A UAV Aerial Image Stitching Method Based on Quality Assessment (Postprint)

Authors: Wang Huan, Zhou Yingyue, Wang Xueyuan, Wang Xinyu

Date: 2022-04-07T00:00:00+00:00

Abstract

To address the issues of misalignment and artifacts in UAV aerial image stitching under large parallax, we propose a method called QEB-U (Quality Evaluation Based for UAV) to find the optimal seam. First, an initial seam is obtained through conventional seam estimation. Then, according to the characteristics of UAV aerial images, a novel quality evaluation function is proposed that comprehensively considers structural similarity, color difference, and texture complexity to evaluate each pixel on the seam. Based on the evaluation results, the difference cost is updated, and the seam is re-estimated. This estimation and evaluation process is repeated iteratively until the seam stabilizes. Finally, the result is generated through gradient fusion. Experimental results demonstrate that the proposed method can effectively avoid misalignment and artifacts in UAV aerial image stitching with large parallax, outperforming several existing UAV aerial image stitching methods. Furthermore, the resulting seam preferentially traverses regions such as roads and woodlands, which better aligns with human visual perception, and shows favorable performance on common image sharpness evaluation metrics.

Full Text

Quality Evaluation-Based Image Stitching for UAV Images

Wang Huan[1], **Zhou Yingyue**[1,2]**†**, **Wang Xueyuan**[1], **Wang Xinyu**[1]
[1]School of Information Engineering, [2]Sichuan Provincial Key Laboratory of Robot Technology Used for Special Environment, Southwest University of Science & Technology, Mianyang, Sichuan 621010, China

Abstract

Aiming to address dislocation and artifact problems in stitching large-parallax UAV aerial images, this paper proposes a method called QEB-U (Quality

Evaluation-Based image stitching for UAV images) to find the optimal stitching seam. First, an initial seam is obtained through conventional seam estimation. Then, according to the characteristics of UAV aerial images, a novel quality evaluation function is proposed to assess each pixel on the seam by comprehensively considering structural similarity, color difference, and texture complexity. Based on the evaluation results, the difference cost is updated and the seam is re-estimated. This estimation and evaluation process repeats until the seam stabilizes, after which the final result is generated through gradient fusion. Experimental results demonstrate that the proposed method can avoid dislocation and artifacts in large-parallax UAV aerial image stitching, outperforming several current UAV image stitching methods. Moreover, the obtained stitching seam preferentially passes through roads, woodlands, and other similar regions, better conforming to human visual perception, and performs well on common image clarity evaluation metrics.

Key words: image stitching; seam estimation; quality evaluation; UAV images

0 Introduction

With the development of UAV remote sensing technology, its research has been widely applied in fields such as geoinformatics, disaster monitoring and assessment, and national defense, particularly in disaster emergency response and investigation. UAV remote sensing can obtain the most intuitive high-resolution images of target areas within a short time, providing visual evidence for decision-making. However, due to limitations in UAV flight altitude and camera performance, it is difficult to capture the entire scene of a region of interest from a single UAV aerial image. Therefore, it is necessary to study techniques for rapidly stitching overlapping UAV aerial images into high-resolution images with a larger field of view.

Image stitching generally involves two approaches: one is alignment technology based on homography transformation to achieve the most accurate possible alignment; the other is seam-guided cutting technology, which seeks a seam that minimizes a predefined energy function. Alignment methods based on homography transformation ultimately aim to align images as precisely as possible, but such methods produce local artifacts or distortions when encountering large-parallax environments. The seam-driven approach to image stitching was first proposed by Gao et al., who used the RANSAC algorithm to obtain several homography matrices, calculated a stitching seam for each, and then selected the best seam based on quality evaluation results. Lin et al. tightly coupled local alignment computation and seam estimation through adaptive feature weighting and effectively reduced iteration time through superpixel feature grouping. Li et al. incorporated human perception nonlinearity and non-uniformity into energy minimization, using saliency weights to simulate objects that the human eye tends to focus on, making stitching results more consistent with human visual perception. Liao et al. noted that previous seam-driven methods evaluated the average performance of seams and could not find seams optimal for human visual perception. Therefore, they evaluated each pixel on the seam more meticulously and introduced a hybrid cost to assess pixels along the seam, then re-estimated the seam based on evaluation results. Their method avoided artifacts and achieved better overall results. Yuan et al. proposed a superpixel-based stitching method for UAV aerial images, significantly improving stitching time and incorporating texture complexity information into the energy function to make seams pass through specific low-texture-complexity regions, making results more consistent with human visual perception. Fang et al. innovated in the fusion stage, proposing a color blending method to eliminate color inconsistency in stitching results and using superpixel segmentation to greatly reduce computational complexity and improve efficiency. Zhao et al. used multiple planes instead of mapping points for optimization, directly applying reconstruction results to image fusion to achieve a new aerial image stitching method with high accuracy and speed. Pham et al. first estimated overlapping regions based on aerial image geolocation information, then restricted feature matching to these regions to avoid unnecessary computation and error accumulation, making stitching more accurate and faster. Ren et al. used aerial image position information for coarse registration to calculate a global transformation matrix and corrected it quickly through a multilayer perceptron, thus effectively combining parametric and feature-based methods to achieve an ideal compromise between stitching speed and accuracy. Guo et al. proposed a global alignment strategy combined with local registration technology to achieve better alignment accuracy while preserving shape. Mo et al. used neural networks for deep feature extraction and matching to implement a new robust aerial image stitching method, providing assistance for other researchers exploring deep learning-based image stitching methods.

Seam-driven methods typically find an optimal stitching seam based on predefined seam quality metrics. However, the quality metrics defined in these methods assess overall seam quality. As demonstrated in the literature, in many cases, the seam with the best overall quality does not necessarily conform to human visual perception. This paper proposes a novel iterative seam estimation method for large-parallax UAV aerial images (QEB-U), where the iteration process is guided by each pixel on the seam. First, a conventional seam estimation obtains an initial seam. Then, according to UAV aerial image characteristics, a new quality evaluation function comprehensively considers structural similarity, color difference, and texture complexity to evaluate each pixel on the seam. Based on evaluation results, the difference cost is updated and the seam is re-estimated. This estimation and evaluation process repeats until the seam stabilizes, after which gradient fusion generates the final result. Experimental results show that the proposed method achieves good results for large-parallax UAV aerial image stitching, with obtained seams preferentially passing through roads, woodlands, and other regions, better conforming to human visual perception and performing well on common image clarity evaluation metrics.

1 Proposed Method

To find the optimal stitching seam conforming to human visual perception, we evaluate each pixel on the seam more meticulously and propose QEB-U, a new method for finding the best seam based on quality evaluation. This approach solves dislocation and artifact problems while making the seam more consistent with human visual perception.

1.1 Method Flow

The flowchart of our proposed method is shown in [Figure 1: see original paper]. First, conventional seam estimation is used to obtain an initial seam, after which quality evaluation is performed on the seam. Previous quality evaluation methods were based on overall seam quality, whereas our method evaluates each pixel on the seam. Pixels that do not meet expectations incur larger costs, while those that meet expectations have smaller costs. Pixels with larger costs should be avoided by the stitching seam, so subsequent seam re-estimation applies greater penalties to them in the energy function. Iteration stops when the seam stabilizes, and the final result is generated through gradient fusion. Key steps are detailed below.

1.2 Conventional Seam Estimation

Let I_0 and I_1 be two preliminarily registered images to be stitched, and $L \in \{0, 1\}$ a label set. For any pixel p in the overlapping region P, the meaning of stitching is to assign a label $l_p \in L$ to each pixel p. The value $l_p = 0$ means the pixel value is copied from I_0, while $l_p = 1$ means it is copied from I_1. Conventional seam searching is a process of energy function minimization, where the energy function is defined as:

$$E(l) = \sum_{p \in P} D_p(l_p) + \sum_{(p,q) \in N} S_{pq}(l_p, l_q)$$

where N represents the four-neighborhood and q is a neighboring pixel of p. The smoothness term S_{pq} for conventional seam estimation is defined as:

$$S_{pq}(l_p, l_q) = \|I_p - I_q\|^2 \text{ if } l_p \neq l_q, \text{ else } 0$$

where the difference cost $d(I_0, I_1) = \|I_0 - I_1\|$ represents Euclidean color difference. The data term $D_p(l_p)$ represents the cost of assigning label l_p to pixel p; further details can be found in [?]. Finally, the seam is estimated by minimizing the energy function via graph cuts.

1.3 Quality Evaluation and Cost Update

1.3.1 Quality Evaluation Combining UAV aerial image characteristics, this paper proposes a new quality evaluation function that comprehensively considers structural similarity, color difference, and texture complexity. First, unaligned parts in overlapping regions are usually structurally inconsistent, so the Structural Similarity Index (SSIM) is introduced into the quality evaluation function to assess the seam. Second, color difference between seam sides is introduced for more detailed seam evaluation. Finally, for UAV aerial images, seams conforming to human visual perception tend to pass through specific regions such as roads, woodlands, and lakes. Based on the fact that these regions have weak texture, texture complexity is introduced into quality evaluation, making pixels in low-texture-complexity regions have smaller costs to guide the seam through these areas. Considering that Gabor filters can well characterize basic texture information, Gabor feature descriptors are used to describe texture complexity. [Figure 2: see original paper] shows some visualization examples of Gabor features, where image brightness represents texture complexity.

For a pixel p_i on the seam, the hybrid cost $E^i(p)$ is defined as:

$$E^i(p) = E^i_{patch}(p) \cdot E^i_{point}(p) \cdot E^i_{gabor}(p) \cdot \lambda$$

where λ is a constant maintaining the overall magnitude. $E^i_{patch}(p) = 1 - SSIM(p_0, p_1)$ represents SSIM between local patches centered at pixel p_i in I_0 and I_1. $E^i_{point}(p) = \|I_0(p_i) - I_0(p)\|^2 + \|I_1(p_i) - I_1(p)\|^2$ represents color difference between pixel p_i and its neighboring pixel p. $E^i_{gabor}(p)$ represents the Gabor feature of pixel p_i.

1.3.2 Updating Difference Cost Typically, pixels on the seam that are unaligned or located in complex regions have large hybrid costs. Therefore, the difference cost $d_i(p)$ is corrected using the pixel' s hybrid cost $E^i(p)$ to obtain a new difference cost:

$$d'_i(p) = \begin{cases} f(E^i(p)) \cdot d_i(p) & \text{if } p \in N_{seam} \\ d_i(p) & \text{otherwise} \end{cases}$$

where N_{seam} is a band-shaped region containing the seam, generated by expanding 5 pixels to the left and right of the seam. σ is a constant controlling the magnitude of change, and ϵ is a reasonable cost threshold. When the hybrid cost $E^i(p)$ is small, the new difference cost $d'_i(p)$ also becomes smaller; if the hybrid cost is large, the new difference cost increases accordingly. The energy function is then recalculated using the new difference cost, and the seam is re-estimated via graph cuts.

The estimation and evaluation process repeats until the seam stabilizes. Here, "stable" means the current seam can be completely contained within the band-shaped region of the previous seam. After obtaining the optimal seam, the final result is generated through gradient fusion.

1.4 Algorithm Pseudocode

Algorithm 1: QEB-U - Input: Two preliminarily registered images I_0 and I_1 - **Output:** Final stitching seam

1. Initialize:
   - (a) Input images I_0 and I_1
   - (b) Set $S_B = \emptyset$ (best seam)
2. Compute difference cost matrix d_I and smoothness term S, and minimize energy function via graph cuts to obtain initial seam S_1
3. While not converged:
   - Expand seam S_1 to band-shaped region B and recompute difference costs d_I' for pixels in B
   - Compute new smoothness term S using d_I' and minimize energy function via graph cuts to obtain new seam S_2
   - If $S_2 \subseteq B_{S_1}$ (current seam stable), break
   - Update $S_1 = S_2$
4. Return final seam $S_B = S_2$

2 Experiments

To verify the effectiveness of our proposed method, we selected publicly available large-parallax UAV aerial images for algorithm performance testing, comprising 290 images from [?] and [?]. For optimal performance, parameters were set as follows: local patch size 21×21 when computing SSIM between patches, $\lambda = 10$ in Equation (4), and $\sigma = 5$, $\epsilon = 0.12$ in Equation (8). Input images were first preliminarily aligned using the method provided in [?], then our iterative seam estimation method was applied to obtain the final seam, with gradient fusion generating the final result. The algorithm was implemented on Matlab R2020a, running on a 2.9GHz eight-core CPU laptop.

2.1 Visual Comparison with Other Stitching Methods

We compared our method with APAP (as-projective-as-possible) [?], SPHP (shape-preserving half-projective) [?], AANAP (Adaptive As-Natural-As-Possible) [?], ELA (parallax-tolerant image stitching based on robust elastic warping) [?], and SPW (single-perspective warps in natural image stitching) [?].

Figure 3: see original paper shows a set of airport UAV aerial images with large parallax. Figure 3: see original paper-(f) display results from APAP, SPHP, AANAP, ELA, and SPW, respectively, while Figure 3: see original paper shows our algorithm' s result. Obvious artifacts can be seen in the wings and tails (red boxes) and clear boundaries (green boxes) in Figure 3: see original paper-(f), whereas our result in Figure 3: see original paper is clear, artifact-free, and more natural.

[Figure 4: see original paper] shows a set of house UAV aerial images with large parallax and results from several methods. Artifacts and distortion blur (red boxes) appear in Figure 4: see original paper-(f), with clear boundaries (green boxes) in Figure 4: see original paper(d)(e), while our result in Figure 4: see original paper is clear, natural, and free of artifacts and distortion.

2.2 Visual Comparison with Other Seam Search Algorithms

To verify the effectiveness of our seam search strategy, we compared it with recent seam search algorithms: Perception-based (PB) [?], Quality evaluation-based (QEB) [?], and Fast and robust (FARSE) [?].

As shown in [Figure 5: see original paper], red lines mark the optimal seam positions detected by each method. The seams found by PB, QEB, and FARSE pass through salient objects such as buildings (green boxes), while our method's seam avoids buildings and other prominent objects, preferentially passing through roads, woodlands, etc., making the stitching effect more consistent with human visual perception.

2.3 Efficiency Comparison with Other Seam Search Algorithms

We also compared stitching time with PB, QEB, and FARSE using three sets of high-definition aerial images with resolution 3680$×$2456 from [?]. shows the detailed data. For 4K high-resolution aerial image stitching, all methods (PB, QEB, FARSE, and ours) require over one minute. Our method' s efficiency is similar to PB and QEB but inferior to FARSE, indicating room for improvement. Future work will consider using superpixel segmentation to reduce iterative computation and improve efficiency.

2.4 Objective Metric Evaluation

For more precise and objective comparison of stitched image quality, we used common image clarity evaluation metrics: Brenner gradient function, Tenengrad gradient function, Laplacian gradient function, gray variance function, and energy gradient function. Larger values from these functions indicate higher clarity and better stitching quality.

The Brenner gradient function is defined as:

$$D_B(f) = \sum_x \sum_y [f(x+2, y) - f(x, y)]^2$$

where $f(x, y)$ represents the grayscale value of pixel (x, y).

The Tenengrad gradient function is defined as:

$$D_T(f) = \sum_x \sum_y [T(x, y)]^2 \text{ for } T(x, y) > T_0$$

where T_0 is a given edge detection threshold, and G_x and G_y are Sobel horizontal and vertical edge detection operator convolutions at pixel (x, y), with $T(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2}$.

The Laplacian gradient function is defined as:

$$D_L(f) = \sum_x \sum_y [L(x, y)]^2 \text{ for } L(x, y) > T_0$$

where $L(x, y)$ is the Laplacian operator convolution at pixel (x, y).

The gray variance function is defined as:

$$D_S(f) = \sum_x \sum_y \frac{|f(x, y) - f(x, y-1)| + |f(x, y) - f(x-1, y)|}{2}$$

The energy gradient function is defined as:

$$D_E(f) = \sum_x \sum_y \{[f(x, y) - f(x+1, y)]^2 + [f(x, y) - f(x, y+1)]^2\}$$

Clarity metrics for results in [Figure 3: see original paper] and [Figure 4: see original paper] were computed and linearly normalized, with results shown in and . Our algorithm achieves higher values than other methods on almost all metrics, indicating it produces clearer stitched images.

3 Conclusion

This paper proposes a quality evaluation-based strategy for finding the optimal stitching seam for UAV aerial images. It not only solves dislocation and artifact problems in large-parallax UAV aerial image stitching but also makes obtained seams preferentially pass through roads, woodlands, and other regions in complex UAV aerial images. Experiments demonstrate that our method is suitable for UAV aerial image stitching, producing results more consistent with human visual perception.

References

[1] Li Deren, Li Ming. Research advance and application prospect of unmanned aerial vehicle remote sensing system [J]. Wuhan University Journal: Geomatics and Information Science, 2014, 39(5): 505-513.

[2] Yuan Y, Fang F, Zhang G. Superpixel-based seamless image stitching for UAV images [J]. IEEE Trans on Geoscience and Remote Sensing, 2020, 59(2): 1565-1576.

[3] Zaragoza J, Chin T J, Brown M S, et al. As-projective-as-possible image stitching with moving DLT [C]// Proceedings of the IEEE conference on computer vision and pattern recognition. 2013: 2339-2346.

[4] Zhang F, Liu F. Parallax-tolerant image stitching [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014: 3262-3269.

[5] Chang C H, Sato Y, Chuang Y Y. Shape-preserving half-projective warps for image stitching [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014: 3254-3261.

[6] Lin C C, Pankanti S U, Natesan Ramamurthy K, et al. Adaptive as-natural-as-possible image stitching [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 1155-1163.

[7] Chen Y S, Chuang Y Y. Natural image stitching with the global similarity prior [C]// European conference on computer vision. Springer, Cham, 2016: 186-201.

[8] Gao J, Li Y, Chin T J, et al. Seam-driven image stitching [C]// Eurographics (Short Papers). 2013: 45-48.

[9] Lin K, Jiang N, Cheong L F, et al. Seagull: Seam-guided local alignment for parallax-tolerant image stitching [C]// European conference on computer vision. Springer, Cham, 2016: 370-385.

[10] Li N, Liao T, Wang C. Perception-based seam cutting for image stitching [J]. Signal, Image and Video Processing, 2018, 12(5): 967-974.

[11] Liao T, Chen J, Xu Y. Quality evaluation-based iterative seam estimation for image stitching [J]. Signal, Image and Video Processing, 2019, 13(6): 1105-1112.

[12] Fang F, Wang T, Fang Y, et al. Fast color blending for seamless image stitching [J]. IEEE Geoscience and Remote Sensing Letters, 2019, 16(7): 1075-1079.

[13] Zhao Y, Liu G, Xu S, et al. Fast georeferenced aerial image stitching with absolute rotation averaging and planar-restricted pose graph [J]. IEEE Trans on Geoscience and Remote Sensing, 2020, 59(4): 3502-3517.

[14] Pham N T, Park S, Park C S. Fast and efficient method for large-scale aerial image stitching [J]. IEEE Access, 2021, 9: 127852-127865.

[15] Ren M, Li J, Song L, et al. MLP-based efficient stitching method for UAV images [J]. IEEE Geoscience and Remote Sensing Letters, 2022.

[16] Guo D, Chen J, Luo L, et al. UAV image stitching using shape-preserving warp combined with global alignment [J]. IEEE Geoscience and Remote Sensing Letters, 2021, 19: 1-5.

[17] Mo Y, Kang X, Duan P, et al. A robust UAV hyperspectral image stitching method based on deep feature matching [J]. IEEE Trans on Geoscience and Remote Sensing, 2021, 60: 1-14.

[18] Pérez P, Gangnet M, Blake A. Poisson image editing [M]// ACM SIGGRAPH 2003 Papers. 2003: 313-318.

[19] Kwatra V, Schödl A, Essa I, et al. Graphcut textures: Image and video synthesis using graph cuts [J]. ACM Trans on Graphics (TOG), 2003, 22(3): 277-286.

[20] Boykov Y, Kolmogorov V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2004, 26(9): 1124-1137.

[21] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE Trans on Image Processing, 2004, 13(4): 600-612.

[22] Gupta R K, Chia A Y S, Rajan D, et al. Image colorization using similar images [C]// Proceedings of the 20th ACM international conference on Multimedia. 2012: 369-378.

[23] Manjunath B S, Ma W Y. Texture features for browsing and retrieval of image data [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1996, 18(8): 837-842.

[24] https://github.com/YuhuaXu/UAV-image-mosaicing-dataset

[25] Li J, Wang Z, Lai S, et al. Parallax-tolerant image stitching based on robust elastic warping [J]. IEEE Trans on Multimedia, 2017, 20(7): 1672-1687.

[26] Liao T, Li N. Single-perspective warps in natural image stitching [J]. IEEE Trans on Image Processing, 2019, 29: 724-735.

[27] Hejazifar H, Khotanlou H. Fast and robust seam estimation to seamless image stitching [J]. Signal, Image and Video Processing, 2018, 12(5): 987-994.

Note: Figure translations are in progress. See original paper for figures.