

SISTEMA ADAPTATIVO PARA REGISTRO AUTOMÁTICO DE IMAGENS AÉREAS COM USO DO MÉTODO AKAZE

Rodrigo Augusto Rebouças¹, Gabriel Fornari², Lamartine Nogueira Frutuoso Guimarães³ e Elcio Hideiti Shiguemori⁴

¹ Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos – SP – Brasil e rodrigo.pressa@gmail.com, ²Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos – SP – Brasil e gabrielfornari@gmail.com, ³Instituto de Estudos Avançados (IEAv), São José dos Campos – SP – Brasil e elcio@ieav.cta.br, ⁴Instituto de Estudos Avançados (IEAv), São José dos Campos – SP – Brasil e guimarae@ieav.cta.br

RESUMO

O registro de imagens tem importância em várias aplicações no sensoriamento remoto. Uma dessas aplicações é a navegação de VANTs baseada em imagens, em que o registro de imagens se faz necessário para calcular a posição do veículo em tempo real. Este trabalho visa aumentar o desempenho do método através da utilização de um sistema adaptativo. A metodologia pode ser dividida em 4 partes: aquisição da imagem; quantificação da imagem com uso de medidas de textura; extração de características com uso do método de pontos de interesse AKAZE; e, por último, registro da imagem. A extração de características é o processo no qual ocorre a adaptação, sendo o parâmetro de limiar ajustado automaticamente de acordo com a quantificação. O controle da adaptação é baseado em lógica nebulosa. Resultados da metodologia mostram um ganho de taxa de acerto para registro de imagem de 40% em relação ao sistema estático, mantendo o tempo de execução.

Palavras-chave – registro de imagem, sistema adaptativo, AKAZE, lógica fuzzy.

ABSTRACT

Image registration is important in several problems concerning remote sensing. One of these problems is related to UAV navigation based on images, where image registration is used to calculate the position of the vehicle in real time. The work seeks to increase the performance of the method using an adaptive system. The methodology can be divided as follows: image acquisition; image quantification through the use of texture measures; feature extraction using AKAZE to extract interest points; and, at last, image registration. The feature extraction is the process where the adaptation occurs, being the threshold parameter automatically adjusted according to the quantification. The adaptation control is based on fuzzy logic. Results of the methodology show a 40% hit rate gain for image registration if compared to the static system, keeping the execution time.

Key words – image registration, adaptive system, AKAZE, fuzzy logic.

1. INTRODUÇÃO

O registro de imagens, tem importância em várias aplicações no sensoriamento remoto. Esse trabalho faz parte de um projeto final para realização de navegação autônoma de

VANT baseada por imagens, onde um passo essencial, é o registro de imagens. O foco do trabalho é apresentar uma abordagem baseada em inteligência artificial que perceba mudanças no ambiente e se adapte, automaticamente, as características das imagens.

A metodologia é formada por 4 partes principais: aquisição da imagem; quantificação da imagem, com uso de medidas de textura; extração de características, com uso do método de Pontos de Interesse AKAZE (*Accelerated-KAZE*) [1]; e registro de imagem. A extração de características é o processo que será adaptativo, com uso de lógica nebulosa, que de acordo com a quantificação da imagem, altera automaticamente o parâmetro de limiar, que é responsável pela quantidade de pontos característicos, que tem o maior impacto no custo computacional do sistema.

2. MATERIAIS E MÉTODOS

A metodologia do sistema de posicionamento é mostrada no fluxograma da Figura 1. O primeiro passo consiste em fazer a aquisição de pares de imagens capturadas pela aeronave durante o voo em 2 tempos distintos, t e $t + \Delta t$. Em seguida, é quantificada a imagem no tempo t com uso de texturas, para extração de informações da imagem. De acordo com essas informações, o sistema fuzzy seleciona automaticamente o subótimo parâmetro de limiar do método de Pontos de Interesse AKAZE. Em um trabalho anterior [2], com uso de Algoritmo Genético (AG), foi gerado esses valores subótimos. Em seguida, é aplicado o método de pontos de interesse, com seu parâmetro de limiar subótimo calculado pelo fuzzy, para detecção e extração de características entre os pares de imagens, com o intuito de estimar o deslocamento entre elas. Em seguida é necessário fazer a correspondência desses pontos, entretanto, alguns casamentos são identificados erroneamente. Para eliminar falsas correspondências, é aplicado o método RANSAC (RANdom SAmple Consensus) [3]. Por último, é realizado o registro de imagem para estimar a posição da aeronave.

2.1. Base de Dados

Os testes são aplicados no *dataset* [4], que contém um voo realizado na cidade de São José dos Campos-São Paulo, com uso da câmera Nikon Coolpix A, com diferentes altitudes. O voo consiste em um conjunto de 867 imagens, contendo áreas de floresta, lago, urbana, vegetação rasteira e mista. A câmera está com visada nadir, ou seja, a câmera aponta para baixo e com uso de uma plataforma *gimbal* para garantir um nivelamento, com isso, diminui erros de projeção de imagens.

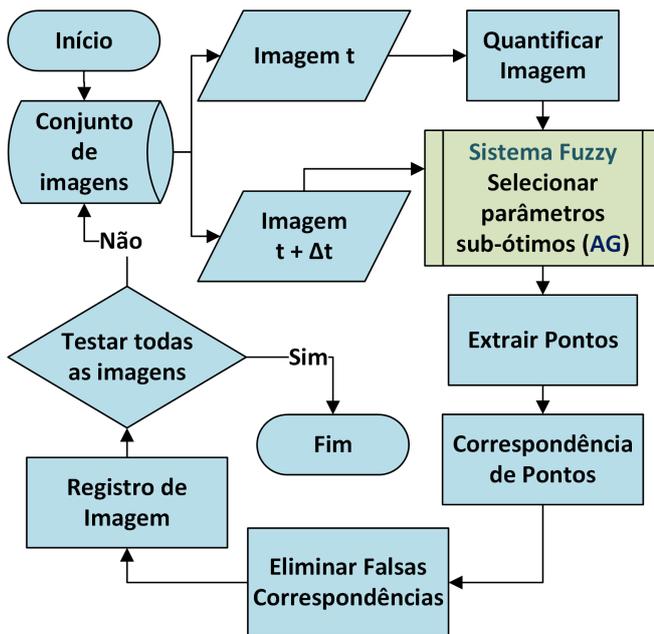


Figura 1: Metodologia adaptativa para posicionamento do VANT.

Para configurar os parâmetros com o método de otimização e criar as regras de inferência, são selecionados 22 pares distintos de imagens com diferentes texturas e altitudes da aeronave do *dataset*, como as apresentadas na Figura 2.

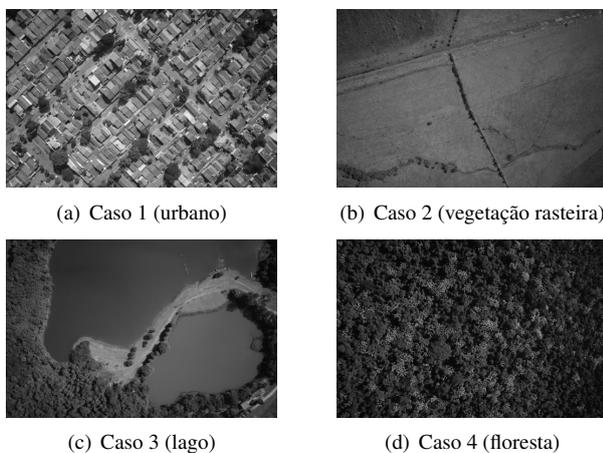


Figura 2: Exemplo de imagens de testes usadas para configuração dos parâmetros.

2.2. Quantificar Imagem

Para quantificar a descrição de uma região, uma abordagem utilizada é a textura pois, com seu uso obtêm-se medidas de propriedades como suavidade, rugosidade e regularidade, entre outras [5]. Para extração de características de texturas, há um método baseado na relação espacial existentes entre os níveis de cinza da imagem, com uso de matrizes de co-ocorrência, onde determinaram 14 descritores estatísticos [6]. Dentre eles, existem seis com maior relevância [7, 8]: energia; entropia; contraste; heterogeneidade; correlação; e homogeneidade.

2.3. Sistema Fuzzy

Para desenvolvimento do sistema adaptativo, faz-se uso da biblioteca FuzzyLite para controlar a lógica nebulosa (*fuzzy*) do sistema. A configuração da lógica nebulosa, é feita com variáveis de entrada e saída, ilustrada na Figura 3. As variáveis de entrada são as medidas de textura contraste e heterogeneidade. A variável de saída é o limiar, parâmetro do AKAZE, pois é o atributo do método de Ponto de Interesse que tem a maior relevância [9], discutidas na seção 4.

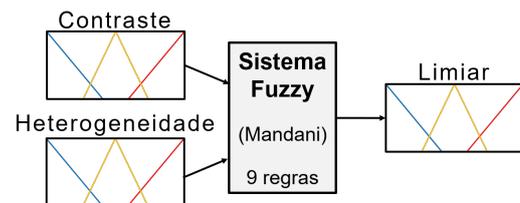


Figura 3: Lógica Nebulosa do sistema adaptativo.

A configuração das variáveis de entrada foi feita da seguinte maneira: o contraste com intervalo de 40 à 1300; a heterogeneidade com intervalo de 250 à 4700. E a variável de saída, o limiar, teve a seguinte configuração:

- Inferência Fuzzy: Mamdani;
- Intervalo: 0,0001 à 0,01;
- Agregação: Soma normalizada;
- Defuzzificação: Centro de área;
- Função de Pertinência: Triangular.

As regras de inferência foram definidas de acordo com uma análise prévia de todo o conjunto de imagens, dividido em termos linguístico, apresentados na Tabela 1.

Tabela 1: Regras de inferência.

	Se	Então
Contraste	Heterogeneidade	Limiar
Baixo	Baixo	Baixo
Baixo	Médio	Médio
Baixo	Alto	Médio
Médio	Baixo	Baixo
Médio	Médio	Médio
Médio	Alto	Alto
Alto	Baixo	Médio
Alto	Médio	Médio
Alto	Alto	Alto

2.4. Pontos de Interesse

O método de Pontos de Interesse tem importância para extrair características nas imagens. Por padrão é dividido em extração de pontos de interesse, correspondência de pontos e eliminação de falsas correspondências [10].

2.4.1. Extrair Pontos

Para extrair pontos, na presente metodologia é aplicada o método de Pontos de Interesse AKAZE (Accelerated-KAZE) [1], pois comparado com outros métodos, como SIFT (Scale Invariant Feature Transform) [11], SURF (Speeded Up Robust Features) [12] e ORB (Oriented FAST and Rotated BRIEF) [13], apresenta um baixo custo computacional e bom desempenho [9]. O AKAZE é uma versão acelerada do KAZE e é baseado em gradiente a partir do espaço de escala não linear, em que utiliza uma descrição binária baseada em uma versão modificada do método LDB (Binary Local Difference) [14], onde gera um vetor de características binária para cada ponto de interesse.

2.4.2. Correspondência de Pontos

Após a extração dos pontos. Com uso do método de busca Força Bruta com KNN (*K-NearestNeighbor*) é realizada a correspondência desses pontos, através da distância de *Hamming* entre os vetores de cada ponto. Os pontos casados, podem ser muito parecidos, e com isso gerar correspondências falsas. Para eliminar isso, na literatura [10], faz uso do método RANSAC (RANdom SAMple Consensus).

2.4.3. Eliminar Falsas Correspondências

O RANSAC (RANdom SAMple Consensus) é um método iterativo para estimar os parâmetros de um modelo matemático, onde tenta utilizar a menor quantidade de pontos possíveis para estimar um modelo [3]. Na metodologia, faz uso das coordenadas dos pontos correspondidos para estimar a matriz de homografia H , uma matriz 3×3 não singular que contém a transformação projetiva do conjunto de pontos $p = (x, y)$ das imagens no tempo t que correspondem com o conjunto de pontos $p' = (x', y')$ das imagens no tempo $t + \Delta t$, como mostrado na Fórmula 1.

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (1)$$

Com isso os pontos que não fazem parte da matriz H , são eliminados e os que se enquadram no modelo são considerados *inliers*.

2.5. Registro de Imagem

O último passo da metodologia é realizar o registro de imagens, para isso, faz-se uso dos *inliers* e da matriz de homografia, onde é possível estimar a posição da imagem no tempo t , correspondida na imagem no tempo $t + \Delta t$, como apresentado na Figura 4.

3. RESULTADOS

O sistema adaptativo é comparado com os parâmetros de Pontos de Interesse padrão da biblioteca e com um valor baixo de limiar, configurado manualmente, a fim de buscar melhor precisão. Na Tabela 2 é apresentada a comparação da média e desvio padrão, em segundos, de todo o *dataset* dos tempos de processamento de execução da metodologia, exceto



Figura 4: Exemplo de registro de imagem.

a parte da aquisição da imagem, e sua respectiva taxa de acerto, que é calculada através da relação entre a quantidade de imagens que o algoritmo encontra a matriz de homografia pela quantidade de imagens totais.

Tabela 2: Comparação do tempo de processamento e taxa de acerto respectivos ao *Dataset*.

Parâmetros	Tempo (s)	Taxa de Acerto (%)
Fuzzy	0,11 ± 0,06	95,4
Padrão	0,10 ± 0,05	68,4
Manual	0,25 ± 0,04	98,5

Para execução do Fuzzy é necessário extrair a textura da imagem, cujo custo computacional é de $0,022 \pm 0,003$ segundos. O tempo tem relevância para a conclusão do trabalho final, que será executado em hardware embarcado, entretanto os testes de comparação, são executados em um Intel Core i7-6500U com 2,5GHz e 8GB de memória RAM.

Na Tabela 3 é apresentada a comparação da quantidade de pontos de interesse, a quantidade de casamentos e a quantidade de *inliers*, essa quantidade tem a importância de ver a relação do número de pontos pelo tempo, apresentada na Tabela 2.

Tabela 3: Comparação quantidade dos pontos de interesse.

Método	Qtd. Pontos	Qtd. Casamentos	Qtd. Inliers
Fuzzy	181 ± 137	27 ± 25	28 ± 18
Padrão	390 ± 414	105 ± 147	95 ± 138
Manual	3199 ± 432	665 ± 535	553 ± 518

4. DISCUSSÃO

Para o sistema adaptativo foram aplicadas 2 medidas de texturas, o contraste, que mede o espalhamento da textura na imagem, é a diferença de intensidade entre um pixel e seu vizinho na imagem e a heterogeneidade, que mede a variância da textura na imagem. Essas medidas foram escolhidas, pois é possível representar as características das imagens e dividi-las em cenários, como mostrado no mapa de calor na Figura 5 as medidas de energia, a entropia, o contraste, a variância, a correlação e a homogeneidade, e o parâmetro limiar do AKAZE. Em um trabalho anterior [2], foi configurado com uso do algoritmo genético, para gerar uma solução subótima o valor de parâmetro de limiar com menor erro e tempo, onde o número de *inliers* varia-se entre 16 à 32 pontos, para garantir a correspondência entre as imagens, pois um valor muito baixo, poderia gerar um erro.



Figura 5: Mapa de calor das texturas e o limiar.

5. CONCLUSÕES

O presente trabalho demonstrou que o sistema adaptativo baseado em lógica nebulosa teve um ganho em termos de esforço computacional de $\approx 40\%$ quando comparado a um parâmetro configurado manualmente e uma diminuição de $\approx 3\%$ em relação a taxa de acerto. Quando comparado ao parâmetro padrão manteve o custo computacional e obteve um ganho na taxa de acerto de $\approx 27\%$.

Dessa forma, pode-se dizer que o sistema adaptativo é uma alternativa válida para realizar registro de imagens, em sistemas embarcados com limitados recursos computacionais, que exigem processamento em tempo real, entretanto é necessário realizar mais testes, para verificação da

confiabilidade da metodologia e a possibilidade de embarcar o sistema.

6. REFERÊNCIAS

- [1] ALCANTARILLA, P. F.; NUEVO, J.; BARTOLI, A. Fast explicit diffusion for accelerated features in nonlinear scale spaces. In: *British Machine Vision Conf. (BMVC)*. [S.l.: s.n.], 2013.
- [2] REBOUÇAS, R. A. et al. Parameters configuration for interest points in images obtained by drone. In: *CEUR Workshop Proceedings*. [S.l.: s.n.], 2018.
- [3] 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, v. 24, n. 6, p. 381–395, 1981.
- [4] ROBERTO, L. *Acurácia do posicionamento e da orientação espacial de veículos aéreos a partir de imagens de câmeras de pequeno formato embarcadas*. Dissertação (Mestrado em Sensoriamento Remoto — Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2017. Disponível em: <<http://urlib.net/sid.inpe.br/mtc-m21b/2017/11.03.11.55>>. Acesso em: 25 mar. 2018.
- [5] GONZALEZ, R. C.; WOODS, R. C. *Processamento digital de imagens*. [S.l.]: Pearson Educación, 2009.
- [6] HARALICK, R. M.; SHANMUGAM, K. et al. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, Ieee, n. 6, p. 610–621, 1973.
- [7] BARALDI, A.; PARMIGGIANI, F. An investigation of the textural characteristics associated with gray level cooccurrence matrix statistical parameters. *IEEE Transactions on Geoscience and Remote Sensing*, v. 33, n. 2, p. 293–304, 1995.
- [8] HALL-BEYER, M. Gcm texture: A tutorial v. 3.0 march 2017. 03 2017.
- [9] ROOS, D. R. *Aprendizado de Máquina Aplicado à Odometria Visual para Estimção de Posição de Veículos Aéreos Não Tripulados*. Dissertação (Mestrado) — Universidade Federal de São Paulo - UNIFESP, São José dos Campos, 2018.
- [10] FRAUNDORFER, F.; SCARAMUZZA, D. Visual odometry: Part ii: Matching, robustness, optimization, and applications. *IEEE Robotics & Automation Magazine*, IEEE, v. 19, n. 2, p. 78–90, 2012.
- [11] LOWE, D. G. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, Springer, v. 60, n. 2, p. 91–110, 2004.
- [12] BAY, H.; TUYTELAARS, T.; GOOL, L. V. Surf: Speeded up robust features. *Computer vision—ECCV 2006*, Springer, p. 404–417, 2006.
- [13] RUBLEE, E. et al. Orb: An efficient alternative to sift or surf. In: IEEE. *Computer Vision (ICCV), 2011 IEEE International Conference on*. [S.l.], 2011. p. 2564–2571.
- [14] YANG, X.; CHENG, K.-T. Ldb: An ultra-fast feature for scalable augmented reality on mobile devices. In: IEEE. *2012 IEEE international symposium on mixed and augmented reality (ISMAR)*. [S.l.], 2012. p. 49–57.