Image Redundancy Filtering for Panorama Stitching

XIN WEI\textsuperscript{1}, WEIQING YAN\textsuperscript{1}, QIANG ZHENG\textsuperscript{1}, MEIQI GU\textsuperscript{1}, KAIQI SU\textsuperscript{1}, GUANGHUI YUE\textsuperscript{2}, AND YUN LIU\textsuperscript{3}

\textsuperscript{1}School of Computer and Control Engineering, Yantai University, Yantai 264005, China
\textsuperscript{2}Health Science Center, School of Biomedical Engineering, Shenzhen University, Shenzhen 518060, China
\textsuperscript{3}College of Information, Liaoning University, Liaoning 110036, China

Corresponding author: Weiqing Yan (wqyan@tju.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61801414, Grant 62072391, and Grant 61901205; in part by the Natural Science Foundation of Shandong Province under Grant ZR2017QF006; in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2019A1515111205; Natural Science Foundation of China under Grant 62001302; and in part by the Natural Science Foundation of Shenzhen under Grant JCYJ20190808145011259.

\section*{ABSTRACT}

In this paper, we designed a novel framework for massive image panorama stitching, which aims to resolve image redundancy, alignment error accumulation and perspective distortion accumulation of the stitching process. First, an iterative method is designed to filter the redundancy images by analysing the similarity relation of adjacent images. Second, to reduce alignment error accumulation, the weight topology graph of filtered images is constructed, which is employed to find the optimal reference image which closes to the central geometrically by the Floyd-Warshall algorithm. Finally, a two-step global alignment strategy is designed to initial align images and perform shape optimization, the first step is that filtered images employ the similarity model to roughly align group by group, the second step is to perform shape optimization further through refining all alignment parameters by the homography model under the anti-perspective energy term, which aims to obtain an optimal solution by balancing the alignment accuracy and the global consistency. Compared with the state-of-the-art methods, the proposed method successfully reduces image redundancy while improving the alignment and reducing perspective distortion for massive image panorama stitching.

\section*{INDEX TERMS}

Image stitching, redundancy image, global consistency, topology estimation.

\section*{I. INTRODUCTION}

Unmanned Aerial Vehicles (UAV) and Aerial remote sensing are the most common ways to quickly acquire high-quality images from the observing region. Since the observing region of single image is limited, we need to combine multiple images with common overlap into a single wide-scale panorama image by panorama image stitching technique\cite{1} for viewing the observing region conveniently. Recently, image stitching technique is widely used in all aspects of computer vision, such as stereoscopic image stitching\cite{2},\cite{3}, video surveillance\cite{4} and virtual reality\cite{5}.

The first critical step of the stitching process is to accurately estimate the transformation relation of all images to the stitching plane, which directly determines the quality of the final stitching result\cite{6}–\cite{8}. If the images are captured from an approximate plane, the homography model is usually employed to align them to achieve a high-accuracy alignment solution, since it has higher degrees of freedom (DoFs) to adjust transformation relation of images\cite{9}–\cite{12}. Considering the stability and efficiency of the stitching process, many approaches utilized feature-based matching method to discover the correspondence relationship and estimate the transformation relation between images\cite{13}–\cite{16} since features (such as point, line or edge features) are easy to extract and reliable, especially point features\cite{17},\cite{18}. Although many studies on image stitching have been presented in the past decade, some challenging problems still need to be further solved, such as image redundancy, error accumulation, the balance between the alignment accuracy and the global consistency.

To obtain a high-accuracy alignment of the stitching result, the bundle adjustment approach\cite{19},\cite{20} is usually employed to optimize the alignment parameters globally by
minimizing the total reprojection errors [21]. In addition, combining and localizing feature and area to align is a robust way to improve accuracy [22]. Many methods divide the alignment process into two steps: obtaining a robust alignment solution first and performing global optimization later. Xing et al. [23] proposed to first obtain an initial alignment by employing the Extended Kalman Filter [24] onto the local region, then all parameters are optimized in the global. To solve the non-linear optimization problem, Kekec et al. [25] first applied the affine model to estimate the initial alignment solution and then applied the homography model to perform global optimization under the global consistency constraint.

For images with parallax, recent state-of-the-art methods show that local alignment model has the advantage to handle parallax since it has more flexible than global alignment by dividing each image into two or more frames and working on each frame. Gao et al. [26] proposed to divide each image into two dominant planes and then use the linear combination of two homography matrices to produce the panoramic image. Zaragoza et al. [27] proposed the as-projective-as-possible (APAP) method to estimate multiple local homography matrices of the grid by a weight method, which can achieve a high-accuracy alignment and reduce the ghosting effect significantly. The references [28], [29] utilized a weighting strategy to combine multiple local homography matrices of non-overlapping with a global similarity matrix, which allows continuously adjusting transformation matrices of the overlapping areas into the non-overlapping areas to achieve transformation smoothly. Zhang et al. [30] proposed to combine the homography model and content-preserving warping to design a hybrid alignment model that provides more flexibility to solve parallax and eliminate local distortion. Seam optimization [31] is also an effective way to improve the quality of the stitching image. Li et al. [32] first proposed an automatic piecewise dynamic program to detect the optimal seamline, and then proposed a cosine distance weight blending method to ensure that the seamline is as invisible as possible. To reduce visible seams, multi-band blending [16] and gradient domain [33] are powerful techniques to seamless blend overlapping areas. To eliminate local misalignment, Li et al. [34] proposed to utilize local deviation correction to optimize the parameters with global similarity constraint. To better remove the outliers, Li et al. [35] employed RBF (radial basis function) approximation to design an outlier-tolerant alignment method that can efficiently remove the outliers by screening out the abnormal weights of RBFs.

However, the above methods are only suitable for stitching dozens of image to a qualified panorama image, which are prone to producing a severe accumulation of perspective distortion for massive images stitching, that means the global consistency of stitching result is broken. To address this problem, Caballero et al. [24] designed a hierarchical model according to the alignment accuracy, where the images with large parallax are solved by the model with less DoFs. Xie et al. [36] proposed a group model according to the alignment accuracy, where the image group with large parallax is connected to the image group with small parallax by the homography model, and then the image group with small parallax is solved by the affine model. Xia et al. [37] proposed to allow continuous transition from the model with less DoFs to the model with more DoFs under the predefined constraint. The essence of these methods is keeping the balance between improving alignment accuracy and reducing distortion accumulation.

For large-view image stitching, the topology relation of images is usually utilized to achieve a global alignment [38], [39]. Elibol et al. [40] combined the fast tentative matching and Minimum Spanning Tree (MST) algorithm to discover the correspondence relation between images, which reduces the number of matching attempts significantly. To select the optimal reference image as the stitching plane, Szeliski [41] proposed a proper solution that is to choose one of the images which is close to the central geometrically as the reference image, since setting the central image as the stitching plane usually produces the least error propagation. To achieve the above supposition, Choe et al. [42] proposed first to calculate the alignment error between each image pair, then select the image with the lowest alignment error accumulation as the reference image by utilizing the graph algorithm. To improve computation speed, Zheng et al. [43] designed a multi-row strategy to scan images and then selected the subset of original image set to stitch.

In this paper, we proposed a novel framework for massive image panorama stitching, which aims to resolve image redundancy, alignment error accumulation and perspective distortion accumulation of the stitching process. The contributions of this paper are summarised as follows: 1) Based on the similarity matrix designed by topological structure information of original image set, we propose an iterative method to filter redundancy images from original image set. Compared with the existing methods, the proposed method greatly improves the stitching efficiency under the condition of obtaining a high-quality stitching result without information loss. 2) A weight equation is designed to build the topology graph of filtered image set, moreover, to find the optimal reference image by the Floyd-Warshall algorithm, which can reduce alignment error accumulation. 3) A two-step global alignment strategy is proposed to align and perform perspective optimization, which can improve the alignment accuracy while reducing distortion accumulation for massive images stitching.

The rest of this paper is organized as follows: In Section II, we describe the proposed image redundancy filtering for panorama stitching in detail, then the experimental results are presented in Section III. Finally, the conclusion of our work is stated in Section IV.

II. THE PROPOSED METHOD

In this section, we introduce the proposed framework for massive image panorama stitching in detail, which aims to speed up the stitching process by filtering redundancy
images meanwhile obtain a high-quality panoramic image without information loss. The proposed framework includes three modules: image redundancy filtering, optimal reference image selection, and alignment and shape optimization. First, in image redundancy filtering, we designed an iterative method based on similarity analysis to filter images from original image set. Second, for reference image selection, error accumulation is unavoidable when all the images are projected to the same stitching plane by a series of transformations, we design an optimal reference image to reduce error accumulation. The optimal reference image is chosen on our designed weight topology graph of filtered image set by the Floyd-Warshall algorithm. Finally, for images with parallax, a high-accuracy alignment is prone to causing severe distortion accumulation. To obtain a trade-off between the alignment accuracy and distortion accumulation, a two-step global alignment strategy is designed to initial align images and perform shape optimization further, here, the first step is that filtered images are roughly aligned between groups by the similarity model, the second step is that all alignment parameters are optimized globally by the homography model under the anti-perspective energy term, which aims to obtain an optimal stitching result. Fig. 1 shows the flowchart of the proposed framework.

To provide an accurate understanding, the vital variables and notations are summarized in Table 1.

### A. IMAGE REDUNDANCY FILTERING

In this section, we remove the redundancy images from original image set meanwhile guarantee the final stitching image without information loss. In massive images stitching, since the overlapping region between adjacent images may be very large, the efficiency of stitching is low when stitching these adjacent images. Therefore, we need to remove these images which have a high overlapping region with their adjacent images. The designed filtering method is achieved through analysing the similarity relation of image pairs based on the topology information and simultaneously considering the connectivity of filtered image set. The process of our filtering method is presented as follows: 1) we find the image pair which has the maximum similarity value in the similarity matrix $s$, if its similarity value is higher than the similarity threshold value, turn to the step 2); otherwise, exit the filtering process. 2) we remove one of the image pair (here, we use A and B denote a matched image pair) from the step 1), according to the designed critical criterion (the following equation (1)); 3) update the similarity matrix $s$ of current image set.

The similarity relation matrix $s$ is defined by the number of matched feature points between two images. $s(i, j)$ denotes the similarity of image $I_i$ and $I_j$, its value is the number of matched feature points between the two images. In this paper, to improve computation efficiency, we employ a feature matching method based on topology analysis from the Ref. [37] to estimate the similarity relation matrix $s$ of original image set.

If the similarity value of two images is large, we attempt to remove one of the two images from the current image set. According to this idea, we design a critical criterion to choose one of them as the redundancy image and then remove it.
from pre-stitched image set. The filter critical is achieved by calculating the sum of similarity of each member of the image pair with their common adjacent image set, respectively. If the image has a smaller similarity value sum, which means it has a suitable overlap area with each common adjacent image. The other one has a relatively more overlap area with some of the common adjacent images and has relatively less overlap area with others of the common adjacent images. Obviously, this one is inappropriate to reserve to produce the final stitching image. Therefore, the critical criterion is denoted as follows:

$$
\min_{p \in \{A, B\}} \left( \sum_{i=1}^{n} \sum_{j=1}^{n} s(i, j) - \sum_{i=1}^{n} s(i, p) - \sum_{j=1}^{n} s(p, j) \right) \quad (1)
$$

where n represents the number of filtered images, A and B denote an image pair with the maximum similarity value. We sum all items of the similarity matrix without someone related with each member of the image pair, respectively. In this paper, we assume that A and B are an image pair with the maximum similarity value. If the sum of the similarity value of image A with other images is smaller than that of B, image A is the redundancy image; otherwise, B is the redundancy image. We apply the iterative approach to find the redundancy image and remove it again and again until the maximum value of the similarity matrix of current image set becomes smaller than or equal to our designed similarity threshold value T. Note: the similarity threshold value T is determined the number of filtered images. In the experiment, we show the relationship between the number of filtered images and the similarity threshold value T.

The pseudo-code of the proposed filtering method is listed in Table 2.

### TABLE 2. Pseudo code of image redundancy filtering.

| Algorithm 1. Image redundancy filtering |
|-----------------------------------------|
| **Input:** original image set O with n images |
| **Output:** filtered image set O' |
| 1: Initialize O' be equal to O, the similarity matrix of image set O' and the similarity threshold value T |
| 2: while (max(s) > T) |
| 3: Compute the image pair A and B corresponding to the similarity value of max(s) according to the equation (1). |
| 4: Remove the image A or B corresponding to the equation (1) from O' |
| 5: Update the similarity matrix s of O' |
| 6: Output the current image set O' |

### B. REFERENCE IMAGE SELECTION

To represent the final stitching image, a natural approach is to select one of pre-stitched images as the reference and then project all of other images into the reference plane. As we know, if an image has no overlap with the reference image, it is projected to the reference plane through using the related intermediate images to continuously transform. In the process of continuously transform, error accumulation is inescapable. To reduce error accumulation, a reasonable solution is regarding the image which makes intermediate images less as the reference image. In other words, the most geometrical central image is the best choice to select as the reference image, which produces the least accumulated errors in the process of aligning all the other images to the reference plane.

Next, we show the process of optimal reference image selection. First, a weight topology graph is constructed based on the similarity matrix of filtered image set obtained in the above section. According to the similarity matrix, we use the edges to link the image pairs whose similarity values are non-zero. There are two existing ways to set the weight of the edge: set the edge as the reciprocal of the total number of matched features between the image pair [33] and set the edge as the alignment error between the image pair [35]. The first way is efficient since it only utilizes the existing condition of matched features, the second way can accurately search the optimal reference image with the least error accumulation of stitching result, but it has to calculate the alignment error between each image pair in advance, which will reduce the stitching efficiency. Inspired by these two ways, we creatively proposed a weight equation which considers the correlation between the number of matched feature points and the alignment accuracy. The weight equation is defined as follows:

$$
w(i, j) = \begin{cases} 
1 & \text{if}(i, j) > 0 \\
\ln(s(i, j) + 1) + \varepsilon & \text{if}(i, j) = 0 
\end{cases} \quad (2)
$$

where s(i, j) denotes the number of matched feature points between I_i and I_j, ε is a constant for regularization. In our experiment, the constant ε is set as 1 empirically, which aims to avoid the denominator of the weight is zero.

According to the weight topology graph, we can find the optimal reference image by finding a node whose sum of weight value of the shortest paths to all the other nodes is minimal. To achieve this mission, the Floyd-Warshall algorithm [44], [45] is employed to calculate the shortest paths from all nodes to all nodes. An n*n size cost matrix C is constructed and C(i, j) records the weight value of the shortest path from I_i to I_j. We choose the image with the least total weight value of the shortest paths from all the other images to it as the reference image.

In the proposed method, the reference image is selected by the weight value of the topology graph. The weight is set according to the number of matched feature points and the alignment accuracy. In this paper, we use Fig. 2 to show the influence of different weights on the final stitching result on our test image set with 61 images. Fig. 2 shows the selection results of different weight equations. Fig. 2(a) shows the topology and the stitching result of the weight with just considering the number of matched feature points. Fig. 2(b) shows the topology and the stitching result of the weight obtained by Eq. (2). From this figure, we can see
the reference image is different by different weight methods. We also measure the alignment accuracy of the stitching image by Root-Mean-Square (RMS) errors. The smaller of RMS, the higher of the alignment accuracy. The RMS of Fig. 2(a) and (b) is 1.69873 and 1.6578, respectively. This demonstrates that the proposed weight method can obtain higher alignment accuracy than that of just considering the number of matched features.

![Image](image.png)

**FIGURE 2.** The reference selection results of different weight strategies: (a) the topology and the stitching result of the weight with just considering the number of matched features; (b) the topology and the stitching result for our weight equation. The reference image is marked with an orange node in the topology and the green rectangle box in the stitching result. The corresponding alignment accuracy of stitching results is 1.69873 and 1.6578, respectively.

### C. IMAGE ALIGNMENT AND SHAPE OPTIMIZATION

The quality of stitching result is determined by both the alignment accuracy and the global consistency. To obtain a high-accuracy stitching image, the homography model is usually used to align images since it has higher degrees of freedom (DoFs). However, it is prone to making a severe accumulation of perspective distortion (in other words, destroy the global consistency) when aligning massive images captured from an approximate plane region. Therefore, we have to well balance the alignment accuracy and the perspective distortion to obtain the optimal solution. Since the similarity model could reduce perspective distortions, we employed it to roughly align, then the homography model is employed to optimize the alignment accuracy under the anti-perspective energy term.

1) ROBUST ALIGNMENT BY GROUP SIMILARITY MODEL

The similarity model has shown that similarity transformation has an advantage in reducing perspective distortions. Inspired by these studies [28], [29], we use it to achieve a robust initial alignment. On the one hand, the four-parameter similarity model can well resist the perspective distortion for aligning the images with parallax. On the other hand, the alignment parameters obtained by the similarity model are qualified as the initial solution for the following shape optimization using the homography model.

Based on the above weight topology graph of filtered image set, the alignment order is ensured by setting the reference image as the root node and performing the breadth-first search, which is used to group filtered image set. After the images are grouped, images are aligned to the reference plane group by group, which has a better effect than the way one by one in reducing the error accumulation of alignment. In the initial alignment process, the alignment error between the aligned images and the following pre-aligned image group, and the alignment error inside the pre-aligned image group are two basic factors we should consider, which will be solved together in the cost function. Let $S_m = \{S_i, i = 1, 2, ...n_1\}$ be the similarity model parameters of the aligned images, $S_{m+1} = \{S_i, i = 1, 2, ...n_2\}$ be the similarity model parameters of the following pre-aligned image group. We can obtain the optimal solution by minimizing the cost function as follows:

$$E(S) = E_1(S|G_m, G_{m+1}) + E_2(S_m|G_m)$$  \hspace{1cm} (3)$$

where the first energy term $E_1(S|G_m, G_{m+1})$ represents the total number of error propagation between image group $G_m$ and $G_{m+1}$, $S = S_m \cup S_{m+1}$, $S$ is the union set of $S_m$ and $S_{m+1}$, it is defined as follows:

$$E_1(S|G_m, G_{m+1}) = \sum_{I_i \in G_m, I_j \in G_{m+1}} \sum_{k=1}^{M_{i,j}} \left\| \tilde{p}_{I_i,I_j}^k(S_i) - \tilde{p}_{I_i,I_j}^k(S_j) \right\|^2$$  \hspace{1cm} (4)$$

In the equation above, the second energy term $E_2(S_m|G_m)$ represents the total number of error propagation inside group in $G_m$, it is defined as follows:

$$E_2(S_m|G_m) = \sum_{I_i \in G_m} \sum_{k=1}^{M_{i,j}} \left\| \tilde{p}_{I_i,I_j}^k(S_i) - \tilde{p}_{I_i,I_j}^k(S_j) \right\|^2$$  \hspace{1cm} (5)$$

where $p_{I_i,I_j}^k(S_i) = r(S_i)p_{I_i,I_j}^k(S_i)$, $r(S_i)p_{I_i,I_j}^k(S_i)$ denotes that apply similarity transformation $S_i$ to the point $p_{I_i,I_j}^k$, $r(x)$ denotes that transform the homogeneous coordinate of a 2D point into the non-homogeneous coordinate, $M_{i,j}$ denotes the number of matched feature points between $I_i$ and $I_j$.

Since the equation (5) is a linear over-determined equation (similar explanation is shown in our previous work [2], [3]), which can be easily solved by Singular Value Decomposition (SVD) solution.

2) SHAPE OPTIMIZATION BY HOMOGRAPHY MODEL

The similarity model parameters estimated group by group are achieved a rough alignment, which aims to well reduce the perspective distortion of the stitching result. Since the DoFs of the similarity model is limited, severe alignment errors...
usually arise when aligning images with parallax, which will produce many misalignments in stitching results. Therefore, we have to perform shape optimization further by increasing alignment accuracy. To increase the alignment accuracy and simultaneously not induce the perspective distortions, we combine two mutually opposing terms into the energy function by a weight way, which is used to balance the contribution of the alignment accuracy and the global consistency to the final stitching result.

The first optimization term is to optimize alignment. In the above section, the similarity model parameters of all the images can be obtained, denoted as \( X = \{ S_I, i = 1, 2, \ldots, n_1 \} \), which are employed as the initial solution for optimizing homography model parameters. The homography model parameters of each image relating to the stitching plane will be optimized in the energy function. The first energy term aims to minimize the sum of feature alignment errors between images, it is defined as follows:

\[
E_1 (H) = \sum_{I_i, I_j \in X} \sum_{k=1}^{M_{I_i I_j}} \left\| p^k_{I_i I_j} (H_{I_i}) - p^k_{I_i I_j} (H_{I_j}) \right\|^2 \tag{6}
\]

Images have more freedom to adjust their transformation relations to the stitching plane, with the increase of parameters of alignment models. It causes an overall improvement in alignment accuracy. Besides, the remaining errors will be evenly distributed under the uniform energy term.

The second optimization term aims to maintain the global consistency by resisting perspective distortions, which provides a constraint to make the optimal homography model parameters close to the previously obtained similarity model parameters. This term is obtained through performing the transition from the similarity model parameters to the homography model parameters. Hence, the energy term is denoted as follows:

\[
E_2 (H|S) = \sum_{I_i, I_j \in X} \sum_{k=1}^{M_{I_i I_j}} \left( \left\| p^k_{I_i I_j} (H_{I_i}) - p^k_{I_i I_j} (S_{I_i}) \right\|^2 + \left\| p^k_{I_i I_j} (H_{I_j}) - p^k_{I_i I_j} (S_{I_j}) \right\|^2 \right) \tag{7}
\]

where the definition of \( p^k_{I_i I_j} (H_{I_i}) \) and \( p^k_{I_i I_j} (H_{I_j}) \) are similar to Eq. (5).

The final energy term is defined by linearly combining the above energy terms defined in Eqs. (6) and (7) as follows:

\[
E = E_1 (H) + \lambda E_2 (H|S) \tag{8}
\]

where \( \lambda \) represents the weight coefficient, which is used to balance the two terms \( E_1 \) and \( E_2 \). Since the constraint is not strict, we should set \( \lambda \) as a relatively small value. In our experiment, \( \lambda \) set as 0.01. As a typical non-linear least squares problem, we employ the Levenberg-Marquardt solution [45] to solve Eq. (8).

### III. EXPERIMENTAL RESULTS AND ANALYSIS

We applied four representative image datasets from the Ref. [37] and open source community to evaluate the proposed method. The first image dataset consists of 61 images, is obtained by an unmanned aerial vehicle (UAV). The second image dataset consists of 56 images. The third image dataset consists of 160 images, is obtained by an unmanned aerial vehicle (UAV). The fourth image dataset consists of 744 images, is taken at an altitude of 780 meters above the urban area. To verify the robustness of the proposed method, the strip-shaped image set and the square-shaped image set are selected from the first image set and second image set, respectively. To demonstrate the superiority of the proposed method, a few comparison experiments are conducted to demonstrate the advantages of the proposed method.

The comparisons are made on the estimated topologies and the final stitching results of the aforementioned datasets generated by different methods.

The number of filtered images is denoted as follows:

\[
N_f = N - \alpha \cdot N \tag{9}
\]

where \( N_f \) represents the number of filtered images, \( N \) represents the number of all images, \( \alpha \) represents the ratio of redundancy images to all images.

In image redundancy filtering, we remove a redundancy image depending on the maximum value of the similarity matrix of current filtered image set. Therefore, once the number of filtered images is determined, we can obtain the current threshold of matched feature points (i.e., the maximum value in the similarity matrix of current filtered image set). To facilitate the understanding, we use Fig. 3 to show the relationship between the maximum value of the similarity matrix of filtered images and the number of filtered images on our test image set with 61 images. In our experiments, \( \alpha \) is set as 0.5 empirically. From this figure, we can see the number of matched feature points is 863 (that is the similar-

![Figure 3. The relationship between the number of the filtered images and the similarity threshold value in our dataset with 61 images. Note: if the maximum value of the similarity matrix of the filtered image set is zero, that means each image in the filtered image set has no overlapping region with other images.](image-url)
In this paper, we use Eqs. (3) and (8) to align images. To evaluate the effectiveness, we conduct one comparison experiment (Fig. 4 and Fig. 5) with the single homography model, the single similarity model, the combination models the similarity model with the group model, and the proposed method. Fig. 4(a) and Fig. 5(a) show that the single homography model makes misalignments and accumulates severe perspective distortions. The single similarity model (Fig. 4(b) and Fig. 5(b)) resists perspective distortion and obtains low alignment accuracy since it is prone to accumulating alignment errors. The group similarity model, which is the first step of the proposed alignment method, can provide a rough alignment and resist perspective distortions by combining the advantage of the similarity model and the group model (see Fig. 4(c) and Fig. 5(c)). Although the group similarity model can just provide a rough alignment, there still exist slight misalignments. The following optimization step of our alignment method can well eliminate misalignments. From Fig. 4(d) and Fig. 5(d), we can see that the proposed method improves alignment and resists perspective distortions.

Fig. 6 shows the results of our image redundancy filtering method. Fig. 6(a) shows the topology estimation of the original image sets obtained by Xia’s method [37]; Fig. 6(b) shows the topology estimation of the filtered image set obtained by the proposed method. Obviously, after redundancy images filtering, the number of the image set which is used to produce the final stitching result decreases significantly, the distribution of the filtered images becomes sparse and uniform, but the member of the filtered image set cover the observing region. In Fig. 6, if the matched feature points of image pair are more than 100, we use the green line to connect, otherwise, we use the grey line to connect. We can see that the topology estimation of the filtered image sets has little or seldom green edges, which means the filtered image set can be connected and the overlapping rate of each two adjacent images is enough to estimate the transformation relation between them.

Fig. 8 shows the comparison results of PTGui, Xia’s method and the proposed method. As shown in the first two rows of Fig. 8, the PTGui and Xia’s method could achieve the global consistency (i.e., reduce global perspective distortions). However, the two methods could cause misalignments in local region (see the first row and second row of each example in Fig. 8). The proposed method considers the local alignment and the global consistency by the group similarity model and the homography model. Compared with these methods, the proposed method has higher accurate alignment than other methods in local alignment accuracy, while providing a good performance in the global consistency (see the third row of each example in Fig. 8).

To show the superiority of the proposed method in alignment accuracy, the Root-Mean-Square (RMS) errors is used to describe the alignment accuracy. The results are shown in Table 3. For panorama image stitching, the traditional approach is to select one of the images as the reference and then warp all of the other images into the reference coordinate system one by one (the method is named as the single model). The group model is that images first are grouped according to the stitching order, then minimize the alignment errors within the image group and between adjacent image groups. As shown in Table 3, we can see that the group similarity model can reduce error accumulation better than the single similarity model. The single homography model can improve alignment accuracy but tend to suffer distortion accumulation (shown in Fig. 5(a)). Based on the group similarity and the homography model, the proposed method employs the homography model to increase alignment accuracy by our energy function Eq. (8). As shown in Table 3, the proposed method produces smaller RMS than other methods. In addition, the remaining errors will be evenly distributed under the uniform energy term, which does not induce distortion.
FIGURE 5. The stitching results on the strip-shaped image set for different alignment models: (a) the result obtained by the single homography model; (b) the result obtained by the single similarity model; (c) the result obtained by the group similarity model; (d) the result obtained by the proposed method. Note: the reference image of each stitching result is marked with the green rectangular box.

TABLE 3. RMS of different models.

| Alignment method              | Strip-shaped image set | Square-shaped image set |
|-------------------------------|------------------------|-------------------------|
| Single similarity model       | 2.02059                | 13.7885                 |
| Group similarity model        | 1.95487                | 6.5675                  |
| Single homography model       | 1.81035                | 3.51015                 |
| The proposed method           | 0.987712               | 1.74463                 |

From these experimental results, we conclude that the proposed method increases alignment accuracy for stitching panorama images by the proposed group model and homography optimization.

The final stitching results of the proposed method are compared with those of PTGui and Xia’s method on visual effects, which aims to demonstrate that the proposed method does not cause information loss and does not destroy connection for panorama stitching. As shown in Fig. 8, we can see that our stitching results have a similar visual effect to the stitching results of the other two methods obtained with the original image set on the whole, but our approach generates the final stitching image with much less images. The number of the stitching images generated by PTGui and Xia’s method is separately 61, 56, 160 and 744, that of the proposed method is separately 34, 27, 86 and 375. Given images to be stitched, the desired final image should contain as much information as possible. We applied the information ratio comparison to detect whether our stitching results have more information loss than the other two methods. The information ratio is designed to describe how much useful information an image contains. It can be presented as follows:

\[
I_r = 1 - \frac{N_{\text{blackpixels}}}{N_{\text{Stitched}}} 
\]

\[
N_{\text{Stitched}} = W_{\text{Stitched}} \cdot H_{\text{Stitched}} 
\]

where \(I_r\) represents the information ratio, \(N_{\text{blackpixels}}\) represents the number of useless pixels (such as black pixels or accumulation. From these experimental results, we conclude that the proposed method increases alignment accuracy for stitching panorama images by the proposed group model and homography optimization.

The final stitching results of the proposed method are compared with those of PTGui and Xia’s method on visual effects, which aims to demonstrate that the proposed method does not cause information loss and does not destroy connection for panorama stitching. As shown in Fig. 8, we can see that our stitching results have a similar visual effect to the stitching results of the other two methods obtained with the original image set on the whole, but our approach generates the final stitching image with much less images. The number of the stitching images generated by PTGui and Xia’s method is separately 61, 56, 160 and 744, that of the proposed method is separately 34, 27, 86 and 375. Given images to be stitched, the desired final image should contain as much information as possible. We applied the information ratio comparison to detect whether our stitching results have more information loss than the other two methods. The information ratio is designed to describe how much useful information an image contains. It can be presented as follows:

\[
I_r = 1 - \frac{N_{\text{blackpixels}}}{N_{\text{Stitched}}} 
\]

\[
N_{\text{Stitched}} = W_{\text{Stitched}} \cdot H_{\text{Stitched}} 
\]

where \(I_r\) represents the information ratio, \(N_{\text{blackpixels}}\) represents the number of useless pixels (such as black pixels or
FIGURE 6. The topology comparison of the original image sets and the filtered image sets: (a) the topology of the original image set; (b) the topology of the filtered image set.

white pixels) in the stitching result. $W_{\text{Stitched}}$ and $H_{\text{Stitched}}$ represent the width and height of the stitching result, respectively. The results of the information ratio are presented in Table 4.

As shown in Table 4, the information ratio values of the proposed method are close to those of the other two methods. The results demonstrate that the proposed method can’t result in more information loss than the other
FIGURE 7. The stitching result of different datasets for different stitching methods: (a) the results obtained by PTGui; (b) the results obtained by Xia’s method; (c) the results obtained by the proposed method. Note: the number of the stitching images obtained by PTGui and Xia’s method is 61, 56, 160 and 744, separately. The number of the stitching images obtained by the proposed method is separately 34, 27, 86 and 375. The information ratio of the stitching results obtained by those three methods represents in Table 4.

In addition, a running time comparative experiment is constructed to demonstrate the efficiency of the proposed method. In this paper, we adopt parallel computing to optimize the proposed method. However, Xia’s method takes serial computing to optimize each stage of the stitching process. To verify the effectiveness of the proposed method reasonably, we presented the running time of Xia’s method, the proposed method with serial computing, and the
FIGURE 8. Comparison of the stitching results with PTGui, Xia’s method and the proposed method. From top to bottom in each example: the stitching results obtained by PTGui, Xia’s method and the proposed method.
TABLE 4. The information ratio on different datasets for different methods.

| Stitching method | First image set | Second image set | Third image set | Fourth image set |
|------------------|----------------|-----------------|----------------|-----------------|
| PTGui            | 0.7677         | 0.8596          | 0.7780         | 0.9656          |
| Xia’s method     | 0.7913         | 0.8488          | 0.7193         | 0.9114          |
| The proposed method | 0.7755       | 0.8471          | 0.7596         | 0.9062          |

TABLE 5. The running time on different datasets for different methods.

| Alignment method | First image set | Second image set | Third image set | Fourth image set |
|------------------|----------------|-----------------|----------------|-----------------|
| Xia’s method     | 5529s          | 35371s          | 1958s          | 399000s         |
| The proposed method with serial computing | 631s          | 2029s          | 341s          | 27079s         |
| The proposed method | 469s          | 1831s          | 260s          | 21681s         |

TABLE 6. Average evaluation value of the stitching results.

| Alignment method | First image set | Second image set | Third image set |
|------------------|----------------|-----------------|----------------|
| PTGui            | 3.64           | 3.59            | 3.29           |
| Xia’s method     | 4.52           | 4.43            | 4.11           |
| The proposed method | 4.71          | 4.67            | 4.32           |

IV. CONCLUSION

In this paper, a novel panorama stitching framework based on image redundancy filtering is proposed, which contains three key modules: image redundancy filtering, reference image selection and global alignment. To effectively filter redundancy images from original image set, an iterative method based on similarity analysis is proposed, which can significantly reduce data redundancy and do not cause information loss of the final stitching result. Based on the weight topology graph, we employ the shortest path algorithm to find the optimal reference image, which makes alignment error accumulation minimum. Finally, to obtain a good balance between the global consistency and the alignment accuracy, a two-step global alignment strategy is proposed, which employs the similarity model to robust align between groups and then perform shape optimization further by the homography model under the anti-perspective function. The experimental results, including the comparison of topology structure analysis, visual effect, running time, information loss ratio, qualitative evaluation and the alignment accuracy between different methods, show that the proposed method successfully reduces image redundancy, while improving the alignment and reducing shape distortion for massive images stitching.

REFERENCES

[1] X. Li, R. Feng, X. Guan, H. Shen, and L. Zhang, “Remote sensing image mosaicking: Achievements and challenges,” IEEE Geosci. Remote Sens. Mag., vol. 7, no. 4, pp. 8–22, Dec. 2019.
[2] W. Yan, C. Hou, J. Lei, Y. Fang, Z. Gu, and N. Ling, “Stereoscopic image stitching based on a hybrid warping model,” IEEE Trans. Circuits Syst. Video Technol., vol. 27, no. 9, pp. 1934–1946, Sep. 2017.
[3] W. Yan, G. Yue, J. Xu, Y. Yu, K. Wang, C. Tang, and X. Tong, “Shape-optimizing mesh warping method for stereoscopic panorama stitching,” Inf. Sci., vol. 511, pp. 58–73, Feb. 2020.
[4] T. Xiang, G.-S. Xia, L. Zhang, and N. Huang, “Locally warping-based image stitching by imposing line constraints,” in Proc. 23rd Int. Conf. Pattern Recognit. (ICPR), Dec. 2016, pp. 4178–4183.
[5] K.-Y. Lee and J.-Y. Sim, “Stitching for multi-view videos with large parallax based on adaptive pixel warping,” IEEE Access, vol. 6, pp. 36904–26917, 2018.
[6] E. Zagrouba, W. Barhoumi, and S. Amri, “An efficient image-mosaicing method based on multifeature matching,” Mach. Vis. Appl., vol. 20, no. 3, pp. 139–162, Apr. 2009.
[7] J. Chen, H. Feng, K. Pan, Z. Xu, and Q. Li, “An optimization method for registration and mosaicking of remote sensing images,” Optik, vol. 125, no. 2, pp. 697–703, Jan. 2014.
[8] B. Zitová and J. Flusser, “Image registration methods: A survey,” Image Vis. Comput., vol. 21, no. 11, pp. 977–1000, Oct. 2003.
X. Kang, L. Wu, Y. Wei, B. Yang, and H. Song, “A highly accurate dense approach for homography estimation using modified differential evolution,” Eng. Appl. Artif. Intell., vol. 31, pp. 68–77, May 2014.

Z. Wang, Y. Chen, Z. Zhu, and W. Zhao, “An automatic panoramic image mosaic method based on graph model,” Multimedia Tools Appl., vol. 75, no. 5, pp. 2725–2740, 2015.

N. R. Gracias, S. van der Zwaan, A. Bernardino, and J. Santos-Victor, “Mosaic-based navigation for autonomous underwater vehicles,” IEEE J. Ocean. Eng., vol. 28, no. 4, pp. 609–624, Oct. 2003.

Y. Xu, J. Ou, H. He, X. Zhang, and J. Mills, “Mosaicking of unmanned aerial vehicle imagery in the absence of camera poses,” Remote Sens., vol. 8, no. 3, p. 204, Mar. 2016.

Y. Chen, J. Sun, and G. Wang, “Minimizing geometric distance by iterative linear optimization,” in Proc. 20th Int. Conf. Pattern Recognit., Aug. 2010, pp. 1–4.

W. Mou, H. Wang, G. Seet, and L. Zhou, “Robust homography estimation based on non-linear least squares optimization,” in Proc. IEEE Int. Conf. Robot. Biomimetics (ROBIO), Dec. 2013, vol. 2014, no. 1, pp. 372–377.

L. Zhang, Y. Li, J. Zhang, and Y. Hu, “Homography estimation in omni-directional vision under the $L_\infty$-norm,” in Proc. IEEE Int. Conf. Robot. Biomimetics, Dec. 2010, pp. 1468–1473.

M. Brown and D. G. Lowe, “Automatic panoramic image stitching using invariant features,” Int. J. Comput. Vis., vol. 74, no. 1, pp. 59–73, Apr. 2007.

H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, “Speeded-up robust features (SURF),” Comput. Image. Vis. Understanding., vol. 110, no. 3, pp. 346–359, Jun. 2008.

D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” Int. J. Comput. Vis., vol. 60, no. 2, pp. 91–110, Nov. 2004.

K. Konolige, “Sparse sparse bundle adjustment,” in Proc. Brit. Mach. Vis. Conf., vol. 10, 2010, pp. 1–102.

B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon, “Bundle adjustment—A modern synthesis,” in Proc. Int. Workshop Vis. Algorithms, 1999, pp. 298–372.

M. Li, D. Li, and D. Fan, “A study on automatic UAV image mosaic method for paroxysmal disaster,” ISPRS–Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., vols. 6, pp. 123–128, Jul. 2012.

R. Feng, Q. Du, X. Li, and H. Shen, “Robust registration for remote sensing images by combining and localizing feature- and area-based methods,” ISPRS J. Photogramm. Remote Sens., vol. 151, pp. 15–26, May 2019.

C. Xing, J. Wang, and Y. Xu, “A robust method for mosaicising sequence images obtained from UAV,” in Proc. 2nd Int. Conf. Inf. Eng. Comput. Sci., Dec. 2010, pp. 1–4.

F. Caballero, L. Merino, J. Ferruz, and A. Ollero, “Homography based Kalman filter for mosaic Building. Applications to UAV position estima- tion,” in Proc. IEEE Int. Conf. Robot. Autom., Apr. 2007, pp. 2004–2009.

T. Kekec, A. Yildirim, and M. Unel, “A new approach to real-time mosaicing of aerial images,” Robot. Auto. Syst., vol. 52, pp. 1755–1767, Dec. 2014.

J. Gao, S. J. Kim, and M. S. Brown, “Constructing image panoramas using dual-homography warping,” in Proc. CVPR. Jun. 2011, pp. 49–56.

J. Zaragoza, T.-J. Chin, Q.-H. Tran, M. S. Brown, and D. Suter, “As-Projective-As Possible image stitching with moving DLT,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 36, no. 7, pp. 1285–1298, Jul. 2014.

C.-H. Chang, Y. Sato, and Y.-Y. Chuang, “Shape-preserving half-projective warps for image stitching,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Jun. 2014, pp. 3254–3261.

C.-C. Lin, S. U. Pankanti, K. N. Ramamurthy, and A. Y. Aravkin, “Adaptive as-natural-as-possible image stitching,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2015, pp. 1155–1163.

F. Zhang and F. Liu, “Parallax-tolerant image stitching,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2014, pp. 3262–3269.

J. Kang, J. Kim, I. Lee, and K. Kim, “Minimum error seam-based efficient panorama video stitching method robust to parallax,” IEEE Access, vol. 7, pp. 161127–161140, 2019.

X. Li, N. Hui, H. Shen, Y. Fu, and L. Zhang, “A robust mosaicking procedure for high spatial resolution remote sensing images,” ISPRS J. Photogramm. Remote Sens., vol. 109, pp. 108–125, Nov. 2015.

A. Levin, A. Zomet, S. Peleg, and Y. Weiss, “Seamless image stitching in the gradient domain,” in Proc. Eur. Conf. Comput. Vis., 2004, pp. 377–389.

J. Li, P. Jiang, S. Song, H. Xia, and M. Jiang, “As-Alignmed-As-Possible image stitching based on deviation-corrected warping with global similarity constraints,” IEEE Access, vol. 7, pp. 156605–156611, 2019.
MEIQI GU is currently pursuing the M.S. degree with the School of Computer and Control Engineering, Yantai University.

Her research interests include multiview data representation, image processing, and pattern recognition.

KAIQI SU is currently pursuing the M.S. degree with the School of Computer and Control Engineering, Yantai University.

His research interests include 3D object detection, image processing, and pattern recognition.

GUANGHUI YUE received the B.S. degree in communication engineering from Tianjin University, in 2014, and the Ph.D. degree in information and communication engineering from Tianjin University, Tianjin, China, in 2019. He was a Joint Ph.D. Student with the School of Computer Science and Engineering, Nanyang Technological University, Singapore, from September 2017 to January 2019. He is currently an Assistant Professor with the Health Science Center, School of Biomedical Engineering, Shenzhen University. His research interests include bioelectrical signal processing, multimedia quality assessment, 3D image visual discomfort prediction, pattern recognition, and medical image analysis.

YUN LIU received the M.S. and Ph.D. degrees in communication and information engineering from Tianjin University, China, in 2016. She was a Visiting Ph.D. Student with the University of California Berkeley, Berkeley, CA, USA. She is currently a Lecturer with the College of Information, Liaoning University. Her research interests include stereo vision research, pattern recognition, and image quality evaluation.

* * *