Enhancing Band Alignment Performance in Multispectral UAV Images Using the Machine Learning

Kim S.[1], Park M.[2] and Kim T.[3*]

[1]Master Student, Department of Geoinformatic Engineering, Inha University, Republic of Korea

[2] Master Student, Department of Geoinformatic Engineering, Inha University, Republic of Korea

[3]Professor, Department of Geoinformatic Engineering, Inha University, Republic of Korea

*tezid@inha.ac.kr

1. Introduction

As Unmanned Aerial Vehicle (UAV) technology advances, it has become possible to monitor crops in agricultural fields in near real-time, enabling efficient crop management. When acquiring UAV images, the visible light band has limitations in understanding the state of crops and soil. Therefore, acquiring crop information using UAV with multiple bands provides more diverse and accurate information. In a multispectral camera, there are inherent physical distances between each band. To address these physical distances, band alignment is needed to combine multiple band images into a single image.

Automated band alignment involves extracting tiepoints and applying geometric transformation methods to create an aligned image. To obtain the transformation coefficients required for band alignment, accurate tiepoints are essential. The traditional tiepoints extraction algorithms such as Speeded Up Robust Features (SURF) and Oriented FAST and Rotated BRIEF (ORB) exhibit superior performance in matching feature points between same bands. However, when extracting tiepoints between different spectral bands with varying brightness characteristics, it becomes challenging to identify tiepoints. This makes it crucial to extract a large number of accurate tiepoints from all multispectral bands. Tiepoint extraction is particularly evident when dealing with challenging areas, such as agricultural or mountainous regions.

To address this issue, we propose a band alignment method that employs LightGlue (Lindenberger et al, 2023). LightGlue is a deep learning-based tiepoint extraction method that enhances the speed of SuperGlue (Sarlin et al, 2020), which has demonstrated robust performance in tiepoint extraction. Deep learning-based tiepoint extraction methods differ from traditional techniques by avoiding the use of texture-based descriptors~~.~~ and, instead, by using pre-trained parameter setting. These characteristics are expected to result in superior performance compared to traditional methods when extracting tiepoints between different band images.

Experiment results demonstrated that the LightGlue method significantly outperforms traditional algorithms in tiepoint extraction from multispectral images, leading to a substantial increase in the success rate of band alignment.

2. MATERIALS AND METHODS

2.1 Materials

The study utilized two datasets, each comprising approximately 100 UAV images. The UAV images were captured using a rotary-wing aircraft, the Phantom4 Multispectral. The camera of the Phantom4 Multispectral consists of five bands: Red, Green, Blue, NIR, and Red. Dataset 1 was captured in the Gimje area of South Korea and primarily included images of farmland and various crops. Dataset 2 was also captured in the Boryeong area of Korea and consisted of images of forested regions with abundant trees and grass.

2.2 Tiepoint Extraction

In this experiment, we used three tie-point extraction methods in this experiment to compare the performance of tie-point extraction across different bands. The first and second method was the traditional technique of tiepoints, which extracts feature points and descriptors that contain information about the surrounding pixel values. The similarity between these descriptors was then calculated to extract the tiepoints. Specifically, for traditional tiepoint extraction, we employed fast matching techniques such as SURF and ORB.

The other method involved using LightGlue, a deep learning-based matching technique, to extract tiepoints. LightGlue was employed to extract and match feature points in multispectral band images. To enhance the accuracy of the tiepoints extracted by LightGlue, the number of layers was set to the maximum of 9, and the filter threshold was adjusted to 0.9.

2.3 Band Alignment

The experiment procedure for band alignment in this study is as follows. First, the images corresponding to five different bands, such as Red, Green, Blue, NIR, and Red Edge, were individually input into the Module. The NIR band, which has the best feature point matching performance among the 5 bands, was selected as the reference band for multi-band alignment. Feature points and tiepoints were extracted between the reference band image and the images of the other bands.

However, the initially extracted tiepoints included numerous outliers with poor accuracy. These outliers were likely to introduce errors when calculating the band alignment coefficients. Therefore, it was essential to perform a filtering step to retain only high-quality tiepoints. In the method using the SURF or ORB, the quality of tiepoints was assessed based on the Euclidean distances calculated during tiepoint extraction. Tiepoints with lower quality were removed based on a threshold value. The method using LightGlue involved filtering tiepoints by setting a filter threshold during the extraction process. Only the filtered tiepoints were used with RANSAC (Random Sample Consensus) to extract inliers. To ensure the high quality of the band-aligned images, the decision to proceed with band alignment was based on the ratio of inliers to the initially filtered tiepoints.

Using only the high-quality tiepoints extracted in this method, coefficients of a homography transformation were computed for each band. The extracted homography coefficients for each band were used to align the images of the other bands with the NIR band as the reference, resulting in the final aligned image.

3. Results and Discussion

In this study, the performance of the experiment results was evaluated using the following metrics: reprojection error, band-aligned images, the number of tiepoints extracted, and the number of images with successful band alignment across the entire dataset. Reprojection error was measured based on matched points between two images. It was defined as the difference between the position of a point transformed from one image to another point using homography matrix and actual position of the point in the target image. Reprojection error is a metric that can assess the accuracy of the homography coefficients estimated from tiepoints. Both methods produced results with an average reprojection error within 0.8 pixels, indicating precise outcomes. Based on these precise results, the aligned images demonstrated superior performance, as shown in Figure 1.

There was a significant difference in the number of tiepoints extracted between bands using SURF, ORB and LightGlue. In the SURF algorithm, the maximum number of feature points was set as 20,800. However, while a large number of tiepoints were extracted for the green and red edge bands, averaging 1,162 and 11,315 points respectively, the tiepoint extraction between the red and blue bands was notably poor with averages of only 40 and 52 points, respectively. The ORB algorithm, like the SURF algorithm, also extracted an average of 1,124 and 8,444 tiepoints for the Green bands and Red Edge, respectively. However, for the red and blue bands, it showed a significant discrepancy with only 109 and 78 tiepoints, respectively.

In contrast, the number of tiepoints extracted through Lightglue was 4,647 per image. In contrast, the LightGlue method extracted an average of 6,453, 6,325, 7,625, and 6,578 tiepoints for the red, green, blue, and red edge bands, respectively. This resulted in a significantly reduced variation compared to traditional methods.

Based on these results, a significant difference was observed in the success rate of band alignment for all five bands, demonstrating the need for matching all bands. SURF and ORB achieved band alignment success rate of 17% and 16% and within the dataset. In contrast, the LightGlue method demonstrated a significantly higher success rate of 98%.

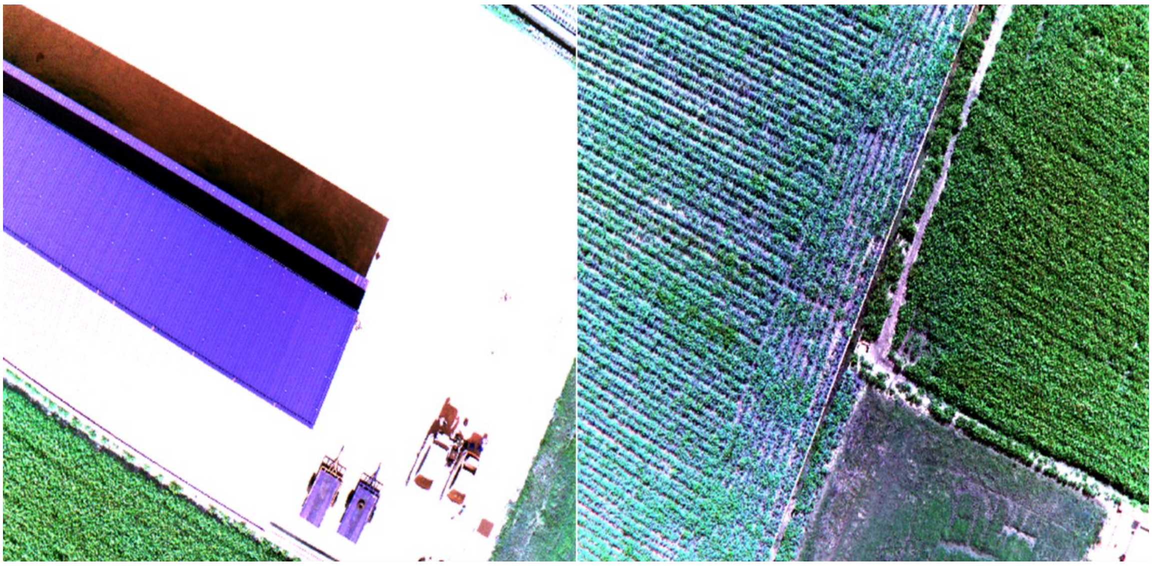

Figure 1: Band alignment image using Lightglue

4. Conclusion and Recommendation

In this study, a method was proposed to address the challenges of extracting tiepoints and aligning multiple bands using traditional approaches by applying the deep learning-based method LightGlue. This study applied LightGlue to a tiepoint extraction process during band alignment and identified improvements compared to the traditional descriptor-based tiepoint extraction method, such as the SURF and ORB algorithm. The experiment results showed that the band alignment method using LightGlue significantly improved the matching success rate between multi-band images, while maintaining high quality in the aligned images, compared to traditional methods. Based on these results, the proposed method is expected to improve the stability of band alignment in challenging agricultural and forested regions when using multispectral images.

Acknowledgments

This work was carried out with the support of "Cooperative Research Program for Agriculture Science and Technology Development (Project No. PJ017042)" Rureal Development Administration, Republic of Korea.

This work was carried out with the support of "Cooperative Research Program for Agriculture Science and Technology Development (Project No. RS-2021-RD009566)" Rureal Development Administration, Republic of Korea.

References

Lindenberger, P., Sarlin, P. E., & Pollefeys, M., 2023. Lightglue: Local feature matching at light speed. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 17627-17638). https://doi.org/10.48550/arXiv.2306.13643

Sarlin, P. E., DeTone, D., Malisiewicz, T., & Rabinovich, A., 2020. Superglue: Learning feature matching with graph neural networks. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 4938-4947). https://doi.org/10.48550/arXiv.1911.11763