

CLASSIFICAÇÃO DE IMAGENS UTILIZANDO IMAGENS RGB E TERMAL OBTIDAS POR ARPS DE PEQUENO PORTE

Marielcio Gonçalves Lacerda¹, Mateus Habermann², Camila Souza dos Anjos Lacerda³, Daniel Rodrigues Roos⁴ Thales Sehn Körting⁵ e Hermann Johann Heinrich Kux⁶

^{1,2,4}Instituto de Estudos Avançados (IEAv), Trevo Coronel Aviador José Alberto Albano do Amarante, nº 1 – Putim, São José dos Campos-SP; ^{1,5,6}Instituto Nacional de Pesquisas Espaciais (INPE), Caixa Postal 515 – 12227-010 – São José dos Campos-SP; ³Instituto Federal do Sul de Minas Gerais - Campus Inconfidentes, Praça Tiradentes, 416, Inconfidentes-MG; ^{1,2}[marielciomgl;habermannmh;danielroosdrr]@fab.mil.br, ²camila.lacerda@ifsuldeminas.edu.br, ^{4,5}[thales.korting;hermann.kux]@inpe.br

RESUMO

Nos últimos anos, verificamos um aumento no uso de Aeronaves Remotamente Pilotadas (ARPs), devido à redução nos custos, e também alavancada pela redução das dimensões e valores de sensores imageadores de altíssima resolução espacial, como os na faixa do visível, que têm sido amplamente utilizados em diversas aplicações do cotidiano, como mapeamento e levantamento cadastral. Porém, em algumas aplicações, a informação termal dos objetos da cena faz-se necessária. Ela pode contribuir para uma melhor caracterização dos alvos e distinção entre alvos que possuem respostas espectrais semelhantes na região do óptico refletido, e distintas na região do óptico emitido. Neste trabalho, explorou-se o conceito de *Data Fusion*, realizando-se a verificação da influência da imagem termal em conjunto com a RGB para melhoria da acurácia obtida na classificação de imagens. Para isso foram utilizadas as imagens obtidas pelas câmeras integradas da ARP DJI Mavic Enterprise Dual, a qual registra as duas imagens (RGB e termal) simultaneamente.

Palavras-chave – Fusão de Dados, Imagem TIR, Imagem RGB, Classificação de Imagens, GEOBIA.

ABSTRACT

In recent years, we have seen an increase in the use of Remotely Piloted Aircraft (ARPs), due to the reduction in costs, and also leveraged by the reduction of dimensions and values of imaging sensors of very high spatial resolution, such as those in the visible range, which has been widely used in various everyday applications, such as mapping and cadastral survey. However, in some applications, the thermal information of the objects in the scene is necessary and can contribute to a better characterization of the targets and to distinguish between targets with similar spectral responses in the region of the reflected optic and different in the region of the emitted optic. In this work, the concept of Data Fusion was explored, verifying the influence of the thermal image fused with RGB to improve the accuracy obtained in the classification of images. For this, the images obtained by the integrated cameras of the ARP DJI Mavic Enterprise Dual were used, which records both images simultaneously.

Key words – Data Fusion, TIR Image, RGB Image, Image Classification, GEOBIA.

1. INTRODUÇÃO

A redução no valor de aquisição, manutenção e operação das Aeronaves Remotamente Pilotadas (ARPs) ocorrida nos últimos anos, tornou essa plataforma aérea uma importante ferramenta na aquisição de dados para a utilização no Sensoriamento Remoto.

Juntamente com a redução dos custos nos voos com ARPs, a evolução tecnológica nos últimos anos permitiu que sensores imageadores tivessem suas dimensões reduzidas, proporcionando sua utilização em ARPs de pequeno porte. Atualmente, os sensores de dimensões reduzidas são capazes de abranger grande parte do espectro eletromagnético. Sensores que abrangem desde o espectro óptico refletido (Visível - RGB e Infravermelho Próximo - NIR), óptico emitido (Infravermelho Termal - TIR), microondas (LIDAR - *Light Detection And Ranging* (Detecção de luz e alcance) e o SAR - *Synthetic Aperture Radar* (Radar de Abertura Sintética). Essas imagens fornecem diferentes informações a respeito do objeto analisado, sendo, portanto, complementares.

Dentre essas imagens, a termal, TIR (*Thermal Infrared*), tem demonstrado ser uma importante fonte de informação para diferentes aplicações como agricultura, geologia, estudos de microclimas urbanos, atividades de vigilância, monitoramento e inspeção de linhas de energia e de prédios, verificação de saúde de plantios, entre outros [1], [2].

A aquisição, processamento e interpretação de imagens termais obtidas em voos com ARPs está se tornando uma fonte útil de informação para diferentes aplicações devido à maior resolução temporal e espacial desses produtos em comparação com as obtidas a partir de satélites [3], [4]. Porém as câmeras de baixo custo possuem precisão baixa em algumas aplicações. Além disso, o baixo contraste das imagens termais torna o processo de fotogrametria impreciso, o que resulta em erros na geração de ortoimagens [5].

Nesse aspecto, a fusão de dados (*Data Fusion*), fusionando as informações RGB e TIR, surge como uma solução para unir as vantagens de ambos os sensores: a alta resolução espacial da RGB e as informações termais dos alvos de interesse, de modo a ser possível a análise e interpretação de imagens de forma mais completa.

Assim, a hipótese principal deste trabalho é que a informação de imagem TIR fusionada com a informação da imagem RGB gera uma classificação de imagens com maior acurácia que a classificação da mesma imagem sem a utilização da informação termal. Para isso, foi realizado

um voo com uma ARP dotada de um sensor termal e visível, obtendo as imagens simultaneamente. As imagens obtidas foram processadas separadamente e no final foram submetidas a classificação de imagens considerando o uso da imagem TIR em um conjunto, e sem o uso da imagem TIR em um segundo conjunto. Os resultados obtidos indicam que as imagens TIR possuem a capacidade de melhorar a acurácia alcançada na classificação.

2. MATERIAL E MÉTODOS

ARP Utilizada. Para a realização deste trabalho foi utilizada a ARP DJI Mavic 2 Enterprise Dual, contendo um sensor imageador na região do visível (RGB) e outro na região do Termal (TIR - *Termal Infrared*). As imagens dos dois sensores são capturadas ao mesmo tempo e estão acopladas com o sistema GNSS da aeronave, logo as imagens RGB e TIR obtidas no aerolevantamento são georreferenciadas.

Área de Estudo. A área escolhida tem dimensões de aproximadamente 250 m x 150 m. As imagens foram obtidas em uma região da cidade de São José dos Campos-SP.

Planejamento e Execução dos Voos. Considerou-se a execução dos voos sobre a área com linhas nos dois sentidos de voo, denominado modo 3D, de forma que todos os pontos fossem imageados em sentidos perpendiculares. O voo foi realizado com altura de 60 metros; recobrimento lateral de 70% e frontal de 80%; e ângulo da câmera de 90° (NADIR). Para recobrir a área de estudo considerada, obteve-se 245 fotografias RGB de 4.056 x 3.040 *pixels*, com resolução espacial (GSD) de aproximadamente 2 cm. E 245 imagens TIR de 640 x 480 *pixels*, com GSD de aproximadamente 5 cm. Os voos foram realizados em outubro de 2022.

Processamento das imagens. Após o aerolevantamento, realizou-se o processamento das imagens em software apropriado. Cada um dos conjuntos de imagens (RGB e TIR) foram processados separadamente, de modo que para cada processamento fosse gerado um conjunto de dados composto por 1 ortomosaico, um Modelo Digital de Superfície (MDS) e um Modelo Digital de Terreno (MDT), além de outros resultados, como nuvens de pontos, que não foram utilizados para a realização deste trabalho. Em complemento às informações geradas, gerou-se o Modelo Digital de Altura (MDA), tomando como referência o MDT e MDS gerados.

Ajuste das Imagens Geradas. Como dito anteriormente, as imagens utilizadas são obtidas ao mesmo tempo, mas por sensores diferentes, com características diferentes. Dentre essas diferenças está a distância focal e dimensões do sensor de captura, fazendo com que as imagens obtidas gerem resultados com GSD diferentes e com uma área coberta total ligeiramente diferente, sendo a área um pouco menor na imagem TIR. Por esse motivo, realizou-se um recorte em todos os produtos gerados com base na imagem TIR.

Além disso, realizando-se a análise em todos os produtos gerados, optou-se pela utilização do MDS, MDT e MDA gerados a partir da imagem RGB.

A Figura 1 apresenta as 5 imagens utilizadas nas etapas seguintes de processamento.

Abordagem GEOBIA. A Análise Orientada a Objetos Geográficas (GEOBIA, do inglês *Geographic Object-Based*

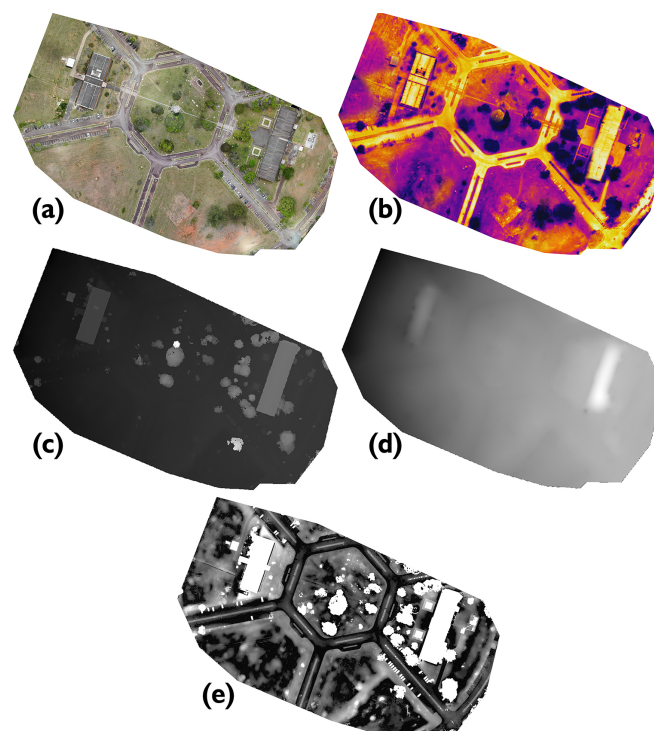


Figura 1: Imagens Recortadas. (a) Mosaico RGB; (b) Mosaico TIR; (c) MDS; (d) MDT; (e) MDA

Image Analysis tem ganhado muito destaque na classificação de alvos complexos, pois utiliza não apenas a informação do *pixel*, mas outros atributos que auxiliam para melhor caracterizar o objeto presente na cena analisada [6] e [7]. Os resultados de classificação baseada em objetos tem apresentado resultados melhores que aqueles obtidos na análise *pixel-a-pixel* [8].

Sendo assim, o primeiro passo dessa abordagem é a segmentação da imagem [9], que é o processo no qual a imagem é subdividida em regiões ou objetos que a compõem [10] e [11]. Nesse trabalho optou-se por realizar esse processo em duas etapas: Segmentação Multiresolução (SMR) seguida da Segmentação por Diferença Espectral (SDE).

A SMR consiste em analisar os objetos da imagem em diferentes níveis de resolução [12], de modo a permitir a separação de objetos pequenos, que são melhor analisados em uma resolução mais fina, dos objetos maiores, que apresentam resultado melhor em uma resolução maior [13]. O algoritmo SMR possui três parâmetros: **escala** (*scale parameter*), que determina a heterogeneidade máxima permitida para os segmentos resultantes; **forma** (*shape*), que determina o peso dos critérios de cor e forma no processo de segmentação; e **compacidade** (*compactness*), que determina o peso dos critérios de compacidade e suavidade [11].

A SDE tem como objetivo mesclar objetos vizinhos de acordo com os valores médios de intensidade da camada. Os objetos de imagem são mesclados caso a diferença entre as intensidades médias da camada estiver abaixo do limiar fornecido como diferença espectral máxima [14]. Sendo assim, o valor definido para o parâmetro de diferença espectral funciona como um filtro de limiar, no qual une os segmentos com a diferença espectral abaixo do valor definido.

Classes e Coleta de Amostras. Após análise do

ortomosaico RGB gerado, identificou-se quais as classes de uso e cobertura do solo estavam presentes na cena em estudo, definindo-se sete classes: **C1.** Vegetação Rasteira; **C2.** Asfalto; **C3.** Vegetação Arbórea; **C4.** Asfalto Pintado; **C5.** Concreto; **C6.** Construções **C7.** Veículos; e **C8.** Solo Exposto.

Após a definição das classes e com o objetivo de realizar uma classificação supervisionada, foram coletadas amostras das classes definidas.

Conjunto de Dados para Classificação. Conforme mencionado anteriormente, a principal vantagem da utilização da abordagem GEOBIA está na possibilidade da utilização de atributos que melhor caracterizam os objetos presentes na imagem [15], [7] e [8]. A determinação dos atributos mais importantes contribui significativamente para a classificação final [16]. Foram criados os arquivos contendo os atributos de cada um dos segmentos gerados. Neste trabalho, foram considerados os seguintes atributos: **Object Features:** > **Layer Values:** Mean, Brightness, Max. Diff., Mode, Standard deviation, Pixel-based (Ratio, Minimum pixel value, Maximum pixel value e Border Contrast), HSI; > **Geometry:** Extent (Area, Border length, Length, Length/Width, Width), Shape (Asymmetry, Border Index, Compactness, Main Direction, Rectangular Fit, Shape Index).

Esse processo foi dividido em duas partes, a primeira considerando todas imagens do projeto, conforme citadas no item 2, e na segunda, as mesmas imagens, mas retiradas as informações referentes a imagem TIR. Sendo assim, o conjunto completo foi composto por 91 atributos ao total, enquanto que o conjunto sem a informação TIR, 67 atributos. Cabe salientar que os conjuntos de dados e os de amostras referentes a cada classificação são compostos dos mesmos atributos.

Classificação. O objetivo da classificação é rotular os segmentos da imagem de acordo com as classes definidas [17] e [18]. Como abordado no Item 1, o objetivo principal deste trabalho é verificar a interferência da informação da imagem TIR no aumento da acurácia da classificação da imagem.

Para realizar essa etapa, foi escolhido o classificador *Random Forest* (RF) ou Floresta randômica, que é um conjunto de árvores de decisão. O RF combina várias árvores aleatórias que votam individualmente no rótulo determinado para cada segmento. O resultado da classificação é a classe que contém mais votos [19].

3. RESULTADOS

As Figuras 2 e 3 apresentam, respectivamente, as imagens classificadas apenas com as imagens RGB e com a fusão de dados (RGB + TIR). Cabe salientar que para ambos os casos foram usados o classificador *Random Forest*.

Realizou-se a avaliação da acurácia pelo método *Cross Validation*, que utiliza o conjunto de amostra particionando em conjunto de treinamento e teste.

A Tabela 1 apresenta a Matriz de Confusão da classificação sem considerar os dados TIR. Realizando-se o cálculo, o valor da Acurácia Global foi de 83,02%.

A Tabela 2 apresenta a Matriz de Confusão da classificação

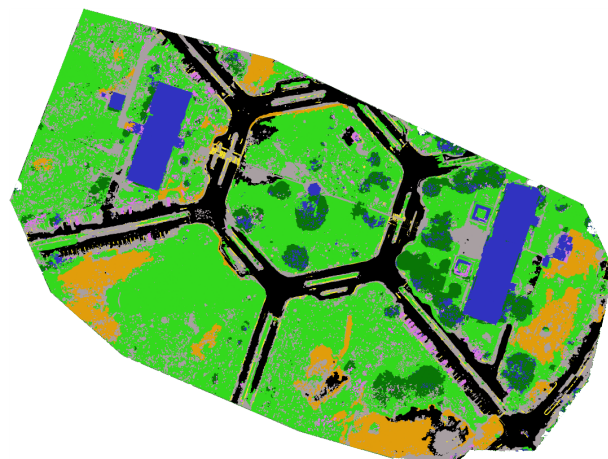


Figura 2: Imagem classificada considerando apenas as informações RGB

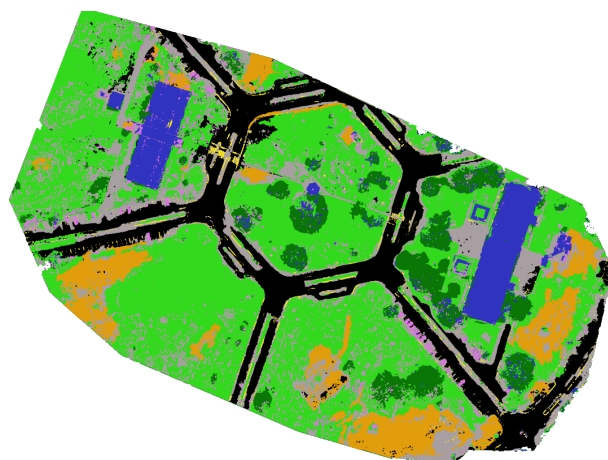


Figura 3: Imagem classificada considerando a fusão de dados (Data Fusion)

	C1	C2	C3	C4	C5	C6	C7	C8
(C1)	5	0	0	0	0	0	0	0
(C2)	0	3	0	0	0	0	0	0
(C3)	2	0	7	0	0	0	0	0
(C4)	0	0	0	8	0	0	0	0
(C5)	1	3	0	0	2	0	0	0
(C6)	0	0	0	0	0	8	1	0
(C7)	0	0	0	0	0	0	10	0
(C8)	0	1	0	0	1	0	0	1

Tabela 1: Matriz de Confusão - RGB

considerando o Data Fusion. Realizando-se o cálculo, o valor da Acurácia Global foi de 88,68%.

	C1	C2	C3	C4	C5	C6	C7	C8
(C1)	5	0	0	0	0	0	0	0
(C2)	0	3	0	0	0	0	0	0
(C3)	1	0	8	0	0	0	0	0
(C4)	0	0	0	8	0	0	0	0
(C5)	1	1	0	0	4	0	0	0
(C6)	0	0	0	0	0	8	1	0
(C7)	0	0	0	0	0	0	10	0
(C8)	0	1	0	0	1	0	0	1

Tabela 2: Matriz de Confusão - Data Fusion

4. DISCUSSÃO

Os resultados obtidos são preliminares e foram realizados para uma pequena região, porém a *Data Fusion* apresenta um grande potencial para utilização na classificação de imagens.

Analisando as imagens classificadas, pode-se verificar que alguns alvos conseguem ser muito melhor distinguidos na classificação utilizando a fusão de dados. Em algumas regiões com solo exposto, por exemplo, nas classificações utilizando apenas as imagens no espectro visível houve muita confusão com a classe vegetação rasteira, enquanto que na classificação que utiliza a fusão de dados, esses erros praticamente não existiram. Outra classe que apresentou uma classificação mais acurada com a utilização da fusão de imagem é a de veículos, que foram identificados com elevada acurácia utilizando as imagens TIR, enquanto houve muita confusão com as classes construções e asfalto no conjunto RGB.

5. CONCLUSÕES

Pelos resultados apresentados, pode-se concluir que a classificação de imagens usando o processo de *Data Fusion*, com a junção das imagens RGB e TIR de uma mesma região analisada, foi possível obter uma acurácia mais elevada que aquela comparada com a classificação utilizando apenas as informações RGB.

6. REFERÊNCIAS

- [1] A. Sledz, J. Unger, and C. Heipke. Thermal ir imaging: Image quality and orthophoto generation. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-1, 2018.
- [2] E. Adamopoulos, M. Volinia, M. Girotto, and F. Rinaudo. Three-dimensional thermal mapping from irt images for rapid architectural heritage ndt. *Buildings*, 10(10), 2020.
- [3] Adam Dlesk, Karel Vach, and Karel Pavelka. Transformations in the photogrammetric co-processing of thermal infrared images and rgb images. *Sensors*, 21(15), 2021.
- [4] A. Fryskowska-Skibniewska, P. Delis, M. Kedzierski, and D. Matusiak. The conception of test fields for fast geometric calibration of the flir vue pro thermal camera for low-cost uav applications. *Sensors*, 22(7), 2022.
- [5] K. Ribeiro-Gomes, D. Hernandez-Lopez, J. F. Ortega, R. Ballesteros, T. Poblete, and M. A. Moreno. Uncooled thermal camera calibration and optimization of the photogrammetry process for uav applications in agriculture. *Sensors*, 17(10), 2017.
- [6] Agnes Silva Araujo, Isabela Lago Silva, Beatriz Pereira, and Alfredo Pereira Queiroz. Análise comparativa dos parâmetros de segmentação e regras de classificação das vias pavimentadas de marília - sp. *Caderno de Geografia*, 30(60):112–127, 2020.
- [7] R. R. Antunes. Integração da análise de imagem baseada em objetos (obia) e aprendizagem de máquina em ambiente de computação distribuída. *Revista Tecnologias em Projeção*, 10(1):89–104, 2019.
- [8] M. G. Lacerda. Abordagem geobia para imagens vhr obtidas por aeronaves remotamente pilotadas e sensores satelitais com o uso de classificadores individuais e *Ensemble*, 2020. Dissertação (Mestrado em Ciências e Tecnologias Espaciais. Área de Sensores e Atuadores Espaciais), ITA (Instituto Tecnológico de Aeronáutica, São José dos Campos - SP, Brasil.
- [9] F. Leonardi. Abordagens cognitivas e mineração de dados aplicadas a dados ópticos orbitais e de laser para a classificação de cobertura do solo urbano, 2010. Dissertação (Mestrado em Sensoriamento Remoto), INPE (Instituto Nacional de Pesquisas Espaciais), São José dos Campos-SP, Brasil.
- [10] M. G. Lacerda, C. S. Anjos, É. H. Shiguemori, and A. J. Damião. Influência dos parâmetros de segmentação de imagem em ortofotomosaicos confeccionados a partir de fotografias obtidas por aeronaves remotamente pilotadas de pequeno porte. *Anais do XIX Simpósio Brasileiro de Sensoriamento Remoto*, 1:3323–3326, 2019.
- [11] J. M. J. E. Matias. Análise comparada das potencialidades e limitações operacionais em ambiente e planeamento territorial. caso de estudo: os municípios de catumbela e lobito - angola, 2019. Dissertação (Mestrado em Gestã do Território), FCSH (Faculdade de Ciências Sociais e Humanas - Universidade Nova de Lisboa), Lisboa, Portugal.
- [12] C. Witharana and D. L. Civic. Optimizing multi-resolution segmentation scale using empirical methods: Exploring the sensitivity of the supervised discrepancy measure euclidean distance 2 (ed2). *ISPRS Journal of Photogrammetry and Remote Sensing*, 87:108–121, 2014.
- [13] C. S. Anjos, C. M. Almeida, and L. S. Galvão. Identificação de materiais urbanos por meio de métodos inovadores de classificação de imagens. *Anais do XVII Simpósio Brasileiro de Sensoriamento Remoto - SBSR, INPE*, 1:4377–4384, 2015.
- [14] Definiens. *Definiens Developer7: Reference Book*, volume 1. Definiens Understanding Images, Munich, Germany, 7.0.0.843 edition, 2007.
- [15] S. D. M. Guarda, E. S. Bias, E. E. Sano, E. F. Castejon, P. N. Happ, R. R. Antunes, A. de A. Teixeira, and R. B. Sousa. Análise da qualidade da segmentação e classificação de uma imagem urbana do satélite geoeye-1 baseada em objeto e realizadas nas plataformas postgresql e interimage. *Anuário do Instituto de Geociências - UFRJ*, pages 514–526, 2019.
- [16] W. Chaofan, D. Jinsong, W. Ke, M. Ligang, and A. R. S. Tahmassebi. Object-based classification approach for greenhouse mapping using landsat-8 imagery. *International Journal of Agricultural and Biological Engineering*, 9(1):79–88, 2016.
- [17] C. S. Anjos. Classificação de áreas urbanas com imagens multiespectrais e hiperespectrais utilizando métodos não-paramétricos, 2016. Tese (Doutorado em Sensoriamento Remoto), INPE (Instituto Nacional de Pesquisas Espaciais), São José dos Campos-SP, Brasil.
- [18] Y. Torres, J. J. Arranz, Jorge M. Gaspar-Escribano, A. Haghi, S. Martínez-Cuevas, B. Benito, and J. C. Ojedab. Integration of lidar and multispectral images for rapid exposure and earthquake vulnerability estimation. application in lorca, spain. *International Journal of Applied Earth Observation and Geoinformation*, 81:161–175, 2019.
- [19] F. Livingston. Implementing breimans random forest algorithm into weka. *ECE591Q Machine Learning Conference Papers*, pages 1–5, 2005.