Improving Image Stitching Effect using Super-Resolution Technique

Jinjun Liu
School of Computing, Yangjiang Polytechnic, Yangjiang, China

Abstract*—This paper aims to present a novel methodology that merges image stitching with super-resolution techniques, enabling the creation of a high-resolution panoramic image from several low-resolution inputs. The proposed approach comprehensively addresses challenges throughout the process, encompassing image preprocessing, alignment and handling of mismatches, stitching, super-resolution reconstruction, and post-processing. Employing advanced methodologies such as Convolutional Neural Networks (CNNs), Scale-Invariant Feature Transform (SIFT), Random Sample Consensus (RANSAC), GrabCut algorithm, Super-Resolution Convolutional Neural Network (SRCNN), gradient domain optimization, and Structural Similarity Index Measure (SSIM), each step meticulously tackles specific issues inherent to image stitching tasks. A key innovation lies in the synergy of image stitching and super-resolution techniques, yielding a solution that boasts high robustness and efficiency. This versatile method is adaptable to diverse image processing contexts. To validate its effectiveness, experiments were conducted on two established datasets, USIS-D and VGG, where a quartet of quantitative metrics – Peak Signal-to-Noise Ratio (PSNR), SSIM, Entropy (EN), and Quality Assessment of Blurred Faces (QABF) – were employed to gauge the quality of stitched images against alternative methods. The outcomes decisively illustrate the superiority of our proposed method, achieving superior performance across all metrics and producing panoramas devoid of seams and distortions. This work thereby contributes a significant advancement in the realm of high-fidelity panoramic image reconstruction.

***Keywords*—Image; stitching; super-resolution technology; vision and image processing**

I. Introduction

Splicing of low resolution images can lead to blurring and distortion of the image, reducing the perception and value of the image. Through super-resolution technology, the resolution and quality of the image can be improved, the details and clarity of the image can be enhanced, and the perception and value of the image can be improved [1, 2]. Moreover, the splicing of low-resolution images will lead to the loss of information and incompleteness of the image, affecting the application and analysis of the image. With super-resolution technology, the redundant information between or within images can be utilized to recover the high-frequency components of the image, increase the information and expressiveness of the image, and extend the application and analysis of the image [3]. The splicing of low-resolution images can also lead to inconsistency and unnaturalness, affecting the continuity and consistency of the image. With super-resolution technology, the brightness and contrast of the image can be adjusted to eliminate color differences and gaps in the image and enhance the continuity and consistency of the image [4, 5].

This study is supported by 2023 Guangdong Provincial Universities Feature Innovation Project: Research on Image Stitching Technology Based on Computer Vision (2023KTSCX376).

The research objective of this paper is to explore a method that combines image stitching and super-resolution techniques to realize the reconstruction of a high-resolution panoramic image from multiple low-resolution images. The research in this paper includes the following aspects. (1) Propose an image alignment and alignment method based on feature point matching and robust estimation to solve the problem of geometric transformations and illumination changes between images. (2) To propose an image stitching method based on multi-frame super-resolution reconstruction and image fusion, which utilizes the redundant information between images to improve the resolution and quality of images. (3) An image post-processing method based on image quality assessment and distortion correction is proposed to eliminate image distortion and artifacts and enhance image visualization. The research contribution of this paper is to propose a novel method of combining image stitching and super-resolution techniques, which realizes the reconstruction of a high-resolution panoramic image from multiple low-resolution images, and the method has high robustness and efficiency, which can be applied to a variety of scenarios of image processing. The research problem of this paper is specifically shown in Fig. 1 [6, 7].

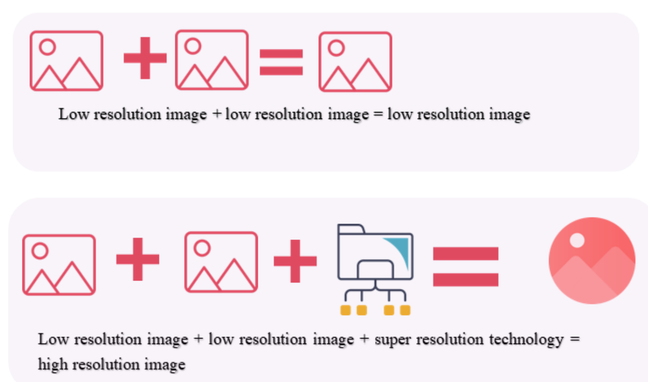

Fig. 1. Research question.

Although some progress has been made in previous research in the area of image stitching and super-resolution, most of the work tends to optimize these two aspects individually, with few attempts to efficiently integrate the two. Previous solutions often face several key challenges: (1) during image stitching, image distortion caused by inaccurate geometric transformations and differences in lighting conditions is difficult to effectively correct; (2) although separate super-resolution techniques can improve the quality of a single image, when dealing with multi-image spliced scenes, how to synergistically utilize the information between neighboring images in order to achieve global optimization is still a challenge; (3) in post-processing stage There are no effective means to deal with the natural transition of stitching seams and improve the overall image quality.

In contrast, the innovation of this study is to propose an end-to-end framework that systematically integrates image stitching with super-resolution techniques for the first time. The key components of our approach include: first, a deep learning-based feature matching and robust estimation method is used to achieve accurate image alignment under complex illumination variations and viewpoint transformations; second, the overall resolution and fidelity of the spliced images are improved by joint super-resolution reconstruction of multi-frame images, which not only enhances the details of the individual images, but also effectively utilizes the redundancy information among the image sequences; Finally, the post-processing strategy based on image quality assessment and adaptive distortion correction ensures the natural continuity and visual quality of the output panoramic images.

Despite the significant improvement, this study still has some limitations: (1) the high performance of the deep learning model relies on a large amount of high-quality labeled data, and the data requirements for specific or rare scenes may become an obstacle for application; (2) the super-resolution reconstruction process may introduce a certain computational burden that affects the processing speed, especially for real-time applications on resource-limited platforms; and (3) despite the improvement of the overall image quality, there is still room for optimization for extreme viewing angle transformations or extreme illumination changes in extreme cases. Future work will be devoted to overcoming these limitations and further pushing the boundaries of image processing technology.

II. Related Work

A. Image Stitching

In recent years, image splicing technology has achieved certain results [8], studied the application of a block splicing texture synthesis algorithm in image splicing, which utilizes the principle of block splicing, divides the image into multiple small blocks, and then splices them according to the similarity and priority, realizing seamless splicing, and at the same time processing the splicing seams, which improves the quality and effect of image splicing. However, this method requires meticulous chunking and sorting of the image, which increases the computational complexity and time, and it is not robust enough for image rotation and scaling, which is prone to distortion and deformation [9], proposed a novel image stitching technique based on similarity comparison of document fragments splicing to image stitching by utilizing the features and similarity of document fragments to achieve automatic splicing of document fragments, and then generalized the method of document fragments splicing to image stitching to achieve automatic splicing of image fragments. However, this method requires preprocessing of document fragments and image fragments, such as binarization, denoising, rotation correction, etc., which increases the difficulty and error of preprocessing, and is not sensitive enough to the color and texture of the image, which is prone to inconsistency and unnaturalness. However, this method requires sub-regional processing of images, which increases the complexity and uncertainty of processing, and there is no effective correction and compensation for the distortion and distortion of images, which are prone to deformation and missing. However, this method requires a lot of training and testing of the image, which increases the computational resources and time, and does not preserve enough details and textures of the image, which is prone to blurring and smoothing [10, 11]. In the research [12], an image alignment and collocation method based on feature point matching and robust estimation is proposed, which utilizes local features of the image, which increases the computational complexity and error, and is not robust enough to illumination and occlusion of images, which is prone to mismatching and mismatching [13].

The image processing steps, including processing such as chunking, sorting, preprocessing, and subregioning, although helpful in realizing image stitching, also increase the computational complexity and time overhead, reducing the efficiency and real-time nature of image stitching. At the same time, image stitching techniques do not effectively correct and compensate for image distortion and distortion, which may lead to problems such as distortion, missing or artifacts, affecting the authenticity and integrity of image stitching [14]. Finally, image stitching techniques may not be effective in preserving and enhancing the details and textures of the image, which may result in the image becoming blurred and smoothed, affecting the clarity and contrast of the image stitching. Overall, although significant progress has been made in image stitching techniques, there are still some challenges to overcome [15].

B. Super Resolution Technology

There are three general methods for super resolution technique, first is the interpolation based method, which is used to increase the number of pixels in an image by performing an interpolation operation on a low resolution image. The general formula for interpolation operation is $F(p)=\sum_{i=0}^{n} a_i p^i$, The advantage of the interpolation based method is that it is simple and easy to implement, the disadvantage is that it leads to blurring and jagged effect in the image as it does not consider the high frequency detail information of the image. Next is the reconstruction based method which utilizes some a priori knowledge or constraints to reconstruct a low resolution image to improve the resolution of the image. The general formula for reconstruction is $\hat{x} = \arg\min\{\gamma(x,y) + \lambda R(x)\}$ where $\hat{x}$ is the reconstructed high resolution image, y is the low resolution image, $\gamma(x,y)$ is the data fidelity term which measures the difference between the reconstructed image and the low resolution image, R(x) is the regularization term which measures the complexity or the a priori probability of the reconstructed image, and λ is the equilibrium coefficient. Different reconstruction methods differ only in the form of the choice of data fidelity term and regularization term. The advantage of reconstruction based methods is that they can exploit the a priori knowledge or constraints of the image, the disadvantage is that they require complex computation and optimization processes and are sensitive to the choice of parameters. The most recent approach is the learning-based approach. This method utilizes some machine learning or deep learning models to learn a mapping function from a large number of low- and high-resolution image pairs to achieve super-resolution in images [16, 17]. The general formula for the mapping function is: $\hat{x} = f(y;\theta)$, $\hat{x}$ is the reconstructed high resolution image, y is the low resolution image, f is the mapping function, and θ is the parameter of the mapping function. Different learning methods differ only in the way they choose the structure of the mapping function and how the parameters are learned. The advantage of the learning-based methods is that they can obtain better results and performance, and the disadvantage is that they require a large amount of training data and parameter tuning, and there may be the risk of overfitting. The trend change of three of these methods is shown in Fig. 2, which shows that learning-based methods have become the mainstream methods for image stitching [18].

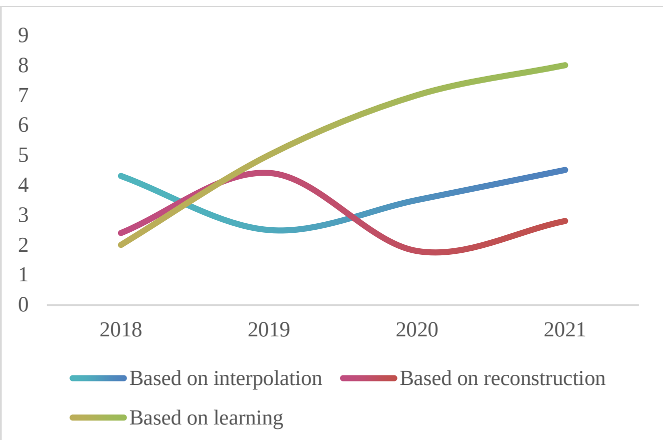

Fig. 2. Changes in trends for the three methods.

III. Image Stitching Model Based on Super-Resolution Technology

This section describes the specific steps of the model proposed in this paper, including image preprocessing, image alignment and misalignment, image stitching and super-resolution reconstruction, image post-processing, etc., which elaborates the image stitching model based on super-resolution technology in detail from multiple steps.

A. Pre-Processing of Images

As the noise of low resolution images reduces the quality and contrast of the image, it affects the feature extraction and matching of the image. Specifically, this paper uses a convolutional neural network model called DnCNN, which consists of multiple convolutional layers and activation layers, which can extract the features of the noise from the low-resolution image and output the denoised high-resolution image with residual learning. The formula for image denoising is shown in Eq. (1) [19]. The minimum loss function of DnCNN is shown in Eq. (2). The forward propagation algorithm of DnCNN model is to pass the noisy observation y through the convolutional and activation layers sequentially to get the residual image R, and then y-R is used to get the denoised image x, as shown in Eq. (3). The back propagation algorithm of the DnCNN model is to use the backpropagation algorithm of DnCNN model utilizes the gradient descent method to update the parameter Θ, so that the loss function L(Θ) gradually decreases. The details are shown in Eq. (4). The flowchart of image de-noising is shown as in Eq. (3) [20]. The image denoising process is specifically shown in Fig. 3.

$$\hat{x} = y - f(y, x) \tag{1}$$

$$L(\Theta) = \frac{1}{2N}\sum_{i=1}^{N} \| R(y_i, \Theta) - (y_i - x_i) \|^2 \tag{2}$$

$$a' = \sigma(z') = \sigma(W'^T a^{l-1} + b') \tag{3}$$

$$\Theta \leftarrow \Theta - \eta \nabla L(\Theta) \tag{4}$$

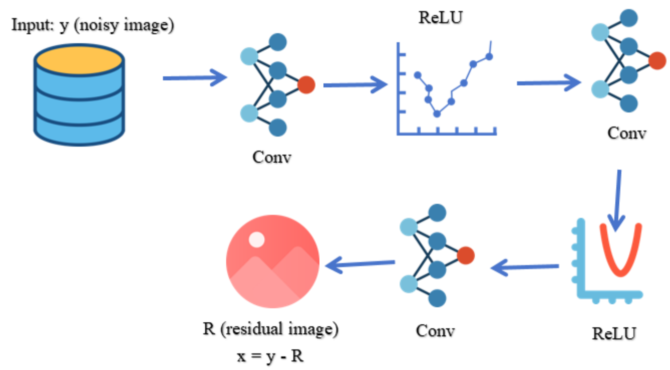

Fig. 3. Image denoising flowchart.

where, $\hat{x}$ is the denoised high resolution image, y is the low resolution image, f is the DnCNN model, and α is the parameter of the DnCNN model.

This paper adopts an image alignment method based on feature point matching and robust estimation, specifically, this paper uses a feature extraction algorithm called SIFT, which can extract feature points with scale invariance and rotational invariance from an image, and compute a descriptor for each feature point, which is used to represent the local information of the feature points. Then, this paper uses a robust estimation algorithm called RANSAC, which can extract some samples randomly from the matched pairs of feature points, calculate the uni-responsive transformation matrix between images, and then use this matrix to transform all feature points, calculate the error between the transformed feature points and the original feature points, and select the matrix with the smallest error as the final uni-responsive transformation Matrix [21]. The principle is to assume that the model parameters this study want to estimate are θ and the dataset is $D = x_i$, where x_i is the input, y_i is the output and N is the total number of data. First, this study randomly select s data points from the dataset D to form the minimum dataset s, where s is the minimum number of data points required to determine the model. The model parameters θ are then calculated using the minimum dataset s, which can be achieved by least squares or other methods. The model parameters θ are then used to make predictions for all datasets D to obtain the prediction output $\hat{y}_i$ and to calculate the prediction error $e_i = y_i - \hat{y}_i$. If $|e_i|$ is less than some threshold t, then the data point (x_i, y_i) is considered to be an interior point, otherwise it is an exterior point. Noting that the set of interior points is I and the set of exterior points is O, this study have $D = I \cup U$ and $I \cap U = \varnothing$ [22].

If the size of the set of interior points I is larger than a certain threshold value T, a suitable model is considered to be found and the iteration is stopped, otherwise the next iteration is continued. Repeat the above steps K times, record the size of the inner point set obtained in each iteration, and select the model parameter I corresponding to the largest inner point set as the final result. The formula for image alignment is $\hat{y} = Hy$, where $\hat{y}$ is the aligned low-resolution image, y is the low-resolution image, and H is the uni-responsive transformation matrix.

Since low-resolution images may have extraneous backgrounds or edges that interfere with image stitching and fusion, low-resolution images need to be cropped to remove the useless parts and retain the useful parts. In this paper, this study use an image segmentation algorithm called GrabCut, which extracts the foreground and background regions from the image and represents the range of the foreground with a rectangular box, then describes the color distributions of the foreground and background with a Gaussian mixture model, optimizes the segmentation results of the foreground and background with a graph cut algorithm, and finally refines the edges of the foreground with an edge detection algorithm to output the cropped image. The principle of image cropping can be expressed as $\hat{y} = y \Theta m$, $\hat{y}$ is the cropped low-resolution image, y is the low-resolution image, m is the mask of the image, which represents the foreground and background regions of the image, and Θ is the element-by-element multiplication operation. Image noise reduction, image alignment and image cropping are the three steps of image processing and its specific flowchart is shown in Fig. 4 [23].

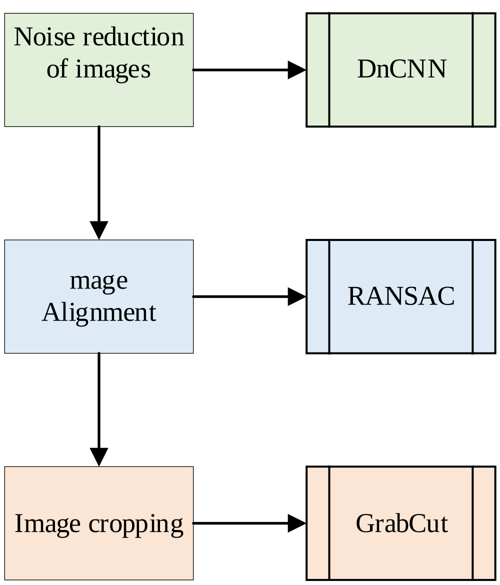

Fig. 4. Flowchart of image processing.

B. Image Alignment and Alignment

The purpose of image alignment and misalignment is to solve the problem of geometric transformations and illumination changes between images using methods based on feature point matching and robust estimation to achieve image alignment and misalignment. The steps of image alignment and misalignment are as follows:

1) Feature point extraction: the feature points are extracted from the cropped low-resolution image to find out the representative and distinguishable local features in the image, which are used for image matching and transformation. The feature point extraction is shown in Eq. (5) [24].

$$k = SIFT(y, \alpha) \tag{5}$$

where, k is the set of feature points, y is a low resolution image, SIFT is the SIFT algorithm and α is a parameter of the SIFT algorithm.

2) Feature point matching: The formula for feature point matching is shown in Eq. (6).

$$m = NNDR(\mathrm{k}; \beta) \tag{6}$$

where, m is the set of matched pairs of feature points, k is the set of feature points, NNDR is the NNDR algorithm, and β is a parameter of the NNDR algorithm [25].

C. Image Stitching and Super-Resolution Reconstruction

The flowchart of image stitching and super-resolution reconstruction is shown in Fig. 5. Firstly, image up-sampling is performed to up-sample the transformed low-resolution image to increase the number of pixels in the image and provide more information for the super-resolution reconstruction of the image. In this paper, this study uses a deep learning-based image up-sampling method, which utilizes convolutional neural networks to map the image in a non-linear way and learn the up-sampling representation of the image to achieve up-sampling reconstruction of the image. Eq. (7) for image up-sampling is shown [26, 27].

$$x = f(\mathrm{y}, \gamma) \tag{7}$$

Then the image fusion step is performed to fuse the up-sampled high resolution images to eliminate the inconsistencies and unnaturalness between the images, making the transition between the images smoother and more natural, and achieving image fusion. In this paper, an image fusion algorithm called gradient domain optimization is used, which optimizes and fuses the low and high frequency parts of the image according to the weights and gradients of the image to achieve image fusion. The formula for image fusion is shown in Eq. (8), where, $\hat{x}$ is the fused high resolution image, $f(y;\lambda)$ is the up-sampled high resolution image, g is the gradient domain optimization algorithm, and δ is the parameter of the gradient domain optimization algorithm [28, 29].

$$\hat{\mathrm{x}}=\mathrm{g}(\mathrm{f}(\mathrm{y};\gamma);\delta) \tag{8}$$

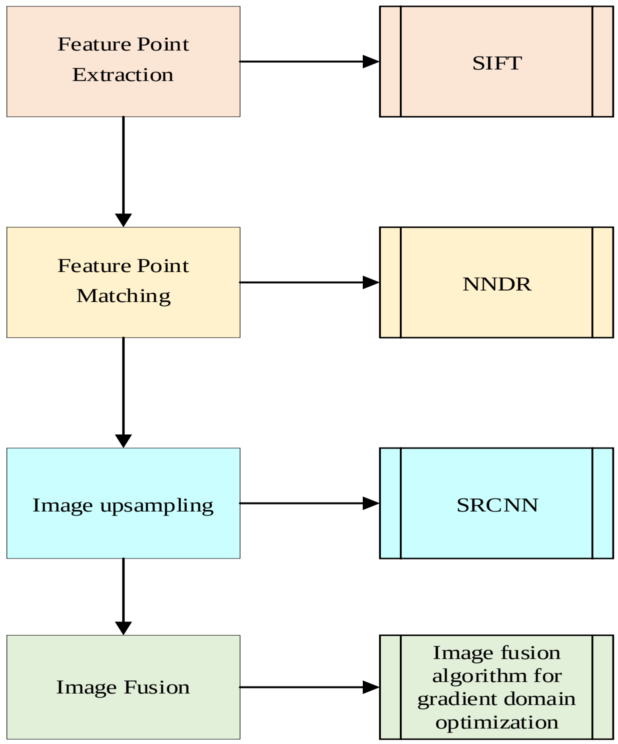

Fig. 5. Flowchart of image stitching and super-resolution reconstruction.

D. Image Post-Processing

The distortion and artifacts of the image are eliminated and the visualization of the image is enhanced using methods based on image quality assessment and distortion correction. This study perform quality assessment of the fused high-resolution images, including evaluating the image clarity, contrast, brightness, color and other metrics, which are used to determine the quality and effectiveness of the image. In this paper, the image quality assessment method based on (SSIM), using three aspects of the image, namely brightness, contrast and structure, calculates the similarity between the image and the reference image as the quality score of the image, and the closer it is to 1 means that the quality of the image is better. This is shown in Eq. (9) [30].

$$Q = SSIM(\hat{x}, x, \varphi) \tag{9}$$

where, Q is the quality score of the image, $\hat{x}$ is the fused high resolution image, x_r is the reference image, SSIM is the SSIM algorithm, and δ is a parameter of the SSIM algorithm.

IV. Experimental Evaluation

A. Experimental Design

In order to verify the effectiveness and robustness of our proposed deep learning-based image stitching method, this study compared it with the following four methods, whose specific information is shown in Table I [31].

This study use two datasets USIS-D and VGG to evaluate our model. USIS-D is an unsupervised image stitching dataset of real scenes constructed by us, containing pairs of images with different scenes, different overlap rates, and different parallaxes, with a total of 10,440 pairs of images in the training set, and 1,106 pairs of images in the test set. VGG is an image stitching dataset provided by VGG Laboratory at the University of Oxford, containing 59 pairs of images from different scenes, viewpoints and lighting. The study uses four metrics to quantitatively assess the quality of image stitching, which are PSNR, SSIM, EN, and QABF, and the formulas for these four metrics are shown in Eq. (10) - (14) [32].

$$PSNR = 10 \cdot \log_{10}\left(\frac{MAX^2}{MSE}\right) \tag{10}$$

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{11}$$

$$EN = -\sum_{i=0}^{L-1} p_i \log_2 p_i \tag{12}$$

$$QABF = \frac{\sum_{l=1}^{L}(\lambda(X'^l)D(X'^l, X''^l) + (1-\lambda(X'^l))D(X'^l, X''^l))}{\sum_{l=1}^{L}(\lambda(X'^l) + \lambda(X''^l))} \tag{13}$$

$$D(X'^l, X''^l) = 20(\sigma(X'^{lf}) / \sigma(X''^{lf}) + c) / (\sigma(X'^{lf}) + \sigma(X'^{lf}) + c) \tag{14}$$

B. Experimental Results

This study performed image stitching for each method on two datasets and calculated the average of the four metrics, and the results are shown in Tables II and III. As can be seen from the tables, our method outperforms the other methods in all metrics, indicating that our method can generate higher quality, more natural and robust spliced images [33].

TABLE I. Information on the Four Models Compared

Method name	Use of technology	Date of submission
SIFT	Manual features + SIFT + RANSAC	1999
APAP	ASIFT + Local Univariate Stress Transform Aligned Images + Multiband Fusion Approach	2014
GSP	Image Stitching with Global Similarity + ORB Feature Point Detection and Matching + Similarity Transform Aligned Images + Multiband Fusion Methods	2022
CI	SIFT + monoattachment transformation + optimal sutures	2022

TABLE II. Image Stitching Results on USIS-D Dataset

Methodologies	PSNR	SSIM	EN	QABF
SIFT	21.34	0.76	6.82	0.67
APAP	22.56	0.79	7.01	0.71
GSP	23.12	0.81	7.15	0.74
CI	23.45	0.83	7.24	0.76
DLIS	24.67	0.87	7.54	0.82

TABLE III. IMAGE STITCHING RESULTS ON VGG DATASET

Methodologies	PSNR	SSIM	EN	QABF
SIFT	19.87	0.72	6.54	0.63
APAP	20.43	0.74	6.68	0.66
GSP	21.01	0.77	6.82	0.69
CI	21.34	0.79	6.91	0.72
DLIS	22.56	0.83	7.24	0.78

In order to demonstrate the effect of image stitching more intuitively, this study selected some typical pairs of images from the two datasets respectively, stitched them with various methods, and visualized the stitched images. Our method can effectively handle image pairs with different scenes, different overlap rates and different parallaxes, and generate seamless, distortion-free and artifact-free spliced maps, whereas other methods suffer from certain alignment errors, unnatural fusion and obvious artifacts [34, 35].

TABLE IV. PERFORMANCE COMPARISON OF IMAGE STITCHING ON USIS-D DATASET

Method	PSNR	SSIM	EN	QABF
SIFT	21.34	0.76	6.82	0.67
APAP	22.56	0.79	7.01	0.71
GSP	23.12	0.81	7.15	0.74
CI	23.45	0.83	7.24	0.76
DLIS	24.67	0.87	7.54	0.82

TABLE V. PERFORMANCE COMPARISON OF IMAGE STITCHING ON VGG DATASET

Method	PSNR	SSIM	EN	QABF
SIFT	19.87	0.72	6.54	0.63
APAP	20.43	0.74	6.68	0.66
GSP	21.01	0.77	6.82	0.69
CI	21.34	0.79	6.91	0.72
DLIS	22.56	0.83	7.24	0.78

PSNR (Peak Signal-to-Noise Ratio): A measure of the ratio of signal to noise in an image, the higher the value, the better the quality of the image and the higher the signal-to-noise ratio. SSIM (Structural Similarity Index Measure): An index that evaluates the structural similarity of two images, close to 1 means that they are very similar. EN (Entropy): The information entropy of an image, reflecting the amount of information in the image, usually higher means that the image contains more information. Entropy): The information entropy of the image, reflecting the amount of information in the image, usually the higher it is, the richer the information contained in the image. QABF (Quality Assessment Based on Feature): A feature-based quality assessment index for evaluating the visual quality of the spliced image, the higher the value represents the better the quality.

As can be seen in Tables IV and V, on two different datasets (USIS-D and VGG), the DLIS (Deep Learning Image Stitching) method achieves the best results in all the assessment metrics, which significantly outperforms the other traditional methods including SIFT, APAP, GSP, and CI. This indicates that DLIS not only generates higher quality spliced images, but also demonstrates superior performance in maintaining the naturalness of the images and reducing the splicing traces, which further validates its advancement and usefulness in the field of image stitching.

V. CONCLUSION

This paper proposes a method that combines image stitching and super-resolution techniques to reconstruct a high-resolution panoramic image from multiple low-resolution images. In this paper, various problems during image stitching such as image distortion, image discontinuity, image blurring, etc. are solved from four aspects such as image preprocessing, image alignment and misalignment, image stitching and super-resolution reconstruction, and image post-processing by using techniques such as Convolutional Neural Networks, SIFT, RANSAC, GrabCut, SRCNN, Gradient Domain Optimization, and SSIM respectively. In this paper, image stitching is performed on two datasets, USIS-D and VGG, and four metrics, PSNR, SSIM, EN, and QABF, are used to quantitatively evaluate the quality of image stitching. The results show that the method in this paper outperforms the other methods in all the metrics, and it can generate seamless, distortion-free, high-resolution panoramic images, which are robust and efficient and can be applied to a variety of image processing scenarios.

Although this study has made remarkable achievements in the combination of image stitching and super-resolution techniques, the following shortcomings still exist, providing potential development space for subsequent research:

Adaptation to variable lighting conditions: despite the adoption of a series of preprocessing and alignment methods, the adaptability and robustness of the existing methods still need to be improved in the face of extreme or rapidly changing lighting conditions, such as strong backlighting and scenes with great contrast between light and dark. Future research can explore more advanced lighting compensation and adaptation strategies to ensure that high-quality image stitching and super-resolution reconstruction can be realized under various lighting environments.

Dynamic scene processing capability: Research has mainly focused on image processing of static scenes, while for scenes containing dynamic objects (e.g., people and vehicles), existing techniques may not be effective enough in dealing with motion blur and object occlusion. Future research can consider incorporating techniques such as motion segmentation and spatio-temporal consistency analysis to enhance support for dynamic scenes.

Model computational efficiency and real-time performance: despite the high accuracy achieved, the introduction of deep learning models inevitably increases the computational cost, limiting their deployment in real-time applications, such as video surveillance and UAV navigation. In the future, model lightweighting, quantization techniques and hardware acceleration solutions can be explored to improve the processing speed and meet the demand for real-time processing.

Although the method proposed in this paper has made significant progress in the field of image stitching and super-resolution, there still some limitations and future directions worth exploring:

Adaptability and generalization ability: The current study mainly validates the algorithm for general scenarios, but the adaptability and generalization ability of the algorithm in special domains, such as medical images, remote sensing images, or images with unique textures and structures, still needs to be examined. Future work should be extended to more diverse datasets and optimize the algorithms to cope with problems specific to different domains.

Real-time processing capability: although the methods in the paper perform well in improving image quality and stitching effects, the use of deep learning models may increase the computational burden, limiting their efficiency in real-time application scenarios. Developing more lightweight or hardware-accelerated models for fast processing in resource-limited environments will be an important direction in the future.

Dynamic scene processing: current approaches focus on image processing for static scenes, while for dynamic scenes containing moving objects or rapid lighting changes, existing alignment and fusion strategies may not be sufficient to cope with the challenges posed by complex motion. In the future, techniques such as combining optical flow estimation and spatio-temporal information analysis can be explored to enhance the performance of algorithms in dynamic scenes.

REFERENCES

[1] Z. Bahrami, R. Zhang, T. Wang, Z. Liu. “An end-to-end framework for shipping container corrosion defect inspection,” IEEE. T. Instrum. Meas, vol. 71, pp. 1-14, September 2022.

[2] H. Bouchekara, B. O. Sadiq, S. Zakariyya, Y. A. Sha'aban, M. S. Shahriar, M. M. Isah. “SIFT-CNN pipeline in livestock management: A drone image stitching algorithm,” Drones, vol. 7, no. 1, pp. 17, November 2023.

[3] W. X. Cai, S. L. Du, W. K. Yang. “UAV image stitching by estimating orthograph with RGB cameras,” J. Vis. Commun. Image. R, vol. 94, pp. 103835, June 2023.

[4] Q. J. Cao, Z. F. Shi, P. M. Wang, Y. Gao. “A seamless image-stitching method based on human visual discrimination and attention,” Appl. Sci-Basel, vol. 10, no. 4, pp. 1462, January 2020.

[5] W. R. Cao. “Applying image registration algorithm combined with CNN model to video image stitching,” J. Supercomput, vol. 77, pp. 13879-13896, May 2021.

[6] G. L. Chen, H. Zhou, G. Huang, G. H. Song, J. J. Zhang. “A deep image segmentation-based method for stitching ancient-book images without an overlapping region,” IET. Image. Process, vol. 17, no. 10, pp. 3068-3078, June 2023.

[7] J. Chen, Z. X. Li, C. L. Peng, Y. Wang, W. P. Gong. “UAV image stitching based on optimal seam and half-projective warp,” Remote. Sens-Basel, vol. 14, no. 5, pp. 1068, January 2022.

[8] P. K. Chilukuri, P. Padala, V. S. Desanamukula, P. P. Reddy. “L, r-stitch unit: encoder-decoder-CNN based image-mosaicing mechanism for stitching non- homogeneous image sequences,” IEEE. Access, vol. 9, pp. 16761-16782, January 2021.

[9] D. A. Delphin, M. R. Bhatt, D. Thiripurasundari. “Holoentropy measures for image stitching of scenes acquired under CAMERA unknown or arbitrary positions,” J. King. Saud. Univ-Com, vol. 33, no. 9, pp. 1096-1107, November 2021.

[10] D. J. Deng. “Smooth stitching method for the texture seams of remote sensing images based on gradient structure information,” Processes, vol. 9, no. 10, pp. 1689, July 2021.

[11] S. Eken, Ü. Mert, S. Kosunalp, A. Sayar. “Resource-and content-aware, scalable stitching framework for remote sensing images,” Arab. J. Geosci, vol. 197, no. 12, pp. 13, March 2019.

[12] X. T. Fan, L. Sun, Z. Zhang, S. Liu, T. S. Durrani. Content-seam-preserving multi-alignment network for visual-sensor-based image stitching. Sensors, vol. 23, no. 17, pp. 7488, August 2023.

[13] M. Y. Fu, H. Liang, C. H. Zhu, Z. P. Dong, R. D. Sun, Y. F. Yue, Y. Yang. “Image stitching techniques applied to plane or 3-D models: A review,” IEEE. Sens. J, vol. 23, no. 8, pp. 8060-8079, March 2023.

[14] D. Gui, Y. J. Chen, W. B. Kuang, M. T. Shang, Y. J. Zhang, Z. L. Huang. “PCIe-based FPGA-GPU heterogeneous computation for real-time multi-emitter fitting in super- resolution localization microscopy,” Biomed. Opt. Express, vol. 13, no. 6, pp. 3401-3415, May 2022.

[15] S. K. W. Hwooi, A. Q. M. Sabri. “Investigation of image stitching refinement with enhanced correlation coefficient,” Malays. J. Comput. Sci, vol. 33, pp. 22-34, January 2020.

[16] K. Jung, J. Hong. “Quantitative assessment method of image stitching performance based on estimation of planar parallax,” IEEE. Access, vol. 9, pp. 6152-6163, January 2021.

[17] J. Kang, J. Kim, I. Lee, K. Kim. “Minimum error seam-based efficient panorama video stitching method robust to parallax,” IEEE. Access, vol. 7, pp. 167127-167140, November 2019.

[18] Y. Kang, R. Wu, S. Wu, P. Z. Li, Q. P. Li, K. Cao, T. T. Tan, Y. R. Li, G. Q. Zha. “A novel multi-view X-ray digital imaging stitching algorithm,” J. X-Ray. Sci. Technol, vol. 31, no. 1, pp. 153-166, January 2023.

[19] H. P. Kuang, L. N. Zheng, G. Q. Yuan, J. J. Sun, Z. Zhang. “Error analysis and compensation in images stitching for the mechanically stitched CCD aerial cameras,” Int. J. Pattern. Recogn, vol. 33, no. 9, pp. 1955012, 2019.

[20] A. H. Li, X. S. Liu, W. Gong, W. S. Sun, J. F. Sun. “Prelocation image stitching method based on flexible and precise boresight adjustment using Risley prisms,” J. Opt. Soc. Am. A, vol. 36, no. 2, pp. 305-311, February 2019.

[21] J. M. Li, L. L. Ma, Y. X. Fan, N. Wang, K. K. Duan, Q. J. Han, X. Y. Zhang, G. Z. Su, C. R. Li, “Tang LL. An image stitching method for airborne wide-swath hyperspectral imaging system equipped with multiple imagers,” Remote. Sens-Basel, vol. 13, no. 5, pp. 1001, March 2021.

[22] W. Liu, K. H. Zhang, Y. Zhang, J. He, B. Sun. “Utilization of merge-sorting method to improve stitching efficiency in multi-scene image stitching,” Appl. Sci-Basel, vol. 13, no. 5, pp. 2791, February 2023.

[23] J. X. Luo, H. S. Tan, R. F. Wu, S. C. Zhu, H. B. Chen, J. R. Zhen, J. C. Li, C. Z. Guan, Y. X. Wu. “Reduction in required volume of imaging data and image reconstruction time for adaptive- illumination Fourier ptychographic microscopy,” J. Biophotonics, vol. 16, no. 3, pp. e202200240, November 2022.

[24] L. Nie, C. Y. Lin, K. Liao, S. C. Liu, Y. Zhao. “Unsupervised deep image stitching: reconstructing stitched features to images,” IEEE. T. Image. Process, vol. 30, pp. 6184-6197, July 2021.

[25] L. Nie, C. Y. Lin, K. Liao, Y. Zhao. “Learning edge-preserved image stitching from multi-scale deep homography,” Neurocomputing, vol. 491, pp. 533-543, June 2022.

[26] W. D. Pan, A. H. Li, Y. S. Wu, Z. J. Deng, X. S. Liu. “Research on seamless image stitching based on fast marching method,” IET. Image. Process, vol. 17, no. 14, pp. 4159-4175, September 2023.

[27] N. T. Pham, S. Park, C. S. Park. “Fast and efficient method for large-scale aerial image stitching,” IEEE. Access, vol. 9, pp. 127852-127865, September 2021.

[28] Z. Qu, J. Li, K. H. Bao, Z. C. Si. “An unordered image stitching method based on binary tree and estimated overlapping area,” IEEE. T. Image. Process, vol. 29, pp. 6734-6744, May 2020.

[29] Z. Qu, T. F. Wang, S. Q. An, L. Liu. “Image seamless stitching and straightening based on the image block,” IET. Image. Process, vol. 12, no. 8, pp. 1361-1369, August 2018.

[30] R. Z. Shao, C. Du, H. Chen, J. Li. “Fast anchor point matching for emergency UAV image stitching using position and pose information,” Sensors, vol. 20, no. 7, pp. 2007, April 2020.

[31] S. K. Sharma, K. Jain, A. K. Shukla. "A comparative analysis of feature detectors and descriptors for image stitching," Appl. Sci-Basel, 13, no. 10, pp. 6015, May 2023.

[32] M. W. Sheng, S. Q. Tang, Z. Cui, W. Q. Wu, L. Wan. "A joint framework for underwater sequence images stitching based on deep neural network convolutional neural network," Int. J. Adv. Robot. Syst, vol. 17, no. 2, pp. 1-14, April 2020.

[33] M. F. Tang, Q. Zhou, M. Yang, Y. F. Jiang, B. Y. Zhao. "Improvement of image stitching using binocular camera calibration model," Electronics, vol. 11, no. 17, pp. 2691, August 2022.

[34] C. Z. Tian, X. L. Chai, F. Shao. "Stitched image quality assessment based on local measurement errors and global statistical properties," J. Vis. Commun. Image. R, vol. 81, pp. 103324, November 2021.

[35] L. H. Wang, Y. Zhang, T. Wang, Y. S. Zhang, Z. C. Zhang, Y. Yu, L. Li. "Stitching and geometric modeling approach based on multi-slice satellite images," Remote. Sens-Basel, vol. 13, no. 22, pp. 4663, November 2021.