



# Review on image-stitching techniques

Zhaobin Wang<sup>1</sup> · Zekun Yang<sup>1,2</sup>

Received: 12 July 2019 / Accepted: 9 March 2020 / Published online: 20 March 2020  
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

## Abstract

Image stitching, or known as image mosaic, is the process that combines images with overlapped areas to form an image with wide view and high resolution. Image stitching technique has been quickly developed these years. It has become an important branch in digital image processing and has wide applications. Many image stitching methods have been proposed. This article takes investigation on some image stitching techniques, including image registration, seam removal and quality assessment. Most existing related methods are introduced. Experiments are done to show the result of some main methods. At last, the advantages and disadvantages of some existing methods are discussed and some future potential work are pointed out.

**Keywords** Image stitching · Seamless · Image fusion · Image registration

## 1 Introduction

A wide-view, high-resolution image is usually required in many fields. At present, there are two main ways to acquire these images, first is to fix the shaft of the camera and taken photos while pivoting; second is to fix the light centre of the camera, moving the lens horizontally and taking photos [1]. To merge these images to a wide-view, high-resolution image, it is necessary to study image stitching techniques.

Image stitching, or also known as image mosaic, is a process that combines images with overlapped areas to form an image with wide view and high resolution [2]. Nowadays, image stitching plays a vital role in digital image processing, making it a popular domain in photographic cartography, computer vision, image processing and computer graphics. It is widely applied in remote sensing, aerospace, virtual reality, medical imaging and so on [3–6].

A well-stitched image should be clear, has smooth edge and high resolution. The main process of image stitching consists of feature matching, registration and seam removal.

Take two images as an example, the stitching process is shown in Fig. 1. Unfortunately, there is not a common method for stitching all kinds of images at present because of the complexity of the image. However, the core problem is to find the overlapped area and determine the position relationship between two images.

Usually, after inputting two images with overlapped areas, feature matching is applied to find the correspond points of the images for stitching. Image registration is done to put the images under the same coordinate system. Seam removal is employed to eliminate the seam, because the outside environments and camera settings (like exposure) when those images are acquired may be different. Therefore, image registration and seam removal are key techniques to image stitching.

With the development of new techniques, demands to time and precision have been refreshed. In recent years, lots of achievements have been made in this field. For example, scale-invariant feature transform (SIFT) and speed-up robust feature (SURF) methods improve the precision of image registration to a certain degree, optimal seam methods make the transition between two images more natural.

In this article, some algorithms about image stitching, including image registration and seam removal, are shown in Sects. 2 and 3, respectively. Next, some quality assessment methods for stitched image are provided in Sect. 4. Discussions are made in Sect. 5. Finally, conclusions are made in Sect. 6.

---

Communicated by Y. Zhang.

✉ Zhaobin Wang  
zhaobin\_wang@hotmail.com

<sup>1</sup> School of Information Science and Engineering, Lanzhou University, Lanzhou, China

<sup>2</sup> Graduate School of Information Science and Technology, Osaka University, Osaka, Japan

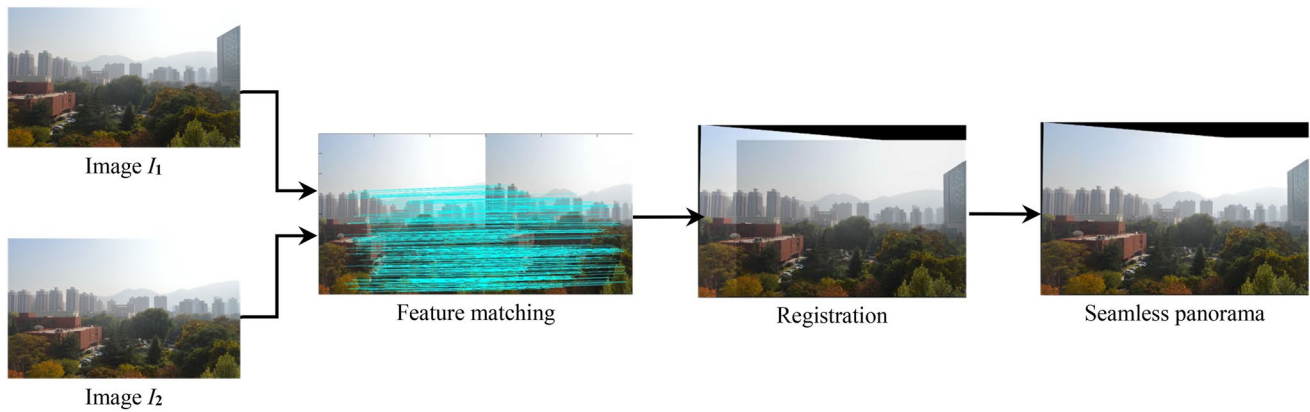


Fig. 1 Basic flow chart of image stitching

## 2 The key to image stitching: image registration

Image registration is a process that combines different images for the same area obtained from different time, different views and different sensors altogether [7]. It can be single modality or multi modalities, rigid or non-rigid. In rigid registration, images are simply supposed to be rotated and translated to be aligned properly. In non-rigid registration, two images cannot be aligned properly without localized stretching of images. The methods commonly used in image stitching are single modality, rigid ones which mainly categorized as method based on regions and method based on features. Method based on regions starts from the pixel value and calculates the intensity difference in the same area between image to be registered and reference image mathematically. In general, such methods have a large calculation amount but low registration accuracy. Method based on features does not make use of pixel value directly, it extracts features in images and matches the overlapped area using those features extracted. This kind of method is of robustness and precision.

### 2.1 Image registration based on regions

#### 2.1.1 Pixel matching algorithm

Basic pixel matching is operated on monochrome images. The algorithm selects one image as a base, and the overlapped area is employed to be registered and determines the matched regions. Some methods use one image as reference and make the overlapped area move inside the reference image. When it moves over one pixel, relativity is calculated. Suppose  $I_1$  is the referrer,  $I_2$  is another image,

the overlapped areas are known as  $\Omega_1$  and  $\Omega_2$ , whose sizes are both  $M \times N$ . The relativity  $R$  can be known as

$$R = \sum_{x=1}^M \sum_{y=1}^N |\Omega_1(x, y) - \Omega_2(x, y)|. \quad (1)$$

Then, the row (column) with maximum relativity is selected for registration. Obviously, the calculate amount is large. For an  $M \times N$  sized overlapped area, the size of  $D$  will be  $M \times N$  as well. It is also difficult for such methods to deal with rotation and scaling.

Figure 2 are two images with overlapped area and Fig. 3 is the stitching result using such method. Note that grey image is employed for calculation but map the results into the RGB domain. In this method, relativity in a window is kept and calculated to find the suitable position to align two images. As the method cannot deal with rotation and scaling, the images are kept as their original size and style. The relativity of Fig. 2a, b has high relativity at the place denoted by an arrow with *high relativity* so it is chosen as the base for aligning the image. The relativity of the place denoted by an arrow with *low relativity* is low and is thus, not chosen as the base. From Fig. 3, we can see that there are obvious discontinuities in the overlapped area. This is because that the lens was rotated into a different angle when photographing the right image (see Fig. 2b), while the matching method cannot deal with rotations in images. To register this image better, some methods that could solve the rotation of the images should be applied.

To reduce calculate amount, the search field is modified to a region and put to practice in [7, 8]. An image stitching method based on cross-correlation is also proposed in [9] to compute the displacements between the source images.

An image stitching method using dynamic time warping (DTW) is proposed in [10]. DTW is a method aligning two



(a) left



(b) right

Fig. 2 Images with overlapped area



Fig. 3 Image stitching using relativity measurement

sequences of time series by matching their elements while minimizing a total distance measure via dynamic programming. The method starts by selecting two images with overlapped area. Then, an  $m \times n$  sub-image is cropped from one of them as template, where features are extracted and normalized into a 1-dimensional sequence. Features of another image are also extracted and compared with DTW method to identify the position. Finally, the stitched image is created.

In [11], two rows with certain interval in the overlapped areas of an image  $I_1$  are used as the template to search the

best matching points in the correspond overlapped area in  $I_2$ . The rows of match are then adjusted to three to improve the searching strategies and make the result more precise. In [12], a technique for stitching biomedical images is presented. This algorithm operates by equilibrating neighbouring edges and forcing the brightness at corners to a common value. Brightness correction is done by a specific matrix before registration in [13] to make the algorithm less sensitive to illumination changes.

In general, this kind of method is simple to realize, but the calculation amount is large. Because the image is often taken with deformations, while no transformation is made in registration process, the result is not so ideal as expected.

### 2.1.2 Image registration based on mutual information

Mutual information (MI) is the scale of dependence between two variables [14]. If two variables are independent, the mutual information is 0. If two variables share strong dependence, the mutual information is large. In recent years, MI is widely applied in image registration [15–17], image retrieval [18], pattern recognition [19] and image stitching [20, 21].

The MI of images  $I_1$  and  $I_2$  is defined as

$$MI(I_1, I_2) = \sum_{I_1, I_2} P(I_1, I_2) \log_{10} \frac{P_{I_1 I_2}(I_1, I_2)}{P_{I_1}(I_1)P_{I_2}(I_2)}, \quad (2)$$

where  $P_{I_1 I_2}(I_1, I_2)$  is the joint probability density,  $P_{I_1}(I_1)$  and  $P_{I_2}(I_2)$  are marginal probability densities.

According to the definition of MI, two images are supposed to be variables, they share large MI if the similarities between the images are large. Thus, the image registration problem now transforms into a problem to search the max value of MI between the template and the sub-image.

To improve the MI algorithm, some concepts, such as NMI (normalized mutual information) [22], LMI (local mutual information) [23], are proposed. A self-similarity weighted graph-based implementation of  $\alpha$ -mutual information ( $\alpha$ -MI) for non-rigid image registration is proposed in [16] and is applied into registering pre-operative magnetic resonance (MR) images to intra-operative ultrasound (US) images.  $\alpha$ -MI is used to detect the similarities between images.

Registration using MI is of high precision, low interference and strong reliability. However, the calculation is very complex, and the robustness to rotation is weak.

### 2.1.3 Image registration based on Laplacian pyramid

Image registration based on Laplacian pyramid was first proposed in [24], where a sequence of low-pass filtered images  $I_1, \dots, I_{L_{dec}}$  can be obtained by repeatedly convolving a small

weighting function with a sample image,  $I_0$ . By this method, image sample density is also decreased with each iteration so that the bandwidth is reduced in uniform one-octave steps. Sample reduction also means that the cost of computation is controlled to a minimum size. The pixel value of  $I_l$  corresponds to the  $5 \times 5$  weighted mean value of transition in  $I_0$  level. Iterate according to this and get a series of low pass filter image (the Gauss low pass filter image series):  $I_0, \dots, I_{L_{dec}}$ . The pixel of the  $l$ th level in image  $I_l(x, y)$  is defined as

$$I_l(x, y) = \sum_{i=1}^5 \sum_{j=1}^5 w(i, j) I_{l-1}(2x + i, 2y + j). \quad (3)$$

$w(i, j)$  is called generating kernel limited by

$$\sum_{i=-2}^2 \sum_{j=-2}^2 w_l(x - 2^l i, y - 2^l j) = 1 \quad (4)$$

A localized algorithm of image stitching based on classic Laplacian method is proposed in [25]. Multi-resolution template, which can calculate multi templates in one scanning process and get the descriptor in different scales, is proposed. Experiments show that the use of structure information limits the decomposition and reconstruction process and reduce the calculation amount.

### 2.1.4 Registration in the transformation domain

In the transformation domain, methods commonly used to register images are Fourier Transformation (FT) and Wavelet Transformation (WT). Here, Fourier Transformation based registration is introduced at first. Then, Wavelet based registration is also introduced.

In 1970, FT was used in image registration [26]. In [27], FT is applied into underwater sonar image stitching. In [28], it is used for stitching computed tomography (CT) image sets.

An image registration method based on WT is proposed in [29]. Dual-tree complex wavelet transformation (DT-CWT) is applied to decompose the images to be registered. Starting at the coarsest level, a first estimate of the transformation vector  $v = [\alpha, t_x, t_y]$  is found, where  $\alpha$  is the rotation angle,  $t_x$  and  $t_y$  are the translation parameters in the  $x$  and  $y$  directions, respectively. Note that in this work, scaling transformation is omitted. Cross correlation is chosen as a matching criterion. Next, edge maps of both low-passed images,  $Emap_1$  and  $Emap_2$ , are extracted. Searching is performed to determine the best initial transformation vector  $v_{init}$ :

$$v_{init} = \arg \max_v c(Emap_1, T(Emap_2)), \quad (5)$$

where  $c(x, y)$  means cross correlation with  $x$  and  $y$ ,  $T(Emap_2)$  denotes the wrapped image using vector  $v$ .

The transformation vector  $v_l$  is found by

$$v_l = \arg \max_v \left[ \frac{I(\text{Re}\{D_1, l\}, T(\text{Re}\{D_2, l\}))}{I(\|Cp\{D_1, l\}\|, T\|Cp\{D_2, l\}\|)} \right]. \quad (6)$$

$\text{Re}\{x, l\}$  [21] and  $Cp\{x, l\}$  are real and complex part of  $x$  in level  $l$ , respectively.  $D_1$  means the decomposing of original image and  $D_2$  means the decomposing of image to be registered. To further reduce the computational burden, an alternative search method proposed can be applied.

Registration in the transformation domain is fast, robust and can solve most deformations in images. However, when the image has few overlapped areas, the result may not be precise enough.

## 2.2 Image registration based on features

### 2.2.1 Corner detecting algorithm

The corner that is usually generated from intersections between lines, is usually the point where there is a sharp difference in brightness or has huge curvature [30, 31]. It is scale-invariant and influenced little from illumination. Therefore, corner is suitable for detection. The operators generally used are Forstner, Moravec and Harris.

Forstner operator was firstly proposed by Wolfgang Forstner in [32], where  $q$  and  $w$  are calculated by

$$q = \frac{4 \det(C)}{(\text{trace}(C))^2} \quad (7)$$

$$w = \frac{\det(C)}{\text{trace}(C)} \quad (8)$$

where  $C$  is the intensity covariance matrix. Suppose  $t$  is the threshold (a certain value). If  $q > t$ , the pixel is one candidate. Then, extreme points of  $q$  are selected as feature. The best candidate window is the centre window of that point. Normal equations are used to calculate the accurate position of the corner, where  $g_x$  and  $g_y$  are Robert gradient in  $x$  and  $y$  directions, respectively:

$$C \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \sum g_x^2 + y \sum g_x g_y \\ x \sum g_x g_y + y \sum g_y^2 \end{bmatrix} \quad (9)$$

In 1980, Moravec operator was formally proposed in [33]. It selected the point with minimum intensity variance in four main directions.

Harris operator was proposed in [34] based on Moravec method [35]. In Harris operator, corner detector  $D_h$  is used to detect corners:



$$D_h(i, j) = \sum_x \sum_y w(x, y) [I(x + i, y + j) - I(x, y)]^2, \tag{10}$$

where  $I(x, y)$  is one pixel in image  $I$ .  $w(x, y)$  is the window function, which could be a rectangle window or Gaussian window. Suppose

$$M = \sum_x \sum_y w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \tag{11}$$

$I_x$  and  $I_y$  are gradient values in  $x$  and  $y$  directions. The detect operator can be approximated as

$$D_h(i, j) = \begin{bmatrix} i & j \end{bmatrix} M \begin{bmatrix} i \\ j \end{bmatrix}. \tag{12}$$

The response value  $R_e$  of the detection is

$$R_e = \det(M) - k[\text{trace}(M)]^2. \tag{13}$$

Usually, the value of  $k$  is 0.04–0.06 [36]. Set  $k = 0$  to the  $k$  value less than a threshold  $t$  and suppress the non-maximum value in a  $3 \times 3$  or  $5 \times 5$  area, whose local maximum point is the corner.

Harris operator is invariant to translation and rotation. However, the value of  $t$  is dependent on image attributes, making no measurement when setting specific threshold. In addition, the points with large eigen values are often assembled in certain range, which will make the distribution of the detected corner in a non-unified way. As more corners can be detected in regions with rich texture, while fewer corners can be detected in regions with less texture, the corners tend to distribute in the place with richer texture information [30, 37].

To improve Harris corner detection, a new response function is given in [38].  $R_e$  is redefined as:

$$R_e = \frac{\det(M)}{(\text{trace}(M))^2 + \varepsilon}, \tag{14}$$

where  $\varepsilon$  is used to prevent the case when  $\text{trace}(M)$  is 0. A very small  $\varepsilon$ , like  $10^{-6}$ , can be taken. In addition, in the redefined  $R_e$ , the selection of  $k$  is not needed.

There are some applications of corner detecting algorithms in image stitching. In [39], an adaptive Harris corner detection for image stitching is given. In [40], Harris corner detector is customized to meet the requirements of viewing a panorama video in real time. In [41], corner detection is adopted to help extract features. In [42], an improved Harris corner detection algorithm is proposed to stitch the images related to traffic accidents.

### 2.2.2 Scale-invariant feature transform (SIFT) method

SIFT method was proposed by David G. Lowe in 1999 and was used in feature detections at first [43, 44]. In 2004, the algorithm was upgraded [45]. SIFT has been used in image registration [46–49] and image stitching [50–54]. SIFT algorithm detects the extreme value in the intensity and scale domain and selects the key point using Gaussian kernel  $G$ . The difference of Gaussian kernel (DoG) in different scale  $\sigma$  is used to detect extreme value:

$$G(x, y, \sigma) = \frac{1}{\sqrt{2\pi\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}}} \tag{15}$$

$$\begin{aligned} DoG(x, y, \sigma) &= [G(x, y, k\sigma) - G(x, y, \sigma)] * I(x, y). \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \tag{16}$$

For every  $L(x, y, \sigma)$ , the amplitude and direction of the key point are then computed. Then, an  $8 \times 8$  window centred by the key point is taken and the 8-direction gradient histogram of every  $4 \times 4$  direction is calculated. The key point descriptor is a 128-dimensional vector.

Key point descriptors are matched using measurements like Euclidian distance. After the matching, random sample consensus (RANSAC) is usually taken to remove mismatches [55].

Then, the image is taken for affine transformations, where one image can be combined with another. Usually, a homography matrix is adopted in this step to describe the class of 2-dimensional planar projective transformations by matrix multiplication. The projective transformations can be defined as a matrix with 8 degrees of freedom called  $H$ :

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = H \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \tag{17}$$

where  $x'$  and  $y'$  can be obtained as

$$\begin{cases} x' = \frac{a_1x + a_2y + a_3}{a_7x + a_8y + 1} \\ y' = \frac{a_4x + a_5y + a_6}{a_7x + a_8y + 1} \end{cases} \tag{18}$$

By affine transformation and finding appropriate parameters, the two images can be combined with their main features matched altogether. The homography matrix is determined by eight parameters and the coordinate of points should be selected to estimate them. However, in pixel-based algorithms, the image registration is employed for selecting a template from one image and measure the relativity in another. Thus, it is unable to take the affine transformation. This is the reason why Fig. 3 has large mismatches.

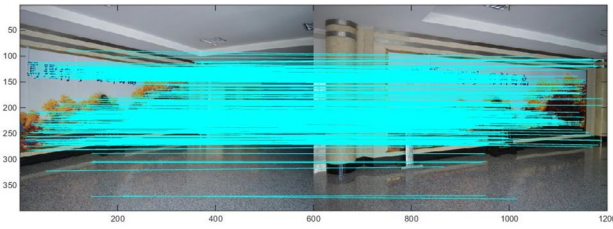


Fig. 4 Matched result with SIFT

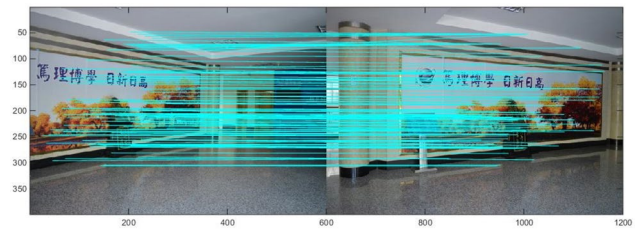


Fig. 6 Matched result with SURF



Fig. 5 Stitched image with SIFT



Fig. 7 Stitched image with SURF

Image stitching with SIFT is in good conditions in many cases. It runs very fast and can deal with translation, rotation and scaling. Figure 4 is the matched result with SIFT. Note that the horizontal lines mean the corresponding points in the two images. Figure 5 is the stitched result of Fig. 2 with SIFT.

SIFT is improved in [56], known as PCA-SIFT. It accepts the same input as the standard SIFT descriptor. PCA [57] is the short for principal component analysis. It is a kind of technique for dimensionality reduction and has been widely applied in computer vision [58–60]. There are mainly three steps for PCA-SIFT: first, an eigen space is computed to express the gradient images of local patches; second, the local image gradient of a given patch is calculated; third, the gradient image vector is projected using the eigen space to derive a compact feature vector.

SIFT descriptor is invariant to image deformations such as translations, rotations and scaling. It is also robust to moderate perspective transformations and illumination variations. However, it cannot meet the demand of calculation speed in some cases. As an improvement, PCA-SIFT consumes less time than SIFT with the cost of lower robustness to scaling.

### 2.2.3 Speed up robust feature (SURF) algorithm

SURF, a robust image recognition algorithm, was firstly proposed in 2006 [61]. It can be used in pattern recognition

and 3D reconstruction. Experiments show that compare with SIFT, SURF will be several times faster [62]. In recent years, studies on image registration [63, 64] and image stitching [65–67] based on SURF demonstrate good results.

SURF takes use of the integral image to convert the second-order Gaussian integral template filter into the modification of the integral image.

Hessian matrix,  $H(x, \sigma)$  in  $x$  at scale  $\sigma$ , is then applied to detect feature points:

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \quad (19)$$

$L_{xx}$ ,  $L_{xy}$ ,  $L_{yy}$  are the second-order derivations in  $xx$ ,  $xy$ ,  $yy$  directions, respectively, which can be acquired from convoluting the image with second-order Gaussian gradient template. In SURF, this template is approximated by box functions, usually a  $9 \times 9$  box function that is equivalent to a Gaussian filter ( $\sigma=1.2$ ). Such approximation is called blob response maps expressed by  $B_{xx}$ ,  $B_{yy}$  and  $B_{xy}$ .

Thus,  $\det(H)$  can be approximated by

$$\det(H) = B_{xx}B_{yy} - w(B_{xy})^2. \quad (20)$$

Next, feature descriptors are created using Haar Wavelet. Integral image is used to simplify the calculation. The descriptor is a 64-dimensional matrix with coordinate, amplitude and rotation angle. Figure 6 is the matching points found with SURF, where the horizontal lines mean the

corresponding points in the two images. Figure 7 is stitching result of Fig. 2 using SURF.

Compare Fig. 6 and Fig. 4, we can see that SURF finds fewer feature points than SIFT. The applications of Hessian matrix and box function make SURF faster than SIFT. However, compare with SIFT, SURF is not robust against large rotations.

Template-convolution speed-up robust features (TSURF), an improvement of SURF, is proposed in [68] to improve the stitching speed while maintaining the precision. The maximum value ( $H_{\max}$ ) of Hessian matrix's determinant minus the secondary value ( $H_{\text{sec}}$ ) is used:

$$H_{\max} - H_{\text{sec}} \geq \theta. \tag{21}$$

In their algorithm,  $\theta$  is set to 0.5 when the image contrast is 0.14. If  $H_{\max} - H_{\text{sec}} \geq \theta$ , the point is assumed as a candidate. The higher the contrast is, the greater the  $\theta$  will be.

Registration based on SURF consumes less time than registration based on SIFT. However, it is less robust than SIFT to rotations in images. In addition, misalignments with SURF are more than using SIFT.

### 2.2.4 KAZE features

KAZE feature is a multi-scale 2-dimensional feature detection and description algorithm in nonlinear scale spaces proposed in 2012 [69], where 2-dimensional features are detected and described in a nonlinear scale space by nonlinear diffusion filtering. In KAZE feature description, the nonlinear scale space up to a maximum evolution time is built at first. Then, the 2-dimensional features, which exhibit the max value of the scale-normalized determinant of the Hessian response through the nonlinear scale space, are detected. Finally, the main orientation of the key point is calculated to obtain a scale and rotation invariant descriptor.

In KAZE feature description, the scale space is discrete in logarithmic as a series of  $O$  octaves and  $S$  sub-levels. The octave sets, and sub-levels are identified by  $o$ , a discrete octave index, and  $s$ , a sub-level octave index. They are mapped to their corresponding scale  $\varepsilon$  as

$$\varepsilon_i(o, s) = \varepsilon_0 2^{o+\frac{s}{S}}, \quad o \in [0, O - 1], \quad s \in [0, S - 1], \quad i \in [0, N], \tag{22}$$

where  $\varepsilon_0$  is the base scale level,  $N$  is the total number of filtered images.

Then, the scale levels in pixel units are converted into time units with

$$t_i = \frac{\varepsilon_i^2}{2}, \quad i \in [0, N]. \tag{23}$$

An input image is firstly convoluted with a Gaussian kernel with standard deviation  $\varepsilon_0$  to reduce noise and possible

image artefacts. Then, its gradient histogram is computed to obtain the contrast parameter  $k$  in an automatic procedure.

Given the contrast parameter and the set of evolution times  $t_i$ , the non-linear space is built as

$$L^{i+1} = \left( I - (t_{i+1} - t_i) \sum_{l=1}^m A_l L_l \right)^{-1} L_i. \tag{24}$$

To detect the interest points, the response of scale-normalized determinant of the Hessian matrix is computed at multiple scale levels:

$$L_{\text{Hessian}} = \varepsilon^2 (L_{xx} L_{yy} - L_{xy}^2), \tag{25}$$

where  $L_{xx}$  and  $L_{yy}$  are the second-order horizontal and vertical derivations, respectively,  $L_{xy}$  is the second-order cross derivation.

Then, the maximum value in scale and spatial location is found. The extrema are also found in all the filtered images (except the top and the end levels). Each extremum is searched over a rectangular window of size  $\varepsilon_i \times \varepsilon_i$  on the current  $i$ , upper  $i + 1$  and lower  $i - 1$  filtered images.

Finally, the descriptors are generated. The dominant orientation is found in a circular area of radius  $6\varepsilon_i$  with a sampling step of size  $\varepsilon_i$ . First-order derivations  $L_x$  and  $L_y$  are weighted with a Gaussian centred at the interest point for each of the samples in the circular area. Then, the derivation responses are represented as points in vector space. The dominant orientation is found by summing the responses within a sliding circle segment covering an angle of  $\pi/3$ . From the longest vector, the dominant orientation is obtained.

For a detected feature with  $\varepsilon_i$  scale, first-order derivation  $L_x$  and  $L_y$  of size  $\varepsilon_i$  are computed over a  $24\varepsilon_i \times 24\varepsilon_i$  rectangular grid divided into  $4 \times 4$  sub-regions of size  $9\varepsilon_i \times 9\varepsilon_i$  with an overlap of  $2\varepsilon_i$ . The derivation responses in each sub-region are weighted with a Gaussian ( $\varepsilon_l = 2.5\varepsilon_i$ ) centred on the sub-region centre, and summed into a descriptor vector  $dy = (\Sigma L_x; \Sigma L_y; \Sigma |L_x|; \Sigma |L_y|)$ . Then, each sub-region vector is weighted using a Gaussian ( $\varepsilon_2 = 1.5\varepsilon_i$ ) defined over a  $4 \times 4$  mask and centred on the interest key point. When considering the dominant orientation of the key point, each of the samples in the rectangular grid is rotated according to the dominant orientation. The derivations are also computed according to the dominant orientation. Finally, the 64-dimensional descriptor vector is normalized into a unit vector to achieve contrast invariance.

Experiments show that KAZE feature description is robust to noise and blurring. However, it adopts non-linear diffusion in a low level and it costs more time than SURF. To reduce the calculate amount, A-KAZE, the short for accelerate KAZE, is proposed [70]. It uses the thought of fast explicit diffusion (FED) in a pyramidal framework and

binary descriptors [71] to speed up the calculation and is suitable for portable devices.

Once KAZE feature is proposed, many researches are focused on its application. In [72], KAZE feature is used to matching the feature points of frame-patch in videos. In [73], S-AKAZE is proposed as a modification of KAZE to do image matching. In [74], A-KAZE is applied for object recognition. In [75], KAZE is taken to assist image segmentation. In [76], A-KAZE is adopted for image stitching.

### 2.3 Assessment to image registration descriptors

In this section, four main image registration descriptors (Harris, SIFT, SURF, KAZE) are taken to register two images, as shown in Fig. 2. The metrics of the matching points are set as Euclidian distance. For one key point in one image, two correspond points in another image is found. If the ratio between the distance of the second nearest point and the nearest point is less than a threshold, such two points are considered as correct matched points. The ratio threshold of Euclidian distance is set as 0.6. The mismatching points are all removed by the same method (RANSAC) to form correct matching points. The correct rate is the percentage of correct matching points in the matching points found. The results are shown in Table 1.

From the table, we can see that KAZE finds the maximum number of matching points and correct matching points among the four descriptors. SURF finds the minimum number of matching points and correct matching points. Among the four methods, it is SIFT that has the highest correct rate in the matching points found. Thus, it is considered to be the best-performed descriptor.

## 3 The key to image stitching: seam removal

As image registration is the most important issue to a successful stitching, many literatures pay more attention to image registration rather than seam removal. However, seam removal is also a critical issue in image stitching as humans are sensitive to the seam in an image, as shown in Fig. 8.

Seam removal can improve the general quality of the stitched image, and make the information in the stitched image more accurate. Therefore, it is necessary to remove the obvious seam caused by brightness inconsistency of



Fig. 8 Stitched image without seam removal

input images after finishing registration. The seam usually appears at the border of two images, making a bigger pixel variance or gradient difference. Essential seam removal needs to reduce such variance (difference) to an acceptable range, making the transition more smoothly in the overlapped area.

Such process is something like image fusion, a process that fuses images from different sensors into an integrated one using mathematical method to meet certain demands [77]. What is different from image fusion is that seam removal is only the adjustment of brightness and contrast, extra information is not added after such process. Classified from method, there are methods based on pixel weighing, optimal seam method and transformation domain. Seam removal based on pixel weighting performs directly on the overlapped area but cannot get the ideal result when there are large differences in outside environments or camera settings in the two images.

Thus, optimal seam methods are to find an optimal seam by optimizing an objective function that minimizes the difference near the overlapped area. Seam removal in the transformation domain decomposes the image via specific transformation at first and apply different rules to different bands, then the image after seam removal is reconstructed and output.

### 3.1 Seam removal based on pixel weighting

In many references, weighting based seam removal technique [78, 79] is usually employed

$$I(x, y) = a\Omega_1(x, y) + (1 - a)\Omega_2(x, y) \quad (26)$$

where  $a$  is a coefficient between 0 and 1.  $\Omega_1(x, y)$ ,  $\Omega_2(x, y)$  are the pixel values of  $I_1(x, y)$  and  $I_2(x, y)$  in overlapped area  $\Omega$ , respectively. At the beginning,  $a$  is a fixed value, but such setting cannot strengthen the interested parts in an image [80].

Table 1 Assessment to image registration methods

	Harris	SIFT	SURF	KAZE
Matching points found	136	502	106	639
Correct matching points	115	468	91	563
Correct rate	84.56%	93.23%	85.85%	88.11%





Fig. 9 Stitched image with pixel weighting



Fig. 11 Stitched image without seam removal



(a) left



Fig. 12 Stitched image with pixel weighting



(b) right

Fig. 10 Another set of images with overlapped area

Thus,  $a$  is adaptive. When  $a$  varies from 1 to 0, two images can transit smoothly. The selections of  $a$  is shown clearly in [81]. Figure 9 is the seam removal result with pixel weighting. From Fig. 9, we can see that this kind of weighting function is generally rough, making the seam still visible. In some cases, like two images, as shown in

Fig. 10, it will cause *ghost artefacts* and cannot meet the high demands for precision (Fig. 11) (shown in Fig. 12).

### 3.2 Seam removal based on optimal seam methods

The optimal seam method searches for the optimal seam in the overlapped area to create mapping or labelling between pixels in the composite and source images. The optimal seam is found by optimizing an objective function that minimizes the difference near the overlapped area. To eliminate seam more precisely, some optimal seam methods are proposed in [82–88].

An optimal seam finding method is proposed in [82]. Suppose  $I_1$  and  $I_2$  are two images with overlapped areas, the optimal seam is a connected path traversing across the overlap that starts on the first row/column and ends on the last row/column. By adopting a more compact and descriptive data structure called tensor (a symmetric and positive semi-definite matrix  $T$ ), the weights of pixels in the overlapped area are optimized. Riemannian norm is taken to measure the distance between two tensors.

The seam cost function Cost in an image whose height is  $M$  is defined as

$$\text{Cost} = \text{Cost}(I_s) = \sum_{i=1}^M c(I(s_i)), \quad (27)$$

where  $I(s_i)$  is the  $i$ th pixel on seam  $s$ .  $c(i, j)$  is defined as

$$c(i, j) = d(T_1(i, j), T_2(i, j)) + |I_1''(i, j) + I_2''(i, j)|, \quad (28)$$

where  $T_1, T_2$  are tensors of  $I_1$  and  $I_2$ ,  $I''$  is the 2nd derivation of image  $I$ . The optimal seam is defined as

$$s^* = \min_s \text{Cost}(s) = \min_s \sum_{i=1}^M c(I(s_i)). \quad (29)$$

The demos of this method are shown in Figs. 13 and 14.

Some other optimal seam methods also show better results. In [83], optimal seam method is used to handle the ghost artefacts in moving object problems. In [84], an improved seam finding method is proposed by downscaling the overlapped areas to approximate the seam. In [86], a new energy aggregation and traversal strategy is adopted to find an optimal seam for image stitching. In [86], optimal seam lines between adjacent input images are detected via the graph cut energy minimization framework. In [87], the information of image colour, gradient magnitude and texture are fused into graph cuts. In [88], an optimal seam pair is selected by comparing the cross correlations from multiple seams detected by variable cost weights.

### 3.3 Seam removal in the transformation domain

Seam removal in the transformation domain decomposes the image via specific transformation and applies different rules to different frequency bands. By doing seam removal in the transformation domain, the seam in the stitched image can be reduced. The common transformations are Fourier transformation, wavelet transformation and Curvelet transformation.

To remove seam with two images, the common practices are decomposing the images in the transformation domain



Fig. 14 Seam removal using optimal seam method only

at first to achieve  $N$  level of decomposing and coefficients in different frequencies in correspond levels. Then, apply different rules to the high frequency and low frequency bands, respectively, to preserve both the contour and the detail information as the high frequency bands imply the detail, while the low-frequency bands imply the contour. Finally, the image is reconstructed by the inverse transformations.

There is a method for seam removal using Fourier Transformation in [89]. Besides Fourier transformation, Wavelet transformation can also be applied in seam removal [90, 91]. A method suitable for seam removal based on Wavelet transformation is provided in [90], where Wavelet fusion is taken to create a seamless underwater stitching. In [91], Wavelet transformation is used to decompose the images so that the transition of the overlapped area can be smooth enough.

Figures 15 and 16 are the seam removal results of overlapped area using Wavelet transformation. The low frequency bands are obtained using the average value of two parts and the high frequency bands are obtained using the maximum value of two parts. The seam has been greatly reduced compare with Fig. 8, but still more visible than Fig. 11.

Besides Fourier and Wavelet transformation, Curvelet transformation can also be applied to remove seam. Being



Fig. 13 Seam removal using optimal seam method only



Fig. 15 Seam removal with Wavelet transformation (cise)



Fig. 16 Seam removal with Wavelet transformation (rail)

an extension of Wavelet, Curvelet is a non-adaptive technique for multi-scale object representation. It was proposed in 1999 and improved in 2002 [92]. Compare with Wavelet, a curve can be represented by Curvelet in a smoother way [93]. Curvelet transformation has been widely applied in image processing, such as image denoising [94, 95], image enhancement [96] and image fusion [93, 97, 98].

When  $\gamma$  is a curve with second derivations, Curvelet transformation is proved in [99] to have an error bound as

$$\|f - Q_N^C(f)\|_2^2 \leq CN^{-2} \log^{\frac{1}{2}} N, \tag{30}$$

where  $f(x) = g(x)_{I_{\{x_2 \leq \gamma(x_1)\}}}$ , and  $Q_N^C$  is the non-linear approximation of  $N$  items in the Curvelet transformation of  $f$ . The seam removal results with Curvelet transformation are shown in Figs. 17 and 18.

In [93], an image fusion technique based on Curvelet transformation is proposed. Local energy-based fusion rule, which is more effective than single pixel-based fusion rules, are taken into application. In [97], a Curvelet based approach for image fusion is proposed. The image is segmented into small tiles to show detail information. Ridgelet transformation is applied on each of these tiles and used in the fusion



Fig. 17 Seam removal with Curvelet transformation (cise)



Fig. 18 Seam removal with Curvelet transformation (rail)

process. In [98], Curvelet transformation is combined with Genetic Algorithm for medical image fusion.

### 4 Quality assessment of image stitching

It plays a key role to assess a stitched image to improve the stitching algorithm and judge whether the stitching results correspond to the vision of human eyes. Image quality assessment methods contain subjective methods and objective methods [100]. In subjective assessment, experiments are needed to make testers assess the image, while in objective assessment, algorithms are applied to assess the quality of an image. The main algorithms in objective assessment are PSNR (Peak Signal to Noise Ratio) [101] and SSIM (Structural SIMilarity index) methods [102, 103].

PSNR is an engineering term for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. It is usually applied to measure the quality of reconstruction of lossy compression codecs (e.g., for image compression). PSNR is defined by mean squared error (MSE) as

$$PSNR = 10 \log_{10} \left( \frac{MAX_I^2}{MSE} \right), \tag{31}$$

where  $MAX_I$  is the maximum possible pixel value of an image. When the pixels are represented using 8 bits per sample, it is 255. For an  $M \times N$  sized image  $I$  and its noisy approximation  $K$ , MSE is defined as

$$MSE = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (I(x, y) - K(x, y))^2. \tag{32}$$

PSNR approximates human perception of reconstruction quality. In PSNR method, when the reference image is the same with the image under assessment, the PSNR value is  $\infty$  (inf). Typical values for the PSNR in lossy image



**Table 2** Assessment of stitched images using PSNR

	Optimal	Original	Weighted	Wavelet	Curvelet
Cise	$\infty$	32.1528	35.4534	35.7947	36.5353
Rail	$\infty$	37.3452	34.2652	38.0641	38.5621

The data unit in the table is dB

**Table 3** Assessment of stitched images using SSIM

	Optimal	Original	Weighted	Wavelet	Curvelet
Cise	1	0.9594	0.9714	0.9715	0.9771
Rail	1	0.9747	0.9745	0.9765	0.9787

compression are between 30 and 50 dB if the bit depth is eight bits, where higher is better. For 16-bit data, typical PSNR values are between 60 and 80 dB [104].

SSIM is a full reference metric predicting the perceived quality of digital television and cinematic pictures, as well as other kinds of digital images and videos. It is also applied to measure the quality of images. SSIM is calculated on various windows of an image. The measure between two windows  $x$  and  $y$  of with size  $N \times N$  is

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad (33)$$

where  $\mu_x$  and  $\mu_y$  are the mean values of  $x$  and  $y$ ,  $\sigma_x$ ,  $\sigma_y$  are the variances of  $x$  and  $y$ ,  $\sigma_{xy}$  are the covariance of  $x$  and  $y$ .  $c_1$  is set to be  $(k_1L)^2$  and  $c_2$  is set to be  $(k_2L)^2$ .  $L$  is the dynamic range of the pixel-values (usually  $2^8 - 1 = 255$  in eight bit image) and  $k_1 = 0.01$ ,  $k_2 = 0.03$ . SSIM value is between 0 and 1, the higher value means that the image under assessment has higher quality. When the reference image is the same with the image under assessment, the SSIM value is 1. Tables 2 and 3 are assessment results using PSNR and SSIM, respectively.

In tables *cise* denotes the image stitched with Fig. 2 and *rail* denotes the image stitched with Fig. 10. Images with seam removal using optimal seam [55] (for *cise*, it is Fig. 13; for *rail*, it is Fig. 14) are taken as the reference. *original* means image stitched without seam removal (for *cise*, it is Fig. 8; for *rail*, it is Fig. 11). *weighted* means image stitched with pixel weighting [78] to remove seam (for *cise*, it is Fig. 9; for *rail*, it is Fig. 12). *wavelet* means image stitched with Wavelet transformation [90] to remove seam (for *cise*, it is Fig. 15; for *rail*, it is Fig. 16). *curvelet* means image stitched with Curvelet transformation [93] to remove seam (for *cise*, it is Fig. 17; for *rail*, it is Fig. 18). All these images are registered by SIFT, the best-performed descriptor discussed in Sect. 2.3.

PSNR is one of criteria for image assessment. However, PSNR needs reference image, which will sometimes be difficult to obtain. In addition, its results will not correspond to human vision in some cases. This is because the sense of human vision will be affected by many factors such as spatial frequency and brightness. SSIM is calculated based on a sensation model which takes the varying of structural information of an image into consideration. It has strong ability in measuring information loss occurred during the image degradation processes. It is also easy to compute and applicable to various images. However, SSIM also needs the reference image.

In the assessment results, *optimal* is taken for reference. Thus, its PSNR is  $\infty$ , and its SSIM is 1. From the results, we can see that the PSNR and SSIM indices of Curvelet is the highest among the others, meaning its quality is the best. We can also find that there is an obvious ghost artefact in *rail-weighted* (Fig. 12). In PSNR, its score decreased sharply compared with *rail-original*, but in SSIM, its score is almost the same as *rail-original*. This implies the importance of selecting proper seam elimination method, as an improper selection may cause counterproductive results.

An image assessment model based on deep learning is proposed in [105], where the assessment is converted into a 5-level classification system. Levels such as Excellent, Good, Fair, Poor, Bad are set. In this method,  $P(Q|L)$  is the posterior probability based on Bayes function with  $L$  levels of decomposition and intrinsic quality  $Q$ . Given the input image representation  $X$ , the distribution of the intrinsic quality can be obtained by marginal distribution:

$$P(Q|X) = \int P(Q|L)P(L|X)dL \quad (34)$$

$P(Q|X)$  is the quality distribution which represents the evaluations by a population. The numerical measurement of image quality is

$$\text{Quality} = E(P(Q|X)) \quad (35)$$

$E(x)$  is the mean value of  $x$ .

Experimental results show this method is of effectiveness, efficiency, and robustness to small training sets. It corresponds well to human vision. However, this method is hand-crafted, which requires time-consuming hand-tuning. It also fails to express certain eccentric distortions. This means using deep learning to learn more powerful image representation for describing image quality remains a great challenge that has yet to be resolved.

A difference of edge map (DoEM) assessment method is proposed for stitched images in [106]. It detects the mean values of brightness in local edge difference maps and divides them to the entire mean value. It judges whether the difference in brightness plays a major role according to the



ratio. When generating marks, the ratio of brightness mutation and misplacement assessments adjusts dynamically with the variance of differentiate map. DoEM is defined by the following equation:

$$\text{DoEM} = e^{-\frac{\sigma^2}{c_4}} \left( \frac{\mu_e e^{-\frac{\mu_e}{c_1}} + \mu_a e^{-\frac{\mu_a}{c_2}}}{\mu_e + \mu_a} \right) + \left( 1 + e^{-\frac{\sigma^2}{c_4}} \right) e^{-\frac{\sigma^2}{c_3}}, \tag{36}$$

where  $\mu_e$  is the mean value of the transition region in edge difference map,  $\mu_a$  is the entire mean value in the transition region,  $\sigma^2$  is the entire variance in the transition area,  $C_1, C_2, C_3$  and  $C_4$  are four constants.  $C_1$  and  $C_2$  are determined according to the relativity of the mean value difference in the differential map,  $C_3$  and  $C_4$  are on the sub-Gaussian distribution curve. According to numerous experiments, the suggest values are given as  $C_1=80, C_2=50, C_3=600, C_4=256$ . The final mark is normalized between 0 and 1, the higher mark means the better quality of an image. Table 4 is the assessment result with DoEM method.

Experiments show that in *cise*, *curvelet* based method has the best score, while in *rail*, *original* image without seam removal has the best score. We can also see that in *rail-weighted* where a ghost artefact appears, its DoEM score also decreases a lot. This means DoEM could accurately reflects the real quality of stitched image and the performance of image stitching algorithm. However, this metric is not suitable for assessing the stitching sequence.

An approach to assess stitched images is proposed in [107]. SSIM, SAM (spectral angle mapper) and IMR (intensity magnitude ratio) are applied to assess geometric and photometric qualities of stitched images. Suppose  $p_1$  and  $p_2$  are two pixels from two images,  $|p_1|$  and  $|p_2|$  denote their norms, SAM calculates the angle between them using the dot product formula:

$$\text{SAM} = \arccos \frac{p_1 \cdot p_2}{|p_1||p_2|}. \tag{37}$$

If the angle between two vectors is large, SAM will be small and it indicates that the two vectors/image pixels are different.

For intensity quality assessment, IMR, defined as the ratio of magnitudes of two 3D color vectors, is used. If  $p_1$  and  $p_2$  are two pixels,  $|p_1|$  and  $|p_2|$  denote their norms, then IMR is defined as

$$\text{IMR} = \frac{\min(|p_1|, |p_2|)}{\max(|p_1|, |p_2|)}. \tag{38}$$

The algorithmic procedure for IMR calculation is similar to that of colour quality assessment in SAM, except that instead of using SAM, IMR is used. The threshold for significant error in IMR is considered as 2%. This metric can be effectively used to quantitatively determine the quality of stitched images. Its results agree with qualitative analysis and the method is extensible.

In our experiments, we printed the pixel value and the gradient value 25 pixels both sides near the stitching seam. To make the results show more clearly, the figures here are results of one randomly selected column in each image. Figure 19 is the assessment result of *cise* and Fig. 20 is the assessment result of *rail*. Three lines in each figure show three channels (i.e., *R, G* and *B*) in the image, respectively. The full line corresponds to the *R* channel, the break line corresponds to the *G* channel and the dot line corresponds to the *B* channel.

We also calculate the variance of the gradient value in the columns selected in the figure. The variance demonstrates the degree of dispersion of data (pixels). Such data are shown in Tables 5 and 6. *M* means the average value of *R, G* and *B* in a row. It can be calculated by the following equation, where *R, G* and *B* are the mean gradient variance in *R, G* and *B* channel, respectively. The greater variance value shows that the data is more discrete in the set:

$$M = \frac{R + G + B}{3}. \tag{39}$$

From Figs. 19a and 20a, we can see there is a gradient difference between two images, making an obvious stitching seam. Hence, the goal of seam removal is to reduce such differences. When we use weighted, optimal and curvelet for seam removal, we can see the gradient variance has reduced, meaning the seam is going to disappear. We can also see that when we use wavelet, the gradient variance even increases. As the image can be abstracted and approximated by curves, this shows that the representation of curves in wavelet cannot meet the demand of seamless stitching.

By combining the results in figures and tables, we can see that it is better to take optimal seam method for seam removal in our case. However, there is not a method that is suitable for all kinds of seam removal in image stitching, because there are many factors affecting the attribute of the seam.

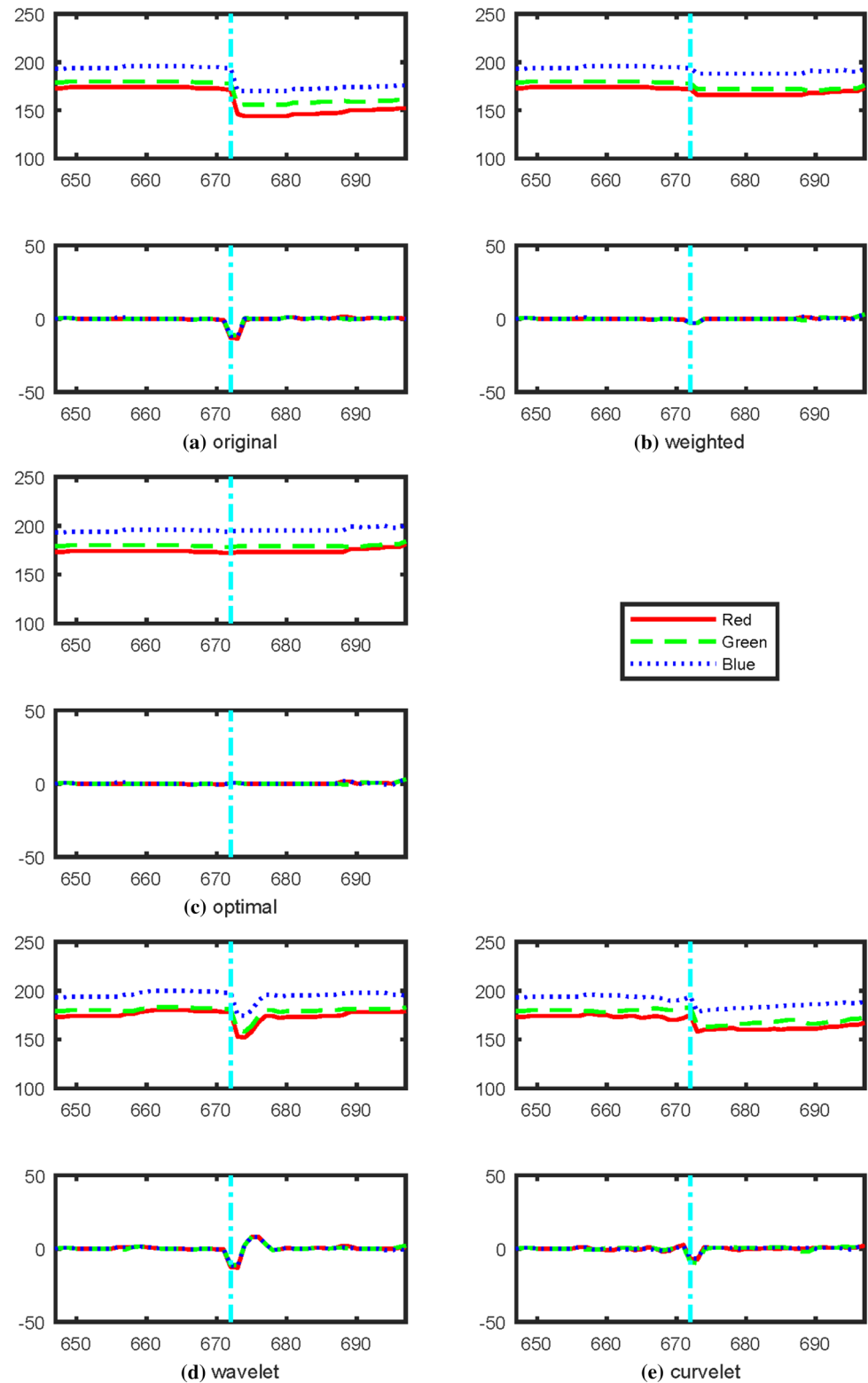
### 5 Discussion

Image stitching has been quickly developing in recent years. Many new methods related to image stitching have been proposed. However, there are still some points that can be improved in the future.

**Table 4** Assessment of stitched images using DoEM

	Original	Weighted	Optimal	Wavelet	Curvelet
Cise	0.8921	0.9051	0.9094	0.9089	0.9111
Rail	0.7961	0.7247	0.7701	0.7733	0.7813

**Fig. 19** Assessment results of *cise*

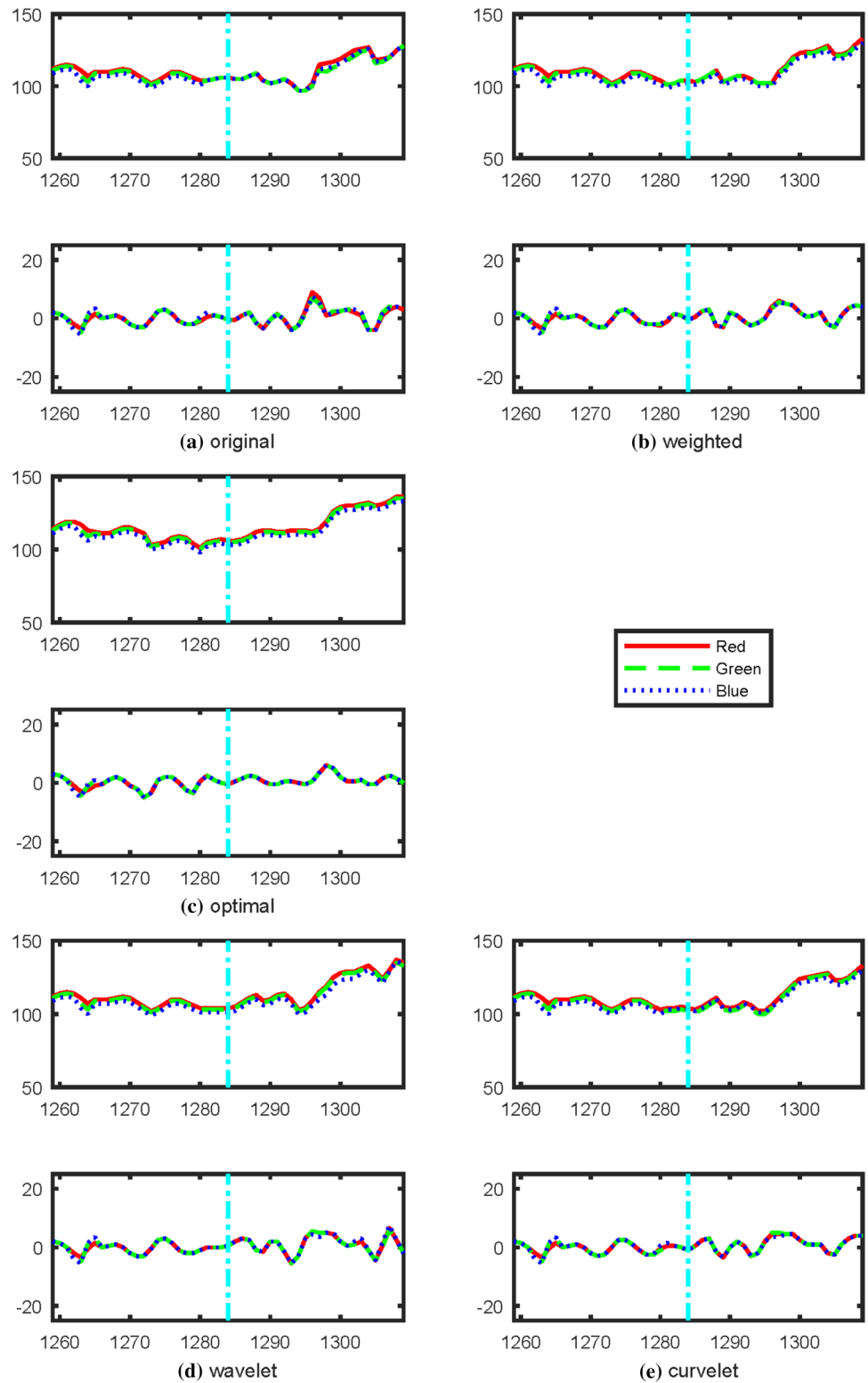


First is image registration with a large number of images. The errors of matching and registering two images will be enlarged when we stitch a large number of images and form it into a whole panorama image. Since the error is not

avoidable, we need to further transform the images to reduce it.

Second is image stitching with 3-dimensional images, which not only have the colour information, but also

**Fig. 20** Assessment results of *rail*



contain the information in the depth dimension. We need to adjust the matching and registration methods to take the depth dimension into consideration. Using point cloud might be one of the solutions.

Third is to develop image stitching for special purposes, including transplanting the algorithm into embedded systems such as cell phones, digital camera, unmanned aircraft, tablet and specific algorithms for geographical, biomedical,

**Table 5** Assessment of *cise* using gradient variance

	Original	Weighted	Optimal	Wavelet	Curvelet
<i>R</i>	7.0802	0.4465	0.1402	9.6902	2.6349
<i>G</i>	5.0749	0.4559	0.0884	8.1084	4.1249
<i>B</i>	5.8182	0.5565	0.2702	8.6565	2.1559
<i>M</i>	5.9911	0.4863	0.1663	8.8184	2.9719

**Table 6** Assessment of *rail* using gradient variance

	Original	Weighted	Optimal	Wavelet	Curvelet
<i>R</i>	6.8508	5.7741	4.2184	6.6125	5.1341
<i>G</i>	6.6459	6.0941	4.4284	7.0702	5.9041
<i>B</i>	7.0737	6.3641	4.5784	7.5349	5.6941
<i>M</i>	6.8568	6.0774	4.4084	7.0725	5.5774

under water images. This can be done by deeply investigating both its application field and the related algorithms.

With the increasing application of panorama image, new requirements are set to image stitching. It has become increasingly important to make the whole stitching process faster and easier. It is also a tendency that the image stitched will become more precise. The stitching algorithms that are designed for specific projects such as geographical, biomedical, under water images will appear. A mature algorithm should achieve a balance between time consuming and precision, making an optimal result that can be accepted by most of audience.

## 6 Conclusion

Image stitching relates to domains such as computer vision, image processing, computer graphics and software developing. The effects of image registration and seam removal play key roles in a successful stitching. In this article, some methods about image registration and seam removal are introduced and analysed. Then, image assessment methods are briefly presented.

In recent years, more and more improved image stitching methods keep appearing. Some algorithms can register images in a fast way, and some combine multi registration algorithms altogether. Many new seam removal methods will be proposed at the same time. Unfortunately, there is not an algorithm suitable for all the stitching, nor is it easy to select a specific method for certain purposes. However, with the development of new techniques, image stitching will be more precise and have wider applications.

**Acknowledgements** This work was jointly supported by National Natural Science Foundation of China (Grant no. 61201421), China Postdoctoral Science Foundation (Grant no. 2013M532097).

## Compliance with ethical standards

**Conflict of interest** We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

## References

- Bhosle, U., Chaudhuri, S., Roy, S.D.: A fast method for image mosaicing using geometric hashing. *Iete J. Res.* **48**(3–4), 317–324 (2002)
- Laraqui, A., et al.: Image mosaicing using voronoi diagram. *Multimed. Tools Appl.* **76**(6), 8803–8829 (2017)
- Kekec, T., Yildirim, A., Unel, M.: A new approach to real-time mosaicing of aerial images. *Robot. Autonom. Syst.* **62**(12), 1755–1767 (2014)
- Xie, X., et al.: A study on fast SIFT image mosaic algorithm based on compressed sensing and wavelet transform. *J. Ambient Intell. Human. Comput.* **6**(6), 835–843 (2015)
- Ye, M.J., et al.: Automatic seamless stitching method for CCD images of Chang'E-1 lunar mission. *J. Earth Sci.* **22**(5), 610–618 (2011)
- Hui, F.M., et al.: An improved landsat image mosaic of Antarctica. *Sci China-Earth Sci* **56**(1), 1–12 (2013)
- Bhat, A.S., et al.: Template matching technique for panoramic image stitching. In: *Modelling symposium (AMS), 7th Asia. IEEE* (2013)
- Dekel, T., et al.: Best-buddies similarity for robust template matching. In: *Proceedings of the IEEE conference on computer vision and pattern recognition* (2015)
- Pilchak, A.L., et al.: Using cross-correlation for automated stitching of two-dimensional multi-tile electron backscatter diffraction data. *J. Microsc.* **248**(2), 172–186 (2012)
- Adwan, S., Alsaleh, I., Majed, R.: A new approach for image stitching technique using dynamic time warping (DTW) algorithm towards scoliosis X-ray diagnosis. *Measurement* **84**, 32–46 (2016)
- Liu, Y.-Y., et al.: A study of image stitching algorithm based on ratio matching. *Electro-Opt. Technol. Appl.* **6**, 17 (2007)
- Legesse, F.B., et al.: Seamless stitching of tile scan microscope images. *J. Microsc.* **258**(3), 223–232 (2015)
- Nasibov, A., H. Nasibov, Hacizade, F.: Seamless image stitching algorithm using radiometric lens calibration for high resolution optical microscopy. In: *Soft computing, computing with words and perceptions in system analysis, decision and control, 2009. ICSCCW 2009. Fifth international conference on. IEEE* (2009)
- Li, W.T.: Mutual information functions versus correlation-functions. *J. Stat. Phys.* **60**(5–6), 823–837 (1990)
- Dame, A., Marchand, E.: Second-order optimization of mutual information for real-time image registration. *IEEE Trans. Image Process.* **21**(9), 4190–4203 (2012)
- Rivaz, H., Karimghaloo, Z., Collins, D.L.: Self-similarity weighted mutual information: a new nonrigid image registration metric. *Med. Image Anal.* **18**(2), 343–358 (2014)
- Gong, M.G., et al.: A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information. *IEEE Trans. Geosci. Remote Sens.* **52**(7), 4328–4338 (2014)



18. Gupta, N., et al.: Extracting information from a query image, for content based image retrieval. In: 2015 eighth international conference on advances in pattern recognition, New York, IEEE, p. 225 (2015)
19. Liu, H., et al.: Feature selection with dynamic mutual information. *Pattern Recogn.* **42**(7), 1330–1339 (2009)
20. Soleimani, S., et al.: Online wear detection using high-speed imaging. *Microsc. Microanal.* **22**(4), 820–840 (2016)
21. Lin, Y., Yu, Q., Medioni, G.: Efficient detection and tracking of moving objects in geo-coordinates. *Mach. Vis. Appl.* **22**(3), 505–520 (2011)
22. Studholme, C., Hill, D.L.G., Hawkes, D.J.: An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recogn.* **32**(1), 71–86 (1999)
23. Klein, S., et al.: Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information. *Med. Phys.* **35**(4), 1407–1417 (2008)
24. Burt, P.J., Adelson, E.H.: A multiresolution spline with application to image mosaics. *ACM Trans. Graph. (TOG)* **2**(4), 217–236 (1983)
25. Wang, C.-Y., Zhou, M.-Q.: A localized pyramid decomposition approach for image mosaic. *J. Northw. Univ. (Nat. Sci. Edn.)* **3**, 12 (2006)
26. Anuta, P.E.: Spatial registration of multispectral and multitemporal digital imagery using fast Fourier transform techniques. *IEEE Trans. Geosci. Electron.* **8**(4), 353–368 (1970)
27. Hurtós, N., Petillot, Y., Salvi, J.: Fourier-based registrations for two-dimensional forward-looking sonar image mosaicing. In: Intelligent robots and systems (IROS), 2012 IEEE/RSJ international conference on. IEEE (2012)
28. Vescovi, R.F.C., Cardoso, M.B., Miqueles, E.X.: Radiography registration for mosaic tomography. *J. Synchron. Radiat.* **24**, 686–694 (2017)
29. Ghantous, M., Ghosh, S., Bayoumi, M.: A multi-modal automatic image registration technique based on complex wavelets. In: Image processing (ICIP), 2009 16th IEEE international conference on. IEEE (2009)
30. Xing, Y.X., et al.: Robust fast corner detector based on filled circle and outer ring mask. *IET Image Proc.* **10**(4), 314–324 (2016)
31. Trajkovic, M., Hedley, M.: Fast corner detection. *Image Vis. Comput.* **16**(2), 75–87 (1998)
32. Förstner, W., Gülch, E.: A fast operator for detection and precise location of distinct points, corners and centres of circular features. In: Proc. ISPRS intercommission conference on fast processing of photogrammetric data (1987)
33. Moravec, H.P.: Obstacle avoidance and navigation in the real world by a seeing robot rover. DTIC Document (1980)
34. Harris, C., Stephens, M.: A combined corner and edge detector. In: Alvey vision conference. Citeseer (1988)
35. Harris, C.G., Pike, J.: 3D positional integration from image sequences. *Image Vis. Comput.* **6**(2), 87–90 (1988)
36. Zhu, M.C., et al.: Efficient video panoramic image stitching based on an improved selection of Harris corners and a multiple-constraint corner matching. *Plos One* **8**(12), e81182 (2013)
37. Zhu, M.C., et al.: A fast image stitching algorithm via multiple-constraint corner matching. *Math. Prob. Eng.* **2013**, 157847 (2013)
38. Fu, Z.X., Wang, L.M.: Optimized design of automatic image mosaic. *Multimed. Tools Appl.* **72**(1), 503–514 (2014)
39. Pan, H., et al.: An adaptive Harris corner detection algorithm for image mosaic. In: Chinese Conference on Pattern Recognition. Springer (2014)
40. Zhou, Z. et al.: Fast image mosaic algorithm based on the improved Harris-SIFT algorithm, In: International Symposium on Photonics and Optoelectronics. SPIE (2015)
41. Jiang, Z., Liu, M.: Fast image mosaic algorithm based on the improved Harris-SIFT algorithm. In: Zhou, Z. (Ed.) International Symposium on Photonics and Optoelectronics (2015)
42. Li, B.P., Guo, C.X., Inc, D.E.P.: Application of image stitching in the scene investigation of traffic accident. In: 2016 international conference on information engineering and communications technology, pp. 127–131 (2016)
43. Lowe, D.G.: Object recognition from local scale-invariant features. In: Computer vision, 1999. The proceedings of the seventh IEEE international conference on. IEEE (1999)
44. Brown, M., Lowe, D.G.: Invariant features from interest point groups. In BMVC (2002)
45. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision* **60**(2), 91–110 (2004)
46. Ni, X.L., et al.: A fully automatic registration approach based on contour and SIFT for HJ-1 images. *Sci. China-Earth Sci.* **55**(10), 1679–1687 (2012)
47. Guo, F., et al.: Automatic retinal image registration using blood vessel segmentation and SIFT feature. *Int. J. Pattern Recognit. Artif. Intell.* **31**(11), 1757006 (2017)
48. Goncalves, H., Corte-Real, L., Goncalves, J.A.: Automatic image registration through image segmentation and SIFT. *IEEE Trans. Geosci. Remote Sens.* **49**(7), 2589–2600 (2011)
49. Wang, V.T., Hayes, M.P.: Synthetic aperture sonar track registration using SIFT image correspondences. *IEEE J. Ocean. Eng.* **42**(4), 901–913 (2017)
50. Min, Z., Jiguo, Z., Xusheng, X.: Panorama stitching based on sift algorithm and Levenberg–Marquardt optimization. *Phys. Procedia* **33**, 811–818 (2012)
51. Zhang, Y.H., Jin, X., Wang, Z.J.: A new modified panoramic UAV image stitching model based on the GA-SIFT and adaptive threshold method. *Memetic Comput.* **9**(3), 231–244 (2017)
52. Qu, Z., et al.: The improved algorithm of fast panorama stitching for image sequence and reducing the distortion errors. *Math. Probl. Eng.* **2015**, 428076 (2015)
53. Zhu, J., Ren, M.W.: Image mosaic method based on SIFT features of line segment. *Comput. Math. Methods Med.* **2014**, 926312 (2014)
54. Wang, F.-B., et al.: Multi-image mosaic with SIFT and vision measurement for microscale structures processed by femtosecond laser. *Opt. Lasers Eng.* **100**, 124–130 (2018)
55. Mills, A., Dudek, G.: Image stitching with dynamic elements. *Image Vis. Comput.* **27**(10), 1593–1602 (2009)
56. Ke, Y., Sukthankar, R.: PCA-SIFT: a more distinctive representation for local image descriptors. In: Computer vision and pattern recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on. IEEE (2004)
57. Jolliffe, I.: Principal Component Analysis. Wiley Online Library, New York (2002)
58. Kshirsagar, V., Baviskar, M., Gaikwad, M.: Face recognition using Eigenfaces. In: Computer research and development (ICCRD), 2011 3rd international conference on. IEEE (2011)
59. Li, B.L., Thomas, G., Williams, D.: Detection of ice on power cables based on image texture features. *IEEE Trans. Instrum. Meas.* **67**(3), 497–504 (2018)
60. Wang, Q.Q., et al.: l(2, p)-norm based PCA for image recognition. *IEEE Trans. Image Process.* **27**(3), 1336–1346 (2018)
61. Bay, H., Tuytelaars, T., Van Gool, L.: SURF: speeded up robust features. *Comp. Vis. ECCV 2006 Pt 1 Proc.* **3951**, 404–417 (2006)
62. Juan, L., Gwun, O.: A comparison of sift, PCA-sift and surf. *Int. J. Image Process. (IJIP)* **3**(4), 143–152 (2009)
63. Lukashovich, P., Zalesky, B., Ablameyko, S.: Medical image registration based on surf detector. *Pattern Recognit. Image Anal.* **21**(3), 519 (2011)

64. Wang, G., et al.: Robust point matching method for multimodal retinal image registration. *Biomed. Signal Process. Control* **19**, 68–76 (2015)
65. Yang, F., Deng, Z.S., Fan, Q.H.: A method for fast automated microscope image stitching. *Micron* **48**, 17–25 (2013)
66. Yang, Z.L., Shen, D.G., Yap, P.T.: Image mosaicking using SURF features of line segments. *Plos One* **12**(3), e0173627 (2017)
67. Tsai, C.M., Shih, F.Y.: An efficient image stitching method for heterogeneous car videos based on bounding boxes of features. *Int. J. Pattern Recognit. Artif. Intell.* **31**(5), 1755008 (2017)
68. Wang, J., et al.: A fast mosaic method for airborne images: the new template-convolution speed-up robust features (TSURF) algorithm. *Int. J. Remote Sens.* **35**(16), 5959–5970 (2014)
69. Alcantarilla, P.F., Bartoli, A., Davison, A.J.: *KAZE Features*. Springer, Berlin (2012)
70. Alcantarilla, P.F., Nuevo, J., Bartoli, A.: Fast explicit diffusion for accelerated features in nonlinear scale spaces. *IEEE Trans. Patt. Anal. Mach. Intell.* **34**, 1281–1298 (2013)
71. Yang, X., Cheng, K.-T.: Local difference binary for ultrafast and distinctive feature description. *IEEE Trans. Pattern Anal. Mach. Intell.* **36**(1), 188–194 (2014)
72. Thanh, T.M., et al.: Robust semi-blind video watermarking based on frame-patch matching. *AEU-Int. J. Electron. Commun.* **68**(10), 1007–1015 (2014)
73. Liu, Y., et al.: S-AKAZE: An effective point-based method for image matching. *Optik-Int. J. Light Electron Opt.* **127**(14), 5670–5681 (2016)
74. Flores-Rodriguez, K.L., Trujillo-Romero, F.: Free form object recognition module using A-KAZE and GCS. *Res. Comput. Sci* **118**, 19–29 (2016)
75. Mukherjee, P., Lall, B.: Saliency and KAZE features assisted object segmentation. *Image Vis. Comput.* **61**, 82–97 (2017)
76. Qu, Z., Bu, W., Liu, L.: The algorithm of seamless image mosaic based on A-KAZE features extraction and reducing the inclination of image. *IEEJ Trans. Electr. Electron. Eng.* **13**(1), 134–146 (2018)
77. Pohl, C., Van Genderen, J.L.: Review article multisensor image fusion in remote sensing: concepts, methods and applications. *Int. J. Remote Sens.* **19**(5), 823–854 (1998)
78. Li, H.Y., et al.: An adaptive image-stitching algorithm for an underwater monitoring system. *Int. J. Adv. Robot. Syst.* **11**, 166 (2014)
79. Li, Z.Y., et al.: A novel image mosaicking algorithm for wireless multimedia sensor networks. *Int. J. Distrib. Sensor Netw.* **9**, 719640 (2013)
80. Zhu, L., et al.: Review on image fusion research. *Transduces Microsyst. Technol.* **33**(2), 14–18 (2014)
81. Zhou, D.-F., He, M.-Y., Yang, Q.: A robust seamless image stitching algorithm based on feature points. *Meas. Control Technol.* **6**, 10 (2009)
82. Zhao, G., Lin, L., Tang, Y.D.: A new optimal seam finding method based on tensor analysis for automatic panorama construction. *Pattern Recogn. Lett.* **34**(3), 308–314 (2013)
83. Abdukholikov, M., Whangbo, T.: Fast image stitching method for handling dynamic object problems in Panoramic Images. *Ksii Trans. Internet Inf. Syst.* **11**(11), 5419–5435 (2017)
84. Jeong, J., Jun, K.: A novel seam finding method using downscaling and cost for image stitching. *J. Sens.* **2016**, 5258473 (2016)
85. Li, M., et al.: A stereo dual-channel dynamic programming algorithm for UAV image stitching. *Sensors* **17**(9), 2060 (2017)
86. Li, L., et al.: A unified framework for street-view panorama stitching. *Sensors* **17**(1), 1 (2017)
87. Li, L., et al.: Optimal seamline detection for multiple image mosaicking via graph cuts. *ISPRS J. Photogram. Remote Sens.* **113**, 1–16 (2016)
88. Lee, D., Lee, S.: Seamless image stitching by homography refinement and structure deformation using optimal seam pair detection. *J. Electron. Imaging* **26**(6), 063016 (2017)
89. Wang, Y., et al.: Microscopic image enhancement based on Fourier ptychography technique. *SPIE Defense + Commercial Sensing*, vol 10990. SPIE (2019)
90. Chen, M., et al.: Underwater image stitching based on SIFT and wavelet fusion. *Oceans—Genova* (2015)
91. Popovic, V., Leblebici, Y.: FIR filters for hardware-based real-time multi-band image blending. In: Kehtarnavaz, N., Carlsohn, M.F. (Eds) *Real-Time Image and Video Processing* (2015)
92. Candes, E.J., Guo, F.: New multiscale transforms, minimum total variation synthesis: applications to edge-preserving image reconstruction. *Signal Process.* **82**(11), 1519–1543 (2002)
93. Srivastava, R., Prakash, O., Khare, A.: Local energy-based multimodal medical image fusion in curvelet domain. *IET Comput. Vis.* **10**(6), 513–527 (2016)
94. Yang, H., et al.: A seismic interpolation and denoising method with curvelet transform matching filter. *Acta Geophys.* **65**(5), 1029–1042 (2017)
95. Gai, S.: Multichannel image denoising using color monogenic curvelet transform. *Soft. Comput.* **22**(2), 635–644 (2018)
96. Math, S.S.P., Kaliyaperumal, V.: Enhancement of SAR images using fuzzy shrinkage technique in curvelet domain. *Sadhana-Acad. Proc. Eng. Sci.* **42**(9), 1505–1512 (2017)
97. Ali, F.E., et al.: A curvelet transform approach for the fusion of MR and CT images. *J. Mod. Opt.* **57**(4), 273–286 (2010)
98. Gattim, N.K., et al.: multimodal image fusion using curvelet and genetic algorithm. *J. Sci. Ind. Res.* **76**(11), 694–696 (2017)
99. Candes, E., Demanet, L.: Curvelets and fourier integral operators. *Comptes Rendus Math.* **336**(5), 395–398 (2003)
100. Wu, Q., et al.: A perceptually weighted rank correlation indicator for objective image quality assessment. *IEEE Trans. Image Process.* **27**(5), 2499–2513 (2018)
101. Huynh-Thu, Q., Ghanbari, M.: Scope of validity of PSNR in image/video quality assessment. *Electron. Lett.* **44**(13), 800–U35 (2008)
102. Wang, Z., Bovik, A.C.: A universal image quality index. *IEEE Signal Process. Lett.* **9**(3), 81–84 (2002)
103. Wang, Z., et al.: Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**(4), 600–612 (2004)
104. Thomos, N., Boulgouris, N.V., Strintzis, M.G.: Optimized transmission of JPEG2000 streams over wireless channels. *IEEE Trans. Image Process.* **15**(1), 54–67 (2006)
105. Hou, W.L., et al.: Blind image quality assessment via deep learning. *IEEE Trans. Neural Netw. Learn. Syst.* **26**(6), 1275–1286 (2015)
106. Guo-ting, W., et al.: Method for quality assessment of image mosaic. *J. Commun.* **8**, 011 (2013)
107. Qureshi, H.S., et al.: Quantitative quality assessment of stitched panoramic images. *IET Image Proc.* **6**(9), 1348–1358 (2012)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.