

# Parameters configuration for interest points in images obtained by Drone

Rodrigo A. Rebouças<sup>1</sup>, Elcio H. Shiguemori<sup>2</sup>, Lamartine N. F. Guimarães<sup>2</sup>

<sup>1</sup> National Institute for Space Research (INPE)  
São José dos Campos – SP – Brazil

<sup>2</sup>Institute of Advanced Studies (IEAv - DCTA)  
São José dos Campos – SP – Brazil

rodrigo.pressa@gmail.com, {elcio, guimarae}@ieav.cta.br

**Abstract.** *Drone use has grown with the use of image processing and computer vision techniques, such as autonomous image navigation, mosaic generation, elevation modeling, 3D reconstruction, and object recognition. In all techniques, an important step is an extraction of features, such as methods of interest points. This work addresses the modes of application of interest points, such as BRISK, ORB, FREAK, AKAZE and LATCH with the parameters configured automatically using the optimization method for images with different textures. This process is one of the pieces of final software that selects the use of a meta heuristic the best parameters automatically according to an input image.*

**Resumo.** *O uso de Drone tem crescido com o uso de técnicas de processamento de imagens e visão computacional, como a navegação autônoma por imagens, geração de mosaico, modelo de elevação, reconstrução 3D e reconhecimento de objetos. Em todas essas técnicas, um passo importante é a extração de características, como os métodos de pontos de interesse. Este trabalho aborda a comparação dos métodos de pontos de interesse BRISK, ORB, FREAK, AKAZE e LATCH com os parâmetros configurados automaticamente com uso de método de otimização para imagens com diferentes texturas. Esse processo é uma das partes de um software final que escolherá com uso de uma meta heurística os melhores parâmetros automaticamente de acordo com a imagem de entrada.*

## 1. Introdução

O uso de Veículo Aéreo Não Tripulado (VANT), mais conhecido como Drone tem crescido atualmente, com isso, aplicações nas áreas de desmatamento [Zhang et al. 2017], sistemas de vigilância [Nijim 2016], agricultura [Otake 2017] e pecuária [Abade et al. 2016]. As técnicas de processamento de imagens e visão computacional, são uma importante ferramenta para se interagir com os sistemas de VANT, seja para navegação autônoma por imagens [Kanellakis and Nikolakopoulos 2017], ou para geração de mosaico [Zhen et al. 2016], modelo de elevação [Domiciano et al. 2018], reconstrução 3D [Shang and Shen 2017] e reconhecimento de objetos [Radovic et al. 2017]. Em todas essas técnicas, um passo importante é a extração de características nas imagens obtidas pela aeronave. Existem diferentes métodos para a extração de características representativas em imagens [Li et al. 2015]. Neste trabalho, é abordado os métodos de Pontos de Interesse. Esses

métodos extraem características locais em imagens e são divididos em 2 partes: o detector, que identifica pontos representativos; e o descritor, que descrevem as características desses pontos [Bradski and Kaehler 2008].

Para o uso dos métodos de pontos de interesse em um sistema embarcado em tempo real, a análise do custo computacional é fundamental, principalmente para o sistema funcionar durante o voo. Para reduzir o tempo, uma das abordagens é alterar os parâmetros do método para gerar uma menor quantidade de pontos para representar a imagem com menor erro e tempo, para isso, uma das técnicas é o uso de método de otimização. Existem na literatura diversos métodos de Pontos de Interesse, entretanto, por se tratar de um sistema em tempo real, os binários são mais indicados, pois possuem um custo computacional menor, BRISK [Leutenegger et al. 2011], ORB [Rublee et al. 2011], FREAK [Alahi et al. 2012], AKAZE [Alcantarilla et al. 2013] e LATCH [Levi and Hassner 2016].

No estado da arte para comparação de Pontos de Interesse [Satnik et al. 2016] [Pusztai and Hajder 2016] [Bureš and Müller 2016] [Cowan et al. 2016] [Wu et al. 2017] faz-se uso de parâmetros padrões utilizados pelas bibliotecas, como do OpenCV [Bradski and Kaehler 2008], entretanto dependendo da imagem, se não configurar os parâmetros, o resultado retornado pode não ser esperado, com uso de método de otimização pretende-se obter resultados de comparações mais justos, pois os parâmetros serão alterados individualmente para cada imagem de teste, buscando uma solução subótima dos parâmetros para obter um erro menor. Essa comparação faz parte de um trabalho em andamento, onde esses parâmetros serão configurados automaticamente de acordo com uma meta-heurística para um sistema de posicionamento da aeronave em tempo real.

## 2. Metodologia

Os procedimentos da metodologia são ilustrados na Figura 1, onde apresenta-se os passos do método de otimização, que faz uso inicialmente do Algoritmo Genético (AG) para testes iniciais, para encontrar uma solução subótima, a fim de escolher os melhores parâmetros dos métodos de Pontos de Interesse para diferentes imagens.

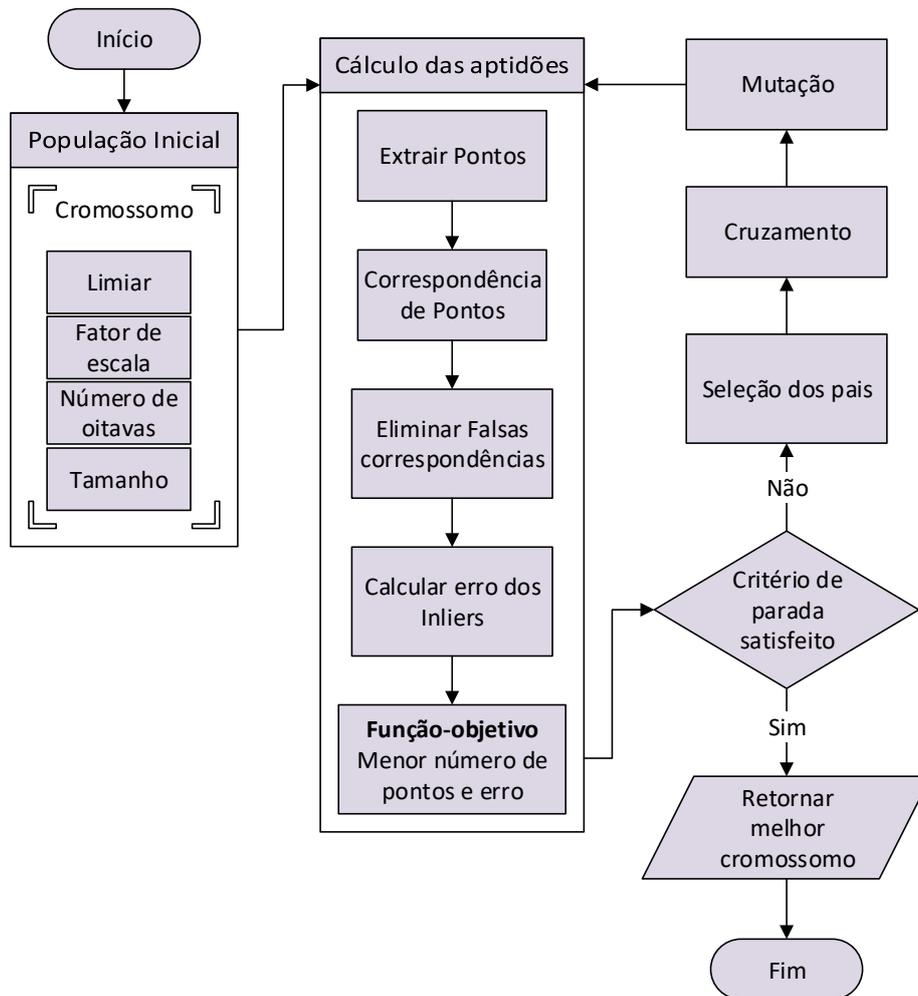
O primeiro passo é fazer a aquisição de pares de imagens capturadas pelo VANT durante o voo nos tempos  $t$  e  $t + \Delta t$ . O algoritmo começa com a inicialização da população, formada por um conjunto de cromossomos, com os seguintes genes:

- Limiar para detecção do ponto de interesse;
- Número de oitavas, representa o número níveis de escala;
- Fator de escala, representa a relação da vizinhança de um ponto para os níveis de escala;
- Tamanho do descritor.

Em seguida, realiza-se o cálculo das aptidões de cada cromossomo de acordo com uma função de custo. Caso o critério de parada seja satisfeito, o algoritmo retorna o cromossomo com o menor custo calculado, caso contrário, pode ser realizado a seleção, o cruzamento e a mutação na população, onde esse processo se repete para  $N$  gerações.

### 2.1. Base de Dados

Os testes são aplicados no *dataset* [Roberto 2017], que contém um voo realizado na cidade de São José dos Campos-São Paulo, com uso da câmera Nikon Coolpix A, com



**Figura 1. Fluxograma do Algoritmo Genético.**

diferentes altitudes. 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. Para configurar os parâmetros com o método de otimização, é selecionado 6 pares distintos de imagens com diferentes texturas e altitudes da aeronave do *dataset*, apresentadas na Figura 2.

## 2.2. Cálculo dos Aptidões

Para realizar o cálculo, faz uso de um dos métodos de Pontos de Interesse para detecção de pontos no par de imagens, que podem ser cantos, quinas ou bordas, chamados detectores. Em seguida, são extraídos vetores de características para cada ponto, utilizando seus pixels vizinhos, chamado descritores. Para fazer a correspondência dos pontos nos pares de imagens, utiliza o método de busca Força Bruta com KNN, onde para cada ponto busca outros  $k$  ( $k=2$ ) pontos mais próximos a ele, utilizando a distancia de *Hamming* entre os vetores de cada ponto, encontrados pelo descritor. Entretanto alguns casamentos podem ser falsos, para eliminar falsas correspondências, no estado da arte, uma das técnicas que pode



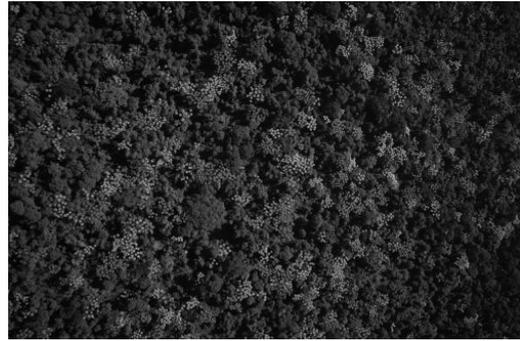
(a) Caso 1



(b) Caso 2



(c) Caso 3



(d) Caso 4



(e) Caso 5



(f) Caso 6

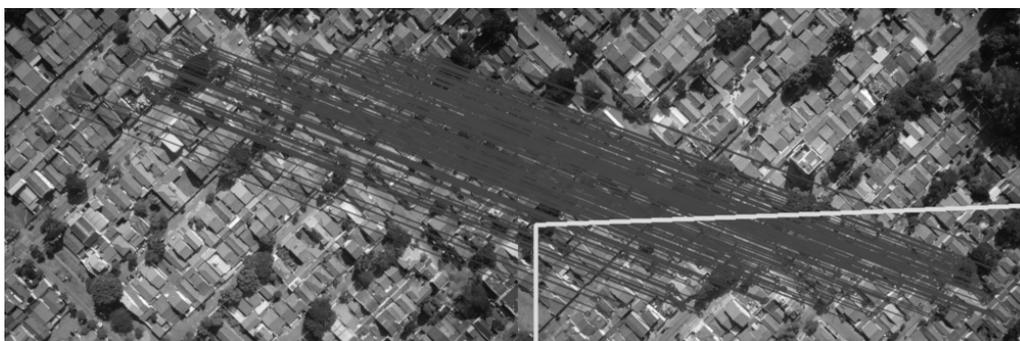
**Figura 2. Imagens de testes para configurar os parâmetros.**

ser aplicada é o RANSAC (RANDOM SAMPLE CONSENSUS) [Fischler and Bolles 1981]. Para o casamento de imagens, o RANSAC faz-se uso das coordenadas dos pontos casados para estimar a matriz de homografia  $H$ , uma matriz  $3 \times 3$  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  $t + \Delta t$ , obtido por:

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

Esse processo é mostrado na Figura 3, onde as linhas escuras representam as correspondências entre os pontos de interesse já filtrados pelo RANSAC e as linhas claras a

transformação projetiva dos pontos.



**Figura 3. Exemplo de casamento de imagens.**

Os pontos que não fazem parte da matriz  $H$ , são eliminados e considerados falsos casamentos, onde o erro  $e$  tende a aumentar com a quantidade de pontos correspondidos incorretamente. O erro pode ser obtido por:

$$e = \sum_{i=0}^m \sum_{j=0}^m \sqrt{V_{i,j}^2 - H_{i,j}^2} \quad (2)$$

Onde,  $m$  é o tamanho das matrizes  $H$  e  $V$ ,  $i$  e  $j$  os índices das matrizes e  $V$  representa a matriz de homografia desejada, obtida através da mediana das matrizes  $H$  de todos métodos de pontos de interesse testados e configurados com os parâmetros de limiar no mínimo e o tamanho no máximo, a fim de obter o maior número de casamentos corretos, pois quanto menor o limiar, maior a quantidade de pontos encontradas e o tamanho implica na quantidade de pixels armazenadas no vetor do descritor, entretanto em um sistema embarcado em tempo real, essa configuração geralmente não é esperada, devido ao custo computacional.

A função objetivo é calculada com uso do erro  $e$  multiplicado pelo tempo para fazer o processo do cálculo das aptidões. Todo o processo é realizado para cada par de imagens da Figura 2 e para cada método de Ponto de Interesse, a fim de obter como saída do GA os parâmetros subótimos em relação ao erro e o tempo de processamento de cada método.

### 2.3. Pontos de Interesse

Toda a metodologia é aplicada para os seguintes métodos de Pontos de Interesse:

- ORB (Oriented FAST and Rotated BRIEF) [Rublee et al. 2011] é baseado no detector FAST (Features from Accelerated Segment Test) e no descritor BRIEF (Binary Robust Independent Elementary Features);
- BRISK (Binary Robust Invariant Scalable Keypoints) [Leutenegger et al. 2011], faz uso do detector AGAST (Adaptive and Generic Accelerated Segment Test) e de um descritor semelhante ao BRIEF, entretanto considera a rotação de um ponto a ser descrito;
- FREAK (Fast Retina Keypoint) [Alahi et al. 2012] é um descritor baseado na retina do sistema de visão humano, similar ao BRISK, porém sua amostragem não

é simétrica, na medida em que se aproxima ao centro do padrão, a densidade de amostragem aumenta, gerando círculos sobrepostos de tamanhos variados;

- AKAZE(Accelerated-KAZE) [Alcantarilla et al. 2013] é uma versão acelerada do KAZE, é baseado em gradiente a partir do espaço de escala não linear, onde utiliza uma descrição binária baseada em uma versão modificada do método LDB (Binary Local Difference) [Yang and Cheng 2012];
- LATCH (Learned Arrangements of Three Patch Codes) [Levi and Hassner 2016] faz uso de comparação de blocos (*patches*), ao invés de pixels e faz uso de trigêmeos em vez de pares de pontos de amostragem.

## 2.4. Parametrização

A parametrização do AG é definida por:

- Cromossomo: 4 atributos;
  - Limiar: 10 à 100;
  - Fator de escala: 1.1 à 2.9;
  - Número de oitavas: 2 à 8;
  - Tamanho: 2, 4, 8, 16, 32, 64.
- População: 1000;
- Gerações: 100;
- Torneios: 50;
- Taxa de Seleção: 98%;
- Taxa de Mutação 30%;
- Taxa de Cruzamento 70%.

## 3. Resultados e Discussão

Na Tabela 1 apresenta a comparação dos erros das matrizes de homografia dos métodos de Pontos de Interesse, com uso da biblioteca opencv [Bradski and Kaehler 2008], utilizando os parâmetros padrões (PD) da biblioteca e com os parâmetros obtidos como resposta do Algoritmo Genético (AG).

**Tabela 1. Comparação dos erros.**

Caso	Método	ORB	BRISK	FREAK	AKAZE	LATCH
1	PD	333,99	4,55	8,13	6,88	-
	AG	0,40	0,30	0,72	0,27	0,32
2	PD	45,20	11,77	15,93	11,73	77,52
	AG	0,39	0,90	0,59	0,59	0,07
3	PD	127,71	123,95	107,87	80,63	369,277
	AG	6,86	1,95	8,38	1,27	6,16
4	PD	433,05	44,02	32,19	40,97	-
	AG	32,18	7,16	1,88	4,74	1,86
5	PD	2,47	1,15	1,45	0,91	4,25
	AG	0,13	0,06	0,08	0,06	0,06
6	PD	4,01	0,76	0,26	0,98	3,52
	AG	0,17	0,17	0,08	0,09	0,07

Pode-se observar que em todos os casos o AG diminuiu o erro e onde os erros com os parâmetros padrões são maiores que 100, por terem muitos casamentos falsos e com isso, encontra-se uma matriz  $H$  errada e onde o erro é “-”, significa que não foi possível encontrar casamentos suficientes para formar a matriz  $H$ , em ambos os casos o AG encontrou a matriz correta.

Na Tabela 2 apresenta a comparação dos tempos dos métodos de Pontos de Interesse, com uso dos parâmetros padrões (PD) e com uso dos parâmetros obtidos pelo Algoritmo Genético (AG).

**Tabela 2. Comparação dos tempos em segundos.**

Caso	Método	ORB	BRISK	FREAK	AKAZE	LATCH
1	PD	0,04	1,14	0,83	0,17	0,07
	AG	0,12	0,07	0,09	0,14	0,31
2	PD	0,02	0,08	0,10	0,10	0,06
	AG	0,14	0,05	0,06	0,09	0,21
3	PD	0,02	0,14	0,14	0,09	0,06
	AG	0,05	0,04	0,06	0,17	0,06
4	PD	0,03	1,06	0,85	0,11	0,06
	AG	0,47	0,04	0,10	0,09	0,86
5	PD	0,03	1,13	0,85	0,16	0,07
	AG	0,17	0,02	0,06	0,09	0,31
6	PD	0,04	2,6	1,9	0,99	0,07
	AG	0,14	0,04	0,11	0,16	0,48

Pode-se observar que o AG na maioria dos casos diminuiu o tempo, contudo, onde não teve melhora no tempo, teve um ganho significativo no erro, como apresentado na Tabela 1. Para os testes iniciais, o algoritmo foi executado em um Intel Core i7-6500U com 2,5GHz e 8GB de memória RAM. O tempo tem importância para a continuação do trabalho, pois será executado em hardware embarcado.

#### 4. Conclusões

Como apresentado nos resultados, pode-se observar que teve uma melhora no tempo e do erro com uso do AG, ao invés de utilizar os parâmetros padrões dos métodos. No estado da arte o BRISK em comparação com os outros binários, tem um custo maior, pois por padrão seus parâmetros são ajustados para encontrar mais pontos, entretanto dentre os métodos testados, na maioria dos casos, o BRISK teve um custo computacional menor com uso do AG. E o LATCH apresentou um erro menor, contudo, é o que obteve um custo computacional maior. Os próximos passos é fazer uso do método de otimização para criar regras, que serão utilizadas em uma meta heurística, para de acordo com uma imagem de entrada, o software escolherá os melhores parâmetros automaticamente.

#### Referências

Abade, A., de Campos, M. D., Porto, L. F., de Farias Coelho, Y., de Moura Sousa, Y., and Nespolo, J. P. (2016). A construção otimizada de um drone para aplicações na agricultura e pecuária de precisão. *Anais da Escola Regional de Informática da Sociedade Brasileira de Computação (SBC)–Regional de Mato Grosso*, 7.

- Alahi, A., Ortiz, R., and Vanderghenst, P. (2012). Freak: Fast retina keypoint. In *Computer vision and pattern recognition (CVPR), 2012 IEEE conference on*, pages 510–517. Ieee.
- Alcantarilla, P. F., Nuevo, J., and Bartoli, A. (2013). Fast explicit diffusion for accelerated features in nonlinear scale spaces. In *British Machine Vision Conf. (BMVC)*.
- Bradski, G. and Kaehler, A. (2008). *Learning OpenCV: Computer vision with the OpenCV library*. "O'Reilly Media, Inc."
- Bureš, L. and Müller, L. (2016). Selecting keypoint detector and descriptor combination for augmented reality application. In *International Conference on Speech and Computer*, pages 604–612. Springer.
- Cowan, B., Imanberdiyev, N., Fu, C., Dong, Y., and Kayacan, E. (2016). A performance evaluation of detectors and descriptors for uav visual tracking. In *ICARCV*, pages 1–6.
- Domiciano, M. A. P., Shiguemori, E. H., Dias, L. A. V., and da Cunha, A. M. (2018). Particle collision algorithm applied to automatic estimation of digital elevation model from images captured by uav. *IEEE Geoscience and Remote Sensing Letters*, (99).
- Fischler, M. A. and Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395.
- Kanellakis, C. and Nikolakopoulos, G. (2017). Survey on computer vision for uavs: Current developments and trends. *Journal of Intelligent & Robotic Systems*, 87(1):141–168.
- Leutenegger, S., Chli, M., and Siegwart, R. Y. (2011). Brisk: Binary robust invariant scalable keypoints. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 2548–2555. IEEE.
- Levi, G. and Hassner, T. (2016). Latch: Learned arrangements of three patch codes. IEEE.
- Li, Y., Wang, S., Tian, Q., and Ding, X. (2015). A survey of recent advances in visual feature detection. *Neurocomputing*, 149:736–751.
- Nijim, M. (2016). Multitasking intelligent surveillance and first response system. In *Technologies for Homeland Security (HST), 2016 IEEE Symposium on*, pages 1–6, Kingsville, USA. IEEE.
- Otake, V. S. (2017). Produtos cartográficos gerados a partir de drones e aplicações na agricultura.
- Pusztai, Z. and Hajder, L. (2016). Quantitative comparison of feature matchers implemented in opencv.
- Radovic, M., Adarkwa, O., and Wang, Q. (2017). Object recognition in aerial images using convolutional neural networks. *Journal of Imaging*, 3(2):21.
- Roberto, L. (2017). Acurácia do posicionamento e da orientação espacial de veículos aéreos a partir de imagens de câmeras de pequeno formato embarcadas.
- Rublee, E., Rabaud, V., Konolige, K., and Bradski, G. (2011). Orb: An efficient alternative to sift or surf. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 2564–2571. IEEE.

- Satnik, A., Hudec, R., Kamencay, P., Hlubik, J., and Benco, M. (2016). A comparison of key-point descriptors for the stereo matching algorithm. In *Radioelektronika (RADIO-ELEKTRONIKA), 2016 26th International Conference*, pages 292–295. IEEE.
- Shang, Z. and Shen, Z. (2017). Real-time 3d reconstruction on construction site using visual slam and uav. *arXiv preprint arXiv:1712.07122*.
- Wu, S., Oerlemans, A., Bakker, E. M., and Lew, M. S. (2017). A comprehensive evaluation of local detectors and descriptors. *Signal Processing: Image Communication*, 59:150–167.
- Yang, X. and Cheng, K.-T. (2012). Ldb: An ultra-fast feature for scalable augmented reality on mobile devices. In *2012 IEEE international symposium on mixed and augmented reality (ISMAR)*, pages 49–57. IEEE.
- Zhang, Y., Atkinson, P. M., Li, X., Ling, F., Wang, Q., and Du, Y. (2017). Learning-based spatial-temporal superresolution mapping of forest cover with modis images. *IEEE Transactions on Geoscience and Remote Sensing*, 55(1):600–614.
- Zhen, Y., Sun, Z., Li, J., and Peng, Y. (2016). An airborne remote sensing image mosaic algorithm based on feature points. In *Instrumentation & Measurement, Computer, Communication and Control (IMCCC), 2016 Sixth International Conference on*, pages 202–205. IEEE.