

DETECÇÃO AUTOMÁTICA DE OBJETOS GEOESPACIAIS POR AGRUPAMENTO DE PONTOS CARACTERÍSTICOS EM IMAGENS OBTIDAS POR DRONES

Leonardo de Souza Vieira¹, Lamartine Nogueira Frutuoso Guimarães², Elcio Hideiti Shiguemori³

¹ Instituto Nacional de Pesquisas Espaciais (INPE), leo76sv@gmail.com, ² Instituto de Estudos Avançados (IEAv), guimarae@ieav.cta.br, ³Instituto de Estudos Avançados (IEAv), elcioks@gmail.com

RESUMO

A utilização de drones em sistemas de sensoriamento remoto tem crescido por ser uma plataforma flexível e de baixo custo se comparado às plataformas tradicionais de imageamento (satélites e aeronaves tripuladas). Além disso, novos conceitos como a GEOBIA (Geographic Object-Based Image Analysis) trouxe novas perspectivas para o sensoriamento remoto. Neste trabalho é proposta uma abordagem que mescla os conceitos da GEOBIA para identificar objetos geoespaciais em imagens capturadas por drones. A abordagem consiste na extração e no agrupamento de um conjunto de pontos característicos da imagem para detecção de diferentes objetos geoespaciais, considerando que os pontos mais próximos pertencem a um objeto específico na imagem. Foram analisados onze diferentes métodos de agrupamento para selecionar o método mais indicado.

Palavras-chave – Métodos de agrupamento, extração de pontos característicos, sensoriamento remoto, veículos aéreos não tripulados

ABSTRACT

The use of drones in remote sensing systems has been increasing since it is a flexible and low-cost platform compared to traditional imaging platforms (satellites and manned aircrafts). Moreover, new concepts like GEOBIA (Geographic Object-Based Image Analysis) have brought new perspective to the remote sensing. The presented approach merges the concepts of GEOBIA to identify geospatial objects in images obtained by drones. A set of close key points of the image is extracted and grouped to detect different geospatial objects. The appropriate clustering method was selected among eleven different ones.

Key words – Clustering Methods, keypoints extraction, remote sensing, unmanned aerial vehicle

1. INTRODUÇÃO

Nos últimos anos, houve um crescimento no estudo e desenvolvimento de *drones*, motivado por avanços tecnológicos da indústria de eletrônicos e computação. Atualmente é possível encontrar uma grande variedade de modelos que são utilizados em diferentes aplicações, sendo o Sensoriamento Remoto (SR) uma de suas principais áreas de atuação. O sensoriamento remoto com uso de VANTs (Veículos Aéreos Não Tripulados) ou *drones* complementa o uso de plataformas tradicionais de imageamento (satélite

e aeronaves tripuladas) através de um modelo flexível e de baixo custo. A utilização destes dispositivos em sistemas de sensoriamento remoto possui vantagens únicas que incluem baixos custos operacionais, controle adaptável de resolução espacial e temporal, coleta de dados de alta intensidade e ausência de risco para as equipes [1].

O crescente desenvolvimento de tecnologias relacionadas ao Sensoriamento Remoto, como é o caso dos *drones*, facilitou a aquisição de imagens aéreas de alta resolução espacial e temporal aumentando consideravelmente o volume de dados geoespaciais a disposição. Este grande volume de dados inviabiliza a interpretação visual, tornando a utilização de sistemas automáticos de análise uma necessidade. Porém, desenvolver sistemas automáticos de interpretação ainda é um grande desafio para pesquisadores da área [2]. Uma etapa fundamental para o desenvolvimento de tais sistemas é identificar, segmentar e classificar de maneira automática diferentes objetos pertencentes a uma grande diversidade de classes em imagens de SR. Este processo está relacionado ao que vem sendo considerado um novo paradigma da Ciência da Informação Geográfica (GIScience) denominada GEOBIA (*Geographic Object-Based Image Analysis*) [3].

Cada vez mais é possível encontrar trabalhos científicos relacionados a GEOBIA. Este tipo de abordagem tem sua fundamentação teórica baseada em um conjunto de métodos clássicos de processamento de imagens: segmentação, classificação e extração de características. Além de métodos de aprendizado de máquina e inteligência computacional. A literatura científica relacionada a GEOBIA abrange o mais variado conjunto de técnicas e métodos relacionados a estas áreas do conhecimento [4].

Em [5] foi desenvolvido um método automático para seleção de marcos baseado no agrupamento (*clustering*) de pontos característicos em imagens aéreas. A abordagem proposta consiste na identificação automática de objetos representativos (marcos) em imagens da região de voo para serem utilizados como pontos de referência durante o voo do VANT. Para uma área urbana, por exemplo, existem diferentes objetos que podem ser tomados como marcos: um monumento, um prédio, um campo de futebol, entre outros. Apesar da técnica ser empregada para a navegação aérea por imagens, a proposta apresentada se relaciona diretamente com os conceitos de GEOBIA. Desta forma, o intuito deste trabalho foi analisar a abordagem utilizando diferentes técnicas de agrupamento, porém, o enfoque da análise se voltou para a identificação de objetos geoespaciais.

A ideia central da abordagem está em extrair um conjunto de pontos característicos na imagem para explorar a relação espacial entre eles, ou seja, considera-se que os pontos característicos mais próximos pertencem a um objeto específico na imagem. Adotando esta relação, é possível

agrupar os pontos característicos e selecionar diferentes objetos dentro da imagem. Para analisar a qualidade do agrupamento, foi utilizada a métrica Silhueta [6]. Foram testados 11 diferentes métodos de agrupamento, incluindo os métodos testados em [5]. A Seção 2 descreverá em detalhes como o procedimento foi realizado.

2. MATERIAIS E MÉTODOS

A abordagem analisada consiste em três etapas principais: extração de pontos característicos, agrupamento dos pontos e segmentação dos objetos. Na Figura 1 é apresentada uma visão geral do funcionamento da abordagem.



Figura 1: Segmentação de objetos por agrupamento e pontos característicos.

A análise de agrupamento é uma abordagem utilizada para análise de dados cujo objetivo é agrupar um conjunto de elementos considerando o grau de semelhança entre eles. Há na literatura diversos métodos de agrupamento baseados em diferentes conceitos: grafos, redes neurais artificiais, computação evolucionária, lógica fuzzy, métodos particionais, métodos hierárquicos, entre outros. Como cada método possui suas vantagens e desvantagens [7], foi selecionado um conjunto heterogêneo destes métodos com intuito de gerar maior grau de variabilidade nos testes. Ao todo foram testados 11 métodos: K-Means [8], X-Means [9], X-Means Adaptado [5], Ward [10], Agglomerative Clustering [11], Birch [12], Mean-Shift [13], DBSCAN [14], Spectral Clustering [15], Gaussian Mixture [16] e Self-Organized Mapping [17].

Para gerar o conjunto de pontos característicos das imagens foi utilizado o extrator e descritor de características Oriented Fast and Rotated Brief (ORB) [18]. Esta escolha se deve a dois fatores principais: o modelo de extração e a escalabilidade. O modelo de extração de ponto característicos do ORB utiliza um modelo de detecção de cantos. Este tipo de detecção é indicado quando a imagem possui um número significativo de objetos construídos pelo homem o que gera muitos cantos para detecção, como é o caso da imagem selecionada para o estudo de caso. Já o segundo fator se relaciona com o custo computacional para geração dos pontos. As imagens utilizadas são imagens grandes e de alta resolução e consomem uma quantidade significativa de memória para aplicação deste tipo de técnica. Outros métodos como o SURF [19] ou SIFT [20] que possuem um

custo computacional mais alto que o ORB, poderiam limitar o número de pontos gerados.

A análise da qualidade do agrupamento teve como orientação a métrica Silhueta. A métrica selecionada utiliza o próprio modelo para qualificar o resultado, ou seja, ela não necessita de um conjunto verdade para avaliar a técnica. A métrica adotada prioriza grupos mais densos e bem separados. Como o conceito central do estudo é utilizar a relação de proximidade entre pontos característicos, priorizar grupos mais densos e bem separados pode gerar um conjunto de objetos mais significativos para o desenvolvimento de um futuro sistema automático de interpretação. A seguir a métrica Silhueta é detalhada.

2.1. Silhueta

A métrica baseada no coeficiente Silhueta pode ser utilizada quando existe uma relação de escala entre os elementos a serem agrupados. Ou seja, o cálculo da distância (como a distância euclidiana) entre os elementos é aplicável. Neste caso, a análise a partir das silhuetas é recomendada quando busca-se encontrar agrupamentos bem definidos e visivelmente separáveis. O coeficiente silhueta $s(i)$ para uma única amostra i do conjunto de dados, pode ser calculado como em 1, ou seja, avaliando a dissimilaridade (ou distância) entre o elemento i em relação tanto aos elementos do grupo ao qual o elemento pertence quanto aos elementos do grupo vizinho mais próximo [6].

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (1)$$

onde: $a(i)$ é a distância média entre o elemento i e todos os elementos pertencentes ao mesmo grupo. E $b(i)$ é a distância média entre o elemento i e todos os elementos pertencentes ao grupo de menor distância média em relação a i .

A pontuação resultante de $s(i)$ está limitada ao intervalo $[-1, 1]$. Desta forma, quando o resultado for próximo de 1, significa que o elemento avaliado possui maior similaridade ao seu grupo do que ao grupo vizinho mais próximo. Já quando a pontuação for próxima de 0, significa que o elemento pode pertencer tanto ao seu grupo quanto ao grupo vizinho mais próximo, ou seja, não há certeza em relação ao agrupamento realizado indicando sobreposições entre os grupos. Se a pontuação for próxima de -1 significa que o elemento tem mais similaridade com o grupo vizinho. Para todo o conjunto de dados, o coeficiente silhueta é dado pela média dos coeficientes de cada amostra, denotado por:

$$S = \frac{1}{N} \sum_{i \in I_N} s(i) \quad (2)$$

onde: N é o número de amostras do conjunto de dados, I_N é o conjunto de amostras (elementos) e $s(i)$ é o coeficiente silhueta para cada amostra do conjunto de dados.

3. RESULTADOS

Nos experimentos realizados, o primeiro passo consistiu em extrair o conjunto de dados para alimentar os modelos. Para isto, foram extraídos de cada imagem 4000 pontos

característicos utilizando o algoritmo ORB. Os parâmetros do detector foram definidos de forma empírica analisando a sensibilidade do detector.

Definido o conjunto de dados, os parâmetros de cada algoritmo foram ajustados considerando o coeficiente Silhueta. Para os ajustes foi realizada uma análise visual para determinar um número mínimo de objetos contidos na imagem e estimar um número mínimo de grupamentos. Considerando o fato de que podem existir mais objetos significativos além dos visualizados, foi definido um intervalo $[min, max]$ de grupamentos para os testes, onde:

- min = número de marcos observados.
- $max = min + 10$

3.1. Estudo de caso

Para realizar a análise foi escolhida um ortomosaico da região de São José dos Campos construído a partir de imagens aéreas obtidas por *drones*. Na região escolhida, foram visualizados pelo menos 15 objetos diferentes, como mostra a Figura 2.

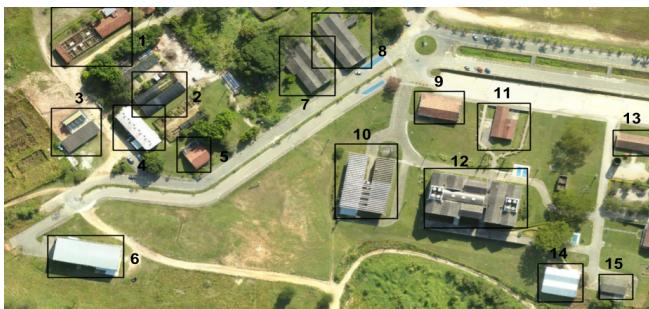


Figura 2: Análise visual para estimar um número mínimo de objetos na região 1.

A partir desta análise foi gerado o intervalo $[min, max] = [15, 25]$ de grupos para relizar o ajuste dos parâmetros dos métodos. Considerando as melhores configurações de cada algoritmo o coeficiente Silhueta foi utilizado para comparação entre os métodos. Os resultados obtidos para as melhores configurações pode ser analisado na Figura 3.

Considerando o gráfico do coeficiente silhueta apresentado na Figuras 3, e os métodos *Spectral Clustering*, *DBSCAN* e *Agglomerative Clustering* obtiveram os piores resultados, sendo que o método *Spectral Clustering* apresentou resultados insatisfatórios de acordo com o coeficiente. Porém, realizando uma análise comparativa geral entre os métodos, como os demais métodos atingiram melhores resultados sem a necessidade de estudos específicos da abordagem, os mesmos foram considerados mais indicados para este tipo de aplicação.

Para os demais métodos, o método *Mean Shift* apresentou o melhor resultado para análise do coeficiente silhueta. Além disso, considerando a automatização do processo, o método *Mean Shift* apresenta uma vantagem em relação aos demais, ele gera o número de agrupamentos automaticamente e não necessita do conhecimento prévio do número de grupos. Neste contexto, determinar um intervalo de k de grupos automaticamente pode não ser uma tarefa trivial. O número de objetos encontrados em uma imagem é variável de região

Coeficiente silhueta para as melhores configurações dos métodos

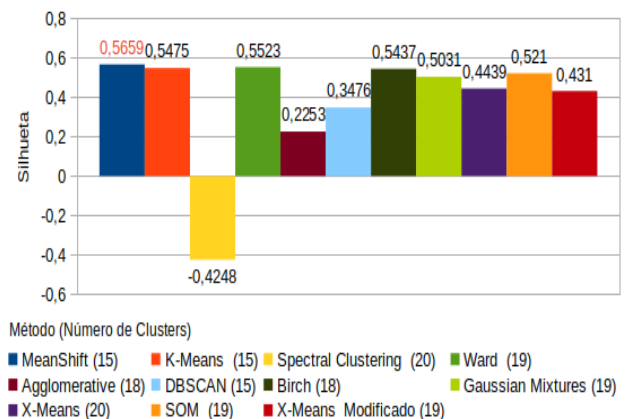


Figura 3: Coeficiente Silhueta para as melhores configurações de cada método dentro do intervalo de parâmetros estipulado para região.

para região. Mesmo no caso apresentado, onde o número de grupos é utilizado para ajustar os parâmetros dos métodos o processo de estimar um número de objetos não é um processo simples de automatizar. Realizadas estas ponderações, o método *Mean Shift* foi considerado o mais indicado para um processo automático de identificação de objetos por agrupamento de pontos característicos. Na Figura 4 é apresentado o resultado da identificação.

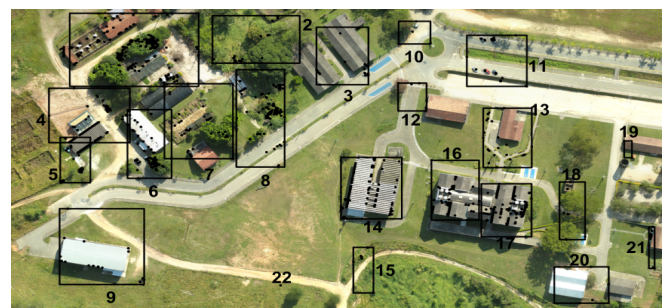


Figura 4: Identificação de objetos utilizando o método Mean Shift.

4. DISCUSSÃO

Considerando o resultado apresentado na Figura 4, a identificação automática apresentou diferenças em relação a análise visual. Os resultados obtidos foram avaliados como promissores, considerando o fato de que a abordagem é não supervisionada. Ou seja, o conjunto de objetos selecionados visualmente não foi adotado como conjunto verdade para identificação. Neste caso estas diferenças já eram esperadas, principalmente em relação ao número de objetos identificados. No modelo de testes realizado já foi considerado que poderiam existir mais objetos do que os observados na análise visual. O método com maior pontuação em relação a métrica identificou 22 objetos ao invés de 15 como foi estipulado visualmente.

A localização dos pontos característicos extraídos na imagem tem influência direta na identificação dos objetos. No

estudo de caso apresentado os pontos característicos extraídos da imagem não necessariamente se localizavam na região dos objetos selecionados em 2 fazendo com que com que alguns objetos não fossem identificados pela falta de pontos característicos na região.

Uma proposta para melhorar os resultados obtidos seria realizar um estudo com foco no método de extração de pontos característicos. Neste trabalho o foco foi avaliar os métodos de agrupamento, porém, pode ser realizado um estudo para ajustar os parâmetros do método de extração de características. Além disso, outros métodos de extração de pontos característicos podem ser analisados.

5. CONCLUSÕES

No artigo apresentado, foram analisados diferentes métodos de análise de agrupamentos para identificação de objetos geoespaciais em imagens obtidas por *drones*. A abordagem consistiu em associar os pontos característicos detectados na imagem através de sua relação espacial. Foi realizada uma avaliação objetiva baseada na métrica silhueta para determinar o melhor método de agrupamento para abordagem. De acordo com a avaliação realizada o método *Mean Shift* foi considerado o método mais indicado para a abordagem.

Os resultados apresentados foram considerados promissores, porém alguns ajustes ainda são necessários para melhorar os resultados obtidos. Como a abordagem tem relação direta com a localização dos pontos característicos extraídos, um próximo passo do estudo seria avaliar diferentes métodos para extração de pontos característicos. Existem diferentes métodos propostos na literatura, assim sendo, outros métodos podem ser testados para avaliação da abordagem.

A abordagem apresentada realiza somente a identificação de objetos sem categorizá-los. Um possível trabalho futuro seria desenvolver um modelo de classificação dos objetos geoespaciais considerando diferentes categorias.

6. REFERÊNCIAS

- [1] TANG, L.; SHAO, G. Drone remote sensing for forestry research and practices. *Journal of Forestry Research*, v. 26, n. 4, p. 791–797, Dec 2015. ISSN 1993-0607. Disponível em: <<https://doi.org/10.1007/s11676-015-0088-y>>.
- [2] CHENG, G. et al. Multi-class geospatial object detection and geographic image classification based on collection of part detectors. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 98, p. 119 – 132, 2014. ISSN 0924-2716. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0924271614002524>>.
- [3] KÖRTING, T. S.; FONSECA, L. M. G.; CÂMARA, G. Geodma—geographic data mining analyst. *Computers Geosciences*, v. 57, p. 133 – 145, 2013. ISSN 0098-3004. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0098300413000538>>.
- [4] BLASCHKE, T. et al. Geographic object-based image analysis – towards a new paradigm. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 87, p. 180 – 191, 2014. ISSN 0924-2716. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0924271613002220>>.
- [5] MELO, A. dos S.; FILHO, P. S.; SHIGUEMORI, E. H. Automatic landmark selection for uav autonomous navigation. In: *29 Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens*. [S.l.: s.n.], 2016. p. 6.
- [6] ROUSSEEUW, P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, v. 20, p. 53 – 65, 1987. ISSN 0377-0427. Disponível em: <<http://www.sciencedirect.com/science/article/pii/0377042787901257>>.
- [7] ROKACH, L.; MAIMON, O. Clustering methods. In: _____. *Data Mining and Knowledge Discovery Handbook*. Boston, MA: Springer US, 2005. p. 321–352. ISBN 978-0-387-25465-4. Disponível em: <https://doi.org/10.1007/0-387-25465-X_15>.
- [8] LLOYD, S. Least squares quantization in pcm. *IEEE Transactions on Information Theory*, v. 28, n. 2, p. 129–137, March 1982. ISSN 0018-9448.
- [9] PELLEG, D.; MOORE, A. X-means: Extending k-means with efficient estimation of the number of clusters. In: *In Proceedings of the 17th International Conf. on Machine Learning*. [S.l.]: Morgan Kaufmann, 2000. p. 727–734.
- [10] JR, J. H. W. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, Taylor & Francis Group, v. 58, n. 301, p. 236–244, 1963.
- [11] MÜLLNER, D. Modern hierarchical, agglomerative clustering algorithms; 2011. *arXiv preprint arXiv:1109.2378*.
- [12] ZHANG, T.; RAMAKRISHNAN, R.; LIVNY, M. Birch: An efficient data clustering method for very large databases. *SIGMOD Rec.*, ACM, New York, NY, USA, v. 25, n. 2, p. 103–114, jun. 1996. ISSN 0163-5808. Disponível em: <<http://doi.acm.org/10.1145/235968.233324>>.
- [13] CHENG, Y. Mean shift, mode seeking, and clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 17, n. 8, p. 790–799, Aug 1995. ISSN 0162-8828.
- [14] ESTER, M. et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In: *Kdd*. [S.l.: s.n.], 1996. v. 96, n. 34, p. 226–231.
- [15] NG, A. Y.; JORDAN, M. I.; WEISS, Y. On spectral clustering: Analysis and an algorithm. In: *Advances in neural information processing systems*. [S.l.: s.n.], 2002. p. 849–856.
- [16] DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, JSTOR, p. 1–38, 1977.
- [17] KOHONEN, T. The self-organizing map. *Proceedings of the IEEE*, v. 78, n. 9, p. 1464–1480, Sept 1990. ISSN 0018-9219.
- [18] RUBLEE, E. et al. Orb: An efficient alternative to sift or surf. In: *2011 International Conference on Computer Vision*. [S.l.: s.n.], 2011. p. 2564–2571. ISSN 2380-7504.
- [19] BAY, H. et al. Speeded-up robust features (surf). *Computer Vision and Image Understanding*, v. 110, n. 3, p. 346 – 359, 2008. ISSN 1077-3142. Similarity Matching in Computer Vision and Multimedia. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1077314207001555>>.
- [20] LOWE, D. G. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, Springer, v. 60, n. 2, p. 91–110, 2004.