

PROPOSTA DE UM MÉTODO DE EXTRAÇÃO DE CARACTERÍSTICAS APLICADO AO PROBLEMA DA ESTIMAÇÃO DA POSIÇÃO DE UM VANT EM NAVEGAÇÃO AUTÔNOMA

JUAN CAMILO FONSECA-GALINDO*, ANDRÉ PAIM LEMOS*

*Programa de Pós-Graduação em Engenharia Elétrica - Universidade Federal de Minas Gerais
Av. Antônio Carlos 6627, 31270-901, Belo Horizonte, MG, Brasil

Emails: juankmilofg@ufmg.br, andrepaaim@ufmg.br

Abstract— This paper proposes a new feature extraction method with applications in autonomous aerial navigation. The proposed method is a modification of the ASIFT algorithm, reducing its time complexity. The proposed method is compared with SIFT and ASIFT for registration of aerial images taken from Unmanned aerial vehicles. The methods are compared based on the number of matched features and the time complexity. The numerical experiments the proposed modification is able to reduce the time complexity of the ASIFT algorithm with a similar accuracy.

Keywords— Autonomous system of air navigation, SIFT, ASIFT, RANSAC and INS.

Resumo— Este trabalho propõe um novo método de extração de características com aplicações em navegação aérea autônoma. O método proposto é uma modificação do algoritmo ASIFT, para reduzir a seu custo de computacional. O método proposto é comparado com SIFT e ASIFT para registro de imagens aéreas tomadas a partir veículo aéreo não tripulado. Os métodos são comparadas com base no número de características correspondentes a o tempo de execução. Os experimentos numéricos mostraram que a modificação proposta é capaz de reduzir a complexidade do algoritmo ASIFT com uma precisão semelhante.

Palavras-chave— Sistema autônomo de navegação aérea, SIFT, ASIFT, RANSAC e INS.

1 INTRODUÇÃO

Atualmente, a implementação de plataformas autônomas como veículos aéreos não tripulados (VANTs), sondas de exploração espacial e mísseis, em tarefas nas quais a intervenção direta de um ser humano não seja necessária ou se mostre inadequada por razões diversas, representa um grande desafio científico e tecnológico. A robótica, eletrônica e visão computacional são áreas em que tais desafios têm grande impacto. Recentemente, tem-se observado um aumento significativo na implementação de VANTs, em aplicações civis e militares, tais como: operações de vigilância, reconhecimento, monitoramento ambiental e agrícola, aerolevantamento, transporte de carga, ataques a alvos no solo, etc (Gonçalves et al., 2015).

Um sistema autônomo de navegação muito aplicado em VANTs é dado pela associação de um sistema inercial de navegação (*Inertial Navigation System* - INS) com dados de um sistema de posicionamento (*Global Navigation Satellite System* - GNSS). O único sistema totalmente operacional em escala mundial é o sistema americano Global Positioning System (GPS). Existem vários problemas na integridade do GPS. Se o sinal de GPS, por algum motivo, se torna indisponível ou corrompido, a solução seria estimar o deslocamento do VANT através do INS. Entretanto, os resultados dessa estimativa divergem rapidamente devido a uma acumulação da medida do viés. O sinal de GPS também se torna pouco confiável quando está operando perto de obstáculos devido às reflexões multi-caminho. Outros

problemas relacionados à confiabilidade do GPS consistem na exposição a interferências, especialmente quando o GPS estiver funcionando em frequências civis, além disso, dispositivos bloqueadores de GPS são encontrados com certa facilidade no mercado. Portanto, os VANTs que utilizam somente o sinal de GPS são vulneráveis a ações maliciosas, o que torna necessário a utilização de um sistema de navegação autônomo capaz de lidar com falhas do GPS de curto e longo prazo.

Uma solução é empregar um sistema de navegação baseado em visão computacional que forneça informação de posicionamento através de reconhecimento de marcos em imagens (*landmarks*), odometria, casamento de imagens, dentre outras metodologias existentes. Os sistemas de navegação baseados em visão computacional utilizam métodos de extração de características e métodos de casamento de imagens para estimar o posicionamento do VANT. O método de extração de características SIFT (*Scale - Invariant Feature Transform*) tem sido amplamente utilizado em aplicações em VANTs, foi testado por (Goltz and Shiguemori, 2008) em aplicações de correspondência entre imagens aéreas que simulam o imageamento de um VANT. Ademais, em (Canhoto et al., 2009), propõe-se um sistema de navegação baseado em um sistema de visão sem dependência do GPS. Esse trabalho apresenta um método para obter uma estimativa de deslocamento de um VANT usando imagens aéreas sequenciais.

Um problema presente nas aplicações de navegação autônoma é a perspectiva da imagem capturada pelo VANT, existe a necessidade de garantir

que as duas imagens estejam nos mesmos ângulos da transformada homográfica e, na prática não pode-se conseguir capturar todas as imagens em visada nadir. O método de extração de características, invariante à transformada afim é ASIFT (*Affine – SIFT*). Esse método já foi aplicado em modelagem 3D (Fernández et al., 2015), casamento de imagens sonar (Ye et al., 2015) e em fotogrametria (González et al., 2014). No entanto, o custo computacional é muito alto. Este artigo expõe uma proposta de um método de extração de características baseado em ASIFT, procurando diminuir custo computacional. Apresenta-se uma análise comparativa dos métodos SIFT, ASIFT e do método proposto, avaliando a robustez e o custo computacional, quando são apresentadas mudanças de perspectiva entre as duas imagens. Os métodos são utilizados para casar imagens de voos com imagens georeferenciadas da região de São José dos Campos, Brasil. As imagens de voo apresentam variações de perspectiva, rotação e translação. Para a calcular a robustez dos métodos, analisa-se a quantidade de pontos obtidos pelo casamento das imagens e o custo computacional, medido através do tempo de processamento.

A seção 2 detalha o problema de navegação baseado em imagens, assim como a proposta e os métodos de extração de características a serem comparadas. A seção 3 os experimentos realizados. Finalmente, a seção 4 apresenta as conclusões e propostas de trabalhos futuros.

2 Revisão de literatura

Com o objetivo de se fornecer uma visão geral de um sistema de navegação baseado em métodos de visão computacional podemos tomar como base o sistema de navegação absoluta ilustrado pela Figura 1 (Sanfourche et al., 2012). O sistema tem dois módulos principais, o módulo de processamento de imagens e o módulo do filtro de Kalman estendido. O módulo de processamento de imagens usa uma câmera óptica que captura imagens do solo. A imagem é processada utilizando métodos de extração de características, como por exemplo SIFT (Lowe, 2004), SURF (Bay et al., 2006) e ASIFT (Morel and Yu, 2009), em seguida, é casada com uma imagem satélite georeferenciada do território voado pelo VANT para indicar seu posicionamento. Técnica como RANSAC (Fischler and Bolles, 1981) é utilizada para remover *outliers* entre as duas imagens. O sistema possui alguns sensores embarcados, como por exemplo, giroscópio, acelerômetro e magnetômetro que entregam medidas inerciais para determinar o estado em termos de posição, velocidade e altitude. Essas medidas são utilizadas junto com o resultado do módulo de processamento de imagens para calcular a posição final da VANT que pode ser calculada usando o filtro Kalman estendido (Anderson and Moore, 1979).

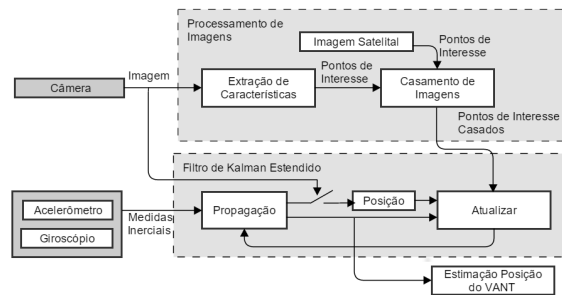


Figura 1: Sistema de navegação baseado em visão computacional. Adaptada de (Sanfourche et al., 2012).

Esta seção apresenta os métodos utilizados no estudo comparativo. São detalhados os métodos de extração de características SIFT, ASIFT e o método proposto.

2.1 SIFT

O método SIFT (Lowe, 2004) extrai características locais ou pontos de interesse de uma imagem. Os pontos de interesse são invariantes à escala, translação e rotação, além de parcialmente invariantes a mudanças de iluminação, afinidade ou projeção em 3D. Em (Lowe, 1999) defini-se que os pontos de interesse são criados para permitir deformações geométricas locais por representação de gradientes em múltiplos planos de orientação com imagens suavizadas em múltiplas escalas. Esses pontos são detectados de forma eficiente através de uma abordagem por etapas de filtragem que identifica os pontos estáveis no espaço. O método SIFT é composto pelas seguintes etapas:

1. Detecção de extremos no espaço/escala: realiza uma busca sobre todas as escalas e localizações mediante o uso da função Diferença de Gaussianas (Lowe, 1999), para identificar candidatos a pontos de interesse invariantes à escala e orientação.
2. Localização dos pontos de interesse: para cada candidato, um modelo é ajustado para determinar a localização, escala e o rádio das curvaturas principais (Harris and Stephens, 1988). Os pontos de interesse são selecionados baseados nestas medidas de estabilidade, rejeitando candidatos com baixo contraste ou localizados ao longo de uma borda.
3. Alocação de orientação: é atribuído a cada ponto de interesse uma direção de acordo com as direções do gradiente e a região que os rodeia. Todas as operações futuras são realizadas sobre os dados da imagem que foi transformada em relação à escala, localização e a orientação atribuída de cada ponto de interesse, proporcionando assim invariância a essas transformações.

4. Descritor dos pontos de interesse: os gradientes locais são medidos na escala selecionada, na região em torno ao ponto de interesse, permitindo níveis significativos de distorção de forma local e as variações de luminosidade.

2.2 ASIFT

O método ASIFT (Morel and Yu, 2009) aborda a detecção de características para a comparação de imagens totalmente invariante à transformada afim. ASIFT complementa SIFT através da simulação de dois parâmetros (θ e ϕ) do modelo da câmera óptica (Figura 2) que definem a orientação do eixo da câmera. Portanto, ASIFT simula três parâmetros da transformada afim (ângulo de longitude (ϕ), ângulo de latitude (θ) e escala (λ)) e normaliza os outros três parâmetros (rotação (ψ) e translação (x_1, y_1)).

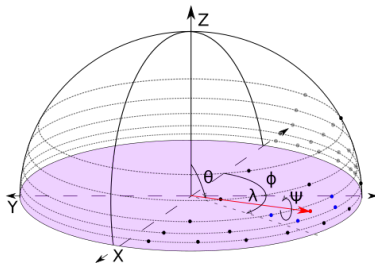


Figura 2: Interpretação geométrica dos parâmetros do modelo da câmera óptica e, representação das simulações realizadas pelo método ASIFT.

Uma interpretação geométrica da transformação afim da imagem capturada pela câmera mostrada na Figura 2 é representada pela equação (1). Onde A é a projeção da imagem frontal u_0 devido ao movimento da câmera.

$$A = \mathbf{H}_\lambda \mathbf{R}_1(\psi) \mathbf{T}_1(\theta) \mathbf{R}_2(\phi) + \mathbf{T}_2(x_1, y_1) \quad (1)$$

$\mathbf{H}_\lambda$ é a componente de escala da imagem capturada pela câmera medida pelo parâmetro λ ($\lambda > 0$). $\mathbf{R}_1$ e $\mathbf{R}_2$ pertencem as rotações do ângulo da câmera em torno do eixo óptico ($\psi \in [0^\circ, 360^\circ]$) e do ângulo entre o eixo óptico e a normal ao plano de imagem ($\phi \in [0^\circ, 180^\circ]$) respectivamente, cada rotação é representada por uma matriz de dimensão 2×2 . $\mathbf{T}_1$ é a o ângulo entre o eixo óptico e a normal ao plano de imagem ($\theta \in [0^\circ, 90^\circ]$), a qual é uma matriz diagonal com dois autovalores, no primeiro $t > 1$ ($\theta = \arccos(1/t)$) e o segundo é 1. $\mathbf{T}_2$ é o movimento inicial da câmera, onde o plano da imagem inicial u_0 e o plano da imagem capturada pela câmera estão em paralelo, representado por uma matriz de dimensão 2×1 , o primeiro termo é x_1 e o segundo é y_1 os quais representam o deslocamento no eixo x e o eixo y respectivamente. A equação (1) pode ser reescrita pela Equação (2).

Os parâmetros da posição inicial da câmera para capturar a vista frontal u_0 na equação (2) são $\lambda = 1$, $t = 1$, $\phi = \psi = 0$ e $x_1, y_1 = 0$.

$$A = \lambda \begin{bmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{bmatrix} \begin{bmatrix} t & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} + \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \quad (2)$$

ASIFT procede da seguinte forma:

1. Cada imagem é transformada pela simulação de distorções causadas pela mudança do eixo óptico da câmera a partir de uma posição frontal, utilizando a equação (2). Esta distorção depende dos ângulos de longitude e latitude. A imagem é submetida a rotações do ângulo longitude seguido pelas inclinações através do parâmetro $t = 1/|\cos\theta|$. (Morel and Yu, 2009) determinou 6 amostras do ângulo θ (n_θ) para cada uma das imagens casadas, a mudança de θ é associada com a função aritmética $1, a, a^2, \dots, a^{n_\theta-1}$, com $a = \sqrt{2}$. O número de amostras do ângulo ϕ é diferente para cada valor da série do ângulo θ , $\Delta\phi = \frac{\pi}{n_\phi}$, no qual $n_\phi = \frac{\text{Round}(10\theta)}{2}$, a função *Round* retorna um valor inteiro arredondado para o número inteiro mais próximo do valor passado como parâmetro. Os pontos apresentados na Figura 2 representam todas as simulações dos ângulos θ e ϕ . Para cada uma das imagens obtidas implementa-se o algoritmo SIFT com a finalidade de conseguir um banco de dados dos possíveis pontos de interesse.
2. O algoritmo de casamento invariante a similaridade exposto em (Lowe, 2004) fundamentada na medida da distância *City Block* descrita em (Everitt et al., 2011) é aplicado ao banco de dados dos possíveis candidatos a pontos de interesse de cada uma das duas imagens. O casamento de dois pontos de interesse de diferentes imagens é encontrado através da identificação do ponto de interesse mais próximo da segunda imagem referente a cada ponto de interesse da primeira.
3. O algoritmo de casamento invariante a similaridades geralmente deixam falsos casamentos. É importante filtrar estes casamentos devido ao grande número de *outliers* obtidos. O critério usado é que os casamentos devem ser compatíveis com uma geometria apipolar.

2.3 Método Proposto

O método de extração de características proposto neste artigo é baseado no método ASIFT, e utiliza os dados fornecidos pelo INS para diminuir a amostragem dos ângulos θ e ϕ , deste modo reduzindo o custo computacional. O INS descreve a orientação de um corpo rígido, como o VANT, girando em um espaço euclidiano tridimensional

mediante o uso dos ângulos Euler. Os ângulos são definidos no sentido positivo dos eixos do sistema do corpo da aeronave. *Roll* (ϕ) em torno do eixo X, *Pitch* (θ) em torno do eixo Y e *Yaw* (ψ) em torno do eixo Z, representados na Figura 3. Portanto, o conhecimento a priori dos ângulos *Pitch* e *Roll* é utilizado para diminuir a amostragem dos ângulos de longitude e latitude do método ASIFT. A imagem territorial só é simulada nos ângulos θ e ϕ mais próximos da orientação real do VANT no momento da captura da imagem. Enquanto, na imagem capturada pelo VANT não é realizada nenhuma simulação dos ângulos θ e ϕ com o objetivo de se garantir que o casamento seja executado em imagens com os mesmos ângulos da transformada homográfica. Na Figura 2 a localização da aeronave é interpretada pelo vetor vermelho, e a otimização proposta diminui a amostragem de ASIFT à simulação dos ângulos mais próximos representados pelos pontos azuis. Além disto, as simulações dos ângulos θ e ϕ , e o algoritmo de casamento são paralelizados.

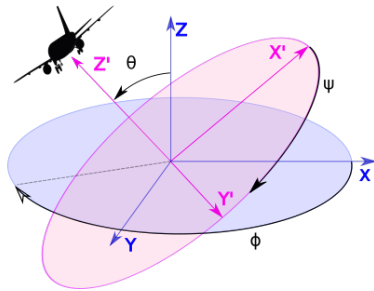


Figura 3: Conversão dos ângulos Euler.

3 Resultados Experimentais

Para o estudo comparativo foram utilizadas imagens de baixa resolução da cidade de São José dos Campos, Brasil, sendo esta mesma região contida na imagem de satélite, proveniente do satélite *Quickbird*. Utilizaram-se nove cenários diferentes com dimensões de 800x800 pixels. Para cada cenário, foram simuladas 6 mudanças do ângulo de latitude, em cada ângulo foram recortadas 36 imagens de 200x200 pixels com variações na latitude, rotação e translação. Aplicou-se um filtro gaussiano 3x3 para gerar uma imagem sintética da captura do VANTs, obtendo um total de 216 imagens para o casamento em cada cenário.

Os cenários foram divididos em três blocos, apresentados na Figura 4. O primeiro bloco foi denominado *Urbano*_{1,2,3}, este bloco tem imagens representativas do setor urbano características homogêneas, contêm muitas características, incluem quarteirões de diferentes tipos de casas delimitadas por ruas, trechos de avenidas principais e rodovias, praças esportivas, e complexos de pontes. O segundo bloco foi nomeado *Google*_{1,2,3}, em diferença com o bloco anterior, este bloco tem imagens representativas do setor urbano característi-

cas não homogêneas, incluem campos de futebol, estádios, armazéns, blocos de edifícios, lotes de construção de cimento e de pasto, diferentes tipos de casas e áreas verdes incluindo árvores. Finalmente o terceiro bloco foi nomeado *Rural*_{1,2,3}, com imagens representativas do setor rural, contêm poucas características, incluem áreas verdes com pasto, diferentes vegetações e algumas ruas não pavimentadas. *Rural*₃ inclui um rio.

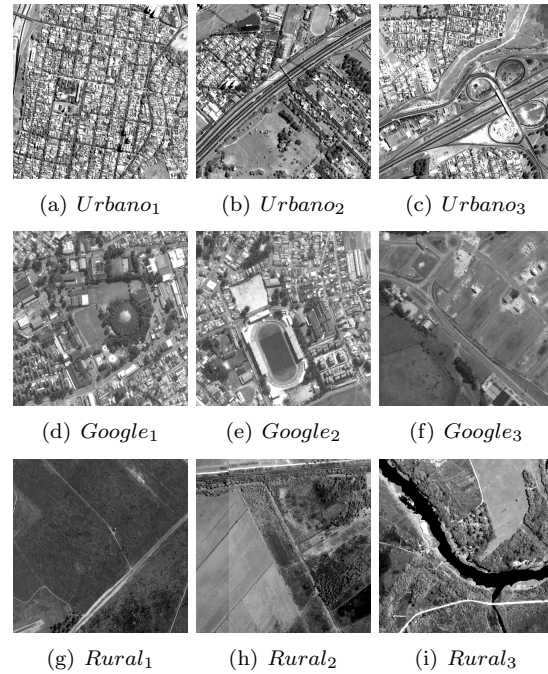


Figura 4: Imagens Satellite.

Baseado na equação 2, utilizou-se o ângulo θ como amostragem da transformada de perspectiva de cada um dos ambientes. A amostragem implementada foi de $0^\circ, 16^\circ, 32^\circ, 48^\circ, 64^\circ$ e 76° . Os valores de rotação, translação e escala foram escolhidos aleatoriamente devido a estes serem invariantes ao método base SIFT.

Para o desenvolvimento dos experimentos, foram implementados os métodos SIFT, ASIFT e o proposto, utilizando OpenCV e OpenMP. O algoritmo ASIFT e foi paralelizado para as simulações dos ângulos de θ e ϕ , para a imagem terrestre foram implementadas as três primeiras amostragens do ângulo θ por ser uma imagem satelital, com visada nadir. Ao passo que para as imagens capturadas pelo VANT foram implementadas as seis amostragens. Adicionalmente, nos três métodos utilizou-se o método RANSAC, para realizar a filtragem dos pontos *outliers*. As simulações foram feitas em um computador com Ubuntu 14.04, processador intel Xenon(R)CPU X5650 @ 2.67GHzx19 e 80 GB de memória RAM.

A Tabela 1 apresenta a média da quantidade de pontos obtidos pelos métodos SIFT (1), ASIFT (2), e o proposto (3). A quantidade de pontos obtidos é proporcional ao número de características

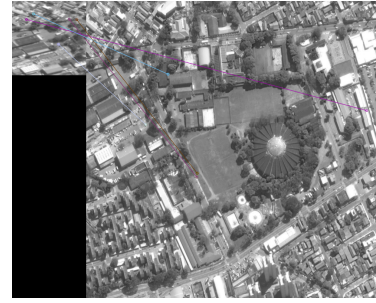
presentes nas imagens, como pode ser visto cenários *Urbanos* (a, b, c), *Google₁* (d, e, f) e *Google₂* (g, h, i) nos quais obtiveram-se mais pontos casados. O método ASIFT e o método proposto demonstraram serem robustos para imagens com perspectivas distintas, dado que esse método obteve casamento de pontos em todas as amostras do ângulo de latitude e em todos os cenários. Já o método SIFT conseguiu bons resultados para ângulos menores a 48° .

Tabela 1: Quantidade média dos pontos obtidos pelos métodos (M.) no casamento da imagem referente ao território completo (C.) e das imagens capturadas pelo VANT.

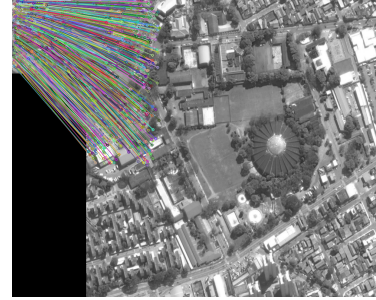
C.	M.	θ					
		0°	16°	32°	48°	64°	76°
a	1	490	477	465	331	7	7
	2	4493	4493	4318	3621	1598	33
	3	755	723	953	1020	1037	423
b	1	294	282	277	218	8	8
	2	2646	2559	2490	2337	1342	66
	3	410	391	510	567	648	294
c	1	299	287	281	207	8	7
	2	2732	2702	2611	2132	959	50
	3	451	426	569	582	605	309
d	1	211	206	196	132	6	6
	2	1600	1552	1555	1236	606	44
	3	263	254	309	297	326	150
e	1	237	228	218	142	6	6
	2	1805	1798	1691	1443	840	74
	3	290	279	345	349	390	170
f	1	62	58	60	43	7	6
	2	374	348	340	289	181	34
	3	74	70	90	87	94	54
g	1	43	39	38	30	7	6
	2	238	218	216	177	76	48
	3	69	62	80	79	64	15
h	1	170	163	158	110	7	7
	2	1283	1220	1190	910	512	39
	3	234	221	288	280	304	132
i	1	270	257	256	196	7	7
	2	2116	1948	2055	1870	754	39
	3	384	368	475	524	521	234

A Figura 5 apresenta a implementação de SIFT, ASIFT e a proposta para casamento da imagem do território *Google₁* com a image tirada pelo VANT com ângulos $\theta = 64^\circ$, $\phi = 0^\circ$ e $\psi = 0^\circ$. As linhas representam a união dos pontos casados nas duas imagens, ilustrando graficamente os resultados apresentados na Tabela 1.

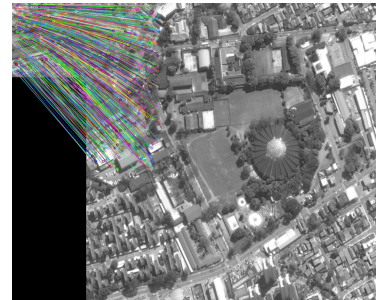
A Tabela 2, ao igual que a Tabela 1, apresenta o tempo de processamento da extração e o casamento de pontos de interesse das duas imagens para cada um dos cenários. O tempo de execução é proporcional ao número de pontos casados, porém independente em cada método. O maior tempo foi nos cenários *Urbanos*, e menor nos *Rurais* e *Google₃* devido número de caracte-



(a) SIFT



(b) ASIFT



(c) Proposta

Figura 5: Casamento da imagem do território do cenário *Google₁* com a imagem tirada pelo VANT com ângulos $\theta = 64^\circ$, $\phi = 0^\circ$ e $\psi = 0^\circ$.

rísticas presentes nas imagens. O método ASIFT apresentou o maior tempo de processamento, por ter um maior número de pontos nas duas imagens. O tempo de processamento de SIFT é o menor de todos os cenários mas só é eficiente para ângulos menores de 48° . Analisando os resultados referentes à metodologia proposta, percebe-se que esta apresenta uma boa relação custo-benefício, para o amostragem do ângulo θ .

4 Conclusões

A análise apresentada neste artigo demonstrou que a eficiência do método SIFT é limitada à imagens com ângulos θ menores na transformada homográfica. Os resultados mostraram a robustez do método ASIFT e o método proposto. Eles foram capazes de casar imagens capturadas pelo VANT com imagens georreferenciadas do território representado pela quantidade de pontos casados pelas imagens em toda a amostragem realizada do ângulo θ , e em todos os cenários, mesmo que estes tenham poucas características. Porém, uma

Tabela 2: M dia do tempo total em segundos da extra  o de pontos de interesse dos m todos (M.) do casamento da imagem referente ao territ rio completo (C.) e das imagens capturadas pelo VANT.

C	M	θ					
		0�	16�	32�	48�	64�	76�
a	1	2,18	2,2	2,21	2,26	2,28	2,22
	2	33	32,8	33,05	34,68	35,5	28,21
	3	2,44	2,44	2,99	2,71	1,95	1,33
b	1	1,65	1,65	1,67	1,69	1,77	1,72
	2	11,52	11,44	11,48	12,25	13,41	11,2
	3	1,73	1,73	2,2	2,02	1,46	1,03
c	1	1,68	1,69	1,7	1,72	1,77	1,75
	2	13,52	13,18	13,14	12,95	13,69	13,14
	3	1,85	1,85	2,34	2,12	1,48	1,05
d	1	1,45	1,5	1,49	1,49	1,56	1,53
	2	5,54	5,35	5,4	5,19	5,44	4,69
	3	1,29	1,3	1,87	1,69	1,18	0,83
e	1	1,5	1,5	1,51	1,52	1,58	1,57
	2	5,89	5,95	5,91	5,75	6,08	4,96
	3	1,31	1,3	1,89	1,73	1,21	0,83
f	1	0,83	0,83	0,84	0,83	0,89	0,89
	2	1,93	1,95	1,97	1,97	1,98	1,99
	3	0,76	0,75	1,34	1,27	0,85	0,58
g	1	0,86	0,86	0,87	0,88	0,91	0,88
	2	1,98	1,99	1,99	1,99	1,99	1,94
	3	0,85	0,85	1,43	1,31	0,85	0,57
h	1	1,26	1,25	1,26	1,27	1,32	1,3
	2	3,84	3,78	3,81	3,76	3,86	3,48
	3	1,31	1,31	1,82	1,64	1,11	0,86
i	1	1,58	1,58	1,6	1,62	1,66	1,63
	2	7,63	7,57	7,69	8,05	8,13	6,62
	3	1,64	1,64	2,13	1,93	1,32	0,9

grande desvantagem do m todo ASIFT   o custo computacional, os resultados demonstraram que o m todo n o pode ser implementado em navega  o aut noma por possuir um tempo de processamento muito alto. Por outro lado, o m todo proposto apresentou um bom custo-benef cio em rela  o   quantidade de pontos obtidos e o tempo computacional. O m todo foi eficiente nos diferentes cen rios e em toda a amostragem do  ngulo θ implementada. Entretanto, recomenda-se utilizar o m todo SIFT para  ngulos θ menores a 48  por ter uma boa quantidade de pontos casados em menor tempo, e o m todo propostos para os  ngulos superiores.

Como propostas de continuidade deste trabalho, pretende-se implementar a proposta junto com t cnicas de odometria para navega  o aut noma em VANT.

Agradecimentos

O presente trabalho foi realizado com o apoio financeiro da CAPES - Brasil.

Refer ncias

- Anderson, B. D. and Moore, J. B. (1979). Optimal filtering. 1979.
- Bay, H., Tuytelaars, T. and Van Gool, L. (2006). Surf: Speeded up robust features, *Computer vision–ECCV 2006*, Springer, pp. 404–417.
- Canhoto, A., Shiguemori, E. and Domiciano, M. (2009). Image sequence processing applied to autonomous aerial navigation, *Signal and Image Processing Applications (ICSIPA), 2009 IEEE International Conference on*, IEEE, pp. 496–499.
- Everitt, B. S., Landau, S., Leese, M. and Stahl, D. (2011). Hierarchical clustering, *Cluster Analysis, 5th Edition* pp. 71–110.
- Fern ndez, J., Gonz lez, D., Rodr guez, P. and Mancera, J. (2015). Image-based modelling from unmanned aerial vehicle (uav) photogrammetry: An effective, low-cost tool for archaeological applications, *Archaeometry* **57**: 128–145.
- Fischler, M. and Bolles, R. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Communications of the ACM* **24**(6): 381–395.
- Goltz, G. and Shiguemori, E. (2008). Aplica  o do algoritmo sift em imagens de navega  o aut noma, *Workshop Anual de Pesquisa e Desenvolvimento do IEAv*, pp. 35–35.
- Gon alves, M., Souza, C. and Shiguemori, E. (2015). Uso de coordenadas geogr ficas para estima  o autom tica de rumo aplicada a navega  o  erea aut noma de ve culos aereos n o tripulados., *Simp sio Brasileiro de Sensoriamento Remoto (SBSR)* **17**: 5598–5605.
- Gonz lez, H., Puente, I., Roca, D., Mart nez, J., Conde, B. and Arias, P. (2014). Uav photogrammetry application to the monitoring of rubble mound breakwaters, *Journal of Performance of Constructed Facilities* p. 04014194.
- Harris, C. and Stephens, M. (1988). A combined corner and edge detector., *Alvey vision conference*, Vol. 15, Citeseer, p. 50.
- Lowe, D. (1999). Object recognition from local scale-invariant features, *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, Vol. 2, Ieee, pp. 1150–1157.
- Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints, *International journal of computer vision* **60**(2): 91–110.
- Morel, J. and Yu, G. (2009). Asift: A new framework for fully affine invariant image comparison, *SIAM Journal on Imaging Sciences* **2**(2): 438–469.
- Sanfourche, M., Delaune, J., Le, G., De, H., Israel, J., Cornic, P., Treil, A., Watanabe, Y. and Plyer, A. (2012). Perception for uav: Vision-based navigation and environment modeling., *AerospaceLab* (4): p–1.
- Ye, X., Li, P. and Zhang, J. (2015). Fully affine invariant matching algorithm based on nonlinear scale space for side scan sonar image, *Mechatronics and Automation (ICMA), 2015 IEEE International Conference on*, IEEE, pp. 2387–2391.