Optimal Region Gradated Fusion Based Automatic Image Stitching for Industrial Systems

Yanlin Huang

Hangzhou Dianzi University

Meilian Zheng (✉ 1060966654@qq.com)

Zhejiang University of Technology

Ziwei Song

Hangzhou Dianzi University School of Computer Science and Technology

Songzhu Mei

National University of Defense Technology

Zebin Wang

Shanghai Institute of Aerospace System Engineering

Gangyong Jia

Hangzhou Dianzi University School of Computer Science and Technology

---

Research

Optimal region gradated fusion based automatic image stitching for industrial systems

Yanlin Huang[1,2], Meilian Zheng[3,4*], Ziwei Song [2], Songzhu Mei [5], Zebin Wang[6] and Gangyong Jia[2]

*Correspondence:
1060966654@qq.com
[3]College of Economics and
Management, Zhejiang University
of Technology, Hangzhou, China
Full list of author information is
available at the end of the article

Abstract

In the process of equipment production in large manufacturing, the continuity of production is becoming more significant. Timely detection of equipment operation faults can ensure production continuity and greatly reduce loss. In this study, the purpose is to use multi-equipment and image stitching algorithm to obtain the complete image of a large-scale production line. An improved image stitching method based on image fusion is proposed in this paper, which mainly solves the technical problems of stitching seams, unnatural effects, and distortion after image transformation in the existing stitching technique. In the image stitching algorithm, the improved fusion algorithm based on the optimal seam and gradated in and out fusion algorithm is used to realize image fusion, including the use of dynamic programming to find the optimal seam and limit the range of fusion based on the optimal seam found. Finally, the gradated in and out fusion algorithm is used to perform fusion calculation within the limited fusion range to complete image stitching. In the end, through the comparison of different dimensional image fusion indicators with the effect of the existing fusion algorithm, the experimental results show that the method in this paper solves the problem of unnatural image stitching effect, enhances the image stitching result, and has the great fusion effect. Therefore, the panorama processed by the image stitching algorithm proposed in this paper can be efficiently processed through the industrial detection module.

Keywords: image stitching; image fusion; optimal seam; gradated in and out fusion algorithm; industrial detection

1 Introduction

At present, detection technology is developing towards online detection and non-contact detection. Image detection technology is widely used in industrial production. Monitoring the parts of large-scale systems can simplify the control of large-scale facilities and ensure safety. In the factory, the operation failure of equipment damage expensive hardware and bring huge losses to the factory. In the manual detection method, the standard consistency is reduced due to subjective reasons, and coupled with the risk of some factories, it is important to observe the system to detect or prevent potential accidents. However, the captured image can only be taken at a limited angle due to the limitation of the field of view of a single camera. There is often large equipment in the factory, which makes it impossible to use only one camera to record the complete production line, and the scenes taken by multiple cameras can only be presented in messy and scattered images. Thus, using multi-equipment and image stitching algorithm to obtain the complete image of a large-scale production line becomes particularly important. Image stitching refers

to a series of small-view and low-resolution images with overlapping areas through a certain stitching technology to generate new panoramic images with high-resolution and large-view [1]. This image contains the information of all previous series of images. Therefore, we can use image stitching to obtain a panoramic view of the production environment, and use the complete panorama for the computer vision processing part to detect.

The application field of image stitching ranges from the simple generation of panoramic images with enlarged fields of view to the seamless synthesis of panoramic images. High-quality panoramas have two important requirements: First, there should be an overlapping area between the images to be stitched, which contains similar features. Based on these features, image registration needs to be completed accurately. Second, the transition area between the stitched images should be smooth and have the smallest difference in color intensity and geometric structure. The quality of image stitching technology [2][3][4][5] directly determines the quality of the panorama. Image fusion is a key step of image stitching to eliminate stitching gaps, ghosting and other bad visual phenomenon [6]. This paper studies the basic theories and key technologies involved in image stitching technology, and analyzes and compares related algorithms [7]. In this paper, an image stitching technique based on the gradated fusion algorithm of the optimal seam is proposed, including the use of feature-based methods to extract and register the feature points of the collected images, and then use the distance algorithm: compare the nearest neighbor distance and the second nearest neighbor distance method to perform rough matching of feature points, then use the sampling consistency algorithm to refine the matching point pairs and calculate the transformation matrix to complete the image registration. For images in non-overlapping areas, if there are areas with missing image information, the image restoration method is used to fill in to ensure the smoothness and consistency of the stitched images. For the transition problem of image stitching in the overlapping area, first, the image convolution operation is used to reduce the problem of significant projective distortion caused by the image transformation. Then, a hybrid fusion method is adopted, which combines the transition smoothing method and the optimal seam method to solve the possible ghosting and double contour problems in the overlapping area of the stitched image, and improve the quality of the image stitching results.

The remainder of this paper is organized as follows: Section 2 reviews the related work. The process of the new image fusion algorithm of the image stitching method is introduced in section 3. Section 4 investigate the performance of our approach and carried out extensive experiments. Next, the experimental results and discussions are presented in Section 5. In Section 6, the conclusion is presented.

2 Related work

Image stitching is an important research topic in the field of computer vision. At present, image stitching technology has been widely used in many fields and has broad application prospects, such as: virtual reality, medical panoramic images [8][9], car driving assistance systems [10], security monitoring system [11], industrial vision-inspecting system [12], remote sensing image processing [13] and other fields. There are two specific steps in image stitching, image registration and image fusion.

Image registration is the first important step of image stitching. It is a process of spatially aligning the image with the overlapping area. The registration method based on image feature matching has received extensive attention and research in the field of image feature processing due to its advantages such as illumination invariance and rotation invariance. Capel et al. proposed Random Sample Consensus (RANSAC) [14] to purify the matching points, which is a classic algorithm in the field of image stitching technology. In 2004, Lowe proposed the Scale Invariant Feature Transformation algorithm (SIFT) [15], which uses the Gaussian scale pyramid method to detect feature points. This algorithm has achieved excellent results in image stitching. Sift has scale and rotation invariance, and is robust to light intensity and noise, which is still widely used and is a milestone in image stitching technology. At the same time, Lowe proposed the sift matching method to use the nearest neighbor distance and the next nearest neighbor distance to further filter the matching points [15], and obtain excellent matching points. In 2007, Lowe used the global homography matrix [16] to complete image registration. The global method can get good registration accuracy in the case of small parallax, but for scenes with big parallax, literature [17] is proposed to align the image with partial blocks of the image to improve the accuracy of image registration, that is the APAP warp locally adapts to different transformations. The previous methods have been successful in terms of parallax, but due to the local homography matrix transformation in the overlapping area, the image caused significant projective distortion in the non-overlapping area.

In addition, in the overlapping area, there are more stitching dislocations, resulting in obvious ghosting. Image fusion [18][19] is used to reduce the stitching seam of overlapping areas of images to achieve natural visual effects, including multi-band fusion and gradient template weighted fusion [20][21]. At present, the weighted average algorithm and the Laplacian fusion algorithm have been widely used in the field of image processing, and fusion algorithms can achieve smooth transition in the overlapping area in a certain extent. However, when dealing with moving objects in the overlapping area, ghost images may still appear, and accurate results cannot be obtained. Using different image weights and dividing regions [22] can eliminate the problem of ghosting, but it produces some distortion in the stitching result image. In this paper, the principle of image stitching is followed, combining smooth transition and optimal stitching seam technology. The optimal stitching seam gathers the color difference, gradient difference and texture complexity, so that we can find the seams that pass through the high similarity area, avoid the salient area, and pass through the specific low texture area. Then use transition smoothing on the selected area around the optimal stitching seam, which reduces ghost to a greater extent and improves the quality of image fusion.

3 Panoramic image stitching method

3.1 Stitching process

The basic process of image stitching can be divided into three parts: Image acquisition, image registration, and image fusion. After the images are obtained by the camera, first preprocess multiple groups of images, then register the images, and finally fuse the image overlapping areas to make the generated panorama looks more

natural. Among them, image registration and image fusion are the most critical steps.

We can use the devices with camera and storage functions to obtain two or more images with overlapping areas. Image preprocessing is a series of processing in order to complete image registration better, which mainly undertakes two tasks. One is to carry out basic operations of digital image processing, such as uniform size, smooth filtering, and enhancement transformation, to solve the influence of light and noise on the image during shooting and transmission. The other is to correct the lens distortion of the image, in order to solve the problem of radial distortion caused by the camera. Image preprocessing can be processed according to the needs and influence of the captured image.

The purpose of image registration is to establish the geometric correspondence between images. It is a registration process to align the overlapping regions of two images by extracting and matching the feature points in the images. The common feature extraction algorithms include Harris corner detection algorithm, sift algorithm, surf algorithm, and orb algorithm. Among them, sift algorithm maintains certain stability for image rotation, scale scaling, brightness change, and projective change. In this paper, the image feature matching algorithm is shown in Algorithm 1, where we can find that sift algorithm is used to extract image feature points, and the KNN matching algorithm is used to match feature points to obtain set C. For set C, the matching pairs filtered by Lowe's algorithm[15] are taken as the initial matching results, and the RANSAC algorithm [14] is used to remove the external points and retain the internal points for accurate filtering. Then, estimation of homography, warping is applied respectively in order to mosaic the images.

The purpose of image fusion is to improve the stitching effect. After the multiple images with overlapping areas are stitched together, the overlapping areas of the stitching will produce blur and ghosting. The purpose of image fusion is to achieve a smooth transition of the stitched overlap areas while keeping the information of the original image as much as possible to avoid information loss. In section 3, the image fusion algorithm proposed in this paper will be introduced in detail.

Algorithm 1 Image Matching Algorithm

Require: image A and image B captured by devices after pre-processing,
        kpsA, kpsB: extract feature points based on sift algorithm,
        featuresA, featuresB: extract feature vectors based on sift algorithm,
        rawmatches: the two most matching feature vectors from all feature vectors,
        kernel_sharp: [[0,-1,0],[-1,5,-1],[0,-1,0]]
Ensure: direct stitched image, overlap area S.
1: **for** each m $\in$ rawMatches **do**
2:    **if** m[0].distance $<$ m[1].distance *0.75 **then**
3:       add m into matches
4:    **end if**
5: **end for**
6: **if** len(matches) $>$ 4 **then**
7:    compute homography and construct warpR based on RANSAC
8: **end if**
9: realize convolution of image based on kernel_sharp,
10: construct direct stitched image

3.2 Image transformation

In order to make the images to be stitched in the same spatial coordinate system, it is necessary to use an image transformation model to calculate the spatial coordi-

nate parameters between feature points. The common image transformation models include rigid transformation, affine transformation, and projective transformation.

The specific transformation form is shown in Table 1. These transformations are also called parametric models, and these models make the image map the coordinates x to a new set of coordinates x'. In Table 1, for rigid transformation, R is an orthogonal rotation matrix. For affine transformation, A is an arbitrary (2×3) matrix. For projective transformation, H is a (3×3) matrix, which is homogeneous [23].

Table 1 2-D Coordinate transformation

Parametric model	Transformation equation	Matrix	DoF	Preserves
Translation	$x'=x+t$	$[I\vert_t]_{2*3}$	2	Orientation
Rigid	$x'=Rx+t$	$[R\vert_t]_{2*3}$	3	Length
Affine	$x'=Ax$	$[A]_{2*3}$	6	Parallelism
Projective	$x'=Hx$	$[H]_{3*3}$	8	Straight lines

Among these transformations, projective transformation [23] has the advantage of not maintaining the parallelism between lines, which is more common than affine transformation and rigid transformation. In the projective transformation of an image, the four corners points of the bounding box do not need to be constrained in their side length direction and can be moved to obtain a more complex projective transformation effect. In this paper, the projective transformation is used to project the image to a new view plane. In homogeneous coordinate system, the projective transformation on the 2D plane can be described in the following nonsingular 3x3 matrix form. A projective transformation is defined as a linear transformation with homogeneous coordinates, and the formula is as follows:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ m_6 & m_7 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = H \begin{bmatrix} X \\ y \\ 1 \end{bmatrix} \tag{1}$$

Where H is a homography matrix, which describes the motion relationship between images in different coordinate systems. The projective transformation matrix H expresses the motion of 8 degrees of freedom, which can cover all kinds of motion forms such as translation, rotation, scale change, etc. After projective transformation, the existing images will be distorted to a certain extent, and then the image distortion will be increased after zooming. In view of this, in order to reduce the distortion of the image after projective transformation and improve the visual quality of the image, this paper uses sharpening processing, specifically, uses the kernel to realize the convolution operation of the image, which can highlight the clarity of the processed image and enhance the blurred level of detail.

3.3 Optimal region gradated image fusion

Due to the influence of various factors in the process of image collection, such as exposure time, intensity, weather conditions, the illumination between the images is uneven, and there will be obvious colors and stitching gaps in the middle of the composite images, and blurring and distortion.

Efros and Freeman [24] proposed to use dynamic programming to solve the optimal seam in the image overlapping area. The optimal seam based on the idea of dynamic programming is an optimization algorithm, which can reduce the stitching phenomenon caused by the gray difference between images. The stitching line should meet the following two points in terms of color intensity and geometric structure [25] :

(1) From the point of color intensity, original images pixel color vary the color difference of pixels on two original images is the smallest.

(2) From the point of view of geometric structure, the structure of the corresponding points on the two original images is the most similar.

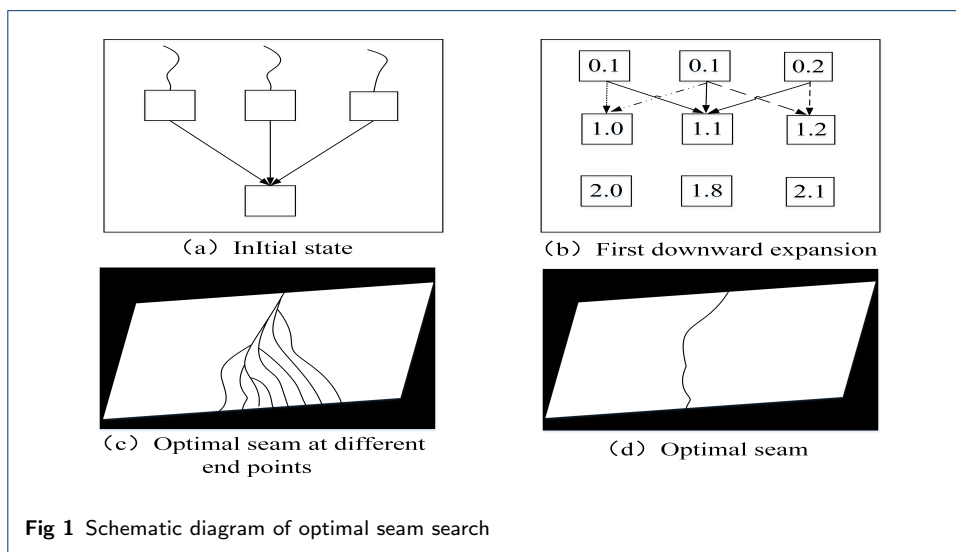The calculation criterion of the optimal seam as shown in equation 2,

$$E(x, y) = E_{color}^2 (x, y) + E_{geometry} (x, y) \tag{2}$$

Among them, $E_{color}(x, y)$ is the value of image color difference, and $E_{geometry} (x, y)$ is the intensity value of image structure difference;

The solution criterion of $E_{geometry} (x, y)$ is as follows:

$$E_{geometry}(x, y) = [S_x \cdot (I_1(x, y) - I_2(x, y))]^2 + [S_y \cdot (I_1(x, y) - I_2(x, y))]^2 \tag{3}$$

Where S_x and S_y denotes the template of 3×3 Sobel operators in X and Y directions respectively.

(a) InItial state

(b) First downward expansion

(c) Optimal seam at different end points

(d) Optimal seam

Fig 1 Schematic diagram of optimal seam search

In the overlapping region, the first row of pixels in the image is taken as the starting point to establish multiple stitching and the one with the least color intensity

and structural strength is found as the optimal seam. Algorithm 2 shows the process of the optimal seam search based on dynamic programming and its schematic diagram shown in Figure 1.

Figure 1(a) is a set of initial stitches and each pixel in the first row is taken as the seam starting point. Figure 1(b) shows the situation when the seam starts to expand downward, considering the points of each next row adjacent to it. The point with the minimum criterion value is the extension point. As shown in Figure 1(d), finally, among the seam obtained as shown in Figure 1(c), the seam with the minimum criterion value is selected as the optimal seam.

The specific steps are as follows:

(1) In the overlapping region, each first line image pixel is taken as the starting point of a suture line and the standard value is taken as the intensity value.

(2) Continue to search the next line, the search directions are 45 degrees below the current point, 45 degrees below the left and right suture line. The three points criterion values are compared with those two adjacent points of the current point by using the calculation criteria of the optimal suture line. The pixel with the smallest criterion value is updated to the current point.

(3) If the last line point found overlapped image, go to step 4), otherwise, return to step (2) and continue the next search.

(4) Among all the seams, the lowest strength value is selected as the optimal seam.

Algorithm 2 Optimal seam algorithm

Require: overlap area S.
Ensure: best seam.
 1: **for** each $i \in S[i]$ **do**
 2:     **while** i is not last one **do**
 3:         $E(i,j) = E\_color + E\_geometry,$
            $j \in$ (left,right,next_left,next_mid,next_right)
 4:     **end while**
 5: **end for**
 6: select min $E(i,j)$ then j=i
 7: return optimal seam: F_line

The optimal seam method described previously is an effective method to achieve a smooth transition of the overlapping parts of the Mosaic image. In this algorithm, a minimum cost stitching line is obtained in the overlapping area for stitching, so as to solve the ghosting caused by the fusion of information of two images to a certain extent. However, in the process of image acquisition by different devices, due to the influence of many external factors, it is easy to cause differences in brightness and tone between images, which will directly affect the quality of the final generated panorama. Therefore, after the calculation of the optimal suture line, we also need to combine the image fusion technology to adjust the image coincidence area, so as to reduce the difference after the merging of the two areas.

In this paper, the optimal seam obtained by the above methods is used to fuse two images by selecting regions on both sides of the suture according to the weight. In this paper, the gradated in and out pixel level image fusion algorithm is used to calculate the fusion on both sides of the seam, which can eliminate the mosaic effect caused by the difference of color brightness to a certain extent.

The gradated in and out algorithm uses different methods to select the weighted coefficient function, and the weight is gradually changed in the process of fusion, so

Fig 2 Schematic diagram of optional fusion area
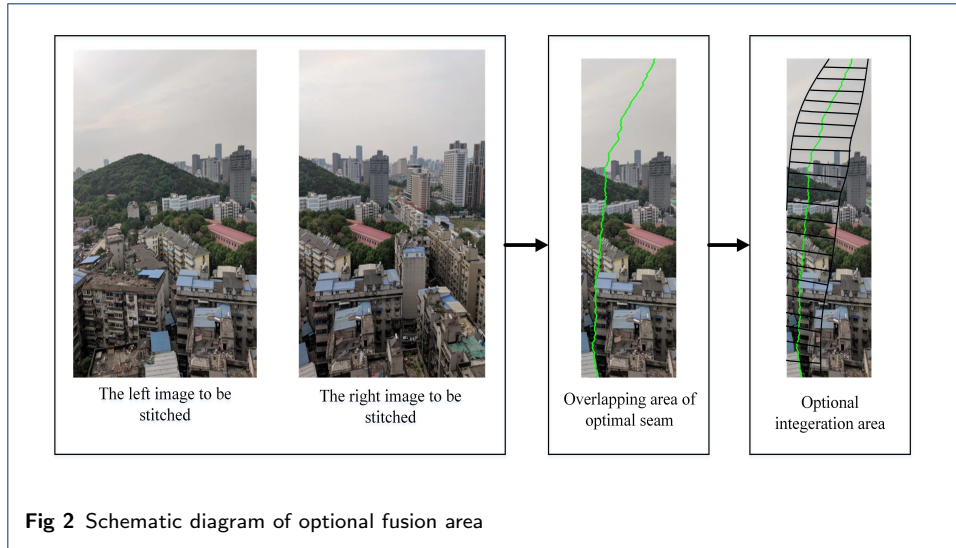
as to obtain smooth fusion results. The calculation formula of fused image I(x, y) is as follows:

$$
I(x,y) = \begin{cases} I_1(x,y) & (x,y) \in I_1 \\ \omega_1 I_1(x,y) + \omega_2 I_2(x,y) & (x,y) \in (I_1 \cap I_2) \\ I_2(x,y) & (x,y) \in I_2 \end{cases} \tag{4}
$$

In formula (4), ω_1 and ω_2 are the corresponding selected weighted coefficients, which satisfy the following requirements: $\omega_1 = (x_r - x_i)/(x_r - x_1)$, $\omega_2 = (x_1 - x_1)/(x_r - x_1)$ and $\omega_1 + \omega_2 = 1$.

Where X_i is the abscissa of the current pixel; X_1 is the left boundary of the overlapping region; and X_r is the right boundary of the overlapping region.

Figure 2 shows the range of the fusion region determined within the overlapping region of the image to be spliced. Specifically, when calculating the pixels, the fusion range is determined as follows:

(1) The left boundary start and right boundary end with a minimum circumscribed matrix of the optimal seam are calculated and the limited fusion region D is determined according to the position of the optimal seam.

(2) Take the current each line seams point x, compare the left and right distance between the points and boundaries. Select the shortest distance between the current point of the seam and the two boundaries as the limited ranged of current line fusion, and the limited boundary is [x-d, x+d].

(3) By traversing, the range of fusion region D is determined according to the d of each row. Basically, it is an irregular region.

In the limited fusion range, the gradated in and out fusion algorithm is used to calculate as shown in Algorithm 3. Specifically, in the determined fusion region D, the pixel values of the images to be fused A and B are set as I_1(x, y) and I_2(x, y),

then the pixel values of the transition processed image are calculated by formula 4 at the limited fusion position. Within the scope of a non-fusion region in overlapping region S, if it is on the left, the pixel is $I_1(\text{x, y})$ of the left image; if it is on the right, the pixel is $I_2(\text{x, y})$ of the left image.

Algorithm 3 Computing fusion zone and calculating pixel value

---

Require: F_line,
    start:left border of overlap area S,
    end: right border of overlap area S.
Ensure: panorama.
 1: **for** each x(i) $\in$ F_line **do**
 2:    **if** x(i)-left ¡ right-x(i) **then**
 3:        d(i) = x(i)-left;
 4:    **else**
 5:        d(i) = right-x(i);
 6:    **end if**
 7: **end for**
 8: zone D={x(i)-d(i),x(i)+d(i)};
 9: **for** each I(i,j) $\in$ D **do**
    I(i,j)=$\omega_1 I_1(i,j)+\omega_2 I_2(i,j)$;
10: **end for**
11: return panorama

---

4 Experiments

In order to illustrate the effectiveness of the algorithm. The experimental environment is as follows: Intel core i7-7700/3.60GHz CPU, 16GB memory, 64-bit Windows 10 operating system, the programming language is Python, running in PyCharm 2020.

4.1 Image data source

The image stitching dataset in this experiment to verify the effectiveness of the algorithm is provided by literature [26], taken by a drone (KC1600) in Hejiang East Village, China. It contains 295 images, each with a resolution of 736x491, and the average flight altitude is 988m.

4.2 Perspective transformation image distortion processing

Figure 3(a)(b) is an adjacent group of images to be stitched in the data set, Figure 3(a) is the left image to be stitched, and Figure 3(b) is the right image to be stitched.

The stitching algorithm proposed in this paper uses the sift algorithm to extract features of the target image and the reference image respectively, and establish a set of feature description points. The feature descriptor subset obtained from the target image and the reference image, including the location, direction and size of the feature points. Then use the KNN matching algorithm to perform feature point matching to obtain the set C, and use the matching pair filtered by the Lowe's algorithm on the set C as the result of the initial matching. Next, projective transformation is applied to the reference image, and the reference image is projected onto the viewing plane of the target image. The image after projective transformation is shown in Figure 3(c). It can be seen from the figure that the projective transformation of the image to be registered after the projective transformation has
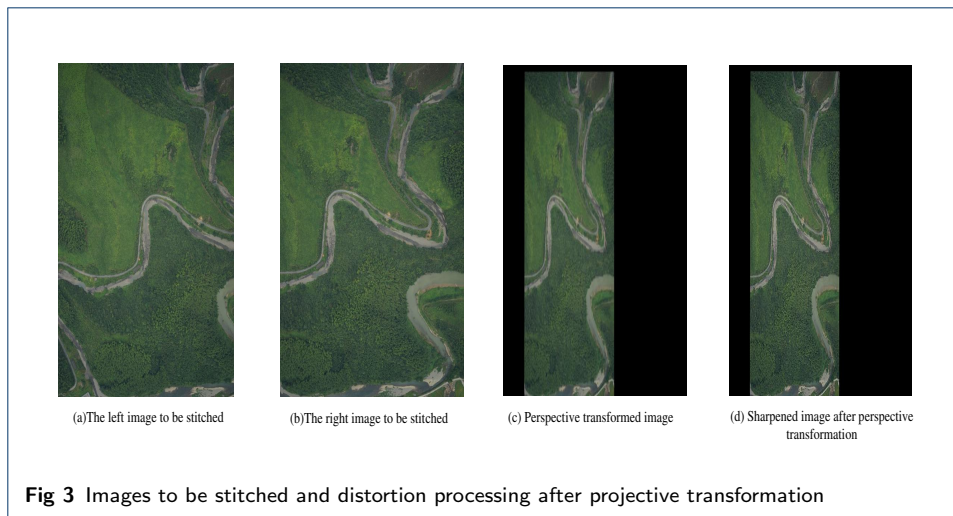
(a)The left image to be stitched   (b)The right image to be stitched   (c) Perspective transformed image   (d) Sharpened image after perspective transformation

Fig 3 Images to be stitched and distortion processing after projective transformation

caused the original image to be stretched and other changes. The rear view image is distorted and looks blurred visually.

In view of this, this article sharpens the transformed view image. According to the Laplacian operator, the template coefficients are deformed on the original basis of the operator and use the mask matrix [[0, -1, 0], [-1, 5, -1], [0, -1, 0]] to realize the convolution operation on the image.

This matrix can add the negative number of the value of the four neighboring points of the center pixel, plus five times the value of the center pixel, which can highlight the location of the sudden change in the brightness value to make the edges, contours, and details of the image clear. The effect of enhancing the image. The processed image is shown in Figure 3(d). After the sharpening process of the convolution operation is performed on the image, it can be seen from the projective that the clarity of the processed view image is highlighted, and the blurred details are enhanced, ultimately improving the quality of the stitched image.

Table 2 Evaluation of the effect of perspective transformation image distortion processing

Image	Image Processing	SF	AG
Figure 3(c)	Perspective transformed image	5.424	2.567
Figure 3(d)	Sharpened image after perspective transformation	8.941	8.095

Evaluation indicators for evaluating the effect of image fusion: Spatial Frequency and Average Gradient, which can be used to measure the clarity of the fused image. The larger the value, the better the image clarity and the better the fusion quality. Table 2 shows the fusion quality evaluation of Figure 4(c) and 4(d). From the test data in Table 2, the spatial frequency of the image has changed from 5.424 to 8.941, and the average gradient has changed from 2.567 to 8.095. This shows that the image sharpness has improved after the image is sharpened by the convolution operation, which helps to improve the quality of the stitched image, enhances the blurred details and helps the subsequent processing of the image.

4.3 Assessment methods of different fusion algorithms

Figure 4(a) shows the result of direct stitching of remote sensing images. It can be seen from the figure that due to the difference in gray scale of remote sensing images, there are obvious splicing seams in directly stitching images (marked with red circles in the figure). Figure 4(b), Figure 4(c) and Figure 4(d) show directly based on the optimal seam stitching algorithm, the directly gradated in and out fusion algorithm and Laplace algorithm can solve the color difference between the overlapping areas around the stitches. However, it can also be seen from the figure that the stitched images finally obtained by the three existing algorithms have black areas on the edges, and there are also black lines that affect the visual effect (marked with red circles in the figure). Because the image stitching is carried out directly using the optimal seam and Laplace fusion algorithm, the pixel-level fusion processing of the re-converged area is not carried out, resulting in the black area of the edge increasing, affecting the visual effect and also affecting the image information. The problem with the directly gradated in and out fusion algorithm is that if there are moving objects in the overlapping area, it is easy to produce ghosting by directly calculating pixel fusion.

(a)The image to be merged after stitching   (b)Directly based on the Optimal seam algorithm

(c)Directly gradated in and out fusion algorithm   (d)Laplace algorithm   (e)Based on optimal seam and gradated in and out fusion algorithm

Fig 4 Effects of different image fusion algorithms
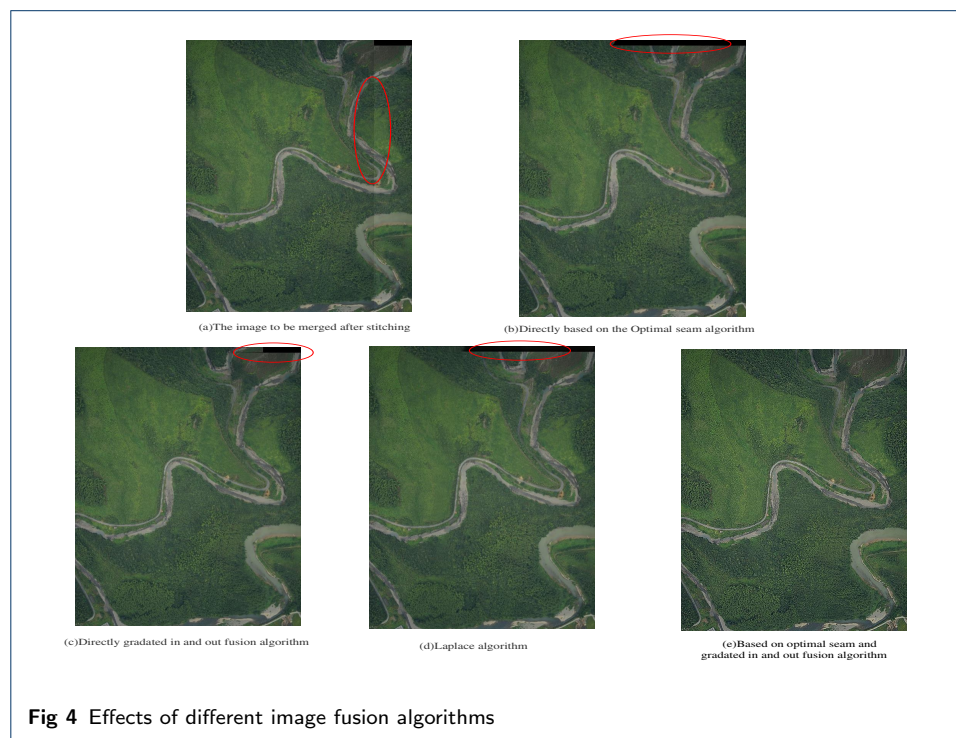
In view of this, an algorithm based on the optimal seam self-selecting area gradual in and out of fusion is proposed in this paper. Pixel fusion calculation is based on the optimal seam to select a small area, while simply doing an image repair process to restore image information for small black areas on the edges. Reduce the impact of distortion problems caused by image transformation by sharpening processing. The final stitching result is shown in Figure 4 (e). It can be seen that the fusion image is not only clearer but also more uniform in transition, contains more image information, and can achieve a good visual effect.

In order to evaluate the quality of the fusion image more objectively, the experimental results of this paper are based on different dimensions of the fusion image quality evaluation index to evaluate and compare the fusion results[**?**], and then verify the effectiveness of the algorithm in this paper. The evaluation indicators include standard deviation, spatial frequency, average gradient, entropy, mutual information. Among them, standard deviation, spatial frequency, and average gradient are indicators based on statistical evaluation, information entropy and mutual information are indicators based on information entropy.

The standard deviation (STD) reflects the degree of dispersion between the image pixels value and the mean value, it's an objective evaluation index to measure the richness of image information. The larger value scattered the gray level distribution of the image, the more information image carries the better the quality fusion image get. The calculation formula is as follows:

$$\text{STD} = \sqrt{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}[H(i,j) - \bar{H}]^2} \tag{5}$$

The spatial frequency (SF) reflects the image gray level change rate. The larger the spatial frequency, the clearer image, and better the quality fused image. The calculation formula is as follows:

$$SF = \sqrt{RF^2 + CF^2} \tag{6}$$

among them,

$$\text{RF} = \sqrt{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}|H(i,j) - H(i,j-1)|^2} \tag{7}$$

$$\text{CF}\sqrt{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}|H(i,j) - H(i-1,j)|^2} \tag{8}$$

The average gradient (AG) is used to measure the clarity of the fusion image. The larger the average gradient, the better image clarity, and better fusion quality. The calculation formula is as follows:

$$\text{AG} = \frac{1}{(M-1)(N-1)}\sum_{i=1}^{M-1}\sum_{j=1}^{N-1}\sqrt{\frac{(H(i+1,j) - H(i,j)^2 + H(i,j+1) - H(i,j))^2}{2}}$$

$$\tag{9}$$

Entropy (EN) is an objective evaluation index that measures how much information an image contains. The larger entropy, the richer information and better the quality of the fused image. The calculation formula is as follows

$$H(A) = -\sum_a P_A(a) \log p_A(a) \tag{10}$$

Where a is the gray value, and P(a) is the gray probability distribution.

Mutual Information (MI) can measure by the degree of similarity between the fused image and the original image, that is, how much information the fused image has obtained from the original image. The larger the mutual information, the more fused image retains source image information and better quality. Mutual information is calculated based on the information entropy H(A) and joint information entropy H(A, B) of the image. The formula for calculating mutual information is as follows:

$$\mathrm{MI(A,B)} = \mathrm{H}(A) + \mathrm{H}(B) - \mathrm{H(A,B)} \tag{11}$$

Table 3 shows the image quality evaluation results based on spatial frequency, standard deviation, entropy, mutual information, average gradient. It shows from the test data in the table that the image to be stitched in Figure 3(a)(b) have been stitched and fused by the algorithm in this paper. The entropy changed from 11.5, 11.6 to 13.3, and the mutual information changed from 15.2, 15.3 to 16.7, which indicates the proposed algorithm makes the fusion image transition uniform, retains more resourceful image information, and achieve better visual effects. Similarly, the standard deviation, the largest stitched result image obtained by the algorithm in this paper, indicates more information and has better quality. Meanwhile, the values of spatial frequency and average gradient have been greatly improved that indicating the improved sharpness of the stitched image obtained by the algorithm. Starting from the above indicators to evaluate the image quality, in this paper, the algorithm can make the image stitched result clearer, best quality, and contain more source image information.

Table 3 Evaluation of the effect of single-case image fusion algorithm

| Different fusion algorithms	STD	SF	AG	EN	MI
DBOOS	23.21	8.68	5.96	11.55	15.21
DGIOF	22.17	9.04	6.37	11.67	15.36
LAP	23.11	8.46	5.73	11.50	15.16
BOSGIOF	**30.94**	**13.38**	**23.48**	**13.36**	**16.72**

5 Results and discussion

In order to avoid the contingency of the experiment, this experiment uses the remote sensing image datasets [26] taken by the drone (KC1600) in Hejiangdong village to

stitch 295 images and use the five above-mentioned indicators to evaluate image quality to compare the effects of different fusion algorithms.

In order to better distinguish and analyze multiple sets of data, firstly, according to the different content contained in the datasets, that is, whether the remote sensing image datasets contain houses, roads, rivers, or only woods, and all datasets are divided into 4 categories. Among them, there are 87 result images for houses, 34 result images for roads, 15 result images for rivers, and 36 result images for forests, for a total of 172 result images.

According to the directly based on the optimal seam algorithm, the directly graded in and out fusion algorithm, the Laplace algorithm, and an algorithm based on the optimal seam gradated in and out of fusion proposed in this paper, the results obtained by these four different fusion algorithms, and calculate the five index values of standard deviation, spatial frequency, average gradient, entropy and mutual information for objective comparison. These evaluation indicators are all positive indicators, the objective evaluation results Q_1 and Q_2 given by the indicator M of any two fused images I_1 and I_2, When the subjective evaluation result in I_1 is better than I_2, $Q_1 > Q_2$. finally, obtained the data index values are classified according to data set categories, and each category takes the average value of the corresponding data for analysis and comparison, as shown in Table 4. In order to compare the fusion effects on different algorithms more intuitively, the data in Table **??** is drawn as a histogram as shown in Figure 5.

Fig 5 Comparison of results of different fusion algorithms
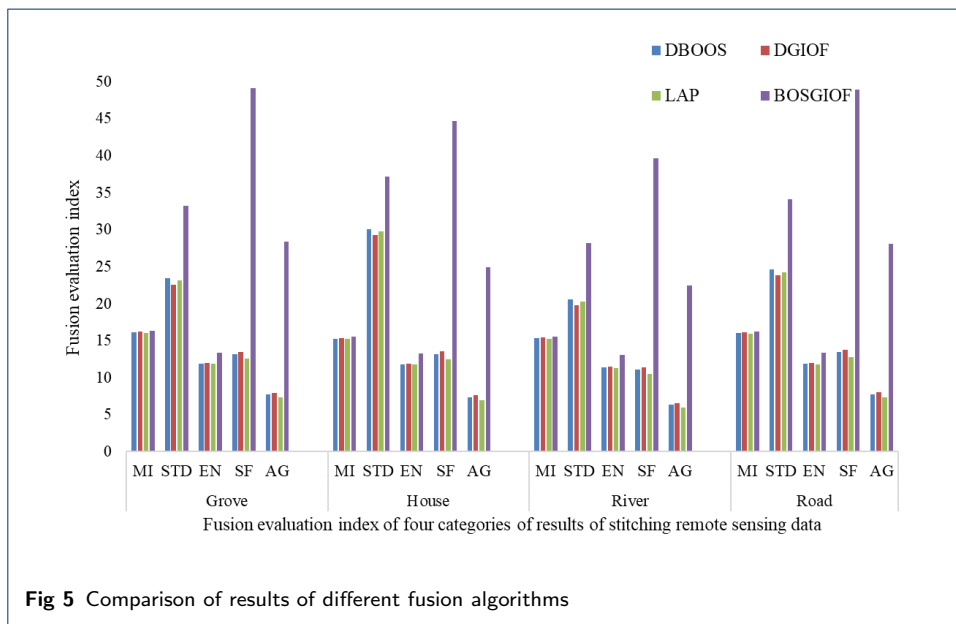
Standard deviation (STD) is used in image quality evaluation. The richness of image information is measured by calculating the difference between the pixel's point and the average value of the pixel. If STD->0, the pixel value of each point of the image has almost no fluctuations above and below in its average, the image is filled with the same gray level. From the distribution of standard deviation in Figure 5, the standard deviation value of the image with house and road information is relatively high, that is, considering the degree of deviation of the pixel gray from

the mean value, the image content is richer. In this paper, some processing is done to solve the problem of image distortion based on the optimal seam and gradated in and out algorithm after optimization. Convolution is performed on the image, which makes the pixel value fluctuate more than the average value, indicating that the image contains richer information.

Spatial frequency (SF) and average gradient (AG), as evaluation indicators based on statistical characteristics, can be used to measure the clarity of an image. According to the distribution of spatial frequency and average gradient shown in Figure 5, the spatial frequency and average gradient value of the proposed fusion algorithm in this paper are four times those of the other three existing fusion algorithms. From the calculation formula of spatial frequency and average gradient given above, we can know that both indicators are calculated based on the difference of adjacent gray values in the image. Figure 6 shows the partial gray values of a certain group of data in the datasets processed by four different fusion algorithms in the vicinity of the stitching area. It can be seen from the figure that the difference between adjacent gray values of the fusion result obtained by the algorithm in this paper fluctuates the most. Therefore, the spatial frequency and average gradient index values of the resulting image are much larger than the index values of the existing fusion algorithm that has not solved the image distortion problem, resulting in the histogram distribution of the spatial frequency and average gradient as shown in Figure 5. The stitching algorithm proposed in this paper performs sharpening of the enhanced image in the registration stage, so its spatial frequency and average gradient index value are greatly improved compared with the other three algorithms. This shows that the sharpness of the stitched image obtained by the algorithm in this paper has been improved, and the blurred part of the image details is enhanced, which is more conducive to subsequent image analysis.

Entropy is an index based on the amount of information, which is mainly calculated by gray value and gray probability distribution and counts the widespreadness of pixel gray distribution in the image. This indicator can reflect the richness of the image. The higher the information entropy value, the wider the gray distribution of image pixels and the richer the information in the image. Judging from the distribution of entropy shown in Figure 5, the entropy value of the image containing forest and road information is the highest in comparison, that is, considering the gray distribution of image pixels, the image content is richer. At the same time, it can be seen from the figure that the entropy values of the results of the existing three fusion algorithms are relatively close, and the entropy value of the splicing processing fusion algorithm proposed in this paper is higher than other three algorithms, indicating that the image pixels obtained by the algorithm of this paper have a wide distribution of gray levels, and the information in the image is more abundant.

Since entropy has nothing to do with the source image, it can only explain the size of the information of the evaluated image, and cannot fully prove the degree of information retention of the stitched image compared to the source image. Qu et al. proposed to use mutual information (MI) to evaluate the quality of image fusion [27]. In the calculation, only the fusion result map and the source image A and the source image B are separately calculated for mutual information, and then

the average weights of the two are added together to obtain the mutual information index values shown in Table 4. Observed from the distribution of mutual information index values shown in Figure 5, due to the large overlapping area of adjacent remote sensing data, each fusion algorithm retains the information of the source image better, and the mutual information index values are very close.

Table 4 Effect evaluation of image fusion algorithm

	STD			
	DBOOS	**DGIOF**	**LAP**	**BOSGIOF**
Grove	23.4	22.5	23.09	**33.2**
House	30.08	29.29	29.8	**37.2**
River	20.6	19.76	20.26	**28.13**
Road	24.57	23.84	24.27	**34.05**
	Spatial_frequency			
	DBOOS	**DGIOF**	**LAP**	**BOSGIOF**
Grove	13.21	13.49	12.57	**49.13**
House	13.2	13.6	12.5	**44.62**
River	11.12	11.36	10.49	**39.62**
Road	13.47	13.77	12.74	**48.93**
	Avg_Gradient			
	DBOOS	**DGIOF**	**LAP**	**BOSGIOF**
Grove	7.7	7.9	7.34	**28.4**
House	7.33	7.6	6.93	**24.94**
River	6.34	6.51	5.98	**22.41**
Road	7.79	8	7.36	**28.06**
	Entropy			
	DBOOS	**DGIOF**	**LAP**	**BOSGIOF**
Grove	11.88	11.98	11.84	**13.4**
House	11.81	11.91	11.75	**13.29**
River	11.39	11.48	11.32	**13.07**
Road	11.88	11.97	11.83	**13.39**
	MI			
	DBOOS	**DGIOF**	**LAP**	**BOSGIOF**
Grove	16.11	16.23	16.07	**16.37**
House	15.27	15.39	15.21	**15.52**
River	15.36	15.46	15.28	**15.57**
Road	15.99	16.09	15.93	**16.25**

Table 5 The objective evaluation index results of image fusion quality in literature [28]

	Direct gradated in and out fusion algorithm	Laplace algorithm	Fusion algorithm proposed in literature [28]
EN	7.183	7.122	7.223
AG	6.693	6.025	7.474
STD	74.929	72.137	74.959

Image fusion algorithm is widely studied by many scholars in the field of image mosaic. An improved fusion algorithm is proposed in the literature [28]. At the same

Fig 6 Comparison of partial pixels in different fusion algorithms
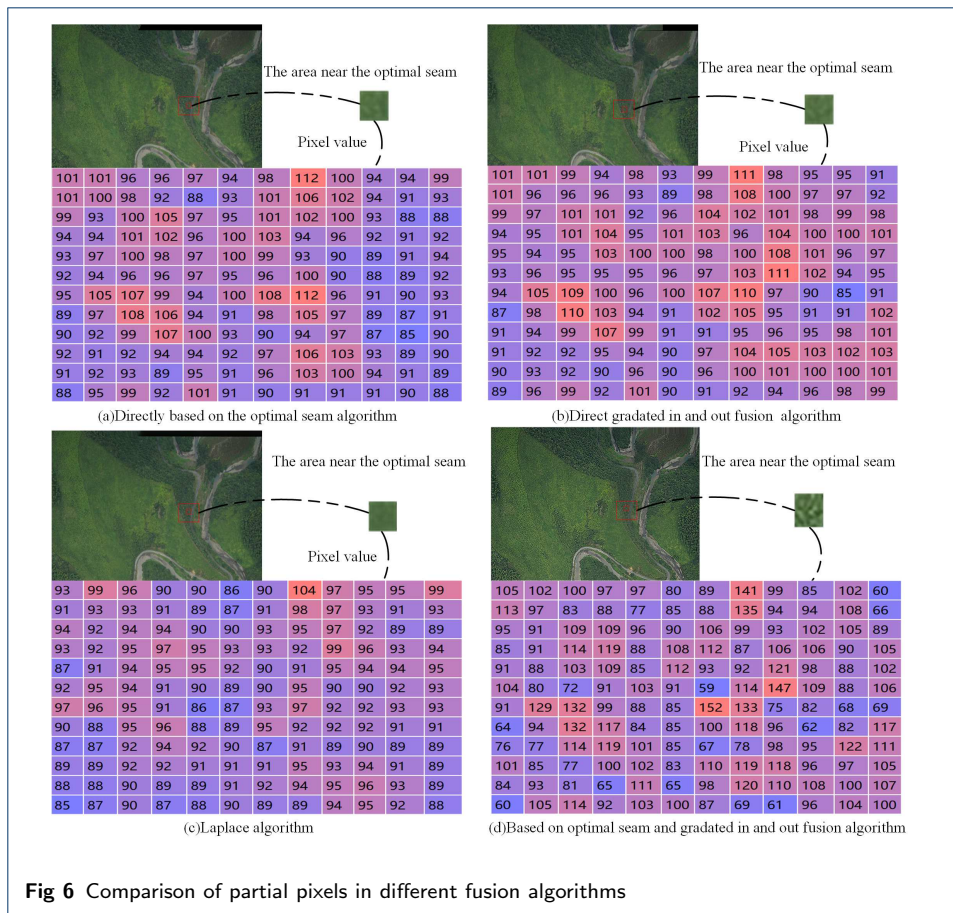
time, it is also objectively quantified and evaluated with the mosaic image under the existing fusion algorithm mentioned above, and the calculated index results are shown in Table 5.

We can know that based on the index of entropy, the fusion algorithm in literature [28] is improved by 0.55% compared to the gradual in and out fusion algorithm, and 1.4% compared with the Laplace algorithm. Our algorithm is 10.6% higher than the gradual in and out fusion algorithm, and 11.64% higher than the Laplacian algorithm. Based on Avg_Gradient, the fusion algorithm in literature [28] is 10.81% higher than the gradual in and out fusion algorithm, and 19.43% higher than the Laplace algorithm. Our algorithm is 72.18% higher than the gradual in and out fusion algorithm, and 74.15% higher than the Laplacian algorithm. Based on STD, the fusion algorithm in literature [28] is 0.04% higher than the gradual in and out fusion algorithm, and 3.76% higher than the Laplace algorithm. Our algorithm is 21.26% higher than the gradual in and out fusion algorithm and 19.9% higher than the Laplacian algorithm.

For the fusion algorithm proposed in this paper, we considered the image distortion and the loss of image edge information due to image transformation and improved the stitched image through image convolution processing and image restoration. The clarity of the output panoramic is improved, and the degree of retention of the source image information has been greatly improved while preserving high-frequency details. And the use of an improved fusion algorithm based on the optimal

seam can better solve the problem of ghosting and double contours in the overlapping area of the image, and improve the quality of the image stitching result.
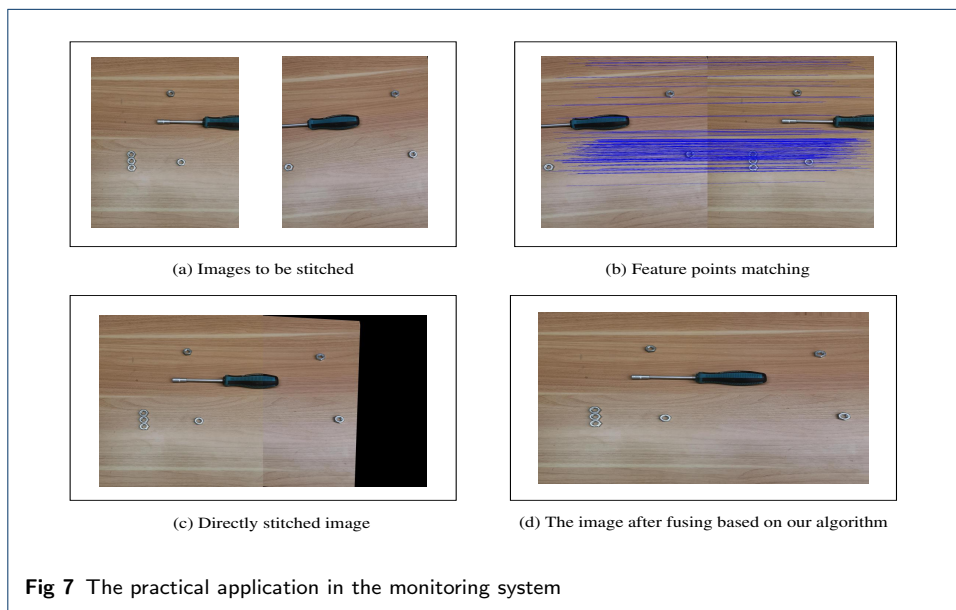

(a) Images to be stitched   (b) Feature points matching

(c) Directly stitched image   (d) The image after fusing based on our algorithm

Fig 7 The practical application in the monitoring system

6 Conclusion

In this paper, based on the research of fusion method in image stitching technology, in order to solve the problem of stitching seam caused by large difference of gray level, a gradated in and out fusion algorithm based on the optimal seam is proposed, which realizes the smooth transition of image overlapping area and effectively eliminates the stitching gap. The basic theory and key technology of image stitch technology are studied in this paper. The relevant algorithms are analyzed and compared, then, the sequence image stitch fusion results of this method are presented, and the subjective comparison and objectivity with the previous algorithms are carried out the evaluation. Through the experimental comparison, it is proved that the algorithm proposed in this paper has obvious advantages compared with the directly based on the optimal seam algorithm, the directly gradated in and out fusion algorithm, and the Laplace algorithm. It can effectively eliminate the seam and make the fused image clearer and have a better visual effect.

The algorithm proposed in this paper can be applied to factory visual inspection. In the process of the actual production line, the position of the workpiece may not be determined. The panorama created by the image stitching algorithm can determine the position of the workpiece and extract more detailed information. As shown in Figure 7(a), in one image, the workpiece is not completely photographed, which will affect the subsequent visual detection. Therefore, the image feature points are extracted and matched for the two images. The resulting image of feature point matching is shown in Figure 7(b), and the resulting image not fused is shown in Figure 7(c). After using the image fusion algorithm in this paper, a complete and clear panorama is obtained as shown in Figure 7(d).

Before using the image stitching algorithm, first, it is necessary to determine the position of the camera, and the angle of view of each camera includes a small

part of the adjacent camera. Then, by finding a similar scene of the input image, homography estimation, image transformation, and image fusion are used to realize the image stitching. After all steps are completed, all these images taken by the camera at the same time are spliced, and finally, a wide-angle clear mosaic panorama is obtained for processing by the visual detection.

Abbreviations
DBOOS: Directly based on the optimal seam algorithm; DGIOF: Direct gradated in and out fusion algorithm; LAP: Laplace algorithm; BOSGIOF: Based on optimal seam and gradated in and out fusion; MI: Mutual Information; EN: Entropy; STD: Standard Deviation; SF: Spatial Frequency; AG: Average Gradient;

Competing interests
The authors declare that they have no competing interests.

Availability of data and materials
Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Authors' contributions
YLH and ZWS conceived the designed the study. YLH made the simulation experiments, write the main manuscript of this work, derived the formulas, and helped enhance the novelty of this paper. ZBW and SZM provided some insights from the work in this paper. MLZ and GYJ reviewed and edited the manuscript. All authors have read and approved the final manuscript.

Author details
[1]HDU-ITMO Joint Institute, Hangzhou Dianzi University, Hangzhou, China. [2]School of Compute Science and Technology, Hangzhou Dianzi University, Hangzhou, China. [3]College of Economics and Management, Zhejiang University of Technology, Hangzhou, China. [4]Financial Information Engineering Technology Research Center, Zhejiang, China. [5]Science and Technology on Parallel and Distributed Processing Laboratory(PDL), National University of Defense Technology, Changsha, China. [6]Shanghai Aerospace System Engineering Research Institute, Shanghai, China.

References
1. Qin, Y., Li, J., Jiang, P., Jiang, F.: Image stitching by feature positioning and seam elimination. Multimedia Tools and Applications, 1–13 (2021)
2. Nocerino, E., Menna, F., Chemisky, B., Drap, P.: 3d sequential image mosaicing for underwater navigation and mapping. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **43**, 991–998 (2020)
3. Bai, Z., Li, Y., Chen, X., Yi, T., Wei, W., Wozniak, M., Damasevicius, R.: Real-time video stitching for mine surveillance using a hybrid image registration method. Electronics **9**(9), 1336 (2020)
4. Chen, X., Xu, W.-W., Yeung, S.-K., Zhou, K.: View-aware image object compositing and synthesis from multiple sources. Journal of Computer Science and Technology **31**(3), 463–478 (2016)
5. Choi, H., Choi, K., Kim, J.: Mosaicing touchless and mirror-reflected fingerprint images. IEEE Transactions on Information Forensics and Security **5**(1), 52–61 (2010)
6. Son, C.-H., Zhang, X.-P.: Layer-based approach for image pair fusion. IEEE Transactions on Image Processing **25**(6), 2866–2881 (2016)
7. Pandey, A., Pati, U.C.: Image mosaicing: A deeper insight. Image and Vision Computing **89**, 236–257 (2019)
8. Hernandez-Mier, Y., Blondel, W., Daul, C., Wolf, D., Guillemin, F.: Fast construction of panoramic images for cystoscopic exploration. Computerized Medical Imaging and Graphics **34**(7), 579–592 (2010)
9. Chen, Y.-c.: Optimal multi-resolution blending of confocal microscope images. Google Patents. US Patent 8,509,565 (2013)
10. Wang, L., Yu, W., Li, B.: Multi-scenes image stitching based on autonomous driving. In: 2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), vol. 1, pp. 694–698 (2020). IEEE
11. Liu, H., Tang, C., Wu, S., Wang, H.: Real-time video surveillance for large scenes. In: 2011 International Conference on Wireless Communications and Signal Processing (WCSP), pp. 1–4 (2011). IEEE
12. Yetiş, H., Karaköse, M.: Adaptive vision based condition monitoring and fault detection method for multi robots at production lines in industrial systems. International Journal of Applied Mathematics Electronics and Computers (Special Issue-1), 271–276 (2016)
13. Yang, T., Li, J., Yu, J., Wang, S., Zhang, Y.: Diverse scene stitching from a large-scale aerial video dataset. Remote Sensing **7**(6), 6932–6949 (2015)

14. Capel, D., Zisserman, A.: Automated mosaicing with super-resolution zoom. In: Proceedings. 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No. 98CB36231), pp. 885–891 (1998). IEEE

15. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International journal of computer vision **60**(2), 91–110 (2004)

16. Brown, M., Lowe, D.G.: Automatic panoramic image stitching using invariant features. International journal of computer vision **74**(1), 59–73 (2007)

17. Zaragoza, J., Chin, T.-J., Brown, M.S., Suter, D.: As-projective-as-possible image stitching with moving dlt. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2339–2346 (2013)

18. Li, X., Feng, R., Guan, X., Shen, H., Zhang, L.: Remote sensing image mosaicking: Achievements and challenges. IEEE Geoscience and Remote Sensing Magazine **7**(4), 8–22 (2019)

19. Wang, W., Ng, M.K.: A variational approach for image stitching i. SIAM Journal on Imaging Sciences **6**(3), 1318–1344 (2013)

20. Uyttendaele, M., Eden, A., Skeliski, R.: Eliminating ghosting and exposure artifacts in image mosaics. In: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, vol. 2, p. (2001). IEEE

21. Ng, M.K., Wang, W.: A variational approach for image stitching ii: Using image gradients. SIAM Journal on Imaging Sciences **6**(3), 1345–1366 (2013)

22. Xiong, Y., Pulli, K.: Mask-based image blending and its applications on mobile devices. In: MIPPR 2009: Remote Sensing and GIS Data Processing and Other Applications, vol. 7498, p. 749841 (2009). International Society for Optics and Photonics

23. Chang, C.-H., Sato, Y., Chuang, Y.-Y.: Shape-preserving half-projective warps for image stitching. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3254–3261 (2014)

24. Efros, A.A., Freeman, W.T.: Image quilting for texture synthesis and transfer. In: Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, pp. 341–346 (2001)

25. Wang, B., Li, H., Hu, W.: Research on key techniques of multi-resolution coastline image fusion based on optimal seam-line. Earth Science Informatics, 1–12 (2019)

26. Xu, Y., Ou, J., He, H., Zhang, X., Mills, J.: Mosaicking of unmanned aerial vehicle imagery in the absence of camera poses. Remote Sensing **8**(3), 204 (2016)

27. Qu, G., Zhang, D., Yan, P.: Information measure for performance of image fusion. Electronics letters **38**(7), 313–315 (2002)

28. Chen, C.: Research and implementation of multi-channel real-time video splicing technology. Master's thesis, Beijing University of Technology (2019)

Figures

Fig 1 Schematic diagram of optimal seam search

Fig 2 Schematic diagram of optional fusion area

Fig 3 Images to be stitched and distortion processing after projective transformation

Fig 4 Effects of different image fusion algorithms

Fig 5 Comparison of results of different fusion algorithms

In the experiment, four different fusion algorithms are used to fuse the same set of image stitching data. Among them, the blue column represents directly based on the optimal seam algorithm, the red column represents direct graded in and out fusion algorithm, the green column represents Laplace algorithm, and the purple column represents a graded in and out fusion algorithm based on optimal seam proposed in this paper.

Fig 6 Comparison of partial pixels in different fusion algorithms

Fig 7 The practical application in the monitoring system