Improvement and Analysis of Panoramic Image Mosaic Technology Based on Mixed Scene

Yuhua Pei

Electronic Engineering College, Heilongjiang University, Harbin, Heilongjiang, 150006, China
20211353@s.hlju.edu.cn

Abstract. Given how quickly augmented reality (AR) and virtual reality (VR) technologies are developing, panoramic image stitching technology is playing an increasingly important role in providing immersive experiences. Especially in complex scenes where natural and urban environments are interwoven, high-quality panoramic image stitching technology is indispensable for creating lifelike virtual experiences. Aiming at the Mosaic problem of panoramic images in mixed scenes, two feature extraction algorithms—Accelerated Robust Feature (ORB) and Scale Invariant Feature Transform (SIFT)—were contrasted in various scenarios. Through experimental design, scenes spanning different distance ranges are selected as test objects to evaluate the stitching effect of the two algorithms under changing the shooting Angle. Moreover, the accuracy of SIFT algorithm in pre-processing image is enhanced according to the existing method, that is, the local contrast of a picture is enhanced by applying the contrast limited adaptive histogram equalization (CLAHE). The selection of images and code implementation in the panoramic image Mosaic of mixed scene are analyzed by the results of the last experiment.

Keywords: Panoramic Image, Mixed Scene, Feature Extraction Algorithm

1 Introduction

With the rapid advancement of virtual reality (VR) and augmented reality (AR) technologies, achieving seamless panoramic image mosaics has become indispensable for delivering an immersive user experience. This necessity is particularly evident in complex scenes where natural and urban landscapes intertwine, demanding flawless image connectivity and adaptability to challenges such as occlusion, lighting variations, and changes in viewing angles. The quality and visual coherence of stitched images directly impact user experience richness and operational efficiency. Despite the availability of numerous stitching algorithms, addressing the complexities of mixed scenarios remains a challenge. Notably, due to the addition of Features from Accelerated Segment Test (FAST) algorithm and Binary Robust Independent Elementary Features (BRIEF) algorithm, Scale Invariant Feature Transform (SIFT) and Oriented FAST and Rotated BRIEF (ORB) algorithms are favored for their robust feature extraction and matching capabilities.

© The Author(s) 2024
Y. Wang (ed.), *Proceedings of the 2024 2nd International Conference on Image, Algorithms and Artificial Intelligence (ICIAAI 2024)*, Advances in Computer Science Research 115,
https://doi.org/10.2991/978-94-6463-540-9_80

Reviewing relevant literature reveals various approaches to panoramic image stitching, including monitoring crops and soil using spectral patterns [1], employing large parallax completion methods in low-altitude urban areas [2], and employing transformations to obtain orthophoto-like panoramas. Additionally, integrating Thin Plate Splines (TPS) transformations for seamless spherical panoramic video generation facilitates the creation of spherical videos using multi-lens panoramic cameras [3].

This study aims to identify a more effective method to solve the challenge of panoramic image stitching in mixed scenes, especially to improve the SIFT algorithm for image pre-processing in close-range scenes and maintain high stitching quality in distant scenes. Close-range shooting often introduces visual differences due to Angle changes, requiring accurate feature matching between different depths of field and complex background details to achieve seamless integration [4]. Through this investigation, the research seeks to propose feasible solutions and contribute to the advancement of panoramic stitching technology, ultimately enhancing virtual experiences in education, entertainment, real estate and more.

In this study, the analysis of mixed scenes and algorithm improvement can be applied to multiple fields, such as VR in ophthalmology [5] and medical image mosaicing system [6]. And the aerial robotic sensor network system with intelligent auto-organization for urban surveillance [7] or the critical role of spacecraft maintenance [8]. In addition, it provides insights into image processing and provides recommendations for the selection and optimization of future algorithms, thereby enriching our understanding of panoramic stitching and enhancing image quality and user experience in complex scenes. Overall, this study highlights the potential of feature extraction and stitching algorithms in advancing VR and AR technologies, highlighting their theoretical and practical implications.

2 Technical Methods and Theoretical Framework

This study examines the effects of several parameters on panoramic image mosaic in mixed situations using the control variable approach. This study verified through a literature search that the evaluation performance indicators utilized were the root-mean-square error, the Spearman rank correlation coefficient, and the Pearson linear correlation coefficient, which could be used to compare the final display effect of the panorama [9]. This phenomenon is analyzed and explained in detail by setting up experiments and in-depth study of specific cases. After that, this study demonstrated the experimental results by systematically collecting, evaluating and integrating the existing research results.

2.1 SIFT

David G. Lowe proposed a feature point detection algorithm based on Gaussian scale space, called SIFT operator [10]. The position, scale, and rotation invariants of the extreme point are retrieved as it locates itself in the spatial scale. Before the implementation of SIFT algorithm, the scale space needs to be obtained through Gaussian ambiguity (Figure 1) .

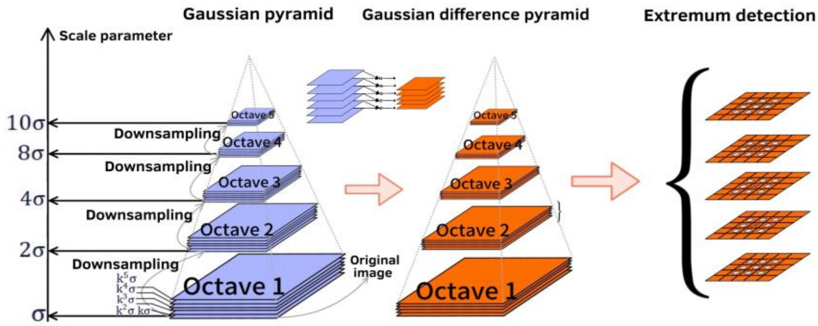

Fig.1 Gaussian pyramid

An illustration of this scale space is a Gaussian pyramid. Then, look for the extreme point in the Difference of Gaussian (DOG) pyramid. Be defined as

$$D(x,y,\sigma) = [G(x,y,k\sigma) - G(x,y,\sigma)] * I(x,y) = L(x,y,k\sigma) - L(x,y,\sigma) \quad (1)$$

The feature points are composed of gradient intensity information in 8 directions and 4×4 subregions, and finally form a 128-dimensional feature vector. After finding SIFT feature points, Random Sample Consensus (RANSAC) algorithm is used for fine matching. RANSAC is a random sampling consensus algorithm with high robustness and easy to use model parameter estimation technique. It is found that there is a geometric transformation relationship between the two images, which can be described as a homologous matrix

$$\mathrm{H}.\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} m_0 m_1 m_2 \\ m_3 m_4 m_5 \\ m_6 m_7 m_8 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2)$$

When the matrix is normalized, it will have eight unknowns and require at least four matrix parameters that are solvable to the matching points. When RANSAC algorithm is faced with a set of data containing "error points", it can train the parameter model by repeatedly selecting a set of data and iterating with a subset, so as to estimate a high-precision parameter model from a large number of observation data, so as to complete the removal of mismatched points. The enhanced SIFT algorithm identifies extremal sites by detecting points and their four adjacent points at the same scale, as well as ten adjacent points at neighboring scales [11]. It then eliminates poor contrast spots and unstable edge response points by fitting the Hessian matrix and DOG function.

In this study, the image quality and the accuracy of feature point matching are improved by adding the steps of image preprocessing [12], grayscale, Contrast Limited Adaptive Histogram Equalization (CLAHE), morphology opening operation and feature point descriptor calculation. CLAHE is a technique used in computer vision to improve image contrast [13]. This process is adaptive, meaning that it adjusts to the local characteristics of the image. The core idea is to prevent excessive amplification of noise by limiting contrast enhancement during histogram equalization. Morphology operation is the process of corrosion operation followed by expansion operation. It is mainly used to remove small objects, break narrow connections, and smooth the edges of objects. Fast Library for Approximate Nearest Neighbors (FLANN) is used to find the closest match between feature point descriptors, using efficient data structures and algorithms to speed up the search process (Figure 2).

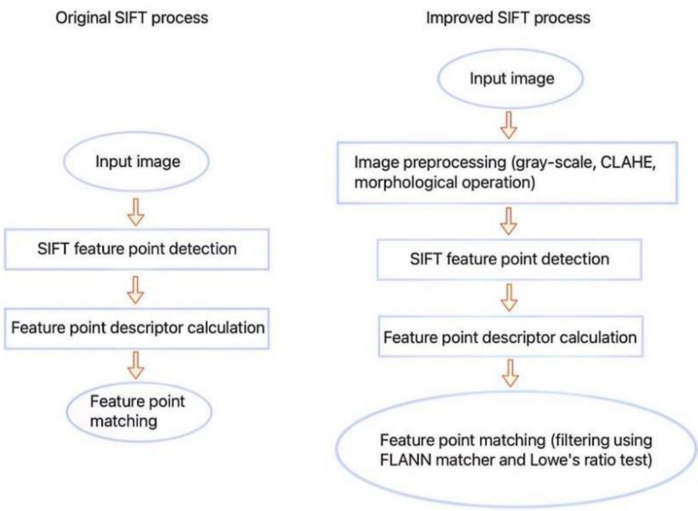

Fig.2 Before and after improvement flow chart

2.2 ORB

ORB uses a FAST keypoint detector to identify corners in the image. The FAST keypoint detector itself does not have scale invariance, but it quickly identifies corner points by examining the circular neighborhood around the pixel. Therefore, ORB applied FAST on the pyramid layers of the image, allowing it to detect key points at different scales. Nowadays, there are many researches on the improvement and application of ORB algorithm with different viewpoints and specific methods [14]. The ORB method mentioned by T.Zhuo is described using the BRIEF descriptor.

$$f(x) = f(x;\, p_1, p_2, s) = \frac{1}{s^2}\left[\sum_{q\in R(p_1,s)} I(q) - \sum_{r\in R} I(r)\right] \tag{3}$$

AdaBoost algorithm is used to continuously iterate p1, p2 and s, and train the feature learner. The training loss function L is as follows:

$$L = \sum_{i=1}^{N} exp[-\gamma l_i \sum_{k=1}^{K} h_k(x) h_k(y)] \tag{4}$$

This descriptor uses Gaussian distribution to select and compare the gray value difference of feature point pairs, which has randomness and can not guarantee the quality of point pairs, which weakens the expression ability of descriptors, resulting in large computation, low detection speed and repetition rate.

3 Result and Discussion

3.1 Experimental Design

In this study, firstly, the results of the SIFT and ORB algorithms for image feature point recognition were compared and analyzed respectively in terms of matching performance, including accuracy and processing time. The SIFT algorithm, whose final results were more suitable for the experiment, was selected for further study. Next, by adding steps of image preprocessing, combining gray-scale conversion, CLAHE, and morphological open operation methods, and using FLANN matching in the process of feature matching, the functions of the SIFT algorithm to improve the accuracy of feature extraction and enhance the robustness of feature matching were improved, and the final results are presented. Finally, the angle of mixed scenes and depth of field factor in image selection were analyzed.

3.2 Result Analysis

Compare the Results of SIFT and ORB Algorithm. In this study, a combination of close-range and long-range image datasets was selected to ensure that the datasets covered complex factors such as occlusion phenomena, diversity of lighting conditions, and multi-angle shooting. By comparing image matching using SIFT and ORB under the same image rotation and reduction conditions, the extraction time and the number of feature points on the rotated or reduced image were obtained. The purpose of this selection was to thoroughly test the algorithm and evaluate its performance when dealing with a variety of challenging scenarios (see Figure 3 and Table 1).

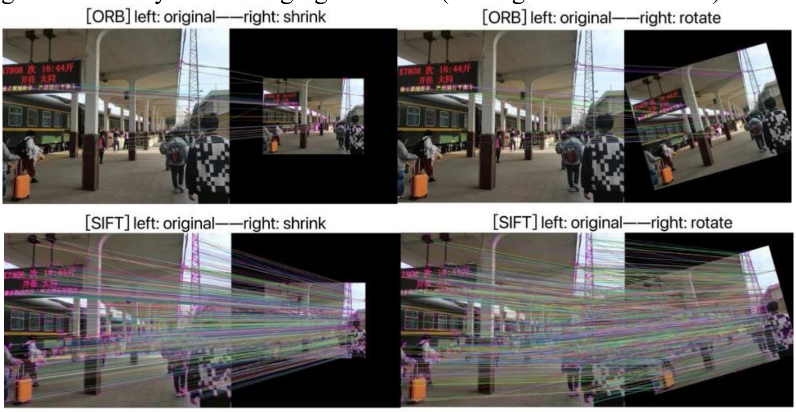

Fig.3 Comparison of SIFT and ORB in special extraction

Table 1 The results of different methods

Met hod	Extraction time (s)	points on the rotated image	points on the shrink image
Sift	1.042	3651	2088
OR B	0.158	500	500

The experimental results demonstrate that the SIFT algorithm has a longer processing time, as noticed in Table 1. But it can extract more feature points at the same time. This is mainly because the SIFT algorithm has strong adaptability to the scale, rotation and brightness changes of the image by constructing image pyramid and Gaussian difference scale space, so it can detect more feature points, but this process is relatively complex, resulting in a long processing time. In contrast, the ORB algorithm achieves faster feature extraction by adopting FAST corner detection and BRIEF descriptors, but may not be as precise in the quantity and quality of feature points [15].

Improvement of Feature Matching Algorithm. In this study, the image quality and accuracy of feature point matching are improved by using feature point descriptor calculation, morphological opening operation, gray-scale, CLAHE and preprocessing steps.

Fig.4 The results before and after improving the feature matching algorithm

Through the observation results (Figure 4), the improved image splicing Angle is more accurate, the overall splicing trace is less obvious, and the edge of the figure and scene is more clear. This is because the gray-scale and de-noising operations in the pre-processing step can reduce the amount of data that the SIFT algorithm needs to process, thereby increasing the speed of the algorithm. Among the improved algorithms in this study, CLAHE and morphological open operation can improve image quality and make SIFT algorithm mo`re efficient and accurate in extracting and describing features. The fast search capability of the FLANN matcher significantly reduces the time required for feature matching, making the entire SIFT algorithm process more efficient.

Compare the Foreground and Near View in the Case of Angle Change. In the experiment of image stitching, this study explores the influence of the change of shooting Angle on the stitching effect of near and far scenes in the image through image stitching of a mixed scene with both near and near. In order to carry out the experiment, the feature points are extracted and matched using the SIFT methodology. For the stitching of near scenes, when the shooting Angle changes significantly, the experiment finds that the stitching effect of some details (such as the head of the character) is not ideal. This is mostly because, while photographing at close range, a change in angle will cause the parallax effect to become more noticeable, resulting in a significant variation between images of the same item taken from different angles. Furthermore, a change in angle may also affect the neighboring scene's occlusion relationship, which could result in the overlay or dislocation phenomena during splicing.

4 Conclusion

Firstly, the performance difference between SIFT algorithm and ORB algorithm in image feature recognition is studied. The results show that SIFT algorithm has strong adaptability to the scale, rotation and brightness changes of images, and can recognize more feature points, but the processing time is longer. The ORB algorithm is faster and suitable for real-time applications. By comparison, the applicability and performance of SIFT and ORB algorithms in different scenarios can be comprehensively evaluated, thus providing a basis for selecting image mosaic algorithms suitable for specific scenarios. Secondly, the SIFT algorithm is improved by introducing image preprocessing steps, including gray-scale, CLAHE, morphology operation, and FLANN matcher. These improvements significantly improve the accuracy of feature extraction and robustness of feature matching, thus optimizing the effect of image Mosaic. In particular, the stitching angle and stitching marks have been significantly improved, making the image more natural and clear after stitching. Finally, the SIFT algorithm is used to observe the impact of Angle change on image stitching, and it is found that the stitching effect of distant scenes has a higher tolerance for Angle change, while close-range scenes are more susceptible to the impact of Angle change, which is mainly caused by the parallax effect and the change of occlusion relationship during close-range shooting.

As a way to assist future research in making more informed choices when maximizing the effect of image stitching, the study's conclusion offers a helpful guide for the selection of algorithms for image stitching technology. Simultaneously, it provides a theoretical foundation for creating panoramic maps and optimizing the visual experience of VR/AR technologies. Future studies may investigate how to combine deep learning and traditional feature extraction algorithms to improve the effect of close-range stitching and reduce the impact of angle changes. In addition, improving the processing speed while ensuring accuracy to meet the higher real-time requirements of the application scenario is also a worthy direction for in-depth research.

References

1. Liu Y. He M. Wang Y., et al. Rapid mosaic of farmland aerial panoramic images based on optimized SIFT algorithm. Trans. Chin. Soc. Agric. Eng., 39(1), (2023).
2. Wang X. Fu H. Yang G. A new method of aerial image mosaic in low-altitude urban areas. Geogr. Geo-Inf. Sci., 3: 34-41,(2020).
3. Huang M. Wu J. Gao J. Seamless spherical video generation for multi-head panoramic camera (MPC). Acta Geod. Cartogr. Sin., 51(5): 703-717, (2022).
4. He W. Wang Z. Wei J., et al. Dual optical target recognition method for collimated images based on BLOB region and edge feature analysis. Acta Photon. Sin., 53(2): 0210001, (2024).
5. Zou X. Sui R. Current status of virtual reality and augmented reality-based training tools in ophthalmology teaching. Basic & Clin. Med., 40(12): 1744-1748,(2020).
6. Win K P, Kitjaidure Y, Hamamoto K. An implementation of medical image mosaicing system based on oriented FAST and rotated BRIEF approach. Appl. Sci., 10: 1800, (2020).
7. Leong W., Martinel N, Huang S., et al. An intelligent auto-organizing aerial robotic sensor network system for urban surveillance. J. Intell. Robot. Syst., 102(2): 33, (2021).
8. Wei S. Zhou Z. Zhang X. Mao L. Jia J. Liu C. Virtual Training System for Spacecraft Maintenance Based on Desktop Virtual Reality. J. Syst. Simul., 33(6): 1358-1363,(2021).
9. Ada; Bai Y.; Yu K.; Yuan H.; Liu Y. The latest development of panoramic image quality evaluation methods. J. Comput. Eng. Appl., 58(24), (2022).
10. Xi S. Li W. Xie J., et al. Feature Point Matching Between Infrared Image and Visible Light Image Based on SIFT and ORB Operators. Infrared Tech., 42(2): 168-175,(2020).
11. Yu Z. Zhang N. Pan Y. Zhang Y. Wang Y. Heterogeneous image matching based on improved SIFT algorithm. Laser Optoelectron. Prog., 59(12): 1211002,(2022).
12. Rymarczyk T, Niderla K, Kozłowski E., et al. Logistic regression with wave preprocessing to solve inverse problem in industrial tomography for technological process control. Energies, 14(23): 8116, (2021) .
13. Anitha V. Efficient framework for fast nearest neighbor search,1-10, (2021).
14. Tao Z. Huang W. Zhang Z., Yao Y., He J. Improved ORB Matching Algorithm Based on Sparse Optical Flow Method, Modular Mach. Tool & Autom. Manuf. Tech., 12: 1-5,(2021).
15. Borman R I, Harjoko A. Improved ORB Algorithm Through Feature Point Optimization and Gaussian Pyramid. Int. J. Adv. Comput. Sci. Appl., 15(2),(2024).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.