



IJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** VI **Month of publication:** June 2024

DOI: <https://doi.org/10.22214/ijraset.2024.63189>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

AI-based Precision Enhancement for Change Detection in Drone/UAV Videos

Hemant Yadav¹, Pradeep Singh²

Militray College of Tele Communication Engineering, Mhow, MP-453441

I. INTRODUCTION

The rapid advancement in drone/UAV technology has led to their extensive use in various domains including surveillance, agriculture, and environmental monitoring. Change detection in UAV videos is crucial for identifying significant alterations in the observed areas. However, challenges such as varying scales, resolutions, and orientations of video frames pose significant hurdles. This paper addresses these challenges by leveraging the SIFT algorithm for panoramic stitching and change detection, followed by the application of the YOLOv8 model for high-precision anomaly detection.

II. RELATED WORK

Existing methods for change detection in UAV videos often suffer from limitations related to scale and orientation discrepancies between video frames. Traditional techniques like frame differencing and pixel-based comparison are inadequate for handling such variations. Recent advances in feature-based methods, particularly those employing SIFT, have shown promise in overcoming these challenges. Additionally, deep learning models like YOLO have revolutionized object detection tasks, making them suitable for high-precision change detection.

III. METHODOLOGY

A. Panoramic Image Stitching Using SIFT

The process of creating panoramic images from drone video frames using SIFT involves several key steps designed to ensure accurate stitching despite challenges such as varying scales, resolutions, and orientations:

- 1) *Feature Detection and Description:* The SIFT algorithm starts by detecting keypoints in each video frame. This involves building a scale-space representation of the image by progressively blurring it and identifying extrema in the Difference of Gaussian (DoG) images. Keypoints are selected based on their stability across scales, and each keypoint is assigned an orientation based on local image gradients. This makes the keypoints invariant to scale and rotation.
- 2) *Feature Matching:* After detecting keypoints, the next step is to find correspondences between keypoints in successive frames. This is done using a descriptor vector for each keypoint, which encodes local gradient information in a 128-dimensional space. Keypoints are matched based on the Euclidean distance between their descriptor vectors. A ratio test is applied to discard ambiguous matches, typically retaining matches where the distance ratio between the closest and the second closest match is below 0.75.
- 3) *Transformation Estimation:* Once keypoints are matched, the geometric transformation (homography) that aligns the frames is estimated. This involves using RANSAC (Random Sample Consensus) to robustly estimate the homography matrix by iteratively selecting random subsets of matched keypoints, computing the transformation, and counting inliers that fit the transformation. This step helps to reject outliers and ensures that only consistent matches contribute to the final transformation.
- 4) *Image Blending:* The final step in panoramic stitching is blending the aligned frames to create a seamless panoramic image. Multi-band blending is employed to minimize visible seams and ensure smooth transitions between frames. This technique involves decomposing images into multiple frequency bands, blending each band separately, and then recombining them to produce the final image. This approach effectively handles variations in exposure and lighting conditions between frames.

B. Challenges in Panoramic Stitching

Creating panoramic images from drone/UAV video frames involves addressing several key challenges:

- 1) *Scaling:* Variations in drone altitude and camera zoom result in frames captured at different scales. The SIFT algorithm's scale invariance property allows it to detect keypoints across a range of scales, making it suitable for matching features in frames taken at different altitudes. The scale-space extrema detection step in SIFT identifies points that are invariant to scale changes, ensuring consistent keypoint detection.

- 2) *Resolution*: Frames from drones may vary in resolution due to different camera settings or post-processing. High-resolution frames provide detailed features but increase computational complexity, while low-resolution frames reduce computational load but may lose detail. Multi-resolution analysis is employed to process images at different resolutions, integrating the results to form a high-resolution panorama.
- 3) *Orientation and Angles*: Drone movements cause variations in camera orientation and angles, complicating frame alignment. The orientation assignment step in SIFT assigns a consistent orientation to each keypoint based on local image gradients, making keypoints rotation invariant. This allows effective alignment of frames taken from different angles. Robust matching and homography estimation techniques ensure accurate alignment despite orientation differences.
- 4) *Lighting and Exposure Variations*: Frames captured at different times or under varying lighting conditions can have significant exposure differences. Multi-band blending techniques help to minimize these differences by blending images in the frequency domain. This approach ensures that high-frequency details are preserved while low-frequency variations, such as lighting differences, are smoothed out.

C. Common Area Detection Using SIFT

Once panoramic images are generated from different video sequences of the same region, SIFT is used to identify common areas:

- 1) *Feature Detection*: SIFT detects keypoints in both panoramic images. This involves constructing a scale-space representation of the images, detecting extrema in the DoG images, and refining keypoint positions to sub-pixel accuracy. Each keypoint is described by a 128-dimensional vector that captures local gradient information, ensuring robustness to scale and rotation changes.
- 2) *Feature Matching*: Correspondences between keypoints in the two panoramic images are established using techniques such as nearest neighbor search. The Euclidean distance between feature vectors is used to find matches. To improve robustness, a ratio test is applied to discard matches that do not pass a certain threshold, typically set to 0.75. This step ensures that only reliable matches are retained.
- 3) *Common Area Identification*: The matched keypoints define the common regions in the panoramic images. RANSAC (Random Sample Consensus) is employed to estimate a homography that best aligns the matched keypoints while rejecting outliers. RANSAC iteratively selects random subsets of matched keypoints, computes the homography, and counts the number of inliers. The homography with the highest inlier count is selected, and the transformation is applied to extract the common area, ensuring geometric consistency.

D. Change Detection Using YOLOv8

The identified common areas are cropped and fed into a pre-trained YOLOv8 model for high-precision change detection:

- 1) *Model Training*: YOLOv8 is trained on a diverse dataset containing various types of changes, such as structural alterations, environmental modifications, and object additions/removals. The model architecture is optimized for real-time object detection, incorporating advanced techniques such as anchor-free detection and feature pyramid networks (FPN). The training process involves optimizing the model's parameters using a large annotated dataset, ensuring that the model can accurately detect changes in different contexts.
- 2) *Inference*: During inference, the common areas extracted from the panoramic images are processed by the YOLOv8 model. The model outputs bounding boxes, class labels, and confidence scores for detected changes. Non-maximum suppression (NMS) is applied to filter out redundant detections and retain only the most confident predictions. The YOLOv8 model's architecture consists of several convolutional layers with skip connections and upsampling layers that enable multi-scale feature aggregation. This ensures that both small and large changes can be detected accurately. The use of anchor-free detection eliminates the need for predefined anchor boxes, reducing computational overhead and improving flexibility.

IV. EXPERIMENTAL RESULTS

Experiments were conducted on UAV video datasets covering urban and rural areas. The performance of the proposed method was evaluated based on precision, recall, and F1-score metrics. The SIFT-based panoramic stitching showed significant improvement in handling scale, resolution, and orientation variances. The YOLOv8 model demonstrated high accuracy in detecting changes within the identified common areas.

A. Dataset and Preprocessing

The datasets used in the experiments consisted of video sequences captured by UAVs over different terrains and environments. Each dataset included videos taken at different times to capture changes in the observed areas. Preprocessing involved frame extraction, resizing, and keypoint detection using SIFT.

B. Performance Metrics

The performance of the panoramic stitching and change detection methods was evaluated using the following metrics:

- 1) *Precision* The ratio of true positive detections to the total number of detections, indicating the accuracy of the detected changes.
- 2) *Recall* The ratio of true positive detections to the total number of actual changes, indicating the completeness of the detected changes.
- 3) *F1-score* The harmonic mean of precision and recall, providing a balanced measure of performance.

C. Results and Analysis

The SIFT-based panoramic stitching method successfully addressed challenges related to scaling, resolution, and orientation, resulting in high-quality panoramic images. The use of multi-band blending ensured smooth transitions and minimized visible seams.

The common area detection using SIFT achieved high accuracy, with the RANSAC-based homography estimation effectively rejecting outliers. The identified common areas were geometrically consistent and suitable for subsequent change detection.

The YOLOv8 model demonstrated robust performance in detecting various types of changes. The use of anchor-free detection and multi-scale feature aggregation contributed to high precision and recall. The experimental results showed that the proposed method outperformed traditional frame differencing and pixel-based comparison techniques, achieving higher F1-scores across different datasets.

V. CONCLUSION

This paper presents a robust method for enhancing precision in change detection in drone/UAV videos by combining SIFT and YOLOv8. The use of SIFT for panoramic stitching addresses key challenges, while YOLOv8 ensures high-precision anomaly detection. Future work will focus on real-time implementation and further optimization of the algorithms to handle more complex scenarios.

VI. ACKNOWLEDGMENTS

The authors, Lt Col Hemant Kumar Yadav and Lt Col Pradeep Singh, would like to thank their Directing Staff guides, Lt Col Gurpreet Singh and Lt Col Aditya Kaushik, for their invaluable guidance and support throughout this research.

REFERENCES

- [1] Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2), 91-110.
- [2] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [3] Mikolajczyk, K., & Schmid, C. (2005). A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(10), 1615-1630.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)