Real-Time Mosaic Method of Aerial Video Based on Two-Stage Key Frame Selection Method

Minwen Yuan, Yonghong Long, Xin Li

College of Railway Transportation, Hunan University of Technology, Zhuzhou, China
Email: 13554723580@163.com

Abstract

A two-stage automatic key frame selection method is proposed to enhance stitching speed and quality for UAV aerial videos. In the first stage, to reduce redundancy, the overlapping rate of the UAV aerial video sequence within the sampling period is calculated. Lagrange interpolation is used to fit the overlapping rate curve of the sequence. An empirical threshold for the overlapping rate is then applied to filter candidate key frames from the sequence. In the second stage, the principle of minimizing remapping spots is used to dynamically adjust and determine the final key frame close to the candidate key frames. Comparative experiments show that the proposed method significantly improves stitching speed and accuracy by more than 40%.

Keywords

UAV Aerial Video, Image Stiching, Key Frame Selection, Overlapping Rate, Remap Error

1. Introduction

Due to its fast, flexible, and convenient operation along with high image resolution, UAV aerial photography technology has become widely used as it continues to popularize and develop [1] [2]. Limited by the flight altitude and the camera's field of view, a single image captured by a UAV often cannot fully cover the target area. Therefore, panoramic stitching of sequential images collected by computers is essential. Studying a fast stitching method for UAV videos offers significant practical value.

In the video, frames that meet criteria for overlap rate are called key frames [3]. They are often used to stitch images, reducing the number of spliced frames and improving computing efficiency [4]. Literature [5] proposed stitching frames in the video that meet specific overlapping criteria to enhance stitching efficien-

cy. Fangbing Zhang *et al.* [6] calculated the frame-by-frame overlap rate used for key frame extraction. Common key frame extraction methods must reference all frames to calculate the overlap rate, making them unsuitable for applications requiring high timeliness. Fadaeieslam *et al.* [7] used a Kalman filter [8] to predict the trajectory of image corners and calculate the overlap between adjacent frames. However, error accumulation makes it difficult to accurately locate image corners, affecting key frame extraction accuracy. Liu Shanlei *et al.* [9] estimated the theoretical overlap between frames using camera's prior knowledge and extracted key frames at fixed intervals. This method fails when such prior knowledge is unavailable, as it cannot extract key frames meeting the overlap criteria. Liu Yong [10] attempted to understand the changing overlap rate of frames over a certain video sequence and select key frames accordingly. This approach requires searching the entire video sequence and establishing a piecewise linear model for the overlap rate, making it unsuitable for real-time video. There is a need for an adaptive key frame extraction algorithm for real-time aerial video.

Simply using the overlap rate threshold cannot guarantee the video splicing effect, as the remapping error between spliced images significantly affects the splicing quality. Existing methods face several problems:

1) Calculating the overlap rate frame by frame is time-consuming, as the rate between adjacent video frames typically exceeds 95%. Direct application of image stitching methods to video can significantly reduce efficiency. To enhance computational efficiency, selecting frames with a suitable degree of overlap for splicing is crucial. Calculating the overlap rate for each frame against reference images is not only lengthy but also impractical for time-sensitive applications.

2) The error in splicing key frames selected solely based on the overlap rate is significant: relying only on the overlap rate cannot ensure a successful video splicing outcome. In aerial videos featuring densely packed buildings or large featureless areas like rivers, the increase in mismatched feature points often leads to splicing deformations and gaps.

In this paper, we propose a two-stage key frame selection strategy that combines key frame rate fitting and key frame remapping error to address the issues of stitching efficiency and accuracy:

1) The overlapping rate is fitted using the Lagrangian interpolation method, and candidate key frames are identified using an empirical threshold. This approach addresses the issue of excessive calculation time for the overlap rate in large volumes of aerial video sequence data when selecting key frames frame by frame.

2) Furthermore, we can identify the key frame by remapping error. We fix the issue of holes and deformation in the panorama resulting from inaccurate or mismatched feature points in overlapping areas between adjacent key frames.

The remaining sections are organized as follows: Section 2 outlines the overall framework for real-time video stitching using a two-stage key frame selection

method. Section 3 details the testing process and analyzes the results. Section 4 concludes the paper.

2. Real-Time Video Stitching Based on Two-Stage Key Frame Selection Method

2.1. Overall Flow of Aerial Video Real-Time Mosaic Framework

Figure 1 presents the framework of a real-time video splicing system based on a two-phase key frame selection method, consisting of key frame selection and splicing fusion phases.

During the key frame selection stage, key frames are initially selected by fitting an overlap rate curve between subsequent video sequences and the current key frame using Lagrange polynomials. Then, key frames are further refined by assessing remapping errors. Finally, the refined key frames are stitched together to create a panoramic view of the aerial video.

2.2. General Flow of Two-Stage Key Frame Selection Methods

The proposed UAV aerial image stitching method is outlined in Figure 2. To enhance splicing efficiency and generate a key frame list, we introduce a key frame selection technique that utilizes Lagrangian interpolation and remapping error. This method filters key frames from the video sequence for splicing, based on two experimental thresholds, and automatically adds the video's first image as the initial key frame. In this selection process, we first fit the overlap rate curve between the subsequent video sequence and the current key frame using Lagrange polynomials, setting an overlap rate threshold. The last image in the sequence exceeding this threshold is considered a candidate key frame. We then calculate the remapping error between this candidate and the current key frame. If the error is below a certain threshold, the candidate is added to the key frame list as the newest key frame. Otherwise, we calculate the remapping error in reverse order from the current to the candidate key frame until an image meeting the error threshold is found and added to the list as the newest key frame.
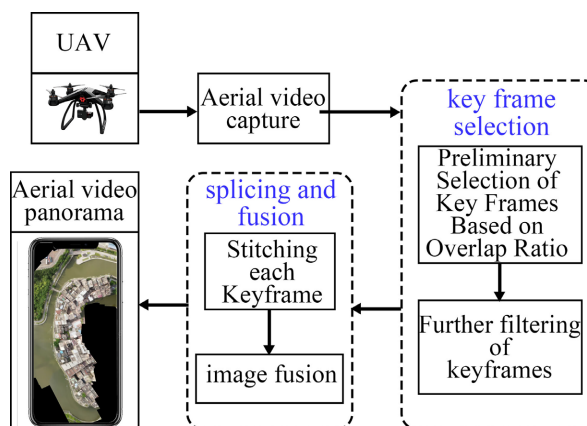

Figure 1. A splicing system framework based on a two-stage key frame selection approach.
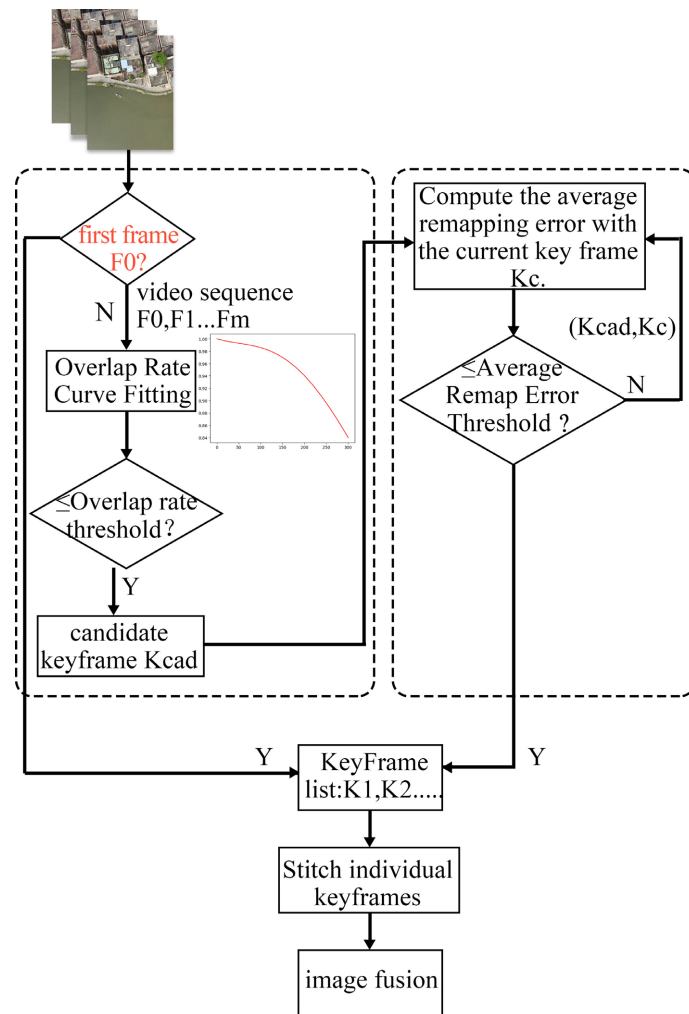
Figure 2. A flowchart of real-time video splicing method based on two-stage key frame selection method.

2.2.1. Improve the Speed of Detecting Key Frames with Overlap Rate Fitting

Inter-frame overlap is a common criterion for selecting key frames, which involves calculating the similarity transformation relationship between two images by identifying matching points between the current frame and the reference frame. This process determines the overlapping region between two frames to calculate the overlap rate [11]. However, due to the high redundancy in neighboring video frames, calculating matching points frame by frame to determine the overlap rate is inefficient. While the UAV flight path is generally fixed, leading to a nearly uniform change in overlap between aerial video images, airflow perturbation can shift the geometric position between neighboring frames, altering their overlap. Despite these perturbations, the high-frequency video acquisition allows for the geometric position change between adjacent frames to remain relatively stable. A mathematical model can thus describe the rule of change in the overlap between frames.

Lagrange interpolation is a high-precision method that produces smooth, os-

cillation-free interpolation results, accurately describing the functional relationship between data points. Its calculation formula is simple, making it easy to implement and particularly suitable for data interpolation requiring smooth outcomes [12]. This paper employs Lagrange interpolation to fit the inter-frame overlap.

Define set F as the video sequence, $F = \{F_1, F_2, \cdots, F_m\}$, K as the list of key frames selected from F, $K = \{K_0, K_1, \cdots, K_n\}$, where $n < m$. Initially, frame 0 of the video sequence is taken as the first key frame K_0, and we also considered the current key frame K_c, where $K_0 = K_c = F_0$. Subsequently, 4 frames are selected from F to calculate the overlap rate between these frames and the current key frame K_c The selection of these 4 frame images is defined in Equation (1), where S_1, S_2, S_3, S_4 are the frame sequence of images with subscripts, $S_1, S_2, S_3, S_4 \in \{1, 2, \cdots, m\}$. $K_c = F_{s0}$, where S_1 is randomly selected.

$$\begin{cases} S_2 = 2S_1 \\ S_3 = S_2 + S_1 \\ S_4 = 2S_2 \end{cases} \tag{1}$$

After calculating the overlap rate y_i between the 4-frame image and the current key frame K_c, a Lagrange polynomial (see Equation (2)) is used to fit the resulting four overlap rates y to the corresponding sequence subscripts S. Here, x_i represents the frame sequence indexes, *i.e.*, S_1, S_2, S_3, S_4 in Equation (1).

$$L_n(x) = \sum_{i=0}^{n} l_i(x) y_i$$

where

$$l_i(x) = \prod_{\substack{j \neq i \\ j=0}}^{n} \frac{x - x_j}{x_i - x_j} \tag{2}$$

Figure 8 displays the change in overlap rate of UAV aerial video sequences, modeled with a Lagrange polynomial. The red curve represents the modeled overlap rate, while the green curve shows the actual overlap rate between the current frame and K_c. When the distance between the current frame and K_c exceeds 300 frames, the overlap rate experiences significant fluctuations, indicating a reduction in the overlap area between frames, making accurate overlap rate derivation and subsequent splicing unfeasible. Therefore, we take S_0, S_1, S_2, S_3, S_4 within 300 frames from K_c, and we choose the interval between S_1 and S_0 to be 75 frames in this paper.

The actual overlap rate decreases with a larger frame interval and fluctuates towards the end. Thus, a range of 300 frames, starting from the index where K_c is located, is examined in ascending order along the fitted overlap rate curve. An overlap rate threshold of $T = 80\%$ is established. The frame that matches 80% of the fitted overlap rate with K_c is selected as the next candidate key frame K_{cad}:

$$T = 0.8$$
$$K_{cad} = F_s, \text{where } S = L^{-1}\left(T \bullet (\max y_i)\right) \tag{3}$$

Using only the selected key frames with a specific overlap rate, the remapping error between neighboring key frames varies around the median remapping error, as illustrated in Figure 4. Adjusting the overlap rate threshold T changes the median remapping error among neighboring images in the list of selected candidate key frames, showing an inverse relationship with the overlap rate threshold. This relationship is depicted in Figure 5, which demonstrates how different overlap rate thresholds, ranging from 0.6 to 0.95, affect the average remapping error between neighboring key frames in the selected group. In this figure, the horizontal axis represents the overlap rate threshold, and the vertical axis shows the average remapping error.

Splicing remains rough when based solely on key frames determined by the overlap rate threshold. Therefore, in addition to fitting the video overlap rate using the Lagrangian interpolation method and selecting key frames through the overlap rate threshold, it is necessary to further determine the optimal key frames.

2.2.2. Controlling Remapping Errors to Improve Splicing Accuracy

The current key frame K_c and the frame F_i detect feature points within an interval of i frames. We match these feature points and screen the matching point pairs, use the RANSAC algorithm [13] to remove mismatched points, ensuring accurate matching and selection of the best matching pairs. We identify the coordinates of each best matching point pairs from the current key frame K_c and frame F_i, and we perform the computation of the remapping coordinates. Assume that the key point coordinates of the current key frame K_c in the best matching point pair are (x, y) and the coordinates of the feature point in frame F_i that matches the key point (x, y) in the key frame K_c are (u, v). The relationship between the key point (x, y) and the remapped coordinates (x', y') of the point in frame F_i is shown in equation (4):

$$s\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{4}$$

Define $\sqrt{(u-x')^2 + (v-y')^2}$ as the remapping error of the key point (x, y). Then the average remapping error of the current key frame K_c with all matching points of frame F_i is:

$$\text{mean\_error} = \frac{\sum_1^n \sqrt{(u-x')^2 + (v-y')^2}}{n} \tag{5},$$

where n is the number of matching point pairs between the current key frame K_c and frame F_i. The remapping error threshold, $T = 4$ pixels, is established based on experimental findings. When a new candidate key frame K_{cap} is acquired during the fitting stage, the average remapping error, mean_error, between K_{cap} and K_c is calculated. If mean_error $\leq$ 4 pix, K_{cap} is added to the key frame list as the latest key frame. Otherwise, starting from K_{cap} to the cur-

rent key frame K_c, we calculate the remapping error mean_error between the current frame and K_c frame by frame until mean_error ≤ 4 pix is satisfied, we and store the current frame into the key frame list as the latest key frame.

2.3. Key Frame Splicing and Fusion

The selected key frame images are stitched together to create a panoramic image through a process that includes feature extraction, feature matching, solving the single-stress transformation matrix [14] and image fusion. Initially, feature points are extracted from each key frame image based on the camera pose. Next, these extracted feature points are matched with corresponding points in other images, using the RANSAC algorithm [15] to enhance the accuracy of feature point matching by addressing issues like noise and mismatches. Subsequently, a uni-responsive transformation matrix is determined from the detected feature points at different scales, establishing the spatial relationship between the images to be stitched. Finally, the Laplace pyramid image fusion algorithm [16] is employed to seamlessly blend the boundaries between images, correcting positional offsets, lens distortions, and luminance differences to ensure a clear image boundary (Figure 3).

Some key frames
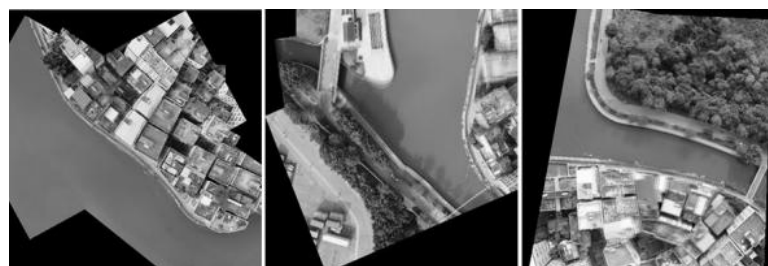
Partial mosaic
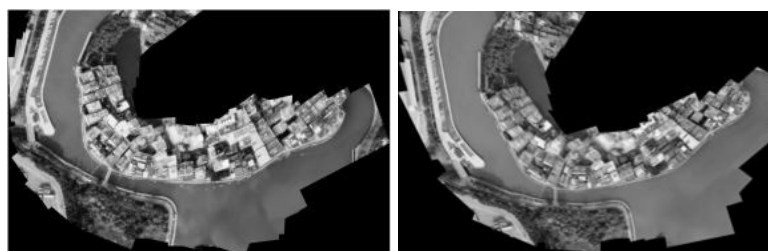
Laplacian pyramid fusion effect and weighted average fusion

Figure 3. Image stitching effect based on two-stage key frame filtering method.

3. Experimental Results

Using only the key frames selected with the overlap rate, the remapping error between neighboring key frames fluctuates above and below the median remapping error, as shown in **Figure 4**. When the overlap rate threshold T is set to a different value, the median remapping error between neighboring images in the list of selected candidate key frames is inversely proportional to the overlap rate threshold, as illustrated in **Figure 5**.

3.1. Overlap Rate Threshold Accuracy Analysis

Figure 6 displays the overlap rate statistics for an image sequence within 400 frames of the current key frame. The horizontal axis represents the frame index, while the vertical axis shows the overlap rate. The initial point on the horizontal axis corresponds to the frame index of the current key frame. When the overlap rate falls below 60%, the error becomes too large for accurate calculation, leading to significant fluctuations in the overlap rate curve. Consequently, this study conducts experiments on an overlap rate range of 60% - 95%, with intervals of 5%.
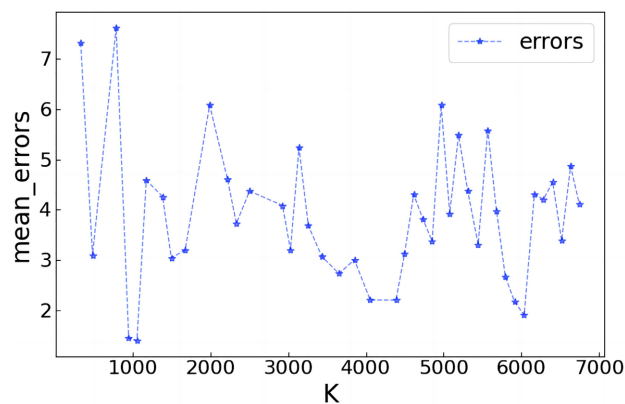

Figure 4. Average remapping error between key frames using overlap rate threshold selection.
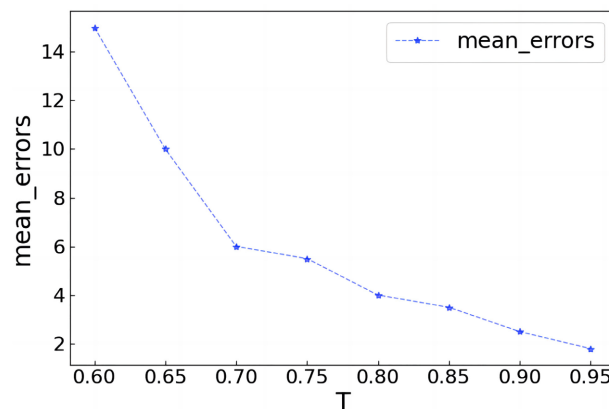

Figure 5. The average remapping error of the key frame for T ∈ [0.6 - 0.95].
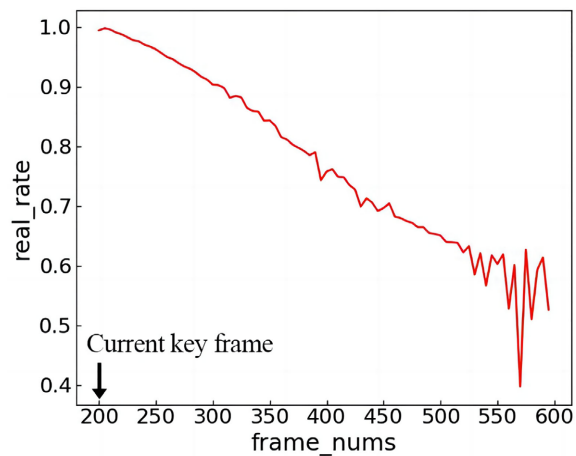
Figure 6. Overlap rate of the image sequence with the current key frame.

Figure 7 presents a comparison of average remapping errors between neighboring key frames at various overlap rate thresholds, using experimental data from a 3-minute and 50-second aerial video measuring 500 × 255 pixels. The horizontal axis represents the frame indexes in the list of key frames K, while the vertical axis shows the average remapping error values. The solid line indicates the median remapping error.

From **Figure 7**, it is observed that an overlap rate threshold of 60% or 65%, while reducing the number of key frames and splicing time, results in an excessive average remapping error, with the largest average error reaching 40 pixels. This means the error between each matched point pair of adjacent key frames averages 40 pixels, with a median error of approximately 15 pixels across all key frames. Such errors lead to noticeable misalignments post-splicing. However, setting the overlap rate threshold to 90% or 95% significantly reduces the remapping error, with a median average error of about 1.8 pixels, but this increases the number of key frames drastically, leading to longer splicing times. With an 80% overlap rate threshold, the median average remapping error is around 4 pixels, and the number of key frames is 42, compared to 172 at 95%. The choice of overlap rate threshold at 80% maintains a higher accuracy than at 60%, with fewer key frames and a manageable average remapping error.

Figure 8 shows the actual overlap rate curve and the fitted overlap rate curve at an 80% overlap rate, with the red line representing the curve fitting effect. The trajectory and the actual overlap rate (green line) within 300 frames from the current key frame show a consistent pattern, indicating that it is effective to use Lagrange polynomials for estimating the inter-frame overlap rate within a certain range. Therefore, setting the overlap rate threshold at 80% and the remapping error threshold at 4 pixels is the optimal choice for balancing splicing accuracy and time efficiency.

3.2. Comparison of Splicing Speed of Different Methods

The experiments conducted in this study were performed using the Python

T = 0.6, 20 key frames

T = 0.65, 24 key frames

T = 0.7, 28 key frames

T = 0.75, 34 key frames

T = 0.8, 42 key frames

T = 0.85, 57 key frames

T = 0.9, 87 key frames
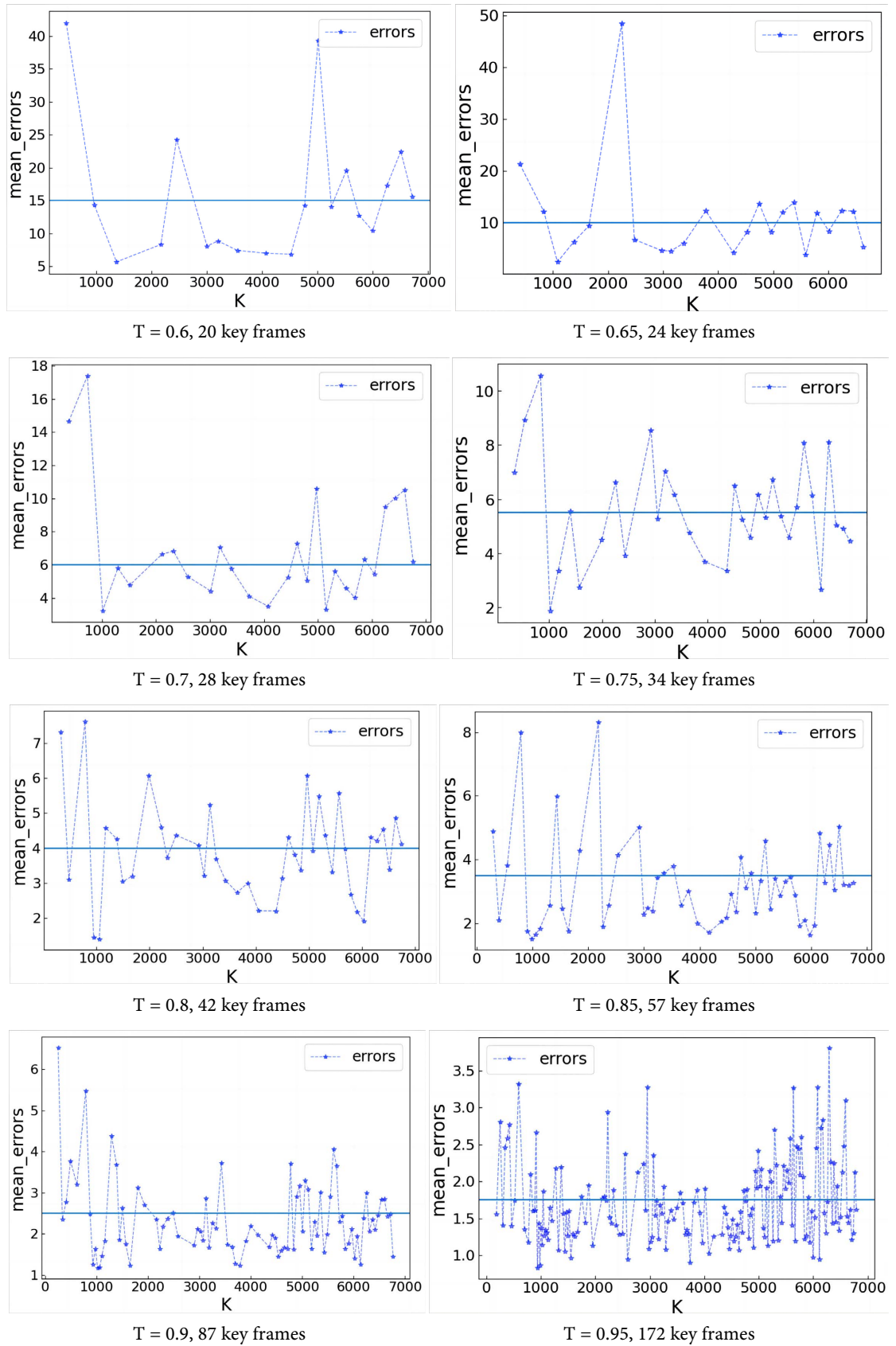
T = 0.95, 172 key frames

Figure 7. Average remapping error between key frames for different overlap rate thresholds T.
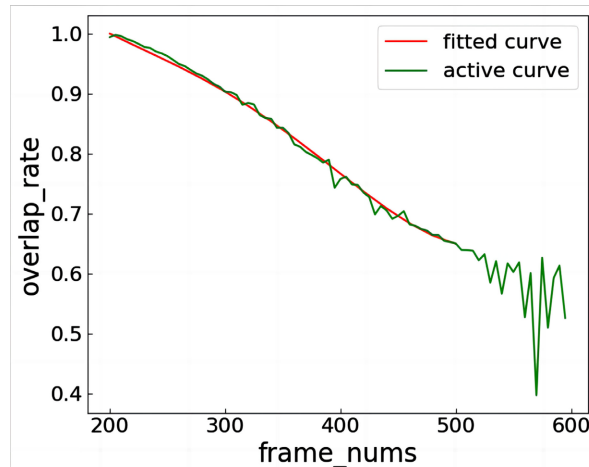
Figure 8. Interpolation of the fitted curve to the actual over-lap rate curve.

platform. The proposed method was compared with five other splicing methods: IORTI [6], inter-frame differencing, NISwGSP [17], and HQPI [18], focusing on key frame extraction and splicing speed. The key frame insertion criteria for IORTI involve two conditions: first, the number of in-points where the current frame matches the latest key frame (N) must be less than a specified value (N1), and second, the overlapping area ratio (P) between the current frame and the latest key frame must exceed a threshold value (P1). For IORTI, the parameters are set as N1 = 300 and P1 = 0.75. A current frame is inserted as the latest key frame in the key frame list when both conditions are met.

Table 1 demonstrates that in the comparison methods, key frame extraction significantly impacts the program's running time. During this phase, IORTI [6] takes the longest, followed by the inter-frame difference method, while the method presented in this paper is the fastest. The other two methods lack a key frame extraction phase. Specifically, the method in this paper is 49% faster than the inter-frame differencing method and 93% faster than IORTI in key frame extraction. In the splicing stage, both NISwGSP and HQPI methods were un-successful. Overall, the method in this paper achieves a 39% and 91% improvement in total video splicing speed compared to the inter-frame differencing method and IORTI, respectively, significantly enhancing the operational efficiency of UAV aerial video splicing.

3.3. Comparison of Splicing Accuracy of Different Methods

The key frames selected by different methods vary. To quantitatively evaluate the splicing accuracy of these key frames, the Root Mean Square Error (RMSE) is used as a metric. RMSE, defined in Equation (6), quantitatively evaluates the splicing effect:

$$\text{RMSN} = \sqrt{\frac{\sum_{1}^{m}\sum_{1}^{n}\left[P(x',y')-Q(u,v)\right]^2}{m\times n}} \tag{6}$$

The remapped coordinates of feature point $P(x, y)$ are $P(x', y')$, and the matching feature point P in the neighboring key frames is $Q(u, v)$. Here, n represents the number of excellent matching point pairs between two neighboring key frames, and m represents the number of key frames. The methods NISwGS and HQPI are not included due to their splicing failures. According to **Table 2**, compared with the IORTI and inter-frame differencing methods, the proposed method has the lowest RMSE value at 15.92 pixels, improving accuracy by 13% and 41% over inter-frame differencing and IORTI, respectively.

3.4. Comparison of Splicing Results

The NISwGSP and HQPI splicing methods both failed. **Figure 9** presents the effects and local detail magnifications of key frame selection methods, including inter-frame differencing, IORTI, and key frame selection based on Lagrangian interpolation and remapping error. It also shows the effects and local detail magnifications of key frames selected by these three methods on the same city and waterside aerial video sequences. All three methods utilize the SIFT algorithm [19] for feature extraction [20] and the Laplace pyramid image fusion algorithm for splicing. SIFT features are local image features that are multi-volume, unique, and information-rich. They remain invariant to scale, rotation, and luminance changes, and maintain some degree of invariance to radial transformation, point-of-view changes, image noise, etc. [21].

The figure illustrates noticeable gaps in the spliced image of the key frame list chosen by IORTI, as highlighted in **Figure 9**, primarily due to feature matching errors. In contrast, the key frame sequence spliced using the inter-frame difference method shows deformation in parts of the image because of inaccurate feature point matching in the overlap area or mismatches, as evident in the zoomed-in picture where the car appears noticeably stretched. The method proposed in this paper effectively avoids these issues by considering the remapping error between neighboring key frames used for splicing, which result in high-quality images with clear details and a splicing effect that integrates all parts of images.

Table 1. Splicing speed and key frame selection speed for different methods.

method	inter-frame difference method	NISwGSP, HQPI	IORTI	Our method
Key frame extraction	104.1 s	/	842.5 s	52.8 s
total time spent	128.8 s	Fail	859.7 s	78.2 s
speed boost	39%	/	91%	/

Table 2. Comparison of RMSE values of key frame sequences of different methods.

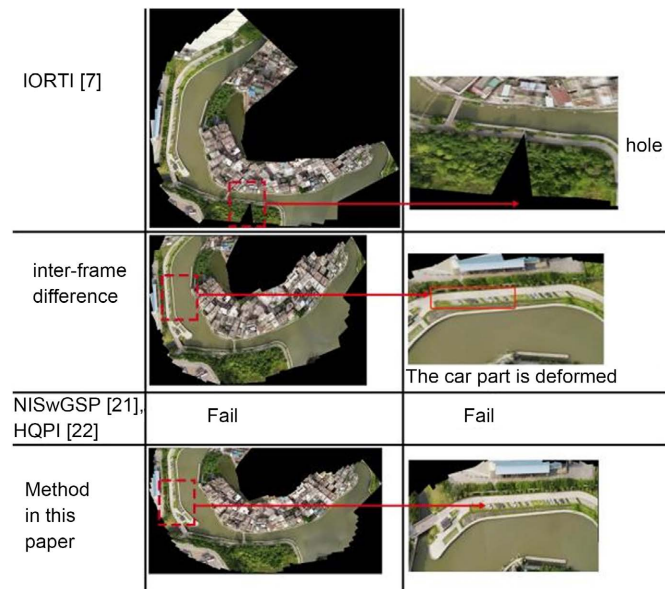| method	inter-frame difference method	NISwGS, HQPI	IORTI	Our method
RMSE	18.25 pix	/	26.77 pix	15.92 pix
Accuracy improvement	13%	/	41%	/

Figure 9. Effect of different methods of splicing.

4. Conclusion

In this paper, we introduce a rapid splicing algorithm for UAV aerial videos, leveraging a key frame selection method combined with Lagrangian interpolation and remapping error analysis. The process of selecting key frames involves two phases. Firstly, candidate key frames are identified by fitting the overlap rate curve between subsequent video sequences and the current key frame using Lagrange polynomials. Subsequently, the most recent key frame is chosen by calculating the remapping error between it and the candidate key frames. This method enhances splicing speed while maintaining quality. Compared to key frame selection methods based on inter-frame differencing and IORTI, our approach improves accuracy by 13% and 41%, respectively, and reduces total splicing time by 39% and 91%, achieving the required balance of accuracy and speed for video splicing.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Li, X.L. and Ling, C.Q. (2022) Application of UAV Remote Sensing Technology in Agricultural Conditions Monitoring. *Modern Agricultural Equipment*, **43**, 45-51.

[2] Liu, Z., Wan, W., Huang, J.Y., *et al.* (2018) Research Progress on Inversion of Key Parameters of Crop Growth Based on UAV Remote Sensing. *Journal of Agricultural Engineering*, **34**, 60-71.

[3] Luo, Y., Li, Y., Li, Z., *et al.* (2021) MS-SLAM: Motion State Decision of Key Frames for UAV-Based Vision Localization. *IEEE Access*, **9**, 67667-67679.
https://doi.org/10.1109/ACCESS.2021.3077591

[4] Zhao, Y., Chen, L., Zhang, X., *et al.* (2021) RTSfM: Real-Time Structure from Motion for Mosaicing and DSM Mapping of Sequential Aerial Images with Low Overlap. *IEEE Transactions on Geoscience and Remote Sensing*, **60**, 1-15.

https://doi.org/10.1109/TGRS.2021.3090203

[5] Wang, Z. and Zhu, Y. (2020) Video Key Frame Monitoring Algorithm and Virtual Reality Display Based on Motion Vector. *IEEE Access*, **8**, 159027-159038.

[6] Zhang, F., Yang, T., Liu, L., *et al.* (2020) Image-Only Real-Time Incremental UAV Image Mosaic for Multi-Strip Flight. *IEEE Transactions on Multimedia*, **60**, 1410-1425. https://doi.org/10.1109/TMM.2020.2997193

[7] Fadaeieslam, M.J., Soryani, M. and Fathy, M. (2011) Efficient Key Frames Selection for Panorama Generation from Video. *Journal of Electronic Imaging*, **20**, 2763-2769. https://doi.org/10.1117/1.3591366

[8] Dong, J. and Liu, H. (2017) Video Stabilization for Strict Real-Time Applications. *IEEE Transactions on Circuits and Systems for Video Technology*, **27**, 716-724. https://doi.org/10.1109/TCSVT.2016.2589860

[9] Liu, S.L., Zhao, Y.D., Wang, G.H., *et al.* (2012) An Automatic Key Frame Extraction Method. *Surveying and Mapping Science*, **37**, 110-112+115.

[10] Liu, Y., Wang, G.J., Yao, A.B., *et al.* (2010) Video Stitching Based on Adaptive Frame Sampling. *Journal of Tsinghua University* (*Science and Technology*), **50**, 108-112.

[11] Ren, C.F. (2014) Research on Key Technologies of Orthophoto Production of Aerial Video Images. Wuhan University, Wuhan.

[12] Liu, J.J. (2018) GPS Satellite Orbit Position Fitting Based on Lagrangian Interpolation Method. *Science and Technology Innovation and Productivity*, **294**, 19-21.

[13] Yang, L., Cao, J., Tang, L., *et al.* (2014) Optimized Design of Automatic Panoramic Images Mosaic. *Multimedia Tools and Applications*, **72**, 503-514. https://doi.org/10.1007/s11042-013-1387-y

[14] Yong, H., Huang, J., Xiang, W., *et al.* (2019) Panoramic Background Image Generation for PTZ Cameras. *IEEE Transactions on Image Processing*, **28**, 3162-3176. https://doi.org/10.1109/TIP.2019.2894940

[15] Zheng, J., Peng, W., Wang, Y., *et al.* (2021) Accelerated RANSAC for Accurate Image Registration in Aerial Video Surveillance. *IEEE Access*, **9**, 36775-36790. https://doi.org/10.1109/ACCESS.2021.3061818

[16] Huang, F.S. and Lin, S.Z. (2019) Multi-Band Image Fusion Rules Comparison Based on the Laplace Pyramid Transformation Method. *Infrared Technology*, **41**, 64-71.

[17] Cheng, Y.S. and Chuang, Y.Y. (2016) Natural Image Stitching with the Global Similarity Prior. *Computer Vision-ECCV* 2016, Amsterdam, 11-14 October 2016, 186-201. https://doi.org/10.1007/978-3-319-46454-1_12

[18] Xiong, Y. and Pulli, K. (2010) Fast Panorama Stitching for High-Quality Panoramic Images on Mobile Phones. *IEEE Transactions on Consumer Electronics*, **56**, 298-306. https://doi.org/10.1109/TCE.2010.5505931

[19] Liu, Y., He, M., Wang, Y., *et al.* (2022) Farmland Aerial Images Fast-Stitching Method and Application Based on Improved SIFT Algorithm. *IEEE Access*, **10**, 95411-95424. https://doi.org/10.1109/ACCESS.2022.3204657

[20] Chang, H.H., Wu, L.G., *et al.* (2019) Remote Sensing Image Registration Based on Modified SIFT and Feature Slope Grouping. *IEEE Geoscience and Remote Sensing Letters*, **16**, 1363-1367. https://doi.org/10.1109/LGRS.2019.2899123

[21] Xiang, Y., Wang, F. and You, H. (2018) OS-SIFT: A Robust SIFT-Like Algorithm for High-Resolution Optical-to-SAR Image Registration in Suburban Areas. *IEEE Transactions on Geoscience and Remote Sensing*, **56**, 3078-3090. https://doi.org/10.1109/TGRS.2018.2790483