos-chave quanto o cálculo de descritores locais. As modificações realizadas incluem um novo cálculo de gradiente adaptado para imagens SAR que também tratam de ruídos em imagens. O detector de pontos-chave utilizado nesse trabalho é baseado no detector de cantos Harris ([HARRIS; STEPHENS, 1988](#)). Para a remoção de falsas correspondências, os autores utilizam o algoritmo RANSAC na etapa final do processo de correspondência de pontos-chave. Neste trabalho, os autores não avaliam a eficiência do algoritmo proposto em relação

ao algoritmo SIFT.

Abordagens para o registro de imagens multiespectrais por meio da realização de correspondência de pontos-chave também são apresentadas nos trabalhos de [Li et al. \(2016\)](#) e [Li et al. \(2015b\)](#). Nestes trabalhos são utilizadas como informações o cálculo dos descritores dos pontos-chave e informações globais contidas em toda imagem. Os resultados experimentais mostram que o método proposto pode, de forma eficiente, estabelecer correspondências de pontos-chave em imagens multiespectrais. Os autores utilizam o método descritor EOH para realizar a descrição dos pontos detectados. Neste trabalho é apontado que a utilização apenas de descritores propostos no estado da arte é insuficiente para realizar correspondências robustas mesmo com a utilização da técnica RANSAC. Para solucionar o problema, os autores utilizam restrições LISK (*Lines Intersecting Segments between Keypoints*), como informação global, para remover correspondências consideradas *outliers* e que podem ser facilmente identificadas.

Com o objetivo de apresentar um método descritor para correspondência de imagens que possuem variações de intensidades não lineares, os autores [Aguilera, Sappa e Toledo \(2015\)](#) propõem um método descritor denominado LGHD (*Log-Gabor Histogram Descriptor*) que descreve a vizinhança de pontos detectados por meio da combinação de informações das imagens nos domínios do espaço e da frequência. A extração dessas informações é realizada por meio dos filtros Log-Gabor. Os autores utilizam três bases de dados para abranger uma grande variedade de imagens que contém variações de intensidades não lineares: imagens de mapas de profundidade, imagens multiespectrais e imagens obtidas por meio de câmeras com e sem flash. Para avaliar o desempenho do método descritor em imagens multiespectrais, os autores utilizam uma base de imagens do espectro infravermelho próximo e uma base de imagens do espectro infravermelho de ondas longas. Este método descritor utiliza 24 filtros Log-Gabor em 16 regiões em torno de um ponto detectado para descrevê-lo, resultando em um vetor de dimensão 384. Para a avaliação do método descritor dentro de um processo de correspondência de pontos-chave, foi utilizado o algoritmo FAST ([ROSTEN; DRUMMOND, 2006](#)), por ser considerado um método de detecção de cantos rápido e robusto. Como crítica, os autores utilizam apenas a métrica de precisão para comparar o método proposto com os métodos SIFT e EOH. A avaliação da revocação e da eficiência do algoritmo, em termos de tempo de processamento, não são apresentadas nesse trabalho. O descritor desenvolvido foi utilizado em pesquisas de odometria visual para imagens multiespectrais, especialmente no trabalho de [Mouats et al. \(2015\)](#).

2.3 Discussão

O número de trabalhos que abordam a correspondência de pontos em imagens multiespectrais têm aumentado significativamente ao longo dos anos, em diferentes áreas. Assim como o presente trabalho, grande parte dos trabalhos aqui apresentados (LI et al., 2016; AGUILERA; SAPPA; TOLEDO, 2015; AGUILERA et al., 2012b) utilizam imagens do espectro visível e imagens do espectro infravermelho visando-se avaliar seus desempenhos. Como o estado da arte mostra, novos métodos descritores têm surgido, o que evidencia a importância de um estudo efetivo nesta área. Fica evidenciado também, por meio da revisão bibliográfica, que os algoritmos responsáveis pela descrição de pontos-chave possui grande importância em um processo de correspondência de pontos em imagens, sendo esta etapa considerada crítica e com considerável custo computacional.

Algoritmos descritores comumente utilizados na literatura e desenvolvidos para imagens do espectro visível (por exemplo, SIFT, SURF, BRIEF, ORB, BRISK e FREAK) foram avaliados e comparados, como nos trabalhos de Heinly, Dunn e Frahm (2012) e Jiang, Xu e Liu (2013), e tais métodos foram posteriormente comparados no âmbito da utilização em imagens multiespectrais, no trabalho de Ricaurte et al. (2014). Como visto na literatura, o algoritmo SIFT se destaca e serve de inspiração para criação de novos métodos descritores ou na utilização em novas abordagens para correspondência de pontos-chave, como visto nos trabalhos de Brown e Susstrunk (2011), Saleem e Sablatnig (2013), Saleem e Sablatnig (2014) e Dellinger et al. (2015).

Em se tratando de métodos desenvolvidos exclusivamente para tratar com imagens multiespectrais, o descritor EOH, descrito nos trabalhos de Aguilera et al. (2012a) e Aguilera et al. (2012b), se destaca e serve de inspiração para outros trabalhos, como em Qin et al. (2014), Aguilera, Sappa e Toledo (2015), Li et al. (2015b), Li et al. (2016), e como visto anteriormente, o descritor EOH é um método que tem sido empregado em pesquisas recentes que abordam o processamento de imagens multiespectrais em geral. Vale lembrar que a abordagem utilizada para a construção do algoritmo LGHD, apresentado no trabalho de Aguilera, Sappa e Toledo (2015) se mostra promissora, uma vez que tal descritor soluciona o problema da definição de limiares do método descritor EOH sem se basear no algoritmo SIFT para sua construção, como é feito nos trabalhos em Dellinger et al. (2015), Saleem e Sablatnig (2014) e Brown e Susstrunk (2011).

Nesse sentido, os descritores EOH e LGHD serviram de base teórica para a construção do descritor desse trabalho. Visando sanar a deficiência do descritor EOH quanto a definição de limiares e o algoritmo LGHD devido sua complexidade e alta dimensionalidade dos descritores (vetores descritores de dimensão 384), esse trabalho propõe um descritor que possua a eficiência do algoritmo EOH, em termos de tempo de processamento, e a eficácia do método descritor LGHD, em termos de precisão e revocação.

Os trabalhos analisados neste capítulo mostram também que o processo de correspondência de pontos-chave em imagens multiespectrais, que utilizam uma abordagem baseada em características, pode ser dividido em etapas bem definidas (detecção, descrição e correspondência) assim como é comumente feito em trabalhos que abordam o processo de correspondência em imagens do espectro visível. Isto contribui para avaliação do desempenho do método aqui proposto, já que é possível avaliar cada uma destas etapas separadamente, tornando assim o processo de avaliação mais simples.

Sobre a metodologia de avaliação, as métricas de repetibilidade, precisão e revocação, definidas nos trabalhos de Mikolajczyk et al. (2005) e Mikolajczyk e Schmid (2005), ainda são utilizadas em grande parte dos trabalhos que abordam a construção de métodos descritores e processos de correspondência de pontos em imagens multiespectrais, como, por exemplo, nos trabalhos de Aguilera, Sappa e Toledo (2015), Qin et al. (2014), Ricaurte et al. (2014) e Aguilera et al. (2012a). Essas métricas também serão utilizadas aqui, para avaliar o sistema proposto, e explicadas na seção [Seção 3.5](#) desse mesmo trabalho.

Dentre os trabalhos descritos, alguns como Li et al. (2016), Li et al. (2015b), Aguilera, Sappa e Toledo (2015) e Aguilera et al. (2012b) apresentam metodologias semelhantes a aqui proposta, no que se refere ao processo de correspondência de pontos-chave. Nestes trabalhos, os algoritmos detectores e descritores são avaliados de forma individual, assim como é feito neste trabalho. Estes estudos foram, portanto, fontes de inspiração para a metodologia utilizada neste estudo. O [Quadro 1](#) exhibe os principais métodos descritores apresentados neste capítulo, listados em ordem cronológica.

Trabalho	Método descritor	Proposto para imagens multiespectrais
(LOWE, 2004)	SIFT	Não
(BAY; TUYTELAARS; GOOL, 2006)	SURF	Não
(CALONDER et al., 2010)	BRIEF	Não
(RUBLEE et al., 2011)	ORB	Não
(LEUTENEGGER; CHLI; SIEGWART, 2011)	BRISK	Não
(BROWN; SUSSTRUNK, 2011)	MSIFT	Sim
(ALAHY; ORTIZ; VANDERGHEYNST, 2012)	FREAK	Não
(AGUILERA et al., 2012b)	EOH	Sim
(SALEEM; SABLATNIG, 2014)	NG-SIFT	Sim
(QIN et al., 2014)	HCGEC	Sim
(DELLINGER et al., 2015)	SAR-SIFT	Sim
(AGUILERA; SAPPA; TOLEDO, 2015)	LGHD	Sim

Quadro 1 – Principais métodos descritores apresentados na revisão bibliográfica.

3 Fundamentação Teórica

Neste capítulo são introduzidos alguns dos principais conceitos e técnicas que fundamentam o desenvolvimento da metodologia proposta nesse trabalho. Para o desenvolvimento de um método descritor de pontos-chave, se faz necessário estudar as principais técnicas de detecção e descrição de pontos-chave comumente utilizadas na literatura.

Serão apresentados aqui, o funcionamento dos métodos detectores e descritores para imagens de espectro visível SIFT, SURF, BRISK, FAST e ORB. Os métodos detectores desses algoritmos foram avaliados nesse trabalho, visando-se definir o método mais adequado a ser aplicado conjuntamente com o descritor proposto para correspondências de pontos em imagens multiespectrais. Também serão apresentados os métodos descritores EOH e LGHD, que inspiraram a construção do descritor desenvolvido nesse trabalho.

O processo de correspondência de pontos-chave em imagens multiespectrais, baseado na extração de características, pode ser dividido em três etapas: detecção, descrição e correspondência. No [Algoritmo 1](#) são mostrados os passos deste processo, em que p_A e p_B consistem nos conjuntos de pontos-chave detectados nas imagens A (por exemplo, obtidas no espectro visível) e imagens B (por exemplo, obtidas no espectro do infravermelho) respectivamente, d_A e d_B consistem nos conjuntos de vetores descritores referentes aos pontos-chave detectados e v é o conjunto que contém pares de pontos correspondidos entre as imagens A e B .

Algoritmo 1: Correspondência de pontos baseada em características

Entrada: Imagem A , Imagem B

Saída: Conjunto v que contém as correspondências dos pontos entre A e B

$p_A \leftarrow \text{Detector}(A)$

$p_B \leftarrow \text{Detector}(B)$

$d_A \leftarrow \text{Descritor}(A, p_A)$

$d_B \leftarrow \text{Descritor}(B, p_B)$

$v \leftarrow \text{Correspondência}(d_A, d_B)$

3.1 Imagens multiespectrais

Uma imagem multiespectral consiste em um conjunto de imagens de um mesmo objeto, capturadas por diferentes faixas de comprimentos de ondas eletromagnéticas, tais como: infravermelho, luz visível, raio X ou qualquer outra faixa do espectro. Cada imagem desse conjunto, é portanto, uma referência para uma banda do espectro eletromagnético. As imagens referente a cada banda podem ser visualizadas em níveis de cinza, ou podem ser combinadas para sintetizar imagens coloridas ([SCHOWENGERDT, 2006](#)).

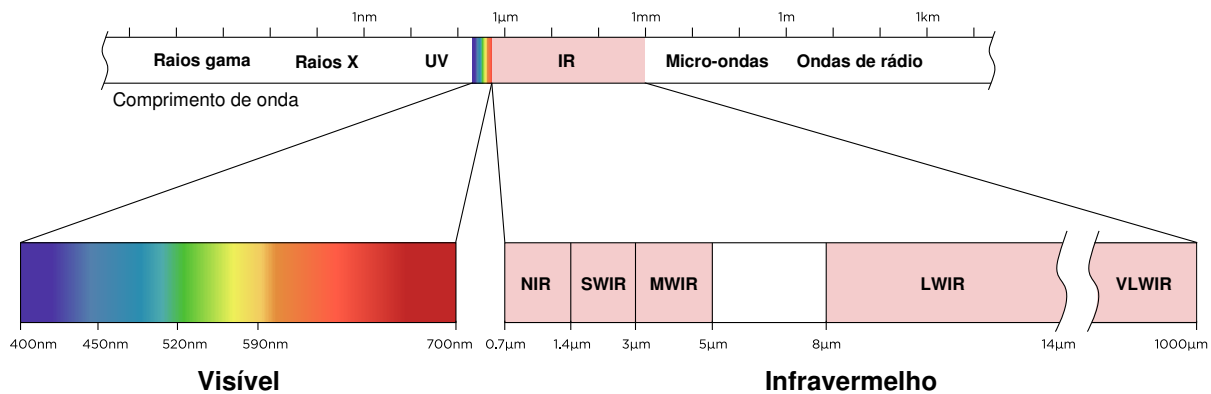


Figura 7 – Espectro eletromagnético.

A [Figura 7](#) ilustra o espectro eletromagnético para diversos tipos de comprimento de onda. A faixa do espectro, referente ao infravermelho, pode ser subdividida de acordo com diferentes faixas de comprimentos de onda ([RICAURTE et al., 2014](#); [AGUILERA et al., 2012a](#); [KRISHNA, 2005](#)), quais sejam: infravermelho próximo (NIR - *Near-Infrared*), infravermelho de ondas curtas (SWIR - *Short-Wave Infrared*), infravermelho de ondas médias (MWIR - *Mid-Wave Infrared*), infravermelho de ondas longas (LWIR - *Long-Wave Infrared*) e infravermelho de ondas muito longas (VLWIR - *Very Long-Wave Infrared*). A [Figura 8](#) ilustra uma imagem multiespectral composta por algumas dessas faixas do espectro eletromagnético.

3.2 Filtros Log-Gabor

Os filtros de Gabor ([GABOR, 1946](#)) consistem em um grupo de *wavelets*¹, em que cada *wavelet* é responsável por uma determinada frequência e orientação. Tais filtros têm

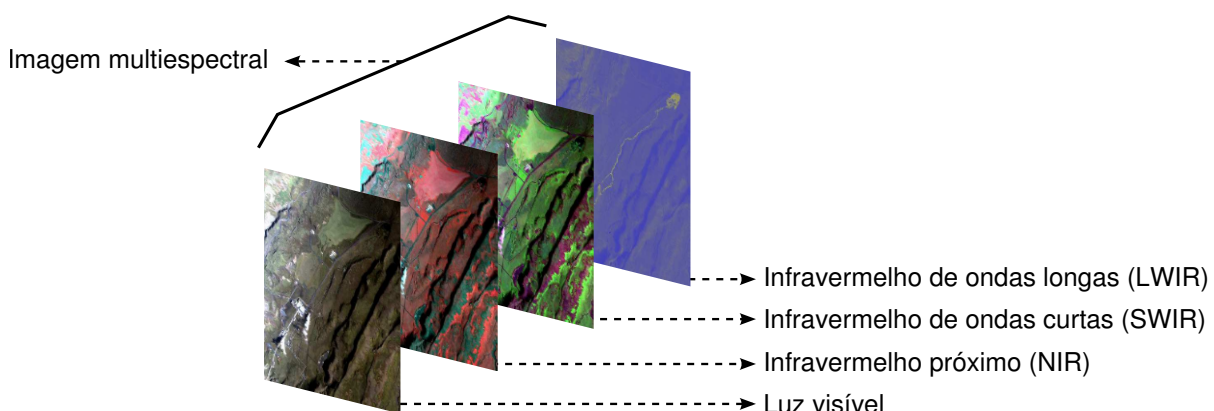


Figura 8 – Exemplo de imagem multiespectral.

¹ *Wavelet*, ou ondaleta, consiste em uma função capaz de representar qualquer outra função como superposições de ondas, de forma a decompô-las em diferentes níveis de frequência e escala ([ANTONINI et al., 1992](#)).

sido comumente utilizados na literatura para extração de características de textura de uma imagem e detecção de bordas (NAVA; ESCALANTE-RAMÍREZ; CRISTÓBAL, 2012). A possibilidade de ajustar a escala e a orientação do filtro Gabor o torna bastante útil para a análise de textura (WALIA; VERMA, 2016).

No entanto, os filtros de Gabor geralmente possuem um componente contínuo² que implica que para uma região constante de uma imagem, a resposta do filtro dependerá do valor de intensidade dessa região. Dessa forma, regiões constantes com valores elevados de intensidade de brilho resultam em respostas elevadas. Esse componente gera um problema, pois geralmente as respostas mais altas são esperadas apenas em bordas para extração de características e não no caso de regiões com alta intensidade de brilho e constantes (WALIA; VERMA, 2016; KOVESI, 1999).

Uma alternativa aos filtros de Gabor são os filtros de Log-Gabor propostos no trabalho de Field (1987). Os filtros Log-Gabor eliminam o componente contínuo dos filtros de Gabor que os torna robustos às variações de iluminação. Além disso, eles têm uma resposta Gaussiana quando visualizada em uma escala de frequência logarítmica, permitindo assim que mais informações sejam extraídas.

A transformação Log-Gabor (WALIA; VERMA, 2016; KOVESI, 2000) no domínio da frequência utilizando coordenadas polares é dada pela Equação (1):

$$G(r, \theta) = \exp\left(-\frac{[\log(r/f_0)]^2}{2\sigma^2}\right) \cdot \exp\left(-\frac{(\theta - m)^2}{2}\right), \quad (1)$$

em que r e θ representam o raio e o ângulo do filtro em coordenadas polares, m consiste no ângulo de orientação do filtro, σ representa o desvio padrão gaussiano do filtro e f_0 representa a frequência central do filtro dada pela Equação (2):

$$f_0 = \lambda k^n, \quad (2)$$

em que λ consiste no comprimento de onda da menor escala do filtro, k representa uma constante multiplicativa de escala entre filtros sucessivos e n consiste a escala do filtro.

Reescrevendo as Equações (1) e (2), tem-se:

$$G(r, \theta, m, n) = \exp\left(-\frac{[\log(\frac{r}{\lambda k^n})]^2}{2\sigma^2}\right) \cdot \exp\left(-\frac{(\theta - m)^2}{2}\right). \quad (3)$$

A transformação Log-Gabor é implementada em uma imagem no domínio da frequência. Por meio da Transformada Rápida de Fourier (FFT), a imagem é transformada do domínio espacial para o domínio da frequência. Então, é realizada a convolução entre a

²Um componente contínuo, ou *DC component*, consiste no valor médio de uma forma de onda quando descrevemos uma onda periódica no domínio do tempo (SCHOUHAMER; PÁTROVICS, 1997).

imagem transformada com o filtro Log-Gabor. A imagem de resposta ao filtro Log-Gabor é obtida por meio do cálculo inverso da Transformada Rápida de Fourier (IFFT), como mostra a [Equação \(4\)](#):

$$I_{m,n} = |\text{IFFT}(\text{IFFT}(I) * G_{m,n})|, \quad (4)$$

em que $I_{m,n}$ consiste na magnitude da imagem de resposta ao filtro Log-Gabor, I representa a imagem de entrada e $G_{m,n}$ representa o filtro Log-Gabor para a orientação m e escala n .

A [Figura 9](#) ilustra imagens de resposta para 12 filtros Log-Gabor dado uma imagem de entrada por meio da utilização de um banco de filtros compostos por 4 orientações ($m = 0, 1, 2, 3$) e 3 escalas ($n = 0, 1, 2$) resultando em um banco de filtros de $4 \times 3 = 12$ filtros. Neste exemplo, foram definidos os parâmetros: $\lambda = 3,0$, $k = 2,0$ e $\sigma = 0,65$.

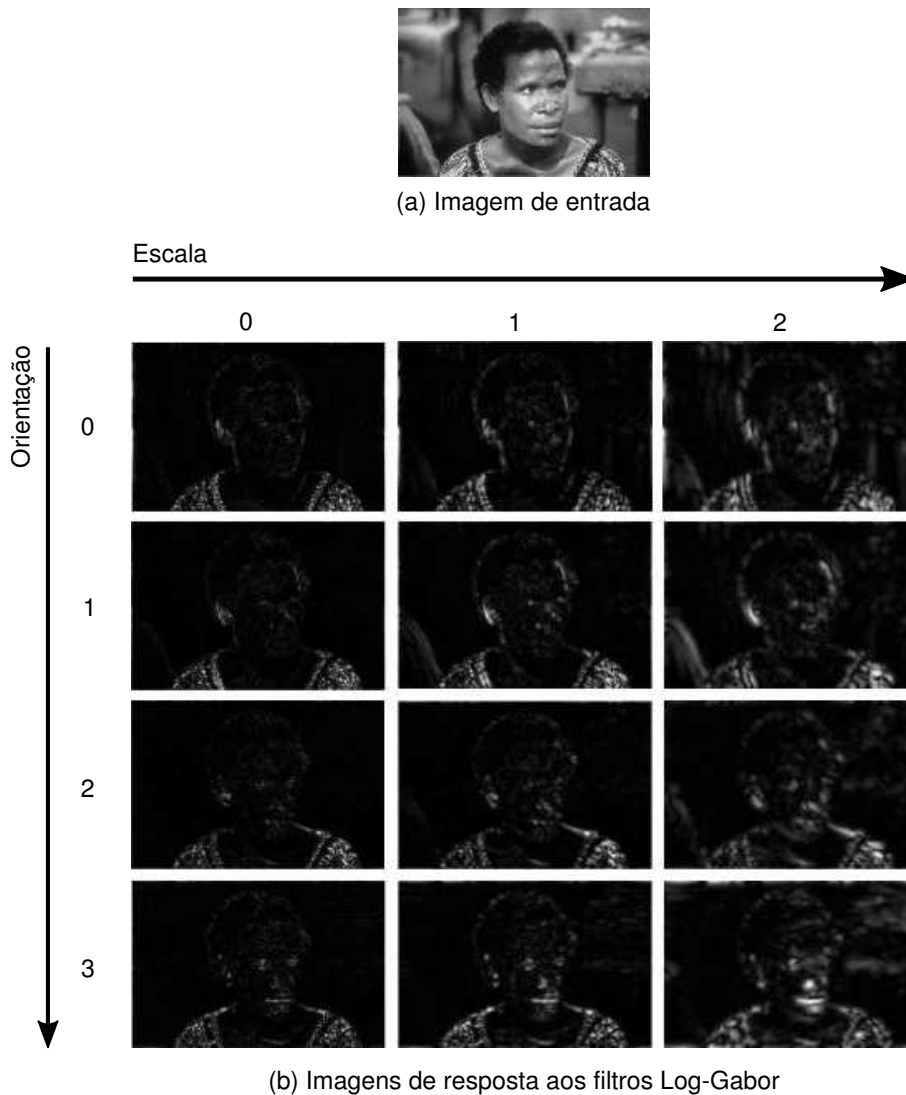
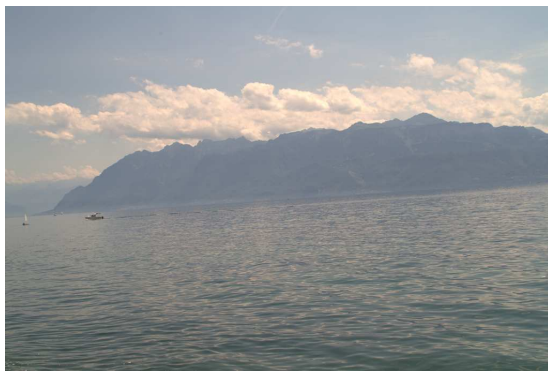


Figura 9 – Exemplo de imagens de resposta aos filtros Log-Gabor.

As imagens de resposta aos filtros Log-Gabor podem, em alguns casos, ter resultados diferentes entre imagens obtidas de espectro diferentes, pois tanto no espectro infravermelho quanto no espectro de luz visível podem ocorrer casos em que as bordas das imagens são confundidas, devido às diferenças na reflexão da luz em diferentes faixas do espectro eletromagnético. Esse fenômeno resulta em uma limitação durante a extração de informações de bordas entre imagens de espectro diferentes. A [Figura 10](#) ilustra essa diferença de bordas extraídas em uma imagem do espectro visível e uma imagem do espectro infravermelho de uma mesma cena. O método descritor proposto neste trabalho não trata essa deficiência ao extrair informações de bordas, sendo esta uma das limitações do método deste trabalho.



(a) Imagem A (Luz visível)



(b) Imagem B (Infravermelho)



(c) Imagem resposta ao filtro Log-Gabor para a Imagem A



(d) Imagem resposta ao filtro Log-Gabor para a Imagem B

Figura 10 – Diferenças na extração de bordas entre uma imagem do espectro visível e uma imagem do espectro infravermelho de uma mesma cena. Em (a) é exibida uma imagem do espectro de luz visível, e em (b) é exibida uma imagem do espectro infravermelho. As imagens (c) e (d) representam as bordas das imagens (a) e (b) respectivamente por meio da utilização de filtros Log-Gabor. É possível observar que as bordas apresentadas em (c) são visivelmente diferentes das bordas apresentadas em (d).

3.3 Detecção e descrição de pontos-chave

São descritos nesta seção os funcionamentos dos principais algoritmos detectores e descritores de pontos-chave utilizados como base para o estudo desenvolvido neste trabalho.

3.3.1 Algoritmo SIFT

O algoritmo SIFT é composto por um algoritmo detector e um algoritmo descritor. Este algoritmo possui quatro etapas principais: (i) detecção de extremos, (ii) localização de pontos-chave, (iii) atribuição da orientação e (iv) descrição de pontos-chave. As duas primeiras etapas correspondem ao seu algoritmo detector e as duas etapas seguintes correspondem ao seu algoritmo descritor.

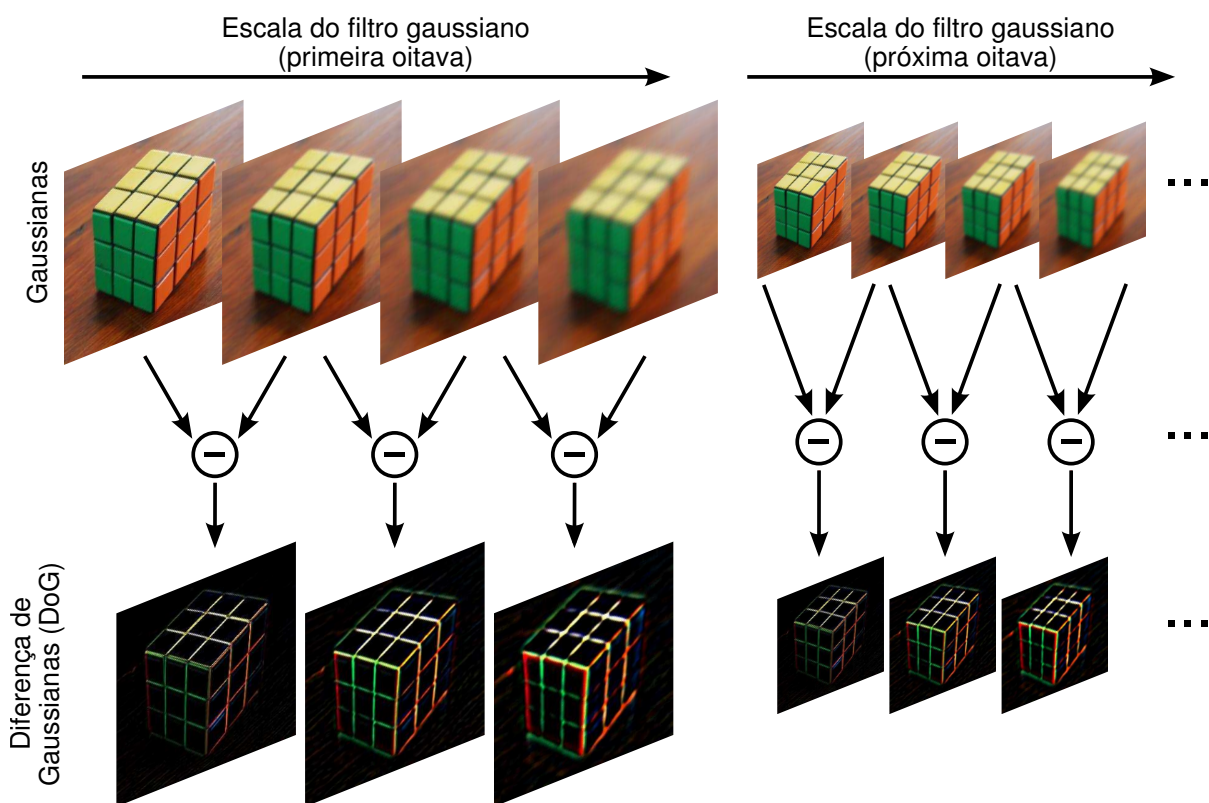


Figura 11 – Pirâmide de escala montada pelo algoritmo SIFT. Para cada oitava, é efetuada a convolução entre a imagem inicial e o filtro Gaussiano repetidamente para produzir um conjunto de imagens Gaussianas, suavizadas em vários níveis. Cada nível de suavização corresponde a uma camada, dentro da oitava. As imagens Gaussianas adjacentes são subtraídas para produzir as imagens *Diferença-de-Gaussianas*. Após cada oitava, o tamanho das imagens Gaussianas são reduzidas pela metade, e então o processo é repetido.

A etapa de detecção de extremos tem o objetivo de encontrar os máximos e mínimos locais por meio de uma filtragem em cascata dada pela função DoG (*Diferença-de-Gaussianas*). Primeiro, uma pirâmide de escala é construída de acordo com a [Figura 11](#). Em várias escalas, é calculada a convolução bidimensional entre a imagem original $I(x,y)$ com o operador DoG, que permite detectar variações de intensidades, como bordas. A função DoG é expressa pela [Equação \(5\)](#):

$$D(x,y,\sigma) = (G(x,y,k\sigma) - G(x,y,\sigma)) * I(x,y), \quad (5)$$

em que σ representa a escala atual do filtro gaussiano, k consiste em uma constante real multiplicativa e $G(x,y,\sigma)$ é a função Gaussiana dada pela [Equação \(6\)](#):

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2/2\sigma^2)}. \quad (6)$$

Após a geração da pirâmide, utilizando-se as imagens resultantes das diferenças de Gaussiana, cada ponto é comparado com seus vizinhos, tanto da imagem atual quanto das imagens adjacentes. Então, faz-se a seleção dos pontos candidatos a serem pontos-chave, como mostrado na [Figura 12](#). Um ponto é selecionado se seu valor de intensidade for maior ou menor do que todos os seus 26 vizinhos.

Na etapa de localização de pontos-chave, todos os pontos candidatos detectados na etapa anterior são avaliados. São rejeitados os pontos extremos com baixo contraste e os pontos os quais são mal localizados ao longo de uma borda. Todos os pontos restantes ao final desta etapa são definidos como pontos-chave.

A etapa de atribuição da orientação é baseada em propriedades locais de cada ponto-chave, o que possibilita alcançar invariância à rotação da imagem. A escala na qual o ponto-

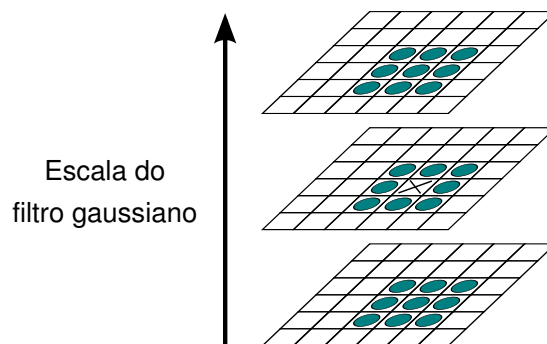


Figura 12 – Detecção de máximos e mínimos locais do algoritmo SIFT. Os máximos e mínimos, das imagens *Diferença-de-Gaussianas*, são detectados por meio da comparação de um pixel (marcado com X) com seus 26 vizinhos (marcados com círculos) em uma região 3x3 da imagem atual e imagens adjacentes.

chave foi detectado é usada para seleccionar a imagem Gaussiana $L(x,y) = G(x,y,\sigma) * I(x,y)$ de escala mais próxima, garantindo que os cálculos sejam realizados em escala invariante. Para os pontos da região que contém o ponto-chave, são calculados a magnitude do gradiente $m(x,y)$, dada pela Equação (7), e a orientação $\theta(x,y)$, dada pela Equação (8). O tamanho desta região é determinada pela escala do ponto-chave.

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2}, \quad (7)$$

$$\theta(x,y) = \tan^{-1} \left(\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)} \right). \quad (8)$$

É construído então um histograma de orientações para os pontos desta região. São definidos 36 valores de $\theta(x,y)$ que abrangem 360° de orientações. Cada amostra na vizinhança do ponto-chave, adicionada ao histograma, é ponderada com base na magnitude $m(x,y)$ e em uma janela gaussiana circular em torno do ponto-chave. Picos na orientação do histograma correspondem a direcções dominantes. É encontrado o maior pico no histograma, e picos que correspondem a valores acima de 80% do maior pico, são usados para definir a orientação do ponto-chave.

Na etapa de descrição dos pontos-chave, são calculados os descritores que representam suas respectivas regiões. Para tanto, é seleccionada a imagem $L(x,y)$ com a oitava referente ao ponto-chave. Em seguida, as orientações dos gradientes são rotacionadas em relação à orientação do ponto-chave. Após isto, é montada uma matriz de 16×16 pixels divididos em 4 regiões em torno do ponto-chave. Em cada região, é calculado um histograma de 8 direcções com base na magnitude dos pixels. É aplicada uma ponderação gaussiana para dar menos ênfase aos gradientes mais afastados do ponto-chave. O descritor então é construído a partir do vetor que armazena este histograma. Este processo é ilustrado na Figura 13. Ao final, o vetor é normalizado para criar um descritor mais robusto à mudanças de iluminação.

3.3.2 Algoritmo SURF

O algoritmo SURF (*Speeded Up Robust Features*) pode ser visto como sendo uma aproximação do algoritmo SIFT. Na etapa de localização de pontos-chave, são criadas escalas de filtros em vez de se reduzir o tamanho da imagem iterativamente, como ocorre na montagem da pirâmide de escala do algoritmo SIFT. Para tanto, é determinado o cálculo aproximado da matriz Hessiana para cada escala diferente.

A matriz Hessiana, dada pela Equação (9), é formada pelas derivadas parciais nas direcções vertical, horizontal e diagonal de uma função. O algoritmo SURF utiliza o determinante da matriz Hessiana para detectar se há variação em uma determinada

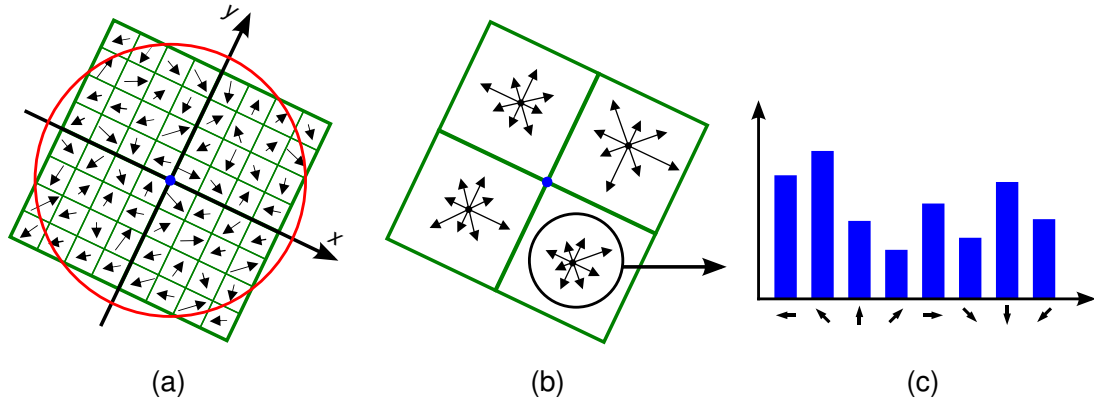


Figura 13 – Construção do descritor SIFT. O descritor SIFT utiliza uma matriz de amostragem de 16x16 pixels divididos em 4 regiões de histogramas com 8 orientações, resultando em um vetor de dimensão 128. Para exemplificar, nesta figura é exibida uma matriz de amostragem de 8x8 pixels divididos em 4 regiões. Em (a) é exibida a matriz de amostragem rotacionada em relação ao ponto-chave. Também é ilustrada a janela circular de ponderação Gaussiana. Em (b) são calculados os histogramas de orientações de cada região. Em (c) é exibido o histograma de 8 direções de uma região.

Fonte: Adaptado de [Lowe \(2004\)](#).

região da imagem. Se o valor do determinante for baixo, dado um limiar, a variação numa determinada região é baixa, caso contrário, a variação é alta.

$$\mathcal{H}(X, \sigma) = \begin{bmatrix} L_{xx}(X, \sigma) & L_{xy}(X, \sigma) \\ L_{xy}(X, \sigma) & L_{yy}(X, \sigma) \end{bmatrix}, \quad (9)$$

em que X representa um ponto da imagem, σ representa a escala, $L_{xx}(X, \sigma)$ é a convolução da derivada parcial da Gaussiana $\frac{\partial^2}{\partial x^2} g(\sigma)$ com a imagem $I(x, y)$ no eixo x , e de forma similar para $L_{yy}(X, \sigma)$, no eixo y , e $L_{xy}(X, \sigma)$, no eixo xy .

Para a realização da aproximação da matriz Hessiana, o algoritmo utiliza filtros de caixa correspondentes a núcleos gaussianos de tamanho 9x9, como ilustrados na [Figura 14](#).

Após a geração da pirâmide de filtros, o algoritmo busca pontos invariantes a escala da mesma forma como ocorre no algoritmo SIFT. Cada ponto é comparado com seus vizinhos, tanto da imagem atual quanto das imagens adjacentes. Então, faz-se a seleção dos pontos candidatos a serem pontos-chave, de forma semelhante àquela exibida na [Figura 12](#).

Na etapa de descrição, o algoritmo utiliza *wavelets* Haar, o que possibilita a determinação do valor do gradiente nas direções x e y . A fim de se obter a orientação dominante, as respostas *wavelet* Haar são computadas para todas as regiões onde se encontra o ponto-chave detectado.

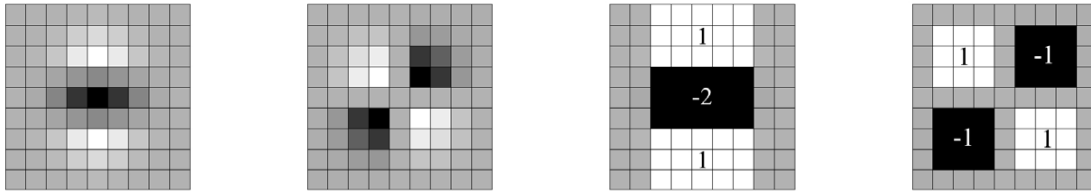


Figura 14 – Aproximações de núcleo Gaussiano do algoritmo SURF. Da esquerda para a direita: Derivadas parciais de segunda ordem na direção x e xy e suas aproximações, utilizando núcleos gaussianos. As regiões em cinza correspondem ao valor zero.

Fonte: Bay, Tuytelaars e Gool (2006).

Uma vez que as respostas *wavelet* são determinadas, a orientação dominante é calculada pela soma de todas respostas horizontal e vertical dentro de uma janela deslizante de orientação. A orientação com o vetor mais longo é selecionada como orientação dominante do descritor. A Figura 15 ilustra o cálculo do descritor SURF. A partir das respostas *wavelet*, o descritor é calculado de acordo com a Equação (10):

$$Descriptor = \left(\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y| \right), \quad (10)$$

em que d_x e d_y são as respostas das *wavelets* Haar ao longo das direções x e y respectivamente.

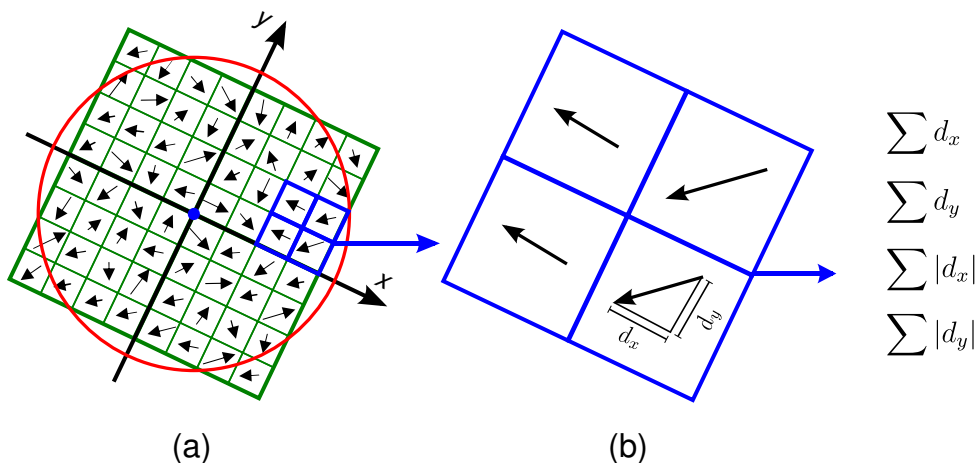


Figura 15 – Construção do descritor SURF. O descritor SURF pode utilizar uma matriz de amostragem de 8x8 pixels ou 16x16 pixels. Em (a) é exibida a matriz de amostragem. Para uma matriz de 8x8, o descritor resulta em um vetor de tamanho 64, já uma matriz de amostragem de 16x16 pixels, resulta em um descritor tamanho 128, formado pelo somatório obtido das respostas *wavelets*, exibidos em (b).

3.3.3 Algoritmo FAST

O algoritmo FAST (*Features from Accelerated Segment Test*), proposto por [Rosten e Drummond \(2006\)](#), consiste em um método de detecção de cantos, que possui um desempenho superior aos métodos que se baseiam na *Diferença-de-Gaussianas*. Para definir se um ponto é realmente um canto, é utilizada uma circunferência de 16 pixels. Se 12 pixels contíguos sobre a circunferência possuem um brilho superior ao brilho do ponto em teste somado a um valor limiar, ou se estes pixels são mais escuros, então o ponto em teste é definido como canto. A [Figura 16](#) ilustra a detecção de canto do algoritmo FAST.

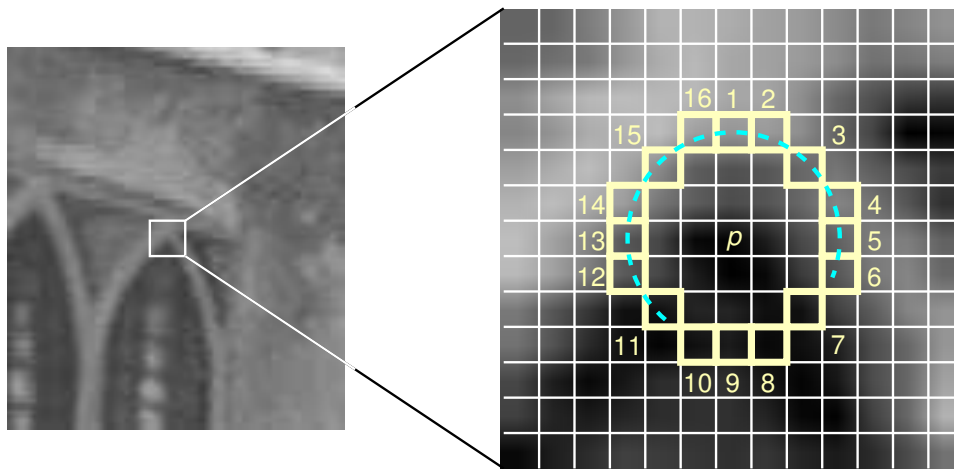


Figura 16 – Detector de cantos FAST. Os quadrados em destaque são pixels usados na detecção de canto. O pixel p , no centro, é um candidato a canto. A circunferência pontilhada passa sobre 12 pixels contíguos os quais são mais claros que o pixel p .

Fonte: [Rosten e Drummond \(2006\)](#).

3.3.4 Algoritmo ORB

Para explicar o algoritmo ORB (*Oriented FAST and Rotated BRIEF*), é necessário fazer uma breve descrição do algoritmo BRIEF (*Binary Robust Independent Elementary Features*), no qual o algoritmo ORB é baseado.

O algoritmo BRIEF, proposto por [Calonder et al. \(2010\)](#), consiste em um descritor binário eficiente que se baseia em testes de diferença de intensidade. Este algoritmo visa ser eficiente em memória e rápido para calcular e realizar a correspondência. Diferentemente dos algoritmos SIFT e SURF, o algoritmo BRIEF não é invariante à rotação da imagem.

A diferença essencial entre o descritor BRIEF e os descritores mencionados anteriormente é que o descritor BRIEF descreve as características em uma cadeia binária em vez de vetores que utilizam uma representação numérica em ponto flutuante. Para construir o descritor BRIEF, é utilizado um conjunto de comparações predefinidas de pares de pontos.

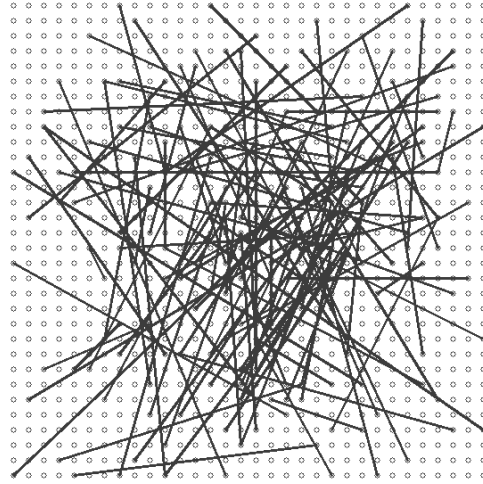


Figura 17 – Padrão de 128 pares de pontos de amostragem do algoritmo BRIEF.

Fonte: Calonder et al. (2010).

A Figura 17 exibe o padrão com a localização dos pares de pontos de amostragem utilizados para as comparações neste algoritmo. Antes de efetuar as comparações, a imagem original é suavizada por meio de um filtro gaussiano, para assim, reduzir possíveis ruídos contidos na imagem.

Considerando-se uma imagem suavizada I e um par de pontos $(a, b) \in I$, o teste binário τ é definido pela Equação (11):

$$\tau(a, b) = \begin{cases} 1, & \text{se } I(a) < I(b) \\ 0, & \text{caso contrário} \end{cases}, \quad (11)$$

em que $I(a)$ consiste na intensidade (brilho) do ponto a e $I(b)$ consiste na intensidade do ponto b .

Um descritor BRIEF é então definido como uma sequência de bits dado pela Equação (12):

$$Descriptor = \sum_{1 \leq i \leq n} 2^{i-1} \tau(a_i, b_i), \quad (12)$$

em que n consiste na quantidade de pares de pontos.

Como dito anteriormente, o algoritmo ORB é baseado na combinação das técnicas FAST e BRIEF. Isto justifica seu acrônimo (*Oriented FAST and Rotated BRIEF*) com o objetivo de cobrir a deficiência da variância de rotação e escala, e ao mesmo tempo manter um bom desempenho. Este algoritmo pode ser entendido como uma melhoria do componente de orientação do algoritmo FAST e a implementação da característica de invariância à rotação ao algoritmo BRIEF. Diferentemente do algoritmo BRIEF, o algoritmo

ORB utiliza um padrão de amostragem que contém 512 pares de pontos, resultando em um descritor de tamanho 512.

Como o algoritmo FAST não detecta características em multiescala, o algoritmo ORB utiliza uma pirâmide de escalas de imagens e aplica o detector FAST em cada nível da pirâmide. Em seguida, os pontos detectados são filtrados com a medida de canto Harris (HARRIS; STEPHENS, 1988) com a finalidade de reduzir a quantidade de pontos ao longo de bordas.

A técnica FAST, aprimorada no algoritmo ORB, foi chamada de oFAST ou *FAST Keypoint Orientation*. Como o detector FAST, originalmente, não inclui um operador de orientação, a modificação proposta adiciona orientação aos pontos-chave obtidos pelo detector FAST por meio do cálculo do centroide de intensidade, que tem como pressuposto a ideia de que a partir do primeiro canto encontrado, haverá uma distância do centro que contém o canto até o seu centroide calculado. Esta distância permite encontrar o ângulo de orientação deste caminho. Com os momentos³ dado pela Equação (13), é calculado o centroide de intensidade, por meio da Equação (14):

$$m_{pq} = \sum x^p y^q I(x, y), \quad p, q = \{0, 1\}, \quad (13)$$

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right). \quad (14)$$

Construindo um vetor do centro que contém o canto para o centroide C , a orientação θ do caminho é simplesmente dada pela Equação (15):

$$\theta = \text{atan2}(m_{01}, m_{10}). \quad (15)$$

No algoritmo ORB, o descritor de características BRIEF foi modificado para possibilitar o tratamento de imagens mesmo que estas sejam apresentadas com diferentes ângulos de rotação. Esta nova modificação foi denominada de rBRIEF ou *Rotation-Aware BRIEF*. Esta modificação considera a orientação θ do ponto-chave, detectado na etapa anterior, que é utilizada para adicionar orientação ao descritor.

O descritor BRIEF é aplicado por meio da utilização do ângulo de orientação θ e de uma matriz de rotação para rotacionar o padrão de amostragem. Dessa forma, o descritor ORB é robusto por manter-se invariante à rotação e escala.

³Momento de imagem, em Visão Computacional e áreas afins, consiste em uma média ponderada específica da intensidade dos pixels da imagem.

3.3.5 Algoritmo BRISK

O algoritmo BRISK (*Binary Robust Invariant Scalable Keypoints*), proposto por [Leutenegger, Chli e Siegwart \(2011\)](#), é um algoritmo de detecção e descrição de pontos-chave que provê invariância tanto na escala quanto na rotação de imagens. Seu detector de pontos-chave é constituído, essencialmente, pelo método FAST.

Na etapa de detecção, para manter a invariância à escala, é montada uma pirâmide de escalas da imagem. A pirâmide de escalas do algoritmo BRISK consiste em n oitavas c_i e n intraoitavas d_i , para $i = \{0, 1, \dots, n - 1\}$, em que, por padrão, os autores definem $n = 4$. As oitavas da pirâmide são formadas por meio do redimensionamento consecutivo da imagem original c_0 pela metade, realizando assim uma subamostragem. Cada intraoitava d_i é localizada entre as camadas c_i e c_{i+1} , como ilustrado na [Figura 18](#).

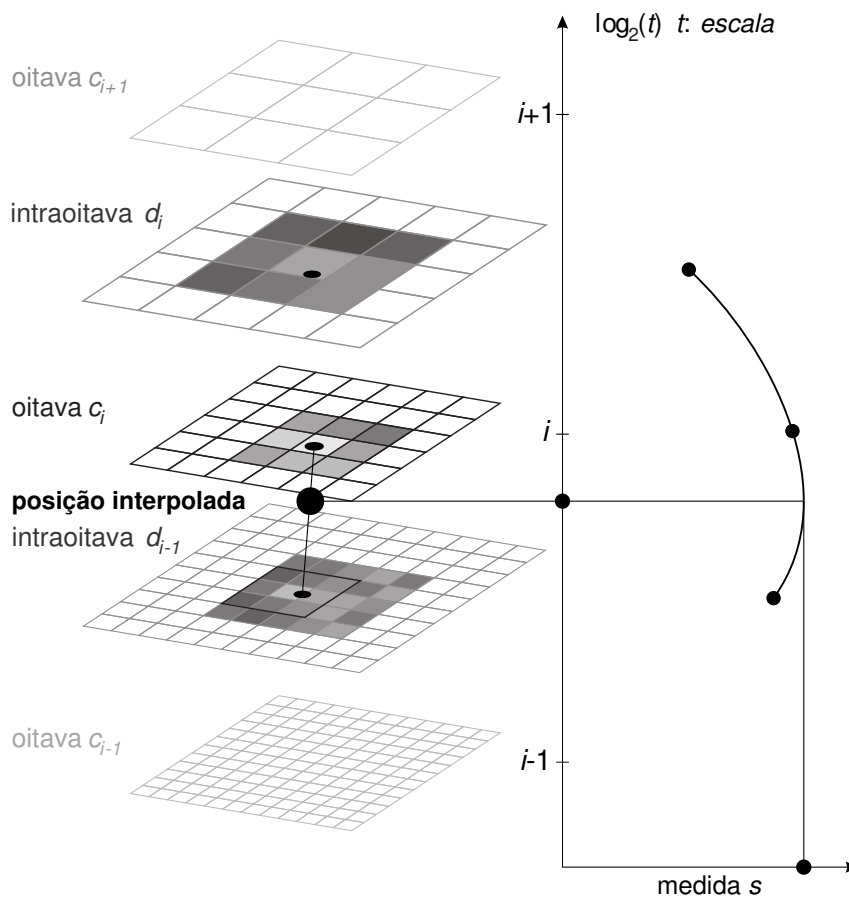


Figura 18 – Pirâmide de escala montada pelo algoritmo BRISK. Um ponto-chave é identificado na oitava c_i por meio da análise dos 8 pontos vizinhos das oitavas inferior e superior. Por meio do valor máximo da medida s , e utilizando as três camadas, é possível determinar a escala real do ponto-chave detectado na posição interpolada.

Após a construção da pirâmide, é aplicado o detector AGAST em cada oitava e intraoitava da pirâmide para identificar potenciais regiões de interesse. Em seguida, os pontos contidos nestas regiões são submetidos a uma supressão de não-máximos. Para a realização da supressão de não-máximos os autores definem a medida s . Esta medida é definida como sendo o limiar máximo em que um ponto é considerado canto por meio do algoritmo FAST. Assim, a supressão de não-máximos consiste em dois critérios: (i) o ponto em questão deve atingir o valor máximo na medida s com respeito aos seus pontos vizinhos da oitava atual; (ii) os valores das medidas s , das oitavas superior e inferior, devem ser menores que o valor da medida s na oitava atual. A Figura 18 exibe também a curva da medida s para as oitavas analisadas. Esta curva também é útil para definir a escala real do ponto-chave detectado. Após a supressão de não-máximos, os pontos detectados são definidos como pontos-chave.

O descritor BRISK, assim como o descritor BRIEF, é composto por uma cadeia binária concatenada, resultante de simples comparações de intensidade. O descritor BRISK faz uso de um padrão utilizado para a amostragem da vizinhança do ponto-chave detectado, ilustrado na Figura 19. Este, consiste em um padrão simétrico que possui pontos de amostragem posicionados em círculos concêntricos em torno do ponto-chave. Cada círculo concêntrico representa uma suavização gaussiana com desvio padrão proporcional ao tamanho do círculo, para redução de ruídos.

No padrão utilizado, os autores definem dois tipos de pares de pontos: pares longos e pares curtos. Pares curtos são pares de pontos de amostragem no qual sua distância euclidiana é menor que um determinado limiar δ_{max} , e pares longos são aqueles no qual sua distância é maior que um certo limiar δ_{min} . Os pares longos são utilizados para determinar a orientação, e os pares curtos são utilizados na comparação de intensidade.

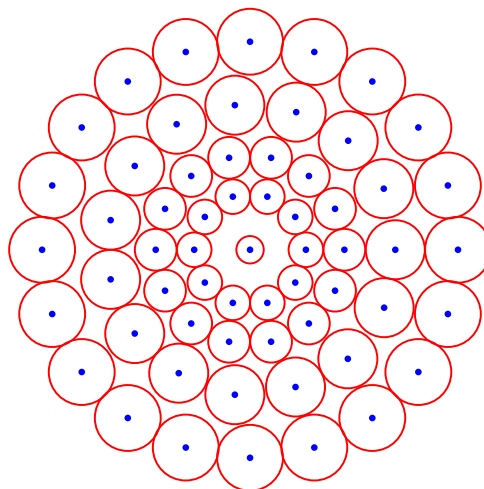


Figura 19 – Padrão de amostragem do algoritmo BRISK.

O algoritmo BRISK estima a orientação do ponto-chave e a rotação do padrão de amostragem por meio da soma de todos os gradientes locais entre todos os pares longos, o que o torna invariante a rotação, retornando o ângulo do ponto-chave. Por meio do conjunto de comparações de intensidade entre todos os pares curtos, o descritor BRISK resulta em uma cadeia binária de tamanho 512.

3.3.6 Algoritmo EOH

O algoritmo descritor EOH (*Edge Oriented Histogram*), proposto por [Aguilera et al. \(2012a\)](#), é baseado no método descritor EHD (*Edge Histogram Descriptor*), definido no padrão MPEG-7 ([SIKORA, 2001](#)). Para realizar o cálculo do descritor, a imagem de entrada é convertida em uma imagem de bordas, por meio da utilização do detector Canny ([CANNY, 1986](#)). Uma vez que a imagem é representada por meio de suas bordas, os descritores são calculados para cada ponto-chave como ilustrado na [Figura 20](#).

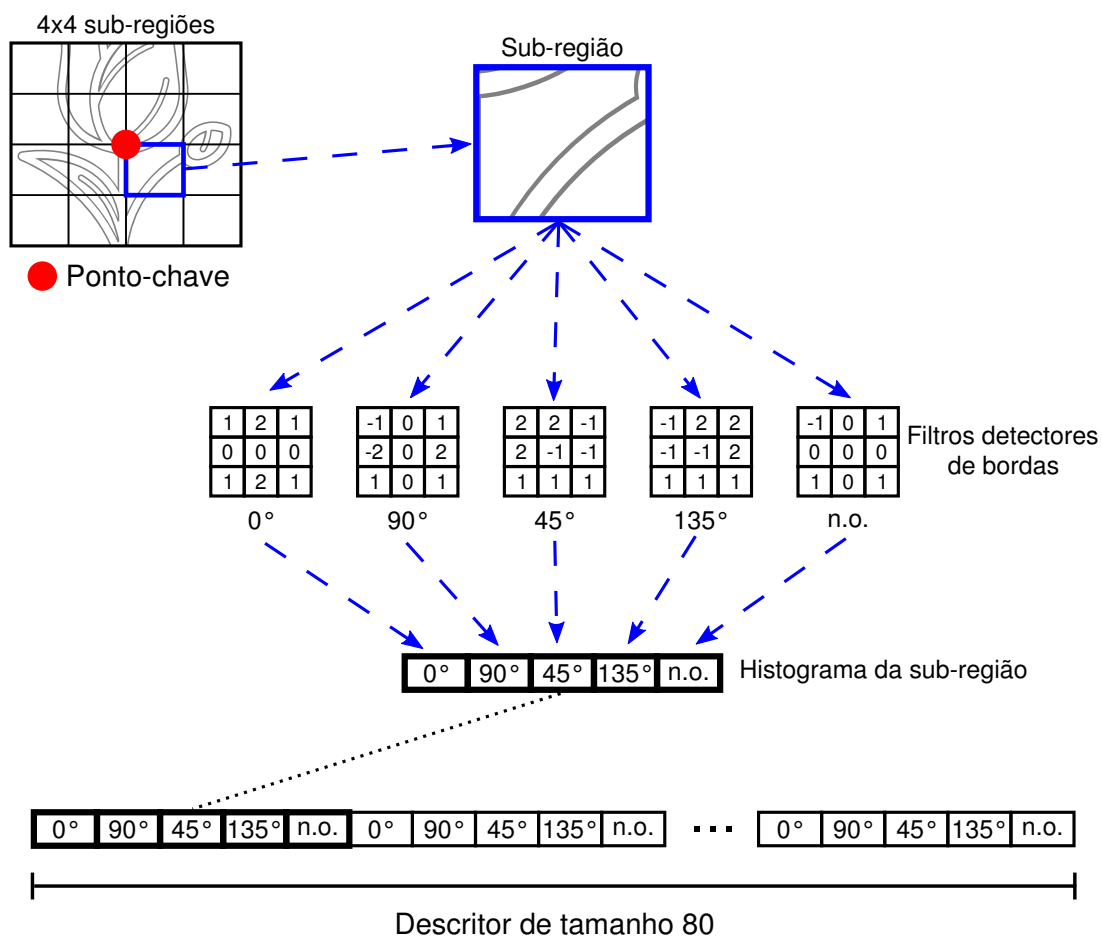


Figura 20 – Construção do descritor EOH.

Fonte: Adaptado de [Aguilera et al. \(2012a\)](#).

Primeiro, uma janela de tamanho $N \times N^4$ é posicionada no centro do ponto-chave detectado. Essa janela é subdividida em 4×4 sub-regiões (total de 16 sub-regiões), e para cada sub-região é calculado o histograma de orientações. O histograma representa a distribuição espacial de quatro filtros de bordas (filtros Sobel) direcionais e um filtro não-direcional (cinco filtros no total). Cada posição do histograma representa os ângulos de 0, 90, 45, 135 graus e, adicionalmente, uma posição do histograma representa nenhuma orientação (n.o.) para corresponder às áreas que não contém bordas.

Todo pixel de cada sub-região contribui para uma posição no histograma de acordo com os cinco filtros: o filtro com o maior valor de resposta é utilizado como critério de votação para uma determinada posição do histograma. Após o processamento de todos os elementos das 16 sub-regiões obtém-se um vetor de tamanho 80 (referente aos 5 filtros para cada uma das 16 sub-regiões). Este vetor final então é normalizado, sendo este o descritor para o ponto-chave detectado.

3.3.7 Algoritmo LGHD

O algoritmo descritor LGHD (*Log-Gabor Histogram Descriptor*), proposto por [Aguilera, Sappa e Toledo \(2015\)](#), foi desenvolvido para realizar a descrição em imagens não lineares, tais como: imagens de mapas de profundidade, imagens multiespectrais e imagens obtidas por meio de câmeras com e sem flash. O descritor LGHD realiza a descrição dos pontos-chave detectados por meio da utilização de filtros Log-Gabor.

Primeiro, é realizada a convolução da imagem de entrada com 24 filtros Log-Gabor, referente à seis filtros de orientações (variando igualmente no intervalo entre $m = 0$ e $m = \pi$) por quatro filtros de escala ($n = 0, 1, 2, 3$). Esse processo resulta em 24 imagens de resposta aos filtros Log-Gabor.

Em seguida, uma janela de tamanho $N \times N^5$ é posicionada no centro do ponto-chave detectado para as 24 imagens. Essa janela é subdividida em 4×4 sub-regiões (total de 16 sub-regiões). Para cada sub-região é calculado um histograma de acordo com a orientação e escala do filtro.

Todo pixel de cada sub-região contribui para uma posição no histograma de acordo com os filtros definidos: o filtro com o maior valor de magnitude determina o filtro dominante. Após o processamento de todos os elementos das 16 sub-regiões obtém-se um vetor descritor de tamanho 384 (referente aos 24 filtros para cada uma das 16 sub-regiões).

O [Quadro 2](#) exibe as principais características dos algoritmos apresentados nesta seção.

⁴No trabalho de [Aguilera et al. \(2012a\)](#) é definido o tamanho da janela como 80.

⁵Assim como no trabalho de [Aguilera et al. \(2012a\)](#), é definido o tamanho da janela como 80.

Algoritmo	Detector	Descritor	Tamanho do descritor	Tipo de descritor
SIFT	Sim	Sim	128	Numérico (ponto flutuante)
SURF	Sim	Sim	128	Numérico (ponto flutuante)
FAST	Sim	Não	(Não se aplica)	(Não se aplica)
ORB	Sim	Sim	64 (512 bits)	Binário
BRISK	Sim	Sim	64 (512 bits)	Binário
EOH	Não	Sim	80	Numérico (ponto flutuante)
LGHD	Não	Sim	384	Numérico (ponto flutuante)

Quadro 2 – Principais características dos algoritmos apresentados.

3.4 Correspondência de descritores

Para determinar o quão similar é um descritor de outro, e assim efetuar a correspondência, é necessário definir um critério de similaridade de acordo com a natureza do descritor. Para descritores que utilizam uma representação numérica em ponto flutuante, como os descritores SIFT, SURF, EOH e LGHD, é usualmente utilizado a norma L_2 (JIANG; XU; LIU, 2013; CALONDER et al., 2010). Para descritores representados por meio de uma cadeia binária, como os descritores ORB e BRISK, é utilizado como critério de similaridade à distância de Hamming, devido sua simplicidade e rapidez (HEINLY; DUNN; FRAHM, 2012).

Sejam os descritores p e q de dimensão n , a norma L_2 e a distância de Hamming são dadas pela Equação (16) e Equação (17) respectivamente:

$$L_2(p, q) = \sqrt{\sum_{i=1}^n (q_i - p_i)^2}, \quad (16)$$

$$H(p, q) = \sum_{i=1}^n |q_i - p_i|. \quad (17)$$

Comparado à correspondência de descritores que utilizam uma representação numérica em ponto flutuante, a correspondência de vetores binários é mais simples, pois a correspondência mais próxima consiste em pares de descritores com a menor distância de Hamming. O número de bits diferentes entre dois descritores mede sua dissemelhança. O uso da distância de Hamming para a correspondência em descritores binários torna este processo eficiente (CALONDER et al., 2010).

Segundo Lowe (2004), para descritores que possuem uma dimensão elevada, assim como o descritor SIFT que possui um vetor de dimensão 128, é difícil a construção de algoritmos, para a correspondência de descritores, que possam identificar o ponto vizinho mais próximo em espaços tridimensionais elevados e que sejam mais eficientes computacionalmente do que a busca exaustiva. Mesmo algoritmos, tais como o *Kd-Tree* (BENTLEY,

1975), não fornecem nenhum ganho de velocidade comparado à busca exaustiva para espaços dimensionais de tamanho maior que 10.

Assim como utilizado na literatura (JIANG; XU; LIU, 2013; YU et al., 2012; LEUTENEGGER; CHLI; SIEGWART, 2011), na etapa de correspondência de descritores deste trabalho, é utilizado o método de busca exaustiva. No método de busca exaustiva, para cada descritor da primeira imagem, busca-se pelo descritor na segunda imagem que possui a menor distância. Este par de descritores então é retornado, representando assim a correspondência de um par de pontos-chave.

3.5 Métricas de avaliação

No processo de avaliação dos algoritmos, para determinar se uma correspondência é correta (correspondência verdadeira), é necessário definir o conceito de erro de sobreposição.

O erro de sobreposição mede o quão duas regiões são de fato correspondentes, dada a matriz de homografia. Neste caso, os pontos da imagem em questão são projetados sobre a outra imagem por meio da matriz de homografia H . O erro de sobreposição é definido como sendo o valor 1 subtraído da razão entre a interseção e a união das regiões que contém o ponto-chave, de acordo com a Equação (18) (MIKOLAJCZYK; SCHMID, 2005):

$$\epsilon = 1 - (A \cap H^T B H) / (A \cup H^T B H), \quad (18)$$

em que A e B são as regiões e H é a matriz de homografia entre o par de imagens.

Assim como no trabalho de Mikolajczyk et al. (2005), assume-se que uma correspondência é verdadeira se o erro de sobreposição for menor que 40%, isto é, $\epsilon < 0,4$. A seguir, serão descritas três das principais métricas utilizadas para a avaliação do processo de correspondência de pontos em imagens, quais sejam: repetibilidade, precisão e revocação. A Figura 21 exibe exemplos de erro de sobreposição entre duas regiões.

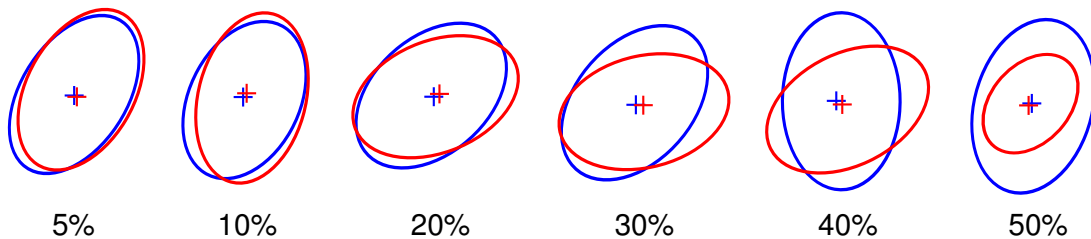


Figura 21 – Exemplos de erro de sobreposição entre duas regiões.

3.5.1 Repetibilidade

A repetibilidade representa a capacidade de detecção dos mesmos pontos na cena sob diferentes ponto de vista, mudanças de iluminação ou ruídos. O cálculo da repetibilidade é dado pela razão entre o número de correspondências verdadeiras e o número mínimo de pontos-chave visíveis em ambas as imagens (LEUTENEGGER; CHLI; SIEGWART, 2011; MIKOLAJCZYK et al., 2005):

$$\text{Repetibilidade} = \frac{|\text{correspondências verdadeiras}|}{\min(|p_A|, |p_B|)}, \quad (19)$$

em que p_A e p_B correspondem aos pontos-chave detectados nas imagens A e B , respectivamente, e que são visíveis tanto em A quanto em B .

A repetibilidade diz respeito somente ao processo de detecção de pontos-chave. A alta taxa de repetibilidade entre duas imagens representa que mais pontos podem ser potencialmente correspondidos, impactando na melhora do resultado da correspondência (JIANG; XU; LIU, 2013).

3.5.2 Precisão

A taxa de precisão, representa a quantidade de correspondências verdadeiras retornadas em relação à quantidade total de correspondências retornadas (correspondências computadas ao final do processo de correspondência de pontos em imagens) e é dado por:

$$\text{Precisão} = \frac{|\text{correspondências verdadeiras retornadas}|}{|\text{correspondências retornadas}|}. \quad (20)$$

3.5.3 Revocação

A revocação representa a quantidade de correspondências verdadeiras retornadas, em relação à quantidade total de verdadeiras correspondências. A quantidade de verdadeiras correspondências é dada pela quantidade de pontos sobrepostos em ambas as imagens nos quais seus erros de sobreposição foram menores que 0,4 (como descrito no início desta subseção). A quantidade de verdadeiras correspondências retornadas são aquelas verdadeiras correspondências que o algoritmo retornou ao final do processo de correspondência. A revocação é representada pela Equação (21):

$$\text{Revocação} = \frac{|\text{correspondências verdadeiras retornadas}|}{|\text{correspondências verdadeiras}|}. \quad (21)$$

As métricas de precisão e revocação dizem respeito ao processo de descrição de pontos-chave. Seus valores podem variar de 0 a 1, e é desejado que se tenha, ao mesmo tempo, alta precisão e alta revocação. A Figura 22 exemplifica o uso destas métricas.

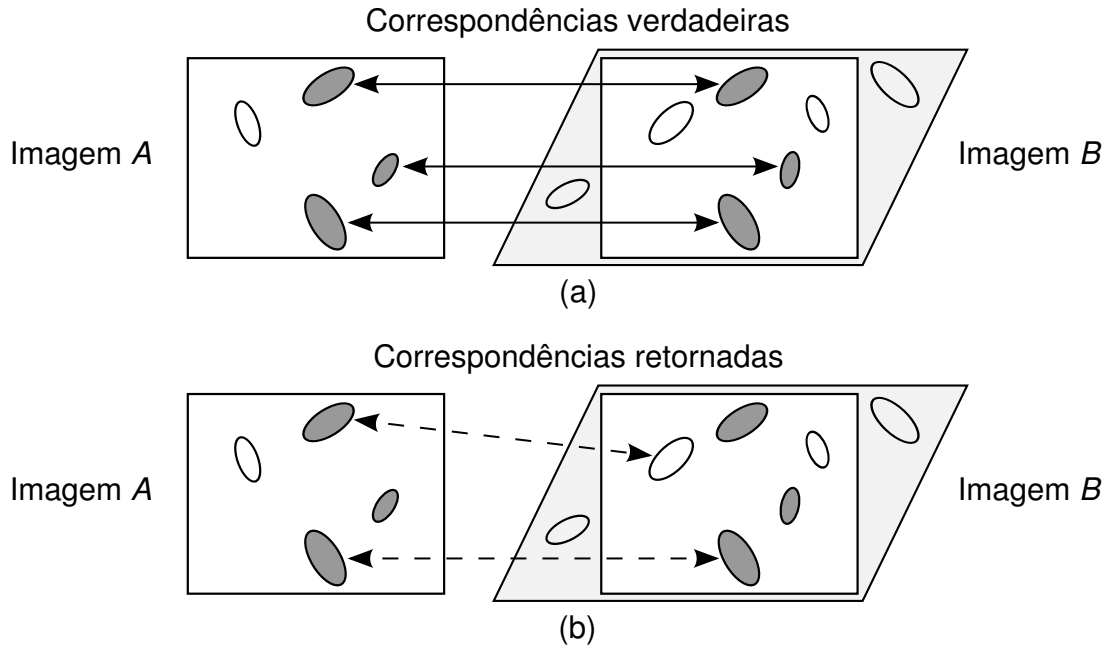


Figura 22 – Exemplo de aplicação das métricas de repetibilidade, precisão e revocação. Cada elipse na figura representa a região no qual se encontra um ponto-chave detectado. Os pontos-chave da Imagem A são projetados, por meio da matriz de homografia, na Imagem B. Em (a) são exibidas 3 correspondências verdadeiras, determinadas por meio do erro de sobreposição. Em (b) é exemplificado um caso em que, ao final da etapa de correspondência, são retornadas 2 correspondências, sendo apenas uma verdadeira. Neste exemplo, tem-se um valor de repetibilidade igual a $3/4$, precisão igual a $1/2$ e revocação igual a $1/3$.

3.5.4 Medida F1

A Medida F1 ou *F-measure* é uma métrica que representa um balanço entre a precisão e a revocação. Quando os valores de precisão e revocação se contradizem, a Medida F1 é usualmente empregada (WANG; ZHU; YUAN, 2013). Assim como a precisão e revocação, seus valores podem variar de 0 a 1, ou seja, quanto mais próximo de 1, maior é a precisão e a revocação. A Medida F1 é representada pela Equação (22):

$$Medida\ F1 = 2 \times \frac{Precisão \times Revocação}{Precisão + Revocação}. \quad (22)$$

4 Metodologia

Este capítulo descreve a metodologia empregada para a realização deste trabalho. A [Figura 23](#) apresenta uma visão geral de um sistema de correspondência de pontos em imagens multiespectrais. Como se pode observar, esse sistema é dividido em três etapas: detecção, descrição e correspondência.

Para a etapa de detecção, foram avaliados os algoritmos detectores SIFT, SURF, FAST, ORB e BRISK. Na etapa de descrição, por sua vez, foram avaliados os algoritmos descritores do estado da arte para imagens multiespectrais EOH, LGHD bem como o método aqui proposto denominado *Multispectral Feature Descriptor* (MFD).

Adicionalmente, foram avaliados os algoritmos descritores SIFT e SURF por serem considerados métodos de referência encontrados em grande parte dos trabalhos apresentados durante a revisão bibliográfica, desenvolvida no [Capítulo 2](#). Como justificado na [Seção 3.4](#) deste trabalho, para a etapa de correspondência de descritores, foi utilizado o método de busca exaustiva.

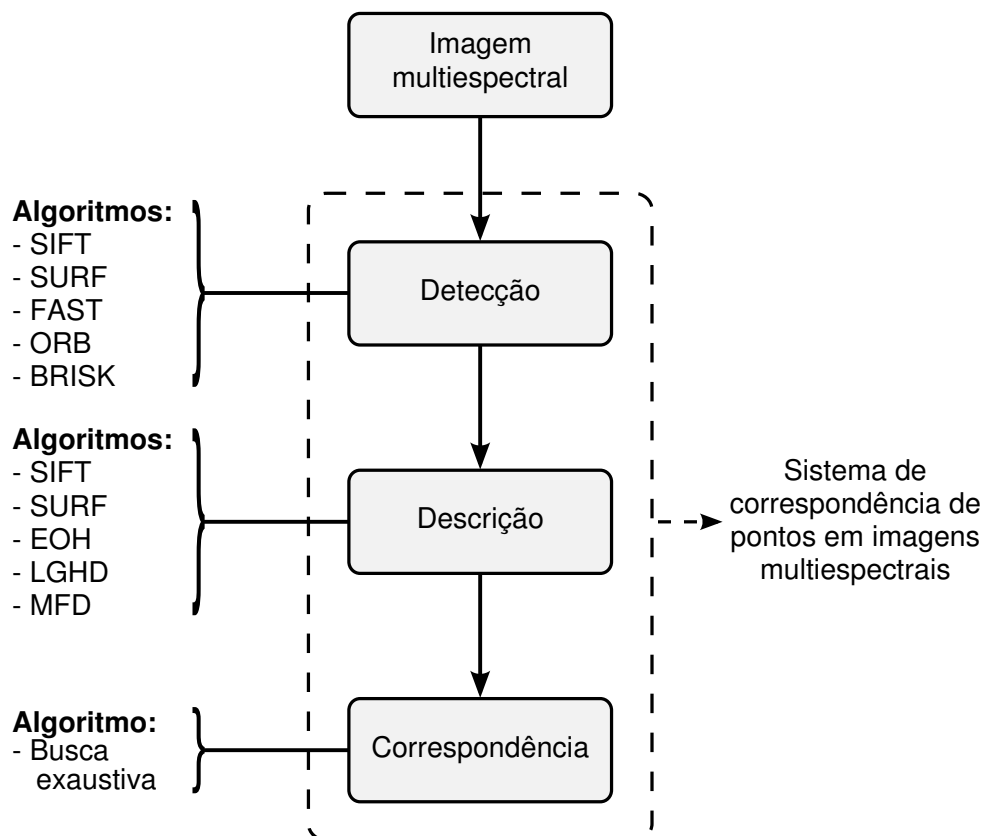


Figura 23 – Sistema de correspondência de pontos em imagens multiespectrais.

Para a implementação dos algoritmos utilizou-se a linguagem de programação C++, com o compilador GCC¹ na versão 5.4.0, e a biblioteca de Visão Computacional OpenCV² (*Open Source Computer Vision Library*) na versão 2.4.13. O OpenCV é uma biblioteca multiplataforma (Linux, Windows, OS X, iOS e Android), totalmente livre ao uso comercial e acadêmico, para o desenvolvimento de aplicações na área de Visão Computacional. Grande parte dos algoritmos apresentados neste trabalho já se encontram disponíveis na versão mais recente desta biblioteca. Estes algoritmos possuem interfaces uniformes, o que facilita a implementação para comparação. Os valores dos parâmetros, utilizados na execução dos algoritmos, seguem aqueles fornecidos pelos autores dos métodos utilizados. No [Apêndice A](#) são exibidos estes valores bem como a descrição de cada parâmetro. É importante ressaltar que os métodos SIFT, SURF, ORB e BRISK são compostos por dois algoritmos: um algoritmo detector e um algoritmo descritor (ver [Quadro 2](#)) que neste trabalho são avaliados de forma separada.

Os experimentos deste trabalho foram divididos em três passos principais. O primeiro passo consistiu em avaliar os algoritmos detectores de pontos-chave com a finalidade de observar o comportamento desses algoritmos em imagens multiespectrais e selecionar o método mais adequado (resultados na [Seção 5.1](#)) a ser utilizado conjuntamente com o descritor MFD. O segundo passo consistiu em realizar a análise dos parâmetros do método descritor MFD, proposto neste trabalho (resultados na [Seção 5.2](#)). O terceiro passo consistiu em avaliar os métodos descritores no processo completo de correspondência, que engloba os algoritmos detectores, descritores e o algoritmo de correspondência utilizado. Para tanto, considerou-se duas bases de dados, detalhadas no [Capítulo 5](#).

A seguir, é apresentado o descritor MFD proposto neste trabalho, bem como descrito os processos metodológicos adotados para se avaliar os algoritmos detectores e descritores aqui estudados.

4.1 *Multispectral Feature Descriptor* (MFD)

Diferentemente de alguns trabalhos recentemente propostos (como, por exemplo, os trabalhos de [Saleem e Sablatnig \(2014\)](#), [Dellinger et al. \(2015\)](#)), dentre outros detalhados no [Capítulo 2](#), o método descritor proposto não se baseia no método SIFT (descrito na [Subseção 3.3.1](#)) e também não utiliza o cálculo de gradientes para a descrição de pontos-chave, que em imagens de espectros diferentes não é desejável. A seguir, serão descritos os detalhes da construção do método descritor proposto neste trabalho, denominado *Multispectral Feature Descriptor* (MFD).

¹Disponível em: <http://gcc.gnu.org/>.

²Disponível em: <http://opencv.org/>.

Como mencionado no [Capítulo 1](#), um dos principais problemas encontrados durante as descrições de pontos em imagens multispectrais se baseia no fato de que as direções dos gradientes de intensidade associados a tais pontos podem variar significativamente entre as imagens obtidas em diferentes faixas do espectro eletromagnético. No entanto, mesmo com essas variações nos gradientes de intensidade, as aparências globais das formas dos objetos, as quais podem, por exemplo, ser representadas pelas suas bordas ou texturas, tendem a permanecer razoavelmente invariantes ([AGUILERA; SAPPA; TOLEDO, 2015](#)). As texturas consistem em padrões visuais em imagens e podem ser caracterizadas pelas propriedades estatísticas dos níveis de cinza dos pixels da imagem ([WALIA; VERMA, 2016](#)).

Os filtros de Gabor (ou *wavelets* de Gabor) são comumente utilizados para a extração de características da textura da imagem e detecção de bordas ([WALIA; VERMA, 2016](#)), e os filtros de Log-Gabor têm melhor capacidade de discriminação do que os filtros de Gabor, uma vez que são independentes do componente contínuo (como descrito na [Seção 3.2](#)) ([NAVA; ESCALANTE-RAMÍREZ; CRISTÓBAL, 2012](#)). De forma semelhante ao método descritor LGHD, no presente trabalho foram adotados os filtros de Log-Gabor para realizar a etapa de extração de textura e bordas da imagem.

O descritor aqui proposto é inspirado nos métodos descritores EOH e LGHD, mas diferentemente da abordagem utilizada no descritor EOH, o MFD não utiliza uma representação intermediária da imagem de entrada, como, por exemplo, um mapa de bordas obtido por meio do detector Canny. Em vez disso, é utilizada a imagem original.

A estrutura utilizada para realizar a descrição de pontos, segue aquela definida pelo padrão MPEG-7³, empregada também no descritor EOH. O padrão MPEG-7 especifica um conjunto de descritores que podem ser utilizados para medir a similaridade entre imagens ou vídeos, dentre eles, o EHD (*Edge Histogram Descriptor*) que descreve a distribuição de bordas por meio de histogramas ([MANJUNATH et al., 2001](#); [WON; PARK; PARK, 2002](#)).

Para realizar o processo de descrição, primeiro é recortada uma janela de tamanho $N \times N$ em torno do ponto-chave previamente detectado. Essa janela é então dividida em 16 sub-regiões. Cada sub-região é uma unidade básica para extrair as informações de bordas. Isto é, para cada sub-região, determina-se qual borda é predominante.

Em seguida, o valor da coluna no histograma, referente à borda predominante, é incrementado. Para o descritor EHD (assim como também no descritor EOH) foram definidos cinco tipos de bordas, representados por cinco filtros direcionais (0, 90, 45, 135 e sem direção).

³O padrão MPEG-7, formalmente conhecido como Interface de Descrição de Conteúdo Multimídia, inclui metodologias padronizadas (descritores, esquemas de descrição e linguagem) que permitem descrições estruturais e detalhadas de informações audiovisuais em diferentes níveis (região, imagem e vídeo), e em diferentes áreas (descrição de conteúdo, gerenciamento, organização, navegação e interação do usuário) ([CHANG; SIKORA; PURL, 2001](#)).

No descritor MFD, foram definidos filtros Log-Gabor direcionais com os mesmos valores de orientação (0, 90, 45 e 135 graus) como definidos no padrão MPEG-7 (cabe lembrar aqui que tais ângulos também foram adotados no descritor EOH, porém com filtros Sobel). Esse conjunto de filtros com tais valores de orientação são capazes de detectar bordas horizontais e diagonais de maneira simples e com uma menor quantidade de filtros (KANOPOULOS; VASANTHAVADA; BAKER, 1988). Como não é possível definir um filtro sem orientação por meio de filtros Log-Gabor esse tipo de orientação não foi empregado (como utilizado no método descritor EOH). A Figura 24 ilustra a extração da sub-região, bem como a aplicação de cada filtro de borda para definir a borda predominante e calcular o histograma.

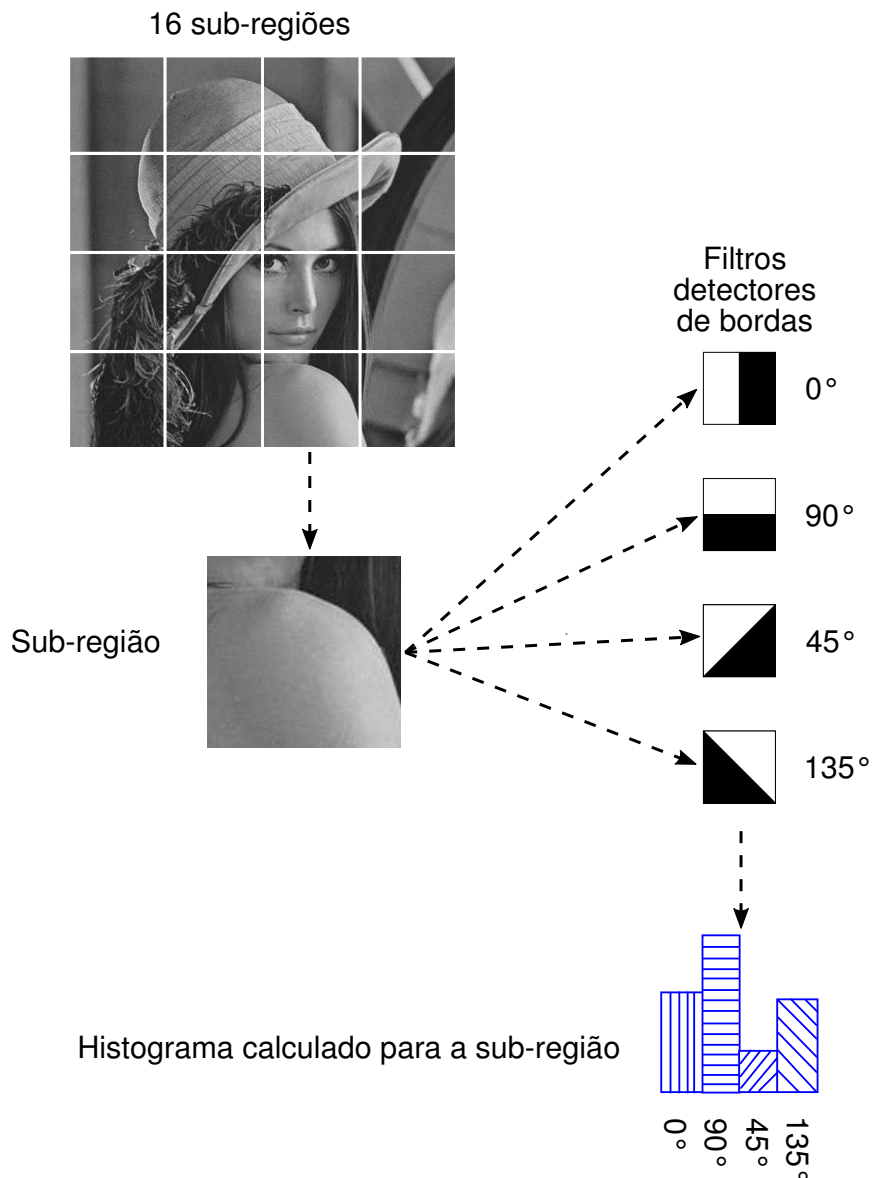


Figura 24 – Extração da sub-região e cálculo do histograma do MFD.

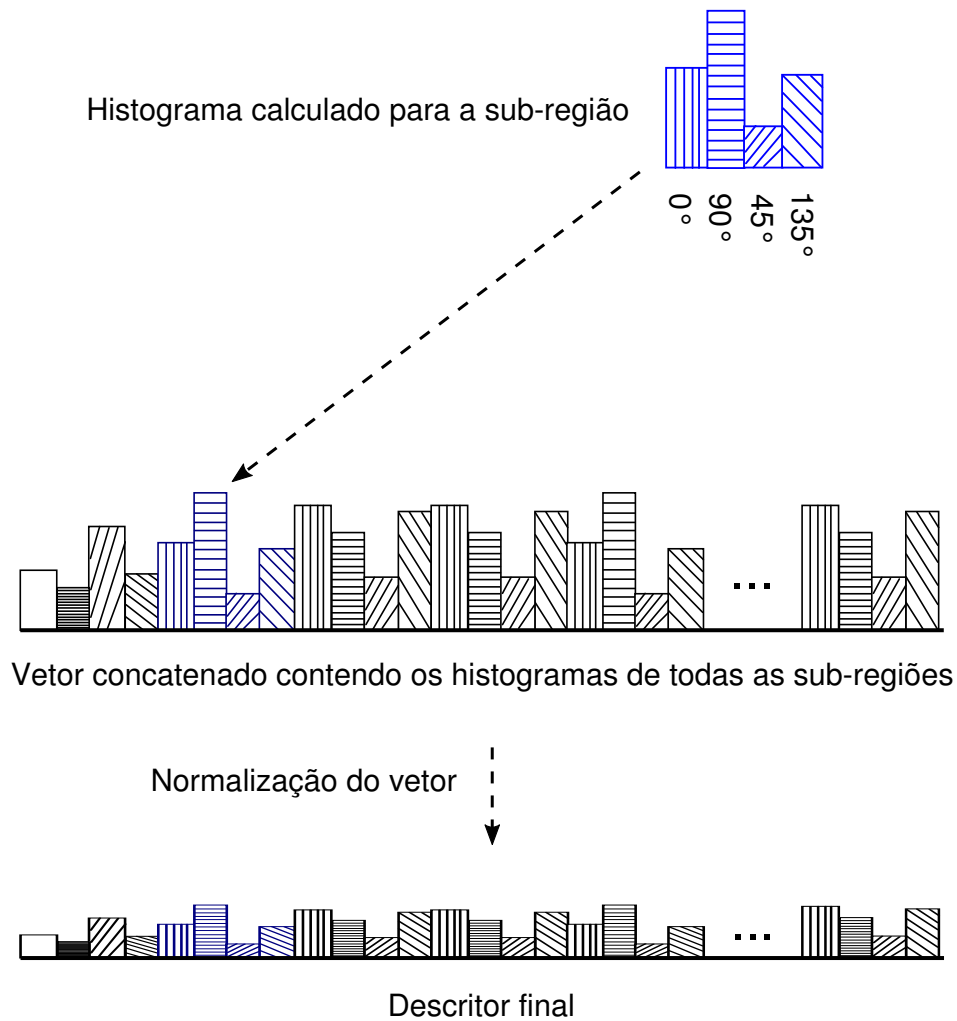


Figura 25 – Montagem final do descritor MFD.

Posteriormente, com a finalidade de se obter uma melhor eficácia para o método, foram avaliadas diferentes quantidades de filtros de orientação, variando a quantidade de filtros entre um à quinze filtros e posteriormente mensurado a precisão e a revocação do método descritor. Os valores de ângulos foram igualmente divididos entre 0 à 180 graus, para cada conjunto de filtros. Como os filtros de Log-Gabor permite variar a escala, foram avaliados também diferentes quantidades de filtros de escala (os resultados dessas avaliações encontram-se na [Seção 5.2](#)).

Neste trabalho foram definidos os valores das constantes para os filtros Log-Gabor, como sugeridos no trabalho de [Walia e Verma \(2016\)](#), por apresentarem melhores resultados na extração de informações de texturas ao utilizar os filtros Log-Gabor em descrição de imagens. Os valores definidos foram: $\lambda = 3,0$, $\sigma = 0,65$ e $k = 1,6$ (ver [Equação \(3\)](#) para detalhes de cada constante).

Após cada filtro de Log-Gabor ser aplicado em cada sub-região e calculado o histograma dessa mesma região, os valores dos histogramas de cada sub-região são concatenados em um só vetor, conforme ilustrado na [Figura 25](#). Esse vetor, que contém os

histogramas de cada sub-região é normalizado (divide-se cada valor do vetor pela norma L_2 do próprio vetor). O vetor normalizado consiste no descritor final.

Apesar das semelhanças do método descritor aqui proposto com os métodos EOH e LGHD, cabe aqui destacar algumas das principais diferenças entre tais métodos, quais sejam:

- Tanto o método descritor MFD quanto o método descritor EOH possuem uma estrutura similar para a realização da descrição de pontos (uma vez que ambos os métodos seguem o padrão do método descritor EHD, definido no padrão MPEG-7), no entanto, o método descritor EOH utiliza filtros Sobel para extração de informações de bordas, enquanto o método descritor MFD utiliza filtros Log-Gabor;
- O método descritor MFD utiliza a própria imagem de entrada para realizar a descrição dos pontos-chaves, já no método descritor EOH há uma etapa inicial em que a imagem de entrada é convertida em uma imagem de bordas por meio do detector Canny;
- O método descritor MFD foi proposto exclusivamente para trabalhar com imagens multiespectrais, já o método descritor LGHD foi proposto para trabalhar com imagens que possuem variações de intensidades não lineares de diversos tipos, como, por exemplo, imagens de mapas de profundidade, imagens multiespectrais e imagens obtidas por meio de câmeras com e sem flash;
- Tanto o método descritor MFD quanto o método descritor LGHD utilizam filtros Log-Gabor para extrair informações de bordas das imagens, no entanto o LGHD conta com 24 filtros Log-Gabor enquanto que o MFD utiliza apenas 10 filtros Log-Gabor (5 filtros de orientação e 2 filtros de escala, conforme os resultados dos experimentos que serão mostrados na [Seção 5.2](#)).

4.2 Avaliação dos algoritmos detectores

Para a avaliação dos algoritmos detectores, foram utilizados pares de imagens de espectros diferentes, representando assim uma imagem multiespectral (ver [Capítulo 5](#) para mais detalhes). Cada par é composto por uma imagem do espectro visível e uma imagem do espectro infravermelho.

Para cada par de imagens, são detectados os pontos-chave. Os pontos-chave da primeira imagem são projetados na outra imagem e então são calculados os erros de sobreposição das regiões nas quais se encontram sobrepostos os pontos-chave da primeira imagem com os pontos-chave da segunda imagem. A partir do erro de sobreposição, é determinado se a correspondência entre os dois pontos é uma correspondência correta ou não. Após isto, é avaliado a repetibilidade do atual par de imagens. A [Figura 26](#) ilustra a metodologia utilizada para a avaliação dos algoritmos detectores.

Foram avaliados os algoritmos detectores SIFT, SURF, FAST, ORB e BRISK. Nesta

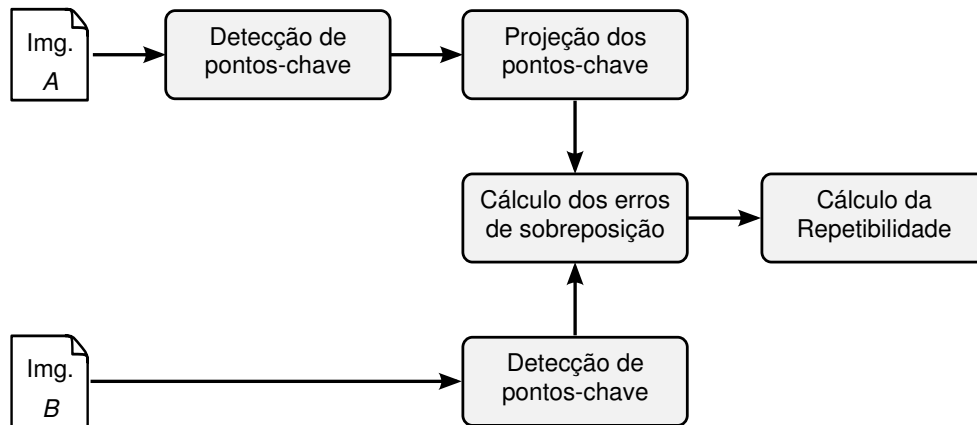


Figura 26 – Metodologia utilizada para avaliação dos algoritmos detectores.

etapa, foi mensurado também o tempo médio de processamento, medido em milissegundos, para 30 execuções em cada par de imagens de cada sequência. O tempo de processamento consiste no tempo de relógio que cada algoritmo detector leva para concluir sua tarefa.

4.3 Avaliação dos algoritmos descritores

Para a avaliação dos algoritmos descritores, foram utilizadas as métricas de precisão, revocação e medida F1 (descritas na [Seção 3.5](#)). Este experimento engloba as etapas de detecção, descrição e correspondência do sistema de correspondência de pontos em imagens. A metodologia utilizada aqui é inspirada nos trabalhos de [Aguilera et al. \(2012a\)](#), [Aguilera et al. \(2012a\)](#) e [Mikolajczyk e Schmid \(2005\)](#).

Para a determinação dos valores de precisão e revocação, são detectados e descritos os pontos-chave no par de imagens. Em seguida, é realizada a correspondência de cada descritor sobre esse par. A [Figura 27](#) ilustra a metodologia utilizada para a avaliação dos algoritmos descritores.

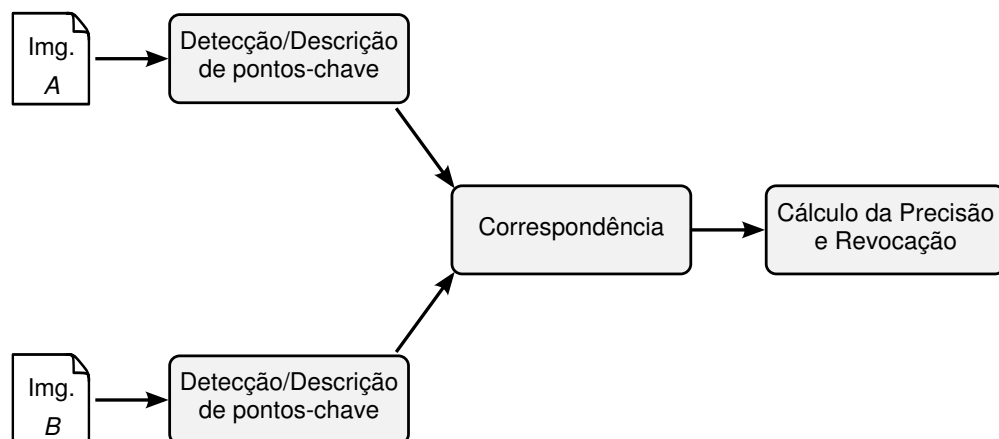


Figura 27 – Metodologia utilizada para avaliação dos algoritmos descritores.

Foram avaliados os algoritmos descritores SIFT, SURF, EOH, LGHD e o método MFD proposto neste trabalho. Nesta etapa, foi fixado um algoritmo detector que, apresentou melhores resultados na etapa anterior para os valores de repetibilidade e de tempo de processamento. Para cada algoritmo descritor, foram realizadas 30 execuções em cada par de imagens e mensurado o seu tempo médio de processamento, em milissegundos.

Assim como no trabalho de [Brown e Susstrunk \(2011\)](#) foram selecionadas 20% das imagens de cada base de dados para realizar o ajuste no descritor (tanto no número de filtros de orientação quanto no número de filtros de escala). O restante da base de dados foi utilizado para a avaliação do método proposto bem como a comparação com outros métodos descritores. Para este processo, repetiu-se 10 ensaios com a divisão desses conjuntos de forma aleatória. O resultado final de precisão e revocação consiste na média dos valores de cada ensaio.

5 Resultados

Neste capítulo são descritos os resultados experimentais obtidos visando-se avaliar a eficácia e a eficiência do método proposto para descrição de pontos-chave em imagens multiespectrais em problemas de correspondência de pontos em imagens. Para tanto, foram utilizadas duas bases de dados principais.

A primeira base de dados, denominada RGB-NIR e proposta por Brown e Susstrunk (2011), é comumente utilizada na literatura (JIN et al., 2016; LI et al., 2015b; LI et al., 2015c; AGUILERA; SAPP; TOLEDO, 2015; SALEEM; SABLATNIG, 2014; WANG; ZHU; YUAN, 2013) e está disponível publicamente¹. Esta base consiste em um conjunto de 477 pares de imagens do espectro visível e do espectro infravermelho (infravermelho próximo - NIR) distribuídas em 9 categorias, quais sejam: *Country* (52 pares), *Field* (51 pares), *Forest* (53 pares), *Indoor* (56 pares), *Mountain* (55 pares), *Old Building* (51 pares), *Street* (50 pares), *Urban* (58 pares) e *Water* (51 pares). Tais categorias cobrem cenas de países, campos, florestas, ambientes internos, montanhas, edifícios, ruas, áreas urbanas e cenas de água.

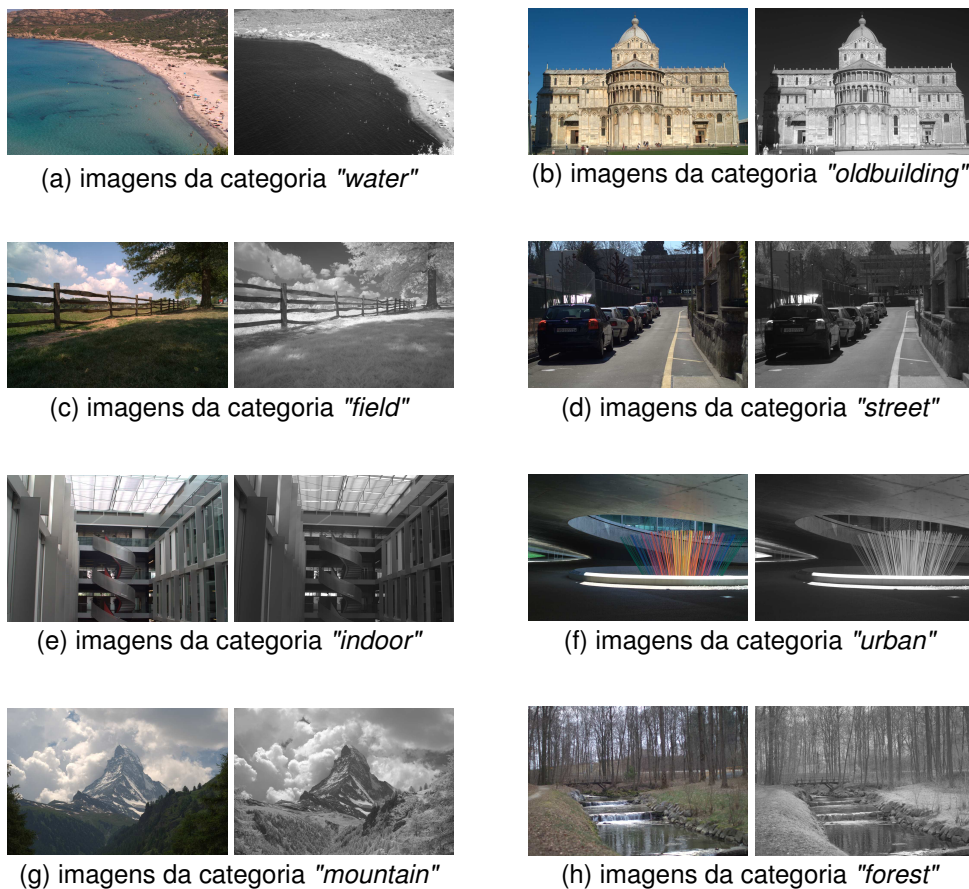


Figura 28 – Exemplos de imagens da base de dados RGB-NIR.

¹Disponível em: <http://ivrg.epfl.ch/supplementary_material/cvpr11/index.html>.

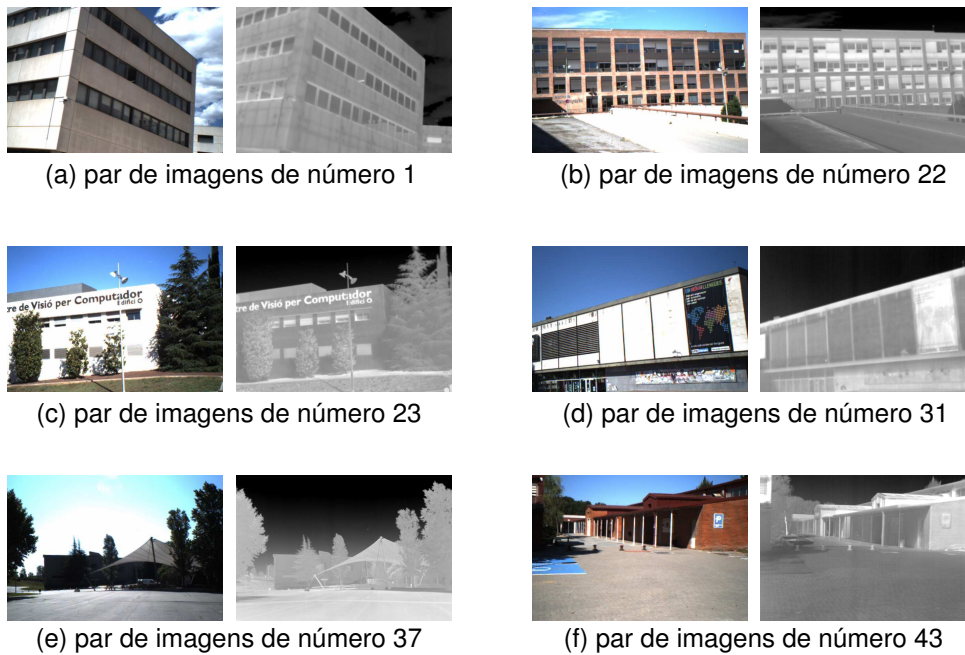


Figura 29 – Exemplos de pares de imagens da base de dados RGB-LWIR.

Todas as imagens já estão retificadas e alinhadas espacialmente. A Figura 28 exibe algumas das imagens de cada categoria desta base de dados.

A segunda base de dados, denominada RGB-LWIR e proposta por Aguilera, Sappa e Toledo (2015), foi utilizada para avaliar a eficácia dos descritores LGHD e EOH no mesmo trabalho dos autores. Esta base também está disponível publicamente² e é composta por 44 pares de imagens do espectro visível e do espectro infravermelho (infravermelho de ondas longas - LWIR). Esta base de dados foi utilizada para complementar a primeira base de dados (RGB-NIR) e verificar a eficácia do método proposto em uma outra faixa do espectro infravermelho, ou seja, infravermelho de ondas longas. A Figura 29 exibe alguns pares de imagens desta base de dados.

Para a realização dos experimentos, foi utilizada uma estação de trabalho com processador Intel Core i5-5200U 2,20GHz e 8 GB de memória RAM, rodando o sistema operacional Ubuntu 16.04 de 64 bits (GNU/Linux com Kernel 4.4.0).

5.1 Avaliação dos Algoritmos Detectores

Nesta seção são apresentados os resultados experimentais do processo metodológico adotado para se avaliar algoritmos detectores de pontos-chave em imagens multi-espectrais e selecionar um método em específico para ser utilizado conjuntamente com o descritor MFD em problemas de correspondência de pontos. Neste grupo de experimentos foram utilizadas as duas bases de dados.

²Disponível em: <<https://owncloud.cvc.uab.es/owncloud/index.php/s/1Wx715yUh6kDAO7>>.

Métrica (valores médios)	Método detector				
	SIFT	SURF	FAST	BRISK	ORB
Repetibilidade	0,344	0,456	0,498	0,373	0,430
Pontos detectados (esp. visível)	472,5	594,1	713,5	1132,0	372,1
Pontos detectados (esp. infravermelho)	466,8	485,8	331,6	914,0	370,2
Tempo (ms)	413,3	187,6	0,1	30,4	13,1

Tabela 1 – Resultados dos métodos detectores.

A [Tabela 1](#) sumariza os valores dos experimentos desta seção (os maiores valores de repetibilidade e de quantidade média de pontos detectados, bem como o valor do menor tempo médio de processamento, foram ressaltados em negrito na tabela). Nesta tabela são exibidos os valores médios de repetibilidade, quantidade de pontos detectados em imagens do espectro visível e quantidade de pontos detectados em imagens do espectro infravermelho de cada método detector para cada imagem das duas bases de dados. Esta tabela também exibe os valores médios do tempo de processamento que cada método detector apresentou para cada imagem das duas bases de dados.

A [Figura 30](#) exibe um gráfico que apresenta a quantidade média de pontos detectados nas imagens (luz visível e infravermelho) de ambas as bases de dados consideradas, por meio dos métodos detectores analisados. Por meio deste experimento, observa-se que os métodos detectores SURF, FAST e BRISK apresentaram diferenças significativas entre a quantidade média de pontos detectados em imagens de espectro visível e em imagens de espectro infravermelho, diferentemente dos métodos detectores SIFT e ORB. Essa diferença mostra que os métodos detectores SURF, FAST e BRISK são sensíveis ao tipo de espectro utilizado. O método detector BRISK apresentou uma maior quantidade de pontos detectados.

A [Figura 31](#) exibe os valores médios de repetibilidade para os métodos detectores aplicados em todas as bases de dados. Para cada imagem foram detectados os pontos-chave e calculado o valor de repetibilidade. Por meio deste experimento, observa-se que o método detector FAST apresentou melhores valores de repetibilidade (repetibilidade média de 0,498). Enquanto os métodos detectores SIFT e BRISK apresentaram menores valores de repetibilidade, com médias de 0,344 e 0,373 respectivamente. Observa-se também, por meio desse experimento, que todos os métodos detectores possuem uma taxa de repetibilidade inferior à 0,5 (ver [Tabela 1](#)). Isso mostra que ambos os métodos detectores possuem certa deficiência ao detectar pontos-chave em imagens multispectrais.

Para verificar a significância estatística destes valores, foi utilizado o teste estatístico de análise de variância (ANOVA, do inglês *Analysis of Variance*) ([MONTGOMERY, 2008](#)). A ANOVA procura verificar se existe uma diferença significativa entre as médias. O resultado obtido indicou que, existem evidências estatísticas com 95% de confiança de que há diferença na média de repetibilidade entre o método detector FAST e os demais métodos

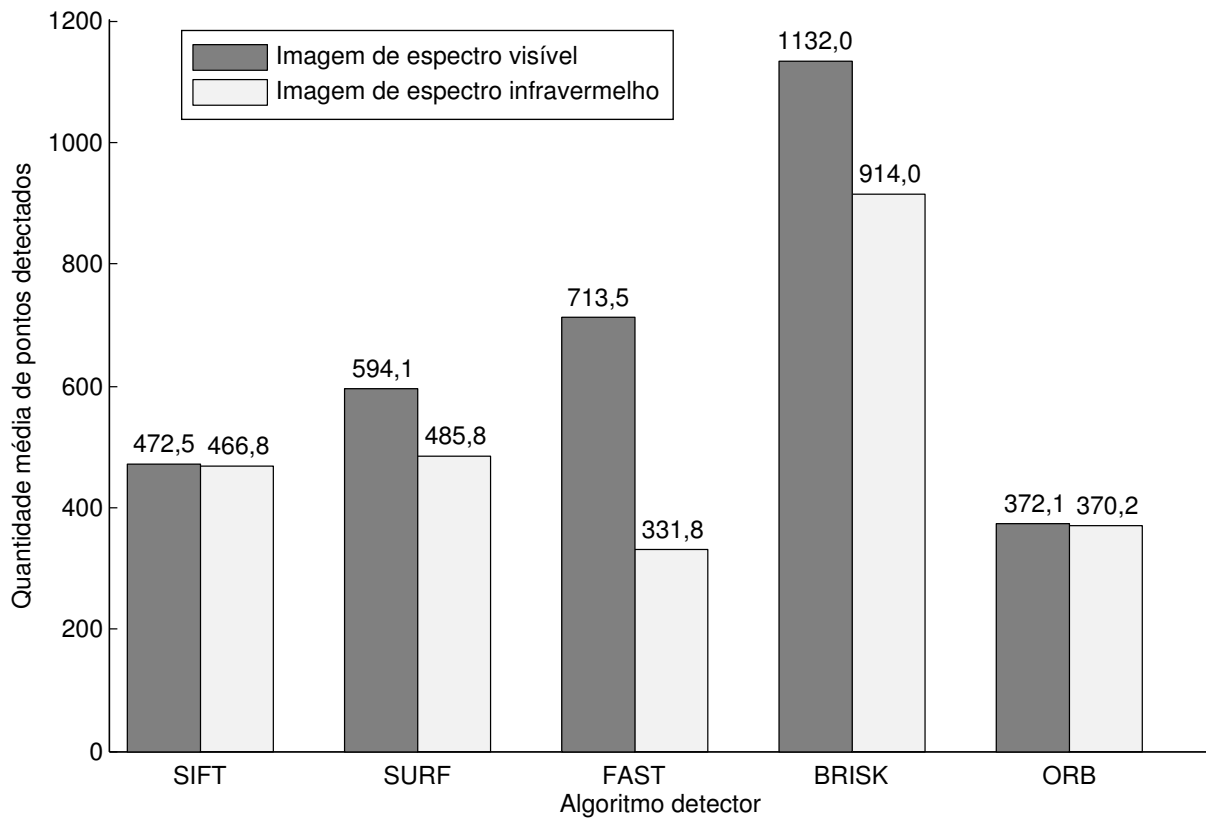


Figura 30 – Quantidade média de pontos detectados em imagens de espectro visível e imagens de espectro infravermelho.

descritores.

Por meio da [Figura 30](#) observa-se que, o método descritor BRISK detecta mais pontos que os demais métodos porém, como apresentado na [Figura 31](#), grande parte

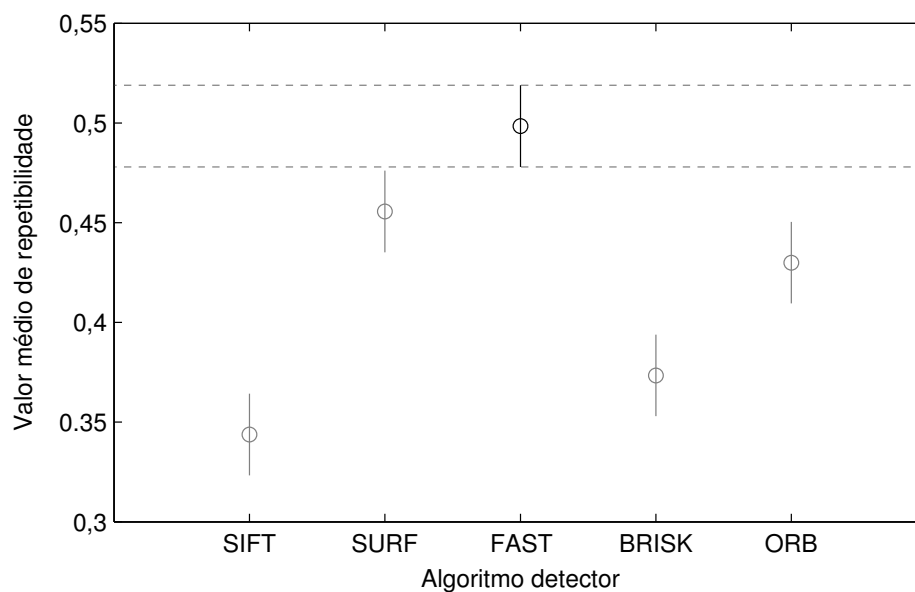


Figura 31 – Valor médio de repetibilidade para os métodos detectores.

destes pontos são detectados incorretamente devido ao fato de que este método apresenta menores valores de repetibilidade.

A [Figura 32](#) exibe o gráfico do tempo médio de processamento de cada método detector. Por meio deste gráfico, observa-se que o método detector FAST apresentou o menor tempo de processamento, já o método detector SIFT apresentou o maior tempo de processamento dentre todos os algoritmos.

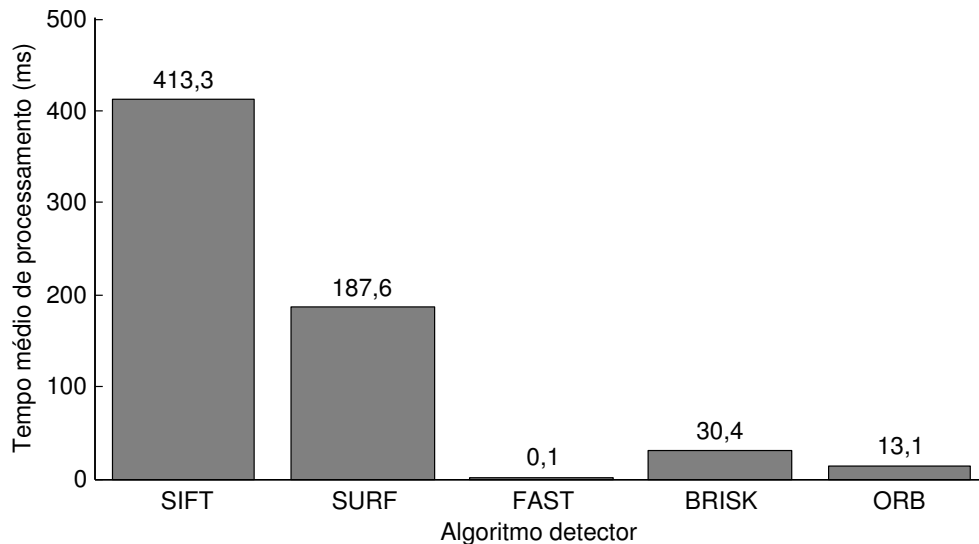


Figura 32 – Tempo médio de processamento para os métodos detectores.

Por meio da [Figura 31](#) e da [Figura 32](#), pode-se concluir que o método detector FAST possui um equilíbrio entre tempo de processamento e desempenho geral de detecção, pois este algoritmo apresentou maiores taxas de repetibilidade com o menor tempo de processamento. Apesar de o método detector FAST detectar uma menor quantidade de pontos em imagens de espectro infravermelho, como exibido na [Figura 30](#), grande parte desses pontos foram detectados de maneira correta devido ao maior valor médio de repetibilidade apresentado por esse método. Sendo assim, definiu-se o método detector FAST como sendo aquele a ser utilizado conjuntamente com o método descritor MFD nos experimentos que avaliam seu desempenho em processos de correspondências de pontos em imagens multiespectrais.

5.2 Análise de Parâmetros do Descritor MFD

Com a finalidade de maximizar a eficácia do método descritor MFD, foram avaliadas diferentes quantidades de filtros em sua construção. Para tanto, foram adicionados filtros de orientação e mensurada a precisão, revocação, medida F1 e tempo de processamento desse método descritor. Para cada conjunto de filtros definidos nesse experimento, os valores de ângulos foram igualmente divididos entre 0 à 180 graus.

Métrica (valores médios)	Quantidade de filtros de orientação				
	1	2	3	4	5
Precisão	0,189	0,627	0,755	0,772	0,799
Revocação	0,014	0,533	0,585	0,600	0,596
Medida F1	0,027	0,576	0,659	0,675	0,683
Tempo (ms)	2406,0	4263,0	6140,0	7925,0	9738,0
Métrica (valores médios)	Quantidade de filtros de orientação				
	6	7	8	9	10
Precisão	0,792	0,801	0,806	0,805	0,808
Revocação	0,589	0,578	0,564	0,557	0,542
Medida F1	0,676	0,671	0,663	0,658	0,649
Tempo (ms)	1147,4	1319,8	1520,0	1695,2	1858,5
Métrica (valores médios)	Quantidade de filtros de orientação				
	11	12	13	14	15
Precisão	0,813	0,811	0,806	0,811	0,810
Revocação	0,530	0,517	0,504	0,498	0,488
Medida F1	0,642	0,632	0,620	0,617	0,609
Tempo (ms)	2039,0	2208,5	2386,9	2562,8	2742,0

Tabela 2 – Valores médios de precisão, revocação e medida F1 para a variação na quantidade de filtros de orientação do método MFD.

A Tabela 2 resume os valores desse experimento (os maiores valores das métricas de precisão, revocação e medida F1, bem como o valor do menor tempo médio de processamento, foram ressaltados em negrito na tabela). Nesta tabela são exibidos os valores médios de cada métrica para a execução do método descritor MFD variando-se a quantidade de filtros de orientação. Os valores médios dessas métricas foram obtidos durante a correspondência de pontos-chave em cada par de imagens das duas bases de dados.

A Figura 33 exibe os valores médios de precisão, revocação e medida F1 para a variação na quantidade de filtros de orientação do método descritor MFD. Por meio deste

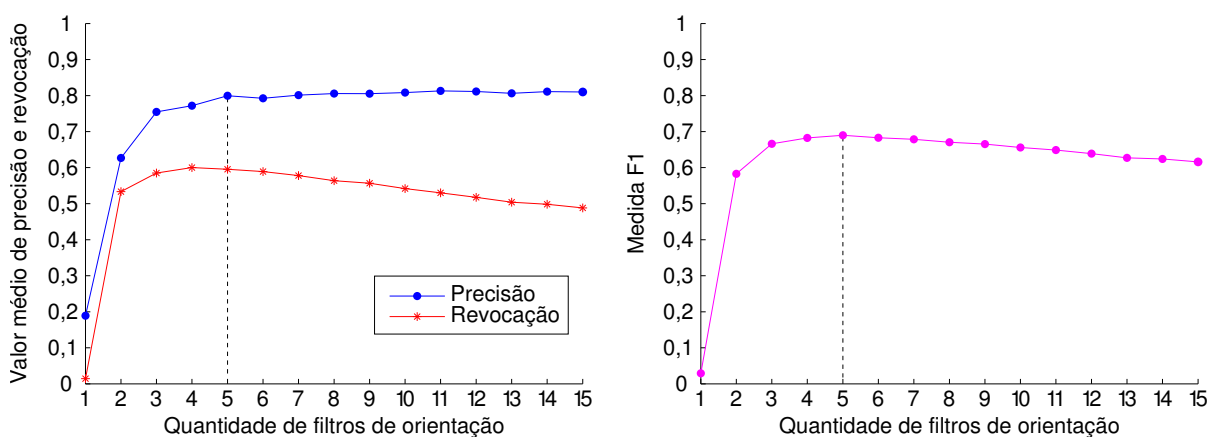


Figura 33 – Valores médios de precisão, revocação e medida F1 para a variação na quantidade de filtros de orientação do método descritor MFD.

experimento, observa-se que a partir de 5 filtros, o método começa a ter um desempenho de revocação cada vez menor. Portanto, para 5 filtros de orientação o método descritor obtém um maior valor de precisão e um maior valor de revocação. Essa informação pode ser visualizada na curva da medida F1, da mesma [Figura 33](#).

A [Figura 34](#) exibe os valores de tempo médio de processamento, em milissegundos, para a variação na quantidade de filtros de orientação do método descritor MFD. Observa-se por meio desse experimento que a inclusão de filtros no método descritor provoca um aumento linear no tempo de processamento devido uma maior quantidade de operações de convolução que o método deve utilizar ao aplicar um determinado filtro na imagem.

É importante ressaltar que a inclusão de filtros no método não apenas provoca o aumento no número de operações (ao aplicar mais um filtro), como também há um aumento no tamanho final do vetor descritor. Para cada filtro adicionado, tem-se um aumento de 16 valores de ponto flutuantes, devido ao histograma de 16 colunas que representam as 16 sub-regiões da janela de imagem (ver [Figura 24](#)). Portanto, para 5 filtros de orientação obtém-se um vetor descritor de tamanho 80.

Após fixado os 5 filtros de orientação, e ainda com a finalidade de se obter uma melhor eficácia para o método descritor MFD, avaliou-se o comportamento do descritor ao adicionar filtros de escala. A [Tabela 3](#) sumariza os valores desse experimento, e a [Figura 35](#) exibe os valores médios de precisão, revocação e medida F1 ao avaliar os filtros de escala (os maiores valores das métricas de precisão, revocação e medida F1, bem como o valor do menor tempo médio de processamento, foram ressaltados em negrito na tabela). Nesta tabela são exibidos os valores médios de cada métrica para a execução do método descritor MFD variando-se a quantidade de filtros de escala. Os valores médios dessas métricas foram obtidos durante a correspondência de pontos-chave em cada par de imagens das

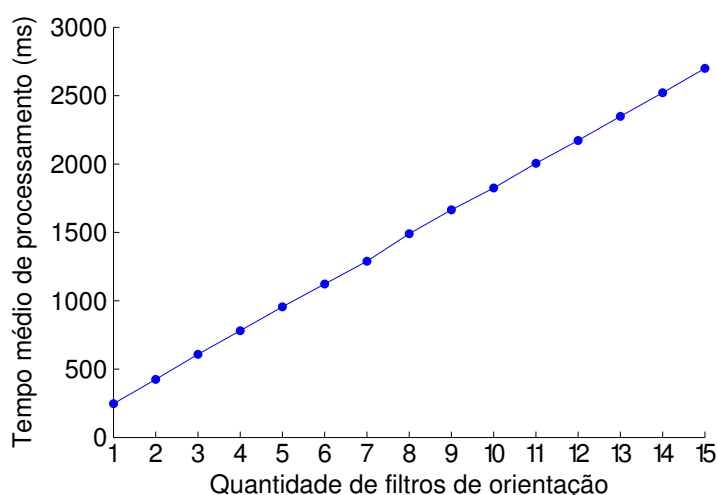


Figura 34 – Tempo médio de processamento para a variação na quantidade de filtros de orientação do método MFD.

Métrica (valores médios)	Quantidade de filtros de escala					
	1	2	3	4	5	6
Precisão	0,799	0,814	0,815	0,807	0,809	0,814
Revocação	0,596	0,623	0,588	0,558	0,541	0,536
Medida F1	0,683	0,706	0,683	0,660	0,649	0,646
Tempo (ms)	980,9	1904,1	2864,4	3808,8	4740,7	5636,7

Tabela 3 – Valores médios de precisão, revocação e medida F1 para a variação na quantidade de filtros de escala do método MFD.

duas bases de dados.

Por meio desse experimento, observa-se que a partir de 2 filtros, o método começa a ter um desempenho de revocação cada vez menor. Observa-se também que não há um ganho significativo no valor de precisão ao adicionar filtros com escalas diferentes. Portanto, para 2 filtros de escala o método descritor obtém maior valor de precisão com um maior valor de revocação. Essa informação pode ser visualizada na curva da medida F1, da mesma [Figura 35](#).

A [Figura 36](#) exibe os valores de tempo médio de processamento, em milissegundos, para a variação na quantidade de filtros de escala do método descritor MFD. Observa-se por meio desse experimento que a inclusão de filtros de escala no método descritor possui um comportamento semelhante ao comportamento encontrado na [Figura 34](#). O tempo cresce linearmente, porém, para cada filtro de escala adicionado são adicionados novos 5 filtros de orientação. Sendo assim, o tempo de processamento é aumentado em média em 5 vezes, e o vetor descritor final também é aumentado em 5 vezes.

De acordo com os experimentos mostrados nesta seção, recomenda-se a implementação do método descritor MFD com 5 filtros de orientação ($m = 0, 1, 2, 3, 4$, da [Equação \(3\)](#))

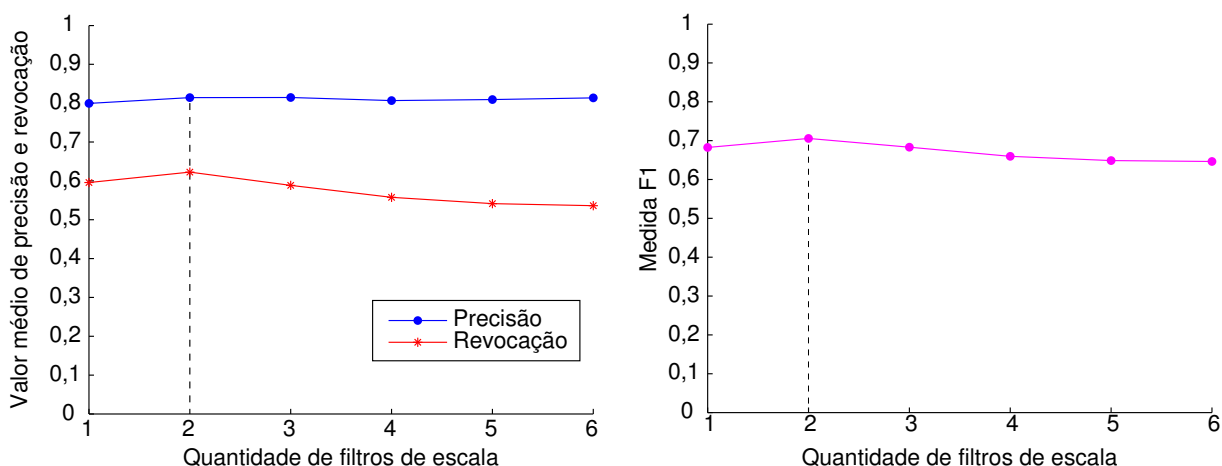


Figura 35 – Valores médios de precisão, revocação e medida F1 para a variação na quantidade de filtros de escala do método descritor MFD.

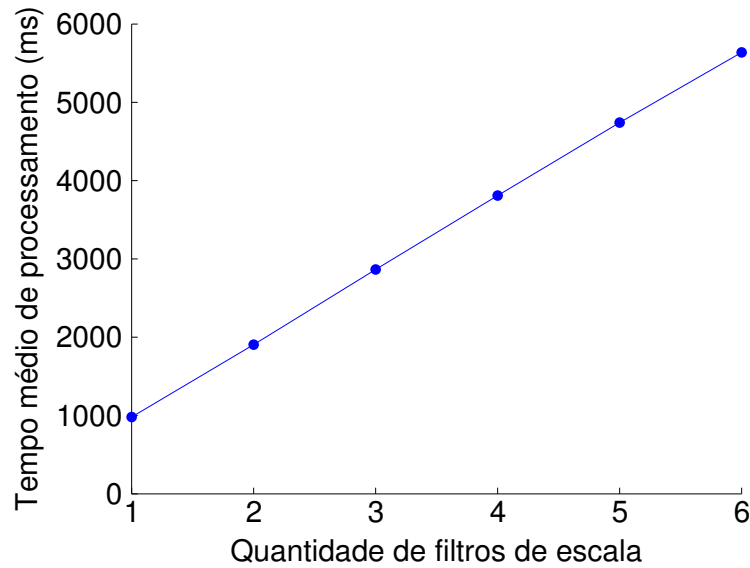


Figura 36 – Tempo médio de processamento para a variação na quantidade de filtros de escala do método MFD.

e 2 filtros de escala ($n = 0, 1$ da [Equação \(3\)](#)). Portanto o vetor descritor final terá 160 valores de pontos flutuantes (16 sub-regiões $\times$ 5 filtros de orientação $\times$ 2 filtros de escala).

O [Quadro 3](#) exibe uma comparação no tamanho dos descritores dos métodos de descrição para imagens multiespectrais que foram estudados aqui e que serão comparados na próxima seção.

Método descritor	Tamanho do descritor
EOH	80
LGHD	384
MFD	160

Quadro 3 – Tamanhos dos descritores dos métodos de descrição para imagens multiespectrais.

5.3 Avaliação de Desempenho do Descritor MFD

O descritor MFD e os descritores utilizados como referências (do inglês, *baselines*) foram avaliados em cada base de dados de forma separada, com a finalidade de se obter uma melhor compreensão dos resultados dos métodos ao utilizar imagens de infravermelho perto (refente à base de dados RGB-NIR exibida neste mesmo capítulo) e imagens de infravermelho de ondas longas (referente à base de dados RGB-LWIR).

A [Tabela 4](#) sumariza os resultados dos experimentos realizados na base de dados RGB-NIR (os maiores valores das métricas de precisão, revocação e medida F1 foram ressaltados em negrito na tabela). Nesta tabela são exibidos os valores médios de cada

métrica, obtidos durante a correspondência de pontos-chave em cada par de imagens da base de dados RGB-NIR.

Métrica (valores médios)	Método descritor				
	SIFT	SURF	EOH	LGHD	MFD
Precisão	0,529	0,394	0,690	0,798	0,814
Revocação	0,446	0,273	0,330	0,520	0,623
Medida F1	0,484	0,323	0,446	0,630	0,706

Tabela 4 – Resultados dos métodos descritores (base de dados RGB-NIR).

A [Figura 37](#) exibe os valores de precisão e revocação obtidos pelos métodos descritores. Por meio deste experimento, observa-se que o algoritmo SURF obteve o menor desempenho tanto na precisão quanto na revocação. Apesar de o método descritor SIFT apresentar um valor de medida F1 maior que o método descritor EOH, o método descritor SIFT obteve um menor desempenho de precisão. Os métodos descritores LGHD e MFD obtiveram maiores valores de precisão e revocação em relação aos demais métodos.

Apesar de o método descritor MFD aparentar ser mais eficaz que o método LGHD, por apresentar resultados de precisão e revocação ligeiramente melhores, foi realizado um teste estatístico para verificar se essa diferença na média possui uma significância estatística.

A [Figura 38](#) exibe um teste de comparação par a par para a média de precisão, por meio dos resultados obtidos do teste ANOVA. O resultado obtido indicou que não há

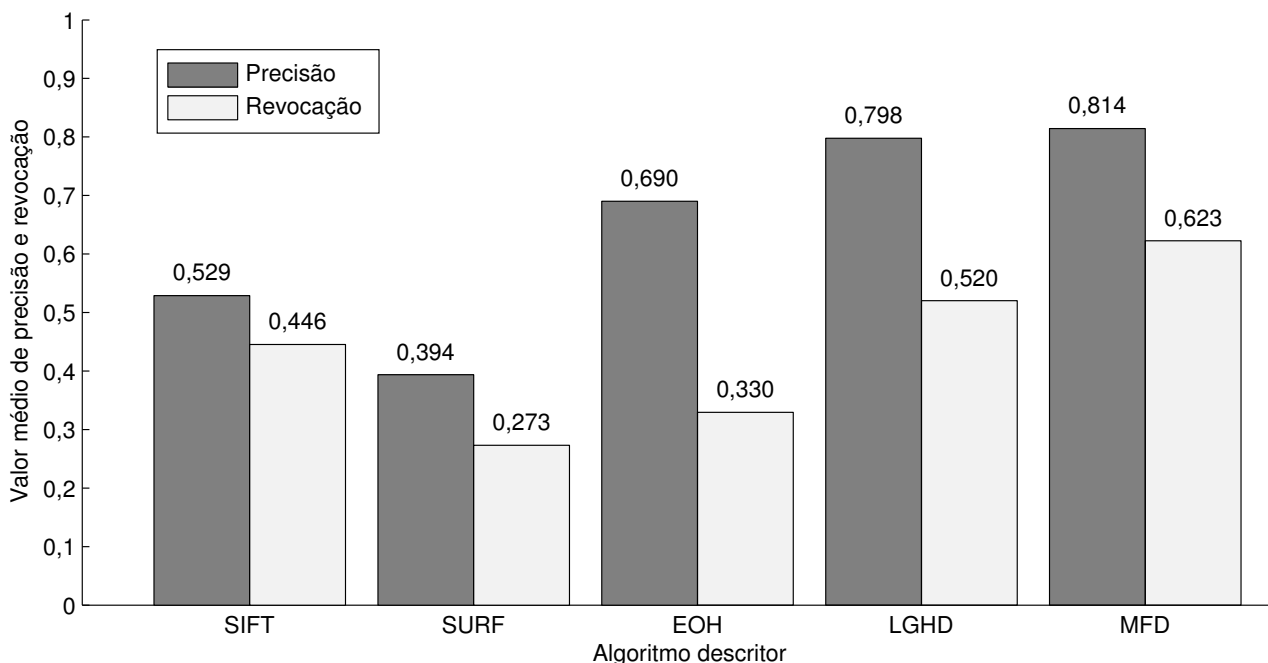


Figura 37 – Valores médios de precisão e revocação para os métodos descritores (para a base de dados RGB-NIR).

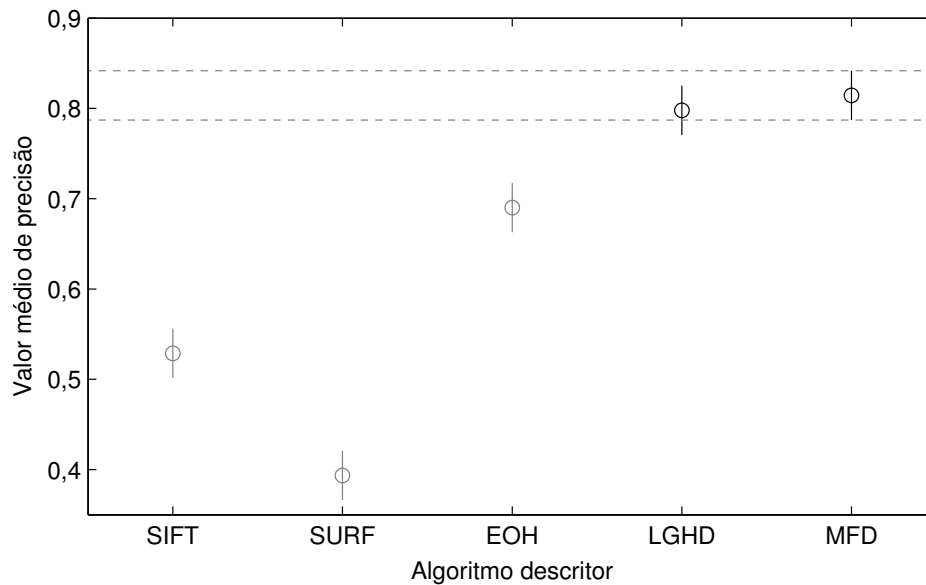


Figura 38 – Resultado de precisão para os métodos descritores (para a base de dados RGB-NIR).

evidências estatísticas com 95% de confiança de que há diferença na média de precisão dos métodos descritores LGHD e MFD. Portanto, não se pode dizer que o método descritor MFD é superior ao método descritor LGHD em termos de precisão.

A [Tabela 5](#), por sua vez, exibe os valores de precisão, revocação e medida F1 obtidos pelos métodos descritores para a base de dados RGB-LWIR (os maiores valores das métricas de precisão, revocação e medida F1 foram ressaltados em negrito na tabela). Nesta tabela são exibidos os valores médios de cada métrica, obtidos durante a correspondência de pontos-chave em cada par de imagens da base de dados RGB-LWIR.

Por meio deste experimento, observa-se que os métodos descritores SIFT e SURF obtiveram menores desempenhos tanto na precisão quanto na revocação. Os métodos descritores EOH, LGHD e MFD obtiveram maiores valores de precisão mas com resultados de revocação próximos de 0,1. O método descritor MFD obteve maiores valores de precisão e revocação, comparado aos demais métodos descritores. Esses valores podem ser visualizados também na [Figura 39](#).

Assim como foi realizado para a base de dados RGB-NIR, para a base de dados RGB-LWIR também foi realizado um teste estatístico para verificar se as diferenças nas

Métrica (valores médios)	Método descritor				
	SIFT	SURF	EOH	LGHD	MFD
Precisão	0,051	0,007	0,185	0,269	0,345
Revocação	0,038	0,005	0,062	0,081	0,094
Medida F1	0,044	0,005	0,093	0,124	0,148

Tabela 5 – Resultados dos métodos descritores (base de dados RGB-LWIR).

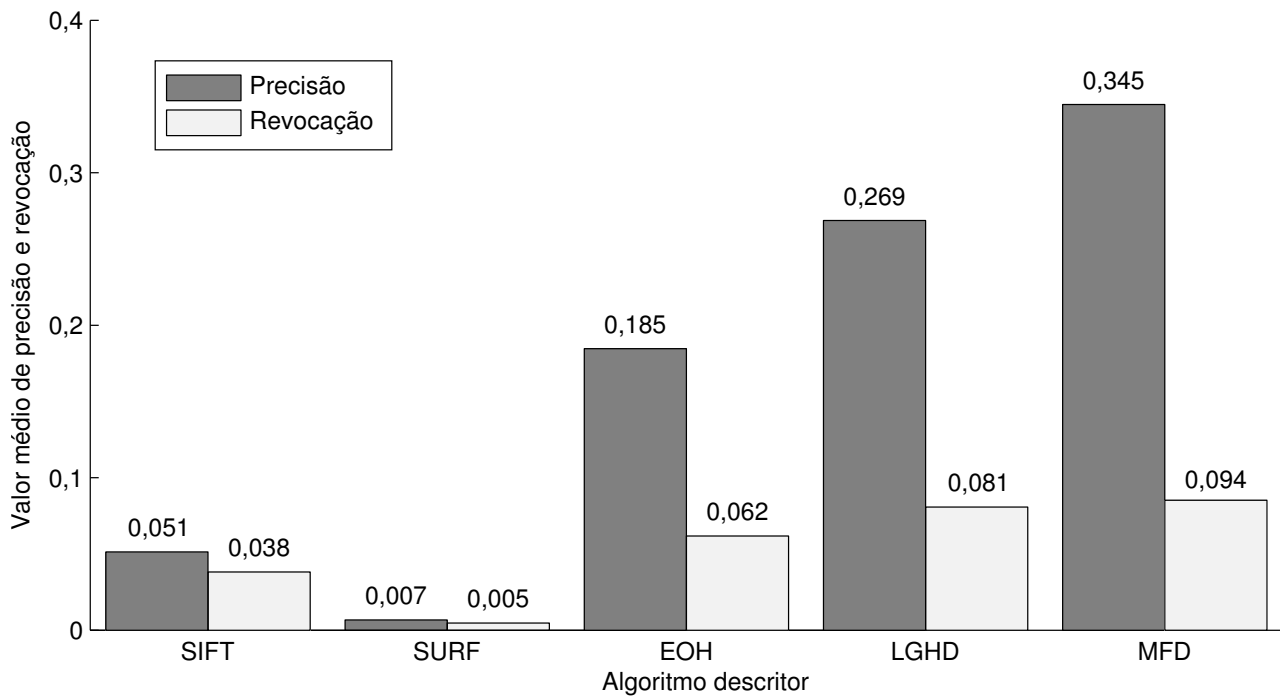


Figura 39 – Valores médios de precisão e revocação para os métodos descritores (para a base de dados RGB-LWIR).

médias de precisão entre os métodos descritores EOH, LGHD e MFD possuem uma significância estatística, como exibido na [Figura 40](#). O resultado obtido indicou que não há evidências estatísticas com 95% de confiança de que há diferença na média de precisão

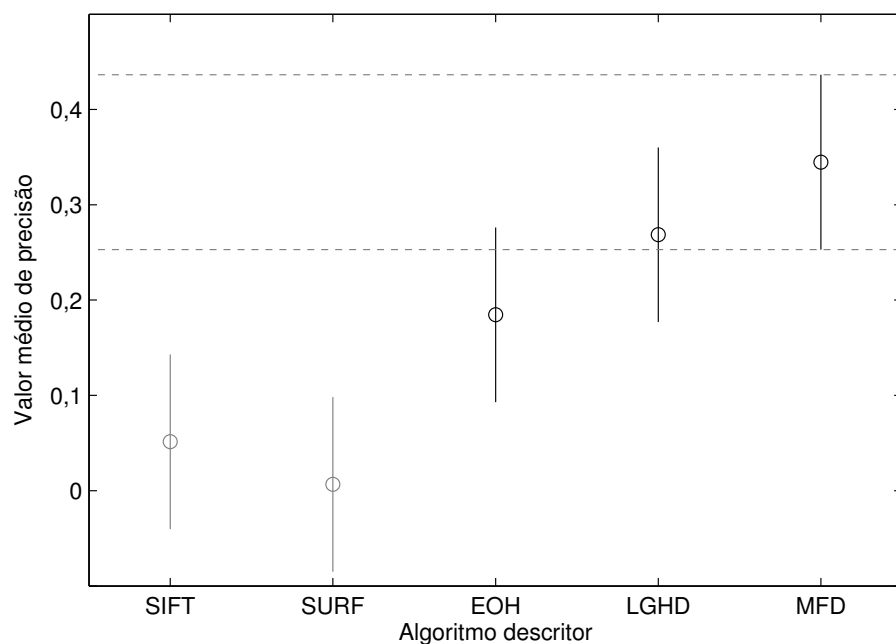


Figura 40 – Valores médios de precisão para os métodos descritores (para a base de dados RGB-LWIR).

dos métodos descritores EOH, LGHD e MFD. Portanto, não se pode afirmar que, para esta base de dados RGB-LWIR, há diferença de desempenho relativo à precisão para tais métodos.

Observa-se também, por meio dos experimentos realizados na base de dados RGB-LWIR, que, em geral, todos os métodos aqui avaliados obtiveram valores de precisão e revocação menores que os obtidos na base de dados RGB-NIR. Como mostram as [Figura 37](#) e a [Figura 39](#), para a base de dados RGB-NIR, o método descritor MFD obteve valor de precisão de 0,814, já para a base de dados RGB-LWIR, esse mesmo método obteve um valor de precisão de 0,345.

Apesar dessa diferença nos resultados entre bases de dados que utilizam imagens do espectro infravermelho próximo e imagens do espectro infravermelho de ondas longas, não se pode concluir que os métodos aqui avaliados possuem melhor aplicação para imagens multiespectrais que possuem imagens do espectro infravermelho próximo, comparado às imagens multiespectrais que possuem imagens do espectro infravermelho de ondas longas. Pois, há outros fatores que poderiam ter influenciado nos valores de precisão e revocação, tais como: diferenças nos tamanhos das imagens, ruídos introduzidos durante a captura dessas imagens, dentre outros, e esses fatores não foram avaliados nesse trabalho. O que se pode concluir com esses experimentos é que o método descritor MFD possui uma eficácia estatisticamente comparável aos métodos descritores EOH e LGHD.

A [Figura 41](#) exibe o gráfico do valor médio do tempo de processamento dos métodos descritores EOH, LGHD e MFD, executados em todas as bases de dados. Por meio deste gráfico, observa-se que o método descritor LGHD apresentou um elevado tempo de

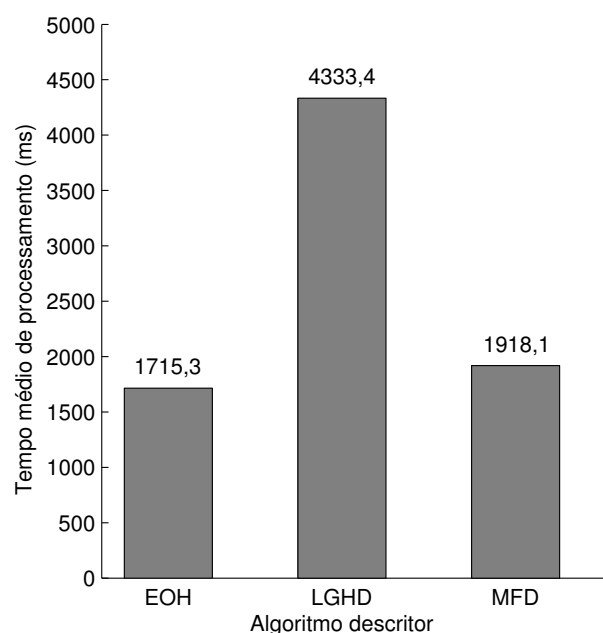


Figura 41 – Valores médios do tempo de processamento dos métodos descritores.

processamento. Isto se deve, principalmente, ao fato da quantidade de filtros Log-Gabor utilizados nesse método. O método descritor MFD possui um tempo de processamento de aproximadamente 42% menor que o método descritor LGHD e um tempo de processamento de aproximadamente 10% maior que o método descritor EOH.

6 Conclusão

Nesse trabalho foi apresentando um novo método descritor de características locais em imagens multiespectrais para ser aplicado em abordagens de correspondência de pontos-chave.

Diferentemente da grande maioria das propostas na literatura, baseadas em métodos descritores projetados exclusivamente para imagens obtidas no espectro de luz visível como, por exemplo, o método descritor SIFT, este trabalho propõe a utilização de um descritor local desenvolvido especialmente para trabalhar com imagens multiespectrais.

Um dos principais problemas encontrados durante as descrições de pontos em imagens multiespectrais se baseia no fato de que as direções dos gradientes de intensidade associados a tais pontos podem variar significativamente entre as imagens obtidas em diferentes faixas do espectro eletromagnético. No entanto, as aparências globais das formas dos objetos tendem a permanecer razoavelmente invariantes.

Deste modo, para a extração das características das imagens multiespectrais, foram adotados os filtros de Log-Gabor por se tratar de uma técnica invariante às diferenças de gradiente, e comumente utilizado na literatura como detector de bordas. Sendo assim, o descritor proposto, denominado *Multispectral Feature Descriptor* (MFD), baseia-se na extração de informações das imagens nos domínios do espaço e da frequência por meio da utilização desses filtros. Cabe lembrar aqui também que, o método descritor proposto nesse trabalho segue a mesma estrutura de descrição definido pelo padrão MPEG-7, que definem metodologias padronizadas para realizar descrições em informações visuais.

Foram utilizadas duas bases de dados compostas por partes de imagens do espectro visível e imagens do espectro infravermelho (infravermelho próximo e infravermelho de ondas longas) para realizar a avaliação dos métodos detectores e dos descritores, incluindo o método descritor MFD. Durante essa avaliação foram utilizadas as métricas de repetibilidade e tempo de processamento para os métodos detectores e as métricas de precisão, revocação e medida F1 para os métodos descritores.

Ao avaliar os métodos detectores SIFT, SURF, FAST, BRISK e ORB, foi mostrado que o método detector FAST apresentou um melhor desempenho de detecção e um menor tempo de processamento, o que fez com que este método fosse selecionado para ser utilizado como o método detector de pontos-chave em imagens multiespectrais para ser utilizado em um sistema de correspondência de pontos em imagens.

Para a construção do método descritor MFD, foram realizados experimentos para determinar uma quantidade de filtros de orientação e filtros de escala que resultasse em

um maior valor de precisão e revocação, e concluiu-se que com 5 filtros de orientação e 2 filtros de escala, o descritor obtém uma melhor desempenho de descrição.

Após a definição da quantidade de filtros, o método descritor MDF foi comparado aos métodos descritores do estado da arte EOH e LGHD. Foi verificado, por meio de uma análise de variância, que o método descritor MDF possui em média uma mesma eficácia que o método descritor LGHD, porém o descritor MDF obteve uma eficiência maior, em termos de tempo de processamento, sendo aproximadamente 42% mais rápido que o método descritor LGHD, para as bases de dados utilizadas. Comparado ao método descritor EOH, o método descritor MDF obteve melhores resultados de precisão, entretanto, obteve um custo computacional similar, com um tempo de processamento de apenas 10% superior.

Foram verificados também que o método descritor MDF supera os métodos descritores SIFT e SURF, em termos de eficácia e eficiência, para realizar a correspondência de pontos-chave em imagens multiespectrais.

6.1 Trabalhos futuros

A seguir são listadas propostas de trabalhos futuros, a serem realizados a partir dos resultados aqui obtidos:

- Avaliar o desempenho do descritor MDF em problemas de correspondências de pontos-chave, utilizando-se imagens obtidas em outras faixas do espectro eletromagnético;
- Avaliar diferentes tamanhos de janelas do descritor MDF para realizar a descrição de pontos-chave, bem como a relação entre a resolução da imagem e os tamanhos dessas janelas;
- Estudar, ou possivelmente propor um algoritmo detector especialmente para imagens multiespectrais;
- Avaliar possíveis transformações ocorridas na cena, como variações de escala e rotação, iluminação e rotação no processo de correspondência de pontos em imagens multiespectrais.

Referências

AGUILERA, C. et al. Multispectral image feature points. **Sensors**, v. 12, n. 9, p. 12661, 2012. ISSN 1424-8220. Citado 9 vezes nas páginas 3, 13, 14, 17, 18, 20, 34, 35 e 47.

AGUILERA, C. et al. A novel sift-like-based approach for fir-vs images registration. In: **Quantitative InfraRed Thermography**. [S.l.: s.n.], 2012. Citado 3 vezes nas páginas 13, 17 e 18.

AGUILERA, C. A.; SAPPA, A. D.; TOLEDO, R. Lghd: A feature descriptor for matching across non-linear intensity variations. In: IEEE. **IEEE International Conference on Image Processing (ICIP)**. [S.l.], 2015. p. 5. Citado 12 vezes nas páginas 1, 2, 3, 9, 13, 16, 17, 18, 35, 43, 49 e 50.

ALAH, A.; ORTIZ, R.; VANDERGHEYNST, P. Freak: Fast retina keypoint. In: **IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. [S.l.: s.n.], 2012. p. 510–517. Citado 3 vezes nas páginas 4, 12 e 18.

ANTONINI, M. et al. Image coding using wavelet transform. **IEEE Transactions on image processing**, IEEE, v. 1, n. 2, p. 205–220, 1992. Citado na página 20.

BALKAYA, C. et al. Real-time identification of disaster areas by an open-access vision-based tool. **Advances in Engineering Software**, v. 88, p. 83–90, 2015. ISSN 0965-9978. Citado na página 7.

BAY, H.; TUYTELAARS, T.; GOOL, L. Surf: Speeded up robust features. In: LEONARDIS, A.; BISCHOF, H.; PINZ, A. (Ed.). **Computer Vision – ECCV 2006**. [S.l.]: Springer, 2006, (Lecture Notes in Computer Science, v. 3951). p. 404–417. Citado 4 vezes nas páginas 4, 11, 18 e 28.

BENTLEY, J. L. Multidimensional binary search trees used for associative searching. **Commun. ACM**, ACM, New York, NY, USA, v. 18, n. 9, p. 509–517, set. 1975. Citado na página 37.

BOKOLONGA, E. et al. A compact multispectral image capture unit for deployment on drones. In: IEEE. **Instrumentation and Measurement Technology Conference Proceedings (I2MTC), 2016 IEEE International**. [S.l.], 2016. p. 1–5. Citado na página 7.

BRITO, D. N. et al. Evaluation of interest point matching methods for projective reconstruction of 3d scenes. **IEEE Latin America Transactions**, IEEE, v. 14, n. 3, p. 1393–1400, 2016. Citado na página 3.

BROWN, M.; SUSSTRUNK, S. Multi-spectral sift for scene category recognition. In: **IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. [S.l.: s.n.], 2011. p. 177–184. ISSN 1063-6919. Citado 5 vezes nas páginas 13, 17, 18, 48 e 49.

CALONDER, M. et al. Brief: Binary robust independent elementary features. In: DANIILIDIS, K.; MARAGOS, P.; PARAGIOS, N. (Ed.). **Computer Vision – ECCV 2010**. [S.l.]: Springer, 2010, (Lecture Notes in Computer Science, v. 6314). p. 778–792. Citado 5 vezes nas páginas 12, 18, 29, 30 e 36.

CAMPO, F. B.; RUIZ, F. L.; SAPPA, A. D. Multimodal stereo vision system: 3d data extraction and algorithm evaluation. **IEEE Journal of Selected Topics in Signal Processing**, v. 6, n. 5, p. 437–446, set. 2012. ISSN 1932-4553. Citado na página 1.

CANNY, J. A computational approach to edge detection. **IEEE Transactions on pattern analysis and machine intelligence**, IEEE, n. 6, p. 679–698, 1986. Citado 2 vezes nas páginas 14 e 34.

CHANG, S.-F.; SIKORA, T.; PURL, A. Overview of the mpeg-7 standard. **IEEE Transactions on circuits and systems for video technology**, IEEE, v. 11, n. 6, p. 688–695, 2001. Citado na página 43.

DEDEEPPYA, P.; SANDHYA, B. Innovations in computer science and engineering: Proceedings of the third IClcse, 2015. In: _____. Singapore: Springer Singapore, 2016. cap. Image Correspondence Using Affine SIFT Flow, p. 137–144. ISBN 978-981-10-0419-3. Citado na página 1.

DELLINGER, F. et al. SAR-sift: A sift-like algorithm for SAR images. **IEEE Transactions on Geoscience and Remote Sensing**, v. 53, n. 1, p. 453–466, jan. 2015. ISSN 0196-2892. Citado 6 vezes nas páginas 7, 9, 15, 17, 18 e 42.

FIELD, D. J. Relations between the statistics of natural images and the response properties of cortical cells. **JOSA A**, Optical Society of America, v. 4, n. 12, p. 2379–2394, 1987. Citado na página 21.

FISCHLER, M. A.; BOLLES, R. C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. **Communications of the ACM**, ACM, New York, NY, USA, v. 24, n. 6, p. 381–395, jun. 1981. Citado na página 7.

GABOR, D. Theory of communication. part 1: The analysis of information. **Electrical Engineers-Part III: Radio and Communication Engineering, Journal of the Institution of**, IET, v. 93, n. 26, p. 429–441, 1946. Citado na página 20.

HARRIS, C.; STEPHENS, M. A combined corner and edge detector. In: MANCHESTER, UK. **Alvey vision conference**. [S.l.], 1988. p. 147–151. Citado 2 vezes nas páginas 15 e 31.

HEINLY, J.; DUNN, E.; FRAHM, J.-M. Comparative evaluation of binary features. In: FITZ-GIBBON, A. et al. (Ed.). **Computer Vision – ECCV 2012**. [S.l.]: Springer, 2012, (Lecture Notes in Computer Science). p. 759–773. Citado 4 vezes nas páginas 12, 13, 17 e 36.

JANNEY, P.; BOOTH, D. Pose-invariant vehicle identification in aerial electro-optical imagery. **Machine Vision and Applications**, Springer, v. 26, n. 5, p. 575–591, 2015. Citado 2 vezes nas páginas 3 e 14.

JIANG, Y.; XU, Y.; LIU, Y. Performance evaluation of feature detection and matching in stereo visual odometry. **Neurocomputing**, Elsevier, v. 120, n. 0, p. 380–390, 2013. Image Feature Detection and Description. Citado 7 vezes nas páginas 3, 12, 13, 17, 36, 37 e 38.

JIN, H. et al. A novel coarse-to-fine method for registration of multispectral images. **Infrared Physics & Technology**, v. 77, p. 219–225, 2016. ISSN 1350-4495. Citado 4 vezes nas páginas 1, 3, 14 e 49.

KANOPOULOS, N.; VASANTHAVADA, N.; BAKER, R. L. Design of an image edge detection filter using the sobel operator. **IEEE Journal of solid-state circuits**, IEEE, v. 23, n. 2, p. 358–367, 1988. Citado na página 44.

KOVESI, P. Image features from phase congruency. **Videre: Journal of computer vision research**, Citeseer, v. 1, n. 3, p. 1–26, 1999. Citado 2 vezes nas páginas 3 e 21.

KOVESI, P. What are log-gabor filters and why are they good? **School of Computer Science & Software Engineering, The University of Western Australia**, 2000. Citado na página 21.

KRISHNA, S. Quantum dots-in-a-well infrared photodetectors. **Infrared physics & technology**, Elsevier, v. 47, n. 1, p. 153–163, 2005. Citado na página 20.

KUPFER, B.; NETANYAHU, N. S.; SHIMSHONI, I. An efficient sift-based mode-seeking algorithm for sub-pixel registration of remotely sensed images. **IEEE Geoscience and Remote Sensing Letters**, v. 12, n. 2, p. 379–383, fev. 2015. ISSN 1545-598X. Citado 2 vezes nas páginas 4 e 8.

LEUTENEGGER, S.; CHLI, M.; SIEGWART, R. Brisk: Binary robust invariant scalable keypoints. In: IEEE. **IEEE International Conference on Computer Vision (ICCV)**. [S.l.], 2011. p. 2548–2555. Citado 8 vezes nas páginas 4, 8, 12, 18, 32, 33, 37 e 38.

LI, Q. et al. Multispectral image alignment with nonlinear scale-invariant keypoint and enhanced local feature matrix. **IEEE Geoscience and Remote Sensing Letters**, v. 12, n. 7, p. 1551–1555, jul. 2015. ISSN 1545-598X. Citado 2 vezes nas páginas 1 e 8.

LI, Y. et al. Robustly building keypoint mappings with global information on multispectral images. **EURASIP Journal on Advances in Signal Processing**, v. 2015, n. 1, p. 1–10, 2015. ISSN 1687-6180. Citado 5 vezes nas páginas 1, 16, 17, 18 e 49.

LI, Y. et al. Building keypoint mappings on multispectral images by a cascade of classifiers with a resurrection mechanism. **Sensors**, v. 15, n. 5, p. 11769, 2015. ISSN 1424-8220. Citado 3 vezes nas páginas 3, 7 e 49.

LI, Y. et al. Establish keypoint matches on multispectral images utilizing descriptor and global information over entire image. **Infrared Physics & Technology**, v. 76, p. 1–10, 2016. ISSN 1350-4495. Citado 6 vezes nas páginas 4, 5, 7, 16, 17 e 18.

LINDBERG, T. Image matching using generalized scale-space interest points. **Journal of Mathematical Imaging and Vision**, v. 52, n. 1, p. 3–36, 2015. ISSN 1573-7683. Citado na página 3.

LOWE, D. Distinctive image features from scale-invariant keypoints. **International Journal of Computer Vision**, Springer, v. 60, n. 2, p. 91–110, 2004. Citado 8 vezes nas páginas 1, 4, 11, 18, 24, 25, 27 e 36.

MANJUNATH, B. S. et al. Color and texture descriptors. **IEEE Transactions on circuits and systems for video technology**, IEEE, v. 11, n. 6, p. 703–715, 2001. Citado na página 43.

MIKOLAJCZYK, K.; SCHMID, C. A performance evaluation of local descriptors. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, IEEE, v. 27, n. 10, p. 1615–1630, 2005. Citado 4 vezes nas páginas 12, 18, 37 e 47.

MIKOLAJCZYK, K. et al. A comparison of affine region detectors. **International Journal of Computer Vision**, Springer, v. 65, n. 1-2, p. 43–72, 2005. ISSN 0920-5691. Citado 4 vezes nas páginas 11, 18, 37 e 38.

MISHRA, B. K. et al. Proceedings of international conference on communication, computing and virtualization (icccv) 2016 face recognition using filtered eoh-sift. **Procedia Computer Science**, v. 79, p. 543–552, 2016. ISSN 1877-0509. Citado na página 14.

MONTGOMERY, D. C. **Design and analysis of experiments**. [S.l.]: John Wiley & Sons, 2008. Citado na página 51.

MOUATS, T. et al. Multispectral stereo odometry. **IEEE Transactions on Intelligent Transportation Systems**, v. 16, n. 3, p. 1210–1224, jun. 2015. ISSN 1524-9050. Citado 2 vezes nas páginas 3 e 16.

NAVA, R.; ESCALANTE-RAMÍREZ, B.; CRISTÓBAL, G. Texture image retrieval based on log-gabor features. In: SPRINGER. **Iberoamerican Congress on Pattern Recognition**. [S.l.], 2012. p. 414–421. Citado 2 vezes nas páginas 21 e 43.

OSBORNE, R. W. et al. Design of an adaptive passive collision warning system for UAVs. **IEEE Transactions on Aerospace and Electronic Systems**, v. 47, n. 3, p. 2169–2189, jul. 2011. ISSN 0018-9251. Citado 2 vezes nas páginas 3 e 7.

QIN, Y. et al. Robust key point descriptor for multi-spectral image matching. **Journal of Systems Engineering and Electronics**, v. 25, n. 4, p. 681–687, ago. 2014. Citado 5 vezes nas páginas 7, 14, 15, 17 e 18.

RICAUARTE, P. et al. Feature point descriptors: Infrared and visible spectra. **Sensors**, v. 14, n. 2, p. 3690, 2014. ISSN 1424-8220. Citado 6 vezes nas páginas 1, 3, 15, 17, 18 e 20.

ROSTEN, E.; DRUMMOND, T. Machine learning for high-speed corner detection. In: LEONARDIS, A.; BISCHOF, H.; PINZ, A. (Ed.). **Computer Vision – ECCV 2006**. [S.l.]: Springer Berlin Heidelberg, 2006, (Lecture Notes in Computer Science, v. 3951). p. 430–443. Citado 2 vezes nas páginas 16 e 29.

RUBLEE, E. et al. ORb: An efficient alternative to sift or surf. In: IEEE. **IEEE International Conference on Computer Vision (ICCV)**. [S.l.], 2011. p. 2564–2571. Citado 4 vezes nas páginas 4, 8, 12 e 18.

SALEEM, S.; SABLATNIG, R. Image analysis and processing – IClap 2013: 17th international conference, naples, Italy, September 9-13, 2013. proceedings, part i. In: _____. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013. cap. A Modified SIFT Descriptor for Image Matching under Spectral Variations, p. 652–661. ISBN 978-3-642-41181-6. Citado 3 vezes nas páginas 7, 14 e 17.

SALEEM, S.; SABLATNIG, R. A robust sift descriptor for multispectral images. **IEEE Signal Processing Letters**, v. 21, n. 4, p. 400–403, abr. 2014. ISSN 1070-9908. Citado 7 vezes nas páginas 7, 9, 14, 17, 18, 42 e 49.

SAXENA, S.; SINGH, R. K. A survey of recent and classical image registration methods. **International Journal of Signal Processing, Image Processing and Pattern Recognition**, v. 7, n. 4, p. 167–176, 2014. Citado na página 1.

- SCHOUHAMER, K.; PÁTROVICS, L. Performance assessment of dc-free multimode codes. **IEEE transactions on communications**, IEEE, v. 45, n. 3, p. 293–299, 1997. Citado na página 21.
- SCHOWENGERDT, R. A. **Remote sensing: models and methods for image processing**. [S.l.]: Academic press, 2006. Citado na página 19.
- SIKORA, T. The mpeg-7 visual standard for content description-an overview. **IEEE Transactions on circuits and systems for video technology**, IEEE, v. 11, n. 6, p. 696–702, 2001. Citado 2 vezes nas páginas 14 e 34.
- SUN, X. et al. Enhancement of spectral resolution for remotely sensed multispectral image. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 8, n. 5, p. 2198–2211, maio 2015. ISSN 1939-1404. Citado na página 7.
- TAN, H. et al. A novel curve lane detection based on improved river flow and ransa. In: **International Conference on Intelligent Transportation Systems (ITSC)**. [S.l.: s.n.], 2014. p. 133–138. Citado na página 7.
- TORABI, A.; MASSÉ, G.; BILODEAU, G.-A. An iterative integrated framework for thermal–visible image registration, sensor fusion, and people tracking for video surveillance applications. **Computer Vision and Image Understanding**, v. 116, n. 2, p. 210–221, 2012. ISSN 1077-3142. Citado 2 vezes nas páginas 3 e 7.
- WALIA, E.; VERMA, V. Boosting local texture descriptors with log-gabor filters response for improved image retrieval. **International Journal of Multimedia Information Retrieval**, Springer, p. 1–12, 2016. Citado 4 vezes nas páginas 21, 22, 43 e 45.
- WANG, Q.; ZHU, G.; YUAN, Y. Multi-spectral dataset and its application in saliency detection. **Computer Vision and Image Understanding**, v. 117, n. 12, p. 1748–1754, 2013. ISSN 1077-3142. Citado 2 vezes nas páginas 39 e 49.
- WON, C. S.; PARK, D. K.; PARK, S.-J. Efficient use of mpeg-7 edge histogram descriptor. **Etri Journal**, Electronics and Telecommunications Research Institute, v. 24, n. 1, p. 23–30, 2002. Citado na página 43.
- WU, Y. et al. A novel point-matching algorithm based on fast sample consensus for image registration. **IEEE Geoscience and Remote Sensing Letters**, v. 12, n. 1, p. 43–47, jan. 2015. ISSN 1545-598X. Citado na página 7.
- YU, Y. et al. A novel algorithm for view and illumination invariant image matching. **IEEE Transactions on Image Processing**, IEEE, v. 21, n. 1, p. 229–240, 2012. Citado 2 vezes nas páginas 1 e 37.

Apêndices

APÊNDICE A – Parâmetros de simulação

Parâmetro	Valor	Descrição
<i>nfeatures</i>	0	Quantidade de melhores características para retornar. As características são classificadas por meio do contraste local. Nesta implementação, o valor nulo faz com que o algoritmo retorne todas as características.
<i>nOctaveLayers</i>	3	Quantidade de camadas em cada oitava. A quantidade de oitavas é calculada automaticamente a partir da resolução da imagem.
<i>contrastThreshold</i>	0,04	Limiar de contraste utilizado para filtrar as características fracas (baixo contraste). Quanto maior for este limiar, menor será a quantidade de características produzidas pelo detector.
<i>edgeThreshold</i>	3,0	Limiar utilizado para filtrar as características ao longo de uma borda. Quanto maior for este limiar, maior será a quantidade de características produzidas pelo detector.
<i>sigma</i>	1,6	O valor de σ da Gaussiana aplicada à imagem de entrada na primeira oitava. Se a imagem for capturada com uma câmera de lente desfocada, pode-se reduzir este número.

Quadro 4 – Parâmetros do algoritmo SIFT.

Parâmetro	Valor	Descrição
<i>hessianThreshold</i>	1500,0	Limiar da detecção Hessiana. Apenas os pontos com resposta Hessiana maior do que este limiar são considerados. Quanto maior for este limiar, menor será a quantidade de características produzidas pelo detector.
<i>nOctaves</i>	4	Quantidade de oitavas na pirâmide de imagens que o detector usará.
<i>nOctaveLayers</i>	3	Quantidade de camadas em cada oitava.
<i>extended</i>	<i>true</i>	Define o tipo de descritor: <i>true</i> para utilizar descritores de dimensão 128, ou <i>false</i> para utilizar descritores de dimensão 64.
<i>upright</i>	<i>false</i>	Define a característica de rotação do descritor: <i>true</i> para não calcular a orientação das características (denominado <i>Upright SURF</i> ou U-SURF), ou <i>false</i> para calcular a orientação.

Quadro 5 – Parâmetros do algoritmo SURF.

Parâmetro	Valor	Descrição
<i>nfeatures</i>	500	Quantidade máxima de características para retornar.
<i>scaleFactor</i>	1,41	Fator de escala da pirâmide de imagens. Este valor deve ser maior que 1. Um valor igual a 2 significa que cada nível da pirâmide possui 4 vezes menos pixels que o nível anterior.
<i>nlevels</i>	8	Quantidade de níveis da pirâmide de imagens.
<i>edgeThreshold</i>	31	Limiar utilizado para filtrar as características ao longo de uma borda. Quanto maior for este limiar, menor será a quantidade de características produzidas pelo detector.
<i>patchSize</i>	31	Define o tamanho do padrão de amostragem do algoritmo descritor BRIEF.

Quadro 6 – Parâmetros do algoritmo ORB.

Parâmetro	Valor	Descrição
<i>threshold</i>	64	Define o limiar da diferença de intensidade de pixels entre o pixel central do círculo definido.

Quadro 7 – Parâmetros do algoritmo FAST.

Parâmetro	Valor	Descrição
<i>thresh</i>	30	Limiar de detecção de pontos-chave do detector de cantos FAST/AGAST. Quanto maior for este limiar, menor será a quantidade de características produzidas pelo detector.
<i>octaves</i>	3	Quantidade de oitavas na pirâmide de imagens que o detector usará.
<i>patternScale</i>	2,0	Fator de escala que pode ser aplicado ao padrão utilizado para a amostragem de pontos.

Quadro 8 – Parâmetros do algoritmo BRISK.

Parâmetro	Valor	Descrição
<i>windowSize</i>	80	Define o tamanho da janela em torno do ponto-chave para realizar a descrição.

Quadro 9 – Parâmetros do algoritmo EOH.

Parâmetro	Valor	Descrição
<i>windowSize</i>	80	Define o tamanho da janela em torno do ponto-chave para realizar a descrição.

Quadro 10 – Parâmetros do algoritmo LGHD.