Automatic Quaternion-Domain Color Image Stitching

Jiaxue Li and Yicong Zhou, *Senior Member, IEEE*

***Abstract*— Taking advantages of the quaternion representation of the color image, this paper proposes a quaternion perceptual seamline detection model to generate the seamline in the quaternion domain. It considers seamline detection as a quaternion-domain color image labeling problem and minimizes the local-area quaternion perceptual difference cost to obtain the optimal seamline. To assess seamline quality effectively, we develop a quaternion perceptual seamline quality measure. Based on the proposed quaternion perceptual seamline detection model and quality measure, we further propose a general framework for automatic quaternion-domain color image stitching (AQCIS). To the best of our knowledge, this is the first attempt to perform color image stitching completely in the quaternion domain. Meanwhile, AQCIS introduces the joint optimization strategy of local alignment and seamline in an iterative fashion. Extensive experiments on challenging datasets demonstrate that our AQCIS achieves superior performance for color image stitching in comparison with state-of-the-art methods.**

***Index Terms*— Quaternion-domain color image stitching, quaternion representation, seamline detection, seamline quality measure.**

I. Introduction

COLOR image stitching is a fundamental yet challenging computer vision task, with the aim of capturing more regions of interest. Due to the limited imaging width of hardware devices, it benefits a wide variety of applications, such as smartphone cameras, stereoscopic media and geographic mapping [1], [2], [3]. However, generating high-quality stitched color images in real-world applications remains a challenge, since it often faces multiple difficulties. Recent advances have targeted stitching color images with large parallax [4], [5], [6] or low textures [7], [8]. Nevertheless, limited efforts are devoted to simultaneously addressing multiple real-world difficulties, including large parallax, moving objects, occlusions, low textures and repetitive textures.

Typically, color image stitching is a process comprised of several independent stages: color image alignment, composition, and various post-processing techniques like image rectangling [9] and color correlation [10]. In particular, seamline detection is the most widely used composition technique that is capable of concealing unavoidable misalignment artifacts, e.g., moving objects and occlusions. Hence, to generate high-quality stitching results, the development of color image stitching technologies primarily involves two key steps, i.e., color image alignment and seamline detection.

Conventionally, an accurate global alignment can be achieved by a single homography model when the real-world scenes are 3D planar or images are captured by a merely rotational camera. However, these assumptions are easily violated leading to parallax errors. To handle parallax, spatially-varying warping methods [11], [12], [13] and mesh-based warping methods [14], [15], [16] were proposed to improve the global alignment quality. However, they are grounded on an underlying assumption that the spatial relationship of matched features should be consistent in different input color images. This causes they cannot align images with large parallax accurately. Furthermore, when there are moving objects and occlusions, it is impractical to achieve a perfect global alignment over the whole overlapping area. Therefore, the following seamline detection plays a critical role in generating visually plausible stitching results.

To remove misalignments, an optimal seamline is supposed to be delineated at a local region where aligned images have minimal differences. Therefore, the difference map that characterizes alignment errors in the overlapping area is a pivotal factor for seamless composition. Traditional seam-cutting methods calculated the difference map generally in terms of colors, gradients and edges [17], [18], [19]. However, it was observed in [20] that seamlines calculated from such difference maps may not have the best perceptual stitching quality. To correlate well with human perception, recent studies tried to introduce the characteristics of human visual system to discriminate perceptual differences making the optimized seamline as invisible as possible [20], [21], [22]. They incorporated various feature maps like visual saliency [20], texture complexity [21] and scene depth [22], into the traditional difference map to highlight the importance of perceptual differences in specific local regions. Alternatively, some nonlinear functions such as Gaussian [16] and Sigmoid [20] were also utilized to mimic the human perception of color differences.

These advanced seamline detection methods exhibit perceptually good stitching performance. However, several issues remain to be solved. First, they treat color images as gray images or simply aggregate the color values stored in RGB

Manuscript received 22 April 2023; revised 13 October 2023 and 24 December 2023; accepted 24 January 2024. Date of publication 8 February 2024; date of current version 12 February 2024. This work was supported in part by the Science and Technology Development Fund, Macau, under Grant 0049/2022/A1; and in part by the University of Macau under Grant MYRG2022-00072-FST. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Claudio R. Jung. *(Corresponding author: Yicong Zhou.)*

The authors are with the Department of Computer and Information Science, University of Macau, Macau, China (e-mail: lijiaxue7@gmail.com; yicongzhou@um.edu.mo).

This article has supplementary downloadable material available at https://doi.org/10.1109/TIP.2024.3361688, provided by the authors.

Digital Object Identifier 10.1109/TIP.2024.3361688

channels as a scalar. As a result, the color information or the structural correlations among different color channels are completely ignored. Correspondingly, color differences cannot be explored sufficiently in the calculated difference map. Second, most methods optimize the seamline by calculating the difference cost of only its adjacent pixels. In the low-textured and repetitive-textured scenes, neighboring misaligned pixels probably have the same or approximate color values, causing their differences not to be distinguished. Third, the seamline is often optimized independently assuming that the image alignment step has been finished. However, the estimated global alignment does not necessarily guarantee the derived seamline also has the best perceptual stitching quality, even with the most precise geometric fitting [23].

In view of the dependence of seamlines on the local alignment, seam-driven image stitching methods have been developed [23], [24], [25], [26]. Specifically, they first estimated multiple local alignment proposals to generate different seamlines and then evaluated them based on a specific seamline quality metric. Afterward, the local alignment with the best seamline quality is selected for the final color image stitching. To improve the seamline quality, an iterative evaluation-re-estimation strategy was further adopted in [25] and [26] to refine local alignment proposals and the difference map, respectively. Compared with the methods following the traditional pipeline, seam-driven methods tend to search for a local alignment in the overlapping area to support the best perceptual seamline generation. Consequently, they are more flexible to obtain the desirable stitching results in challenging scenarios.

Nevertheless, it remains difficult to achieve the final optimal color image stitching performance. First, seamlines derived from local alignment proposals are still optimized independently, such that the selected best seamline cannot be yet ensured to be the optimal one for the final color image stitching. On the other hand, seamlines are preferred to pass through local regions with few textures. However, they estimate local alignment proposals using feature-based alignment methods. Thus, the selected local alignment may suffer from biased estimation due to insufficient features, degrading seamline stitching performance potentially. Finally, most seam-driven methods utilize traditional seamline detection and evaluation methods. This also limits their effectiveness in color image stitching.

Recently, with advantages of the quaternion representation in color image processing [27], [28], [29], a pixel-based alignment model was proposed to align color images in the quaternion domain [30]. Instead of processing the color information as simple scalars, they represented each color pixel as a unit quaternion number to encode the color image into a pure quaternion matrix. When estimating the alignment model, they can not only simultaneously process RGB color values, but also fully explore the correlated structure information among different color channels. Compared with feature-based alignment methods, they obtain superior alignment accuracy, especially in low-textured scenarios.

Motivated by the above issues, we propose an automatic quaternion-domain color image stitching method to generate high-quality stitched color images despite facing multiple real-world difficulties. In color image stitching, both color image alignment and seamline detection share the same goal, i.e., to stitch well-aligned color images seamlessly. For this implementation, we argue that they play complementary roles in advancing the final color image stitching performance, and thus they should be jointly optimized in a unified framework towards the same optimal stitching goal. Meanwhile, to fully utilize the color information of color images for this joint optimization, we perform the whole color image stitching process in the quaternion domain. Specifically, the quaternion rank-1 alignment (QR1A) model [30] is adopted to accurately estimate the local alignment. The contributions of our work are summarized as follows:

1) We propose a quaternion perceptual seamline detection (QPSD) model to generate the seamline in the quaternion domain. Using the quaternion representation of color images, we formulate QPSD as a quaternion-domain color image labeling problem and minimize the local-area quaternion perceptual difference cost to obtain the optimal seamline.
2) To quantitatively evaluate the seamline quality in the quaternion domain, we propose a quaternion perceptual seamline quality (QPSQ) measure.
3) Using QR1A, QPSD, and QPSQ, we develop a general framework for automatic quaternion-domain color image stitching (AQCIS). *To the best of our knowledge, this is the first work to perform color image stitching completely in the quaternion domain.* Meanwhile, AQCIS introduces the joint optimization strategy to simultaneously learn the optimal local alignment and optimal seamline.
4) We validate the effectiveness of QPSD and AQCIS by ablation studies and comparative experiments. Qualitative and quantitative results demonstrate that AQCIS outperforms the state-of-the-art methods in color image stitching, and it has the capability of simultaneously addressing multiple real-world difficulties, including large parallax, moving objects, occlusions, low textures, and repetitive textures.

The remainder of the paper is organized as follows: The QPSD model and the QPSQ measure are proposed in Section II and Section III, respectively. The AQCIS framework is developed in Section IV. Section V shows experimental settings, comparisons, and results. A discussion is given in Section VI. Finally, we conclude this paper in Section VII.

II. Quaternion Perceptual Seamline Detection

This section first presents the quaternion representation of the color image, then proposes a quaternion perceptual seamline detection model in detail.

A. Quaternion Representation of Color Image

We represent the color image in the quaternion domain, aiming to preserve the whole structural color information from both spatial and spectral dimensions. The quaternions were first proposed by Hamilton [31], describing a four-dimensional

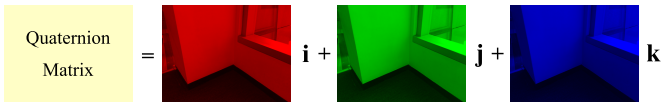


Fig. 1. Illustration for quaternion representation of the color image. The three channels of the color image are encoded into the three imaginary parts of a 2D quaternion matrix.

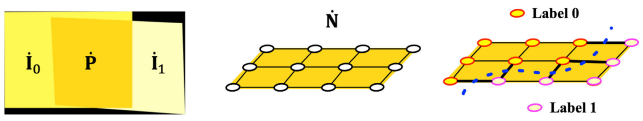


Fig. 2. Aligned color images and relevant notations of the overlapping area. The neighboring system $\dot{\mathbf{N}}$ contains the set of adjacent color pixels in the overlapping area. The seamline is formed naturally when assigning different labels to adjacent color pixels, and the edges in bold denote the corresponding energy cost, i.e., $E_{smooth}(l_{\dot{p}}, l_{\dot{q}})$.

hyper-complex space $\mathbb{H}$. A quaternion number $\dot{q} \in \mathbb{H}$ has the canonical basis $\{1, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$ and is typically written in the form of:

$$\dot{q} = a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}, \tag{1}$$

where $a_0, a_1, a_2, a_3 \in \mathbb{R}$ are quaternion components. Particularly, when $a_0 = 0$, the quaternion $\dot{q}$ becomes purely imaginary and it is referred to as the pure quaternion number. $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are imaginary units satisfying: $\mathbf{ij} = -\mathbf{ji} = \mathbf{k}, \mathbf{jk} = -\mathbf{kj} = \mathbf{i}, \mathbf{ki} = -\mathbf{ik} = \mathbf{j}, \mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1$. Following this rule, quaternion multiplication is defined as:

$$\begin{aligned}\dot{q}_1 \otimes \dot{q}_2 = {} & (a_0b_0 - a_1b_1 - a_2b_2 - a_3b_3) \\ & + (a_1b_0 + a_0b_1 - a_3b_2 + a_2b_3)\mathbf{i} \\ & + (a_2b_0 + a_3b_1 + a_0b_2 - a_1b_3)\mathbf{j} \\ & + (a_3b_0 - a_2b_1 + a_1b_2 + a_0b_3)\mathbf{k},\end{aligned} \tag{2}$$

where $\otimes$ is the Hamilton product denoting the operator of quaternion multiplication. Obviously, quaternion multiplication is usually non-commutative, i.e., $\dot{q}_1 \otimes \dot{q}_2 \neq \dot{q}_2 \otimes \dot{q}_1$. Quaternion addition follows the component-wise scheme as that in the complex space and is defined as:

$$\dot{q}_1 + \dot{q}_2 = a_0 + b_0 + (a_1 + b_1)\mathbf{i} + (a_2 + b_2)\mathbf{j} + (a_3 + b_3)\mathbf{k}. \tag{3}$$

Two commonly used quaternion operators are:

1) Quaternion conjugate: $\overline{\dot{q}} = a_0 - a_1\mathbf{i} - a_2\mathbf{j} - a_3\mathbf{k}$;
2) Quaternion modulus: $|\dot{q}| = \sqrt{\dot{q} \otimes \overline{\dot{q}}} = \sqrt{\overline{\dot{q}} \otimes \dot{q}} = \sqrt{a_0^2 + a_1^2 + a_2^2 + a_3^2}$.

The color image has RGB channels to show its color information. In RGB color space, the three primary colors (i.e., red, green, and blue) are orthogonal bases. Thus, there is an underlying orthogonal constraint on the correlations among RGB channels. For color image representation, we encode the color image into a pure quaternion matrix, which preserves not only RGB data but also the orthogonal property. Mathematically, the quaternion representation of the color image with the size of $M \times N$ is defined as follows:

$$\dot{\mathbf{I}}(x, y) = \mathbf{iR}(x, y) + \mathbf{jG}(x, y) + \mathbf{kB}(x, y), \tag{4}$$

where $\dot{\mathbf{I}}(x, y)$ denotes a color pixel located at (x, y) in the two-dimensional image coordinate system, $\mathbf{R}(x, y)$, $\mathbf{G}(x, y)$ and $\mathbf{B}(x, y)$ are its red, green and blue components, respectively.

As shown in Fig. 1, $\dot{\mathbf{I}} \in \mathbb{H}^{M \times N}$ is a pure quaternion matrix having the same size with the color image, and a one-to-one mapping between the RGB color space and the quaternion space is established. Using the quaternion representation of the color image, we are capable of working on the whole color space to simultaneously process the color information stored in RGB channels. Moreover, the highly correlated information across different color channels can be preserved and explored sufficiently. Specifically, let $\dot{\mathbf{I}}_1$ and $\dot{\mathbf{I}}_2$ stand for two color images with the same size of $M \times N$, then their cross-channel correlation is calculated as [32]:

$$C(x, y) = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \dot{\mathbf{I}}_1(i, j) \otimes \overline{\dot{\mathbf{I}}_2(i - x, j - y)}, \tag{5}$$

where (i, j) is the row and column index of the quaternion matrix.

B. Quaternion Perceptual Seamline Detection Model

Taking advantage of the quaternion representation of the color image, we propose a quaternion perceptual seamline detection (QPSD) model to generate high-quality seamlines.

Notation: We encode the color image as a 2D pure quaternion matrix. In the quaternion domain, we use $\otimes$ to perform quaternion multiplication and $|\cdot|$ to calculate quaternion modulus. When working on quaternion matrices, we perform the operator on each quaternion element, respectively; thus, the output of a quaternion matrix is still a matrix.

1) Quaternion-Domain Color Image Labeling: As shown in Fig. 2, let $\dot{\mathbf{I}}_0$ and $\dot{\mathbf{I}}_1$ be a pair of aligned color images, $\dot{\mathbf{P}}$ be the set of color pixels to be composed in the overlapping area, and $\dot{\mathbf{N}}$ be the set of pixel pairs in the neighboring system. Given a label set $\mathcal{L} = \{0, 1\}$, our goal is to assign each quaternion element $\dot{p} \in \dot{\mathbf{P}}$ a label $l_{\dot{p}} \in \mathcal{L}$, such that the following energy function is minimized:

$$E = \sum_{\dot{p} \in \dot{\mathbf{P}}} E_{data}(l_{\dot{p}}) + \lambda \sum_{(\dot{p}, \dot{q}) \in \dot{\mathbf{N}}} E_{smooth}(l_{\dot{p}}, l_{\dot{q}}), \tag{6}$$

where $E_{data}(l_{\dot{p}})$ is the data term measuring the energy cost of assigning the label $l_{\dot{p}}$ to the color pixel $\dot{p}$. For color pixels $\dot{p} \in \dot{\mathbf{P}}$, the data term is typically set to 0 to guarantee the seamline to be located within the overlapping area. $E_{smooth}(l_{\dot{p}}, l_{\dot{q}})$ is the smoothness term to penalize the spatial incoherence between two adjacent color pixels $(\dot{p}, \dot{q})$ with different labels $(l_{\dot{p}}, l_{\dot{q}})$. By assigning labels to color pixels, we determine the values of each color pixel should be selected from which aligned color image for the final composite color image. For example, if $l_{\dot{p}=0}$, the value of $\dot{p}$ is selected from $\dot{\mathbf{I}}_0$; otherwise, it comes from $\dot{\mathbf{I}}_1$. When different labels are assigned to adjacent color pixels, the seamline will be formed correspondingly. λ is a constant parameter balancing the relative influence between two terms.

To achieve the optimal seamline stitching performance, Eq. (6) is specifically defined to describe differences between aligned (overlayed) color images, and the color difference is the most frequently characterized. The smoothness term is calculated to account for the total differences along the seamline, as shown in Fig. 2. Therefore, by minimizing Eq. (6), we can optimize a seamline with minimal difference cost, such that aligned color images are able to be stitched seamlessly.

2) Quaternion Perceptual Difference Map: The difference map usually plays a critical role in producing seamlines. To generate an optimal seamline in consistent with the human vision system, we define a quaternion perceptual difference map that takes the highly correlated color information and characteristics of the human visual system into account in the quaternion domain.

Specifically, the color information of a natural image is distributed non-uniformly among RGB channels, as presented in Fig. 1. Meanwhile, there are high correlations among the distributed color information of different color channels. This characteristic underlies the color constancy capability of the human visual system [33]. On the other hand, visual saliency could give rise to changes of human attention to the integrated color image information. The difference occurring in the salient areas could easily attract more human attention. With these considerations, we incorporate the quaternion representation into visual saliency, such that the informative areas of a color image can be fully represented in the quaternion domain. We treat this quaternion visual saliency map as a quaternion feature map to highlight the significance of local regions in the quaternion difference map. Mathematically, they are formulated as follows:

$$\dot{\mathbf{D}} = \dot{\mathbf{W}} \otimes (\dot{\mathbf{I}}_0 - \dot{\mathbf{I}}_1), \tag{7}$$

where $\dot{\mathbf{W}}$ is the average quaternion visual saliency map and $\otimes$ denotes the Hamilton product performing the element-wise quaternion multiplication. Following Eq. (2), we can capture the interactions of three imaginary components between $\dot{\mathbf{W}}$ and $\dot{\mathbf{I}}_0 - \dot{\mathbf{I}}_1$. Accordingly, the correlations among different color channels of their preserved color information can be also explored in the quaternion domain. Therefore, it suggests that the color information, including cross-channel correlations, is fully utilized by Eq. (7) to characterize the salient color difference sufficiently. To construct $\dot{\mathbf{W}}$, we first calculate the average visual saliency map [34] of each color channel, then utilize the quaternion representation to encode them as a whole. In this way, both color correlations and visual saliency are preserved and fully explored in the quaternion domain.

Next, we measure the perceptual magnitude of each quaternion element in $\dot{\mathbf{D}}$ to discriminate perceptual color differences. Intuitively, the quaternion modulus operator is adopted to calculate the quaternionic intensity, standing for the geometric distance from the origin in the quaternion coordinate system. It is expected to have a linear relationship with the magnitude of color differences. However, it is inconsistent with human perception. To solve this problem, two characteristics of the human vision system are additionally accounted to the quaternion norm. First, the human eyes cannot discriminate the difference with its magnitude under a certain threshold. Second, the perception process is non-linear. Normally, the response of human eyes is most sensitive to middle gray-level optical signals, but it becomes slow at low and high gray-level stages. Mathematically, we base these two perceptual characteristics on a Sigmoid function defined as follows:

$$f_{sig}(x) = \left[1 + \zeta \exp\left(-\beta(x-\alpha)\right)\right]^{-1}, \tag{8}$$

where α, β and ζ are non-negative constants. In particular, α denotes the threshold at which the discrimination of color differences begins, and β is the maximum curvature of sensitivity to the perceptible middle gray-level differences. The perceptible magnitude is adjusted by ζ, and the perceptual range of this nonlinear discrimination is $[(1+\zeta)^{-1}, 1)$.

Combining Eqs. (7) and (8), we build up the quaternion perceptual difference map as follows:

$$\mathbf{D}_{percept} = f_{sig}\left(\left|\dot{\mathbf{W}} \otimes (\dot{\mathbf{I}}_0 - \dot{\mathbf{I}}_1)\right|\right), \tag{9}$$

3) Quaternion Perceptual Energy Function: Based on the quaternion perceptual difference map, we define the data term and smoothness term in Eq. (6), respectively.

a) Data term: To maintain the continuity of boundary pixels between the overlapping and non-overlapping areas, the seamline is supposed to be constrained within the overlapping area. Thus, we define the data term as follows:

$$E_{data}(l_{\dot{p}}) = \begin{cases} \eta, & \text{if } \dot{p} \in \partial\dot{\mathbf{I}}_{l_{\dot{p}}} \cap \partial\dot{\mathbf{P}} \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

where $\partial\dot{\mathbf{I}}_{l_{\dot{p}}} \cap \partial\dot{\mathbf{P}}$ denotes the common border of $\dot{\mathbf{I}}_{l_{\dot{p}}}$ and $\dot{\mathbf{P}}$, and η is a large constant for penalization. Obviously, Eq. (10) prevents the seamline from cutting through the borders of the overlapping area and also fixes the endpoints of a seamline at the intersections of two partially overlapped images.

b) Smoothness term: In some local regions, e.g., with low textures and repetitive textures, misaligned pixels probably have the same or approximate color values, such that the color differences may not be distinguished sufficiently in $\mathbf{D}_{percept}$. To circumvent this problem, we propose to optimize the seamline by minimizing its local-area quaternion perceptual difference cost. To this end, we first define a local set of quaternion elements centered at $(\dot{p}, \dot{q})$:

$$\dot{\mathbf{O}}(\dot{p}, \dot{q}, t) = \{\dot{s} \,|\, \min(dis(\dot{p}, \dot{s}), dis(\dot{q}, \dot{s})) \le t\}, \tag{11}$$

where $dis(\cdot)$ is an operator calculating the Euclidean distance between two color pixels in the image coordinate system and t is a constant parameter to adjust the size of $\dot{\mathbf{O}}$. Then, we measure the spatial incoherence of $\dot{\mathbf{O}}$ to penalize two adjacent color pixels $(\dot{p}, \dot{q})$ with different labels $(l_{\dot{p}}, l_{\dot{q}})$. Consequently, the smoothness term is defined as follows:

$$E_{smooth}(l_{\dot{p}}, l_{\dot{q}}) = |l_{\dot{p}} - l_{\dot{q}}| \sum\nolimits_{\dot{s} \in \dot{\mathbf{O}}(\dot{p}, \dot{q}, t)} \mathbf{D}_{percept}(\dot{s}). \tag{12}$$

Here, the quaternion perceptual differences of quaternion elements within $\dot{\mathbf{O}}$ are aggregated together to denote the local-area quaternion perceptual difference cost between $(\dot{p}, \dot{q})$. Using $\dot{\mathbf{O}}$ and $\mathbf{D}_{percept}$, we are able to measure the perceptual difference cost from both spatial and spectral dimensions. Particularly, when $t = 1$, the quaternion local set

$\dot{\mathbf{O}}$ contains only $\dot{p}$ and $\dot{q}$, which coincides with the general seam-cutting approaches.

c) Labeling optimization: The final quaternion-domain color image labeling problem is solved by minimizing Eq. (6) that is defined based on Eq. (10) and Eq. (12). Notably, after using the quaternion modulus operator in $\mathbf{D}_{percept}$ to calculate quaternionic intensities, Eq. (6) becomes a real function despite in quaternion variables. Therefore, we can minimize Eq. (6) using the α-expansion move algorithm, and the detailed optimization process can be referred to the literature [35].

III. Quaternion Perceptual Seamline Quality

In the seam-driven image stitching methods, quality evaluation for seamline proposals is an important step, since it directly decides the choice of final color image stitching performance. To generate a visually good stitching result, the seamline quality should also be assessed in consistent with human subjective evaluation. Actually, seamline detection and evaluation are intrinsically related, because they have the same ideal conditions and both depend on the human visual system. Thus, given a pair of aligned color images and a detected seamline, we first compute the quaternion perceptual difference map as a local quality map. It reflects the distribution and magnitude of perceptual alignment errors in the overlapping area. Then, we adopt an average pooling strategy to quantitatively score the quaternion perceptual seamline quality (QPSQ).

Mathematically, let N be the size of the seamline l. We extract a local patch centered at each color pixel along the seamline, and M is the total number of color pixels in each patch. Given a color pixel $\dot{x}_{i,j}$ located at the j^{th} position in the i^{th} patch, we define the QPSQ measure as follows:

$$\text{QPSQ}(l) = \frac{1}{MN}\sum_{i=1}^{N}\sum_{j=1}^{M}\mathbf{D}_{percept}(\dot{x}_{i,j}). \tag{13}$$

Clearly, Eq. (13) is a real function in quaternion variables. The QPSQ score has been normalized to the range of [0, 1]. Generally, the lower score indicates the better seamline quality and better color image stitching performance in the human visual system.

IV. Automatic Quaternion-Domain Color Image Stitching

Using the proposed QPSD model and QPSQ measure, an automatic quaternion-domain color image stitching (AQCIS) framework is developed for high-quality color image stitching. This section first introduces the quaternion rank-1 alignment model and two color image stitching strategies, then describes the AQCIS framework specifically.

A. Quaternion Rank-1 Alignment Model

To fully utilize color image information for alignment, the quaternion rank-1 alignment (QR1A) model was introduced to the framework of AQCIS for estimating accurate local alignment. Specifically, given two partially overlapped color images, their overlapping portions correspond to the same real-world scenes and thus are linearly correlated in theory. Let $\{\dot{\mathbf{Q}}_i \circ \tau\}_{i=1}^{2}$ be the quaternion representations of well-aligned color image overlaps, where $\circ$ denotes image warping operation and $\tau = [\tau_1, \tau_2]$ is the set of corresponding homography transformations. We convert each quaternion matrix into a vector to stack them into a new quaternion matrix $\dot{\mathbf{A}} \circ \tau = [vec(\dot{\mathbf{Q}}_1 \circ \tau_1), vec(\dot{\mathbf{Q}}_2 \circ \tau_2)]$. Ideally, the same image contents could be captured among different well-aligned color images. That means the columns of $\dot{\mathbf{A}} \circ \tau$ are identical, and thus the quaternion rank of $\dot{\mathbf{A}} \circ \tau$ equals to 1.

However, due to various real-world factors such as moving objects and occlusions, there are often unavoidable differences among the columns of $\dot{\mathbf{A}} \circ \tau$. Directly imposing the quaternion rank-1 constraint is too strict to guarantee in practice. To address this issue, we further decompose $\dot{\mathbf{A}} \circ \tau$ into a quaternion rank-1 matrix $\dot{\mathbf{L}}$ and a quaternion sparse matrix $\dot{\mathbf{S}}$. Practically, $\dot{\mathbf{L}}$ stands for the same image scenes, while $\dot{\mathbf{S}}$ accounts for differences among different color images including moving objects and occlusions. Therefore, the QR1A model is formulated as follows:

$$\begin{aligned} \min_{\dot{\mathbf{L}},\dot{\mathbf{S}},\tau} \quad & \|\dot{\mathbf{S}}\|_1 \\ s.t. \quad & \dot{\mathbf{A}} \circ \tau = \dot{\mathbf{L}} + \dot{\mathbf{S}}, \quad rank(\dot{\mathbf{L}}) = 1. \end{aligned} \tag{14}$$

Benefited from the quaternion representation of color images, QR1A achieves superior alignment accuracy, especially in low-textured scenarios. Considering the plausible seamline is usually encouraged to cut through the low-textured areas, however, the common feature-based alignment methods perform poorly due to a lack of features in such local regions. Using QR1A, we can estimate an accurate local alignment in pixel level for seamline detection, circumventing this problem effectively. Moreover, due to the separation of $\dot{\mathbf{S}}$ component, QR1A is robust to aligning color images that involve moving objects and occlusions.

B. Two Color Image Stitching Strategies

1) Full Quaternion-Domain Color Image Stitching: Using quaternion representation of color images, we are capable of processing the color information stored in RGB channels simultaneously and fully exploring their cross-channel correlations for the whole color image stitching process. Both AQCIS and ACIS-QR1A [30] belong to the seam-driven image stitching category. In this category, three fundamental elements influence the final stitching performance: (1) local alignment; (2) seamline; and (3) seamline evaluation metric. In ACIS-QR1A, color images are aligned in the quaternion domain, whereas the seamline detection and evaluation are still operated in the real domain. This causes a deviation of these separated optimization steps and does not guarantee an optimal stitching result. It is thus desirable to develop a unified optimization framework in the full-quaternion domain for better results. *In this regard, AQCIS extends ACIS-QR1A to the full-quaternion domain to improve the optimized image stitching results.*

2) Joint Optimization of Local Alignment and Seamline: Different from the traditional seam-driven methods,

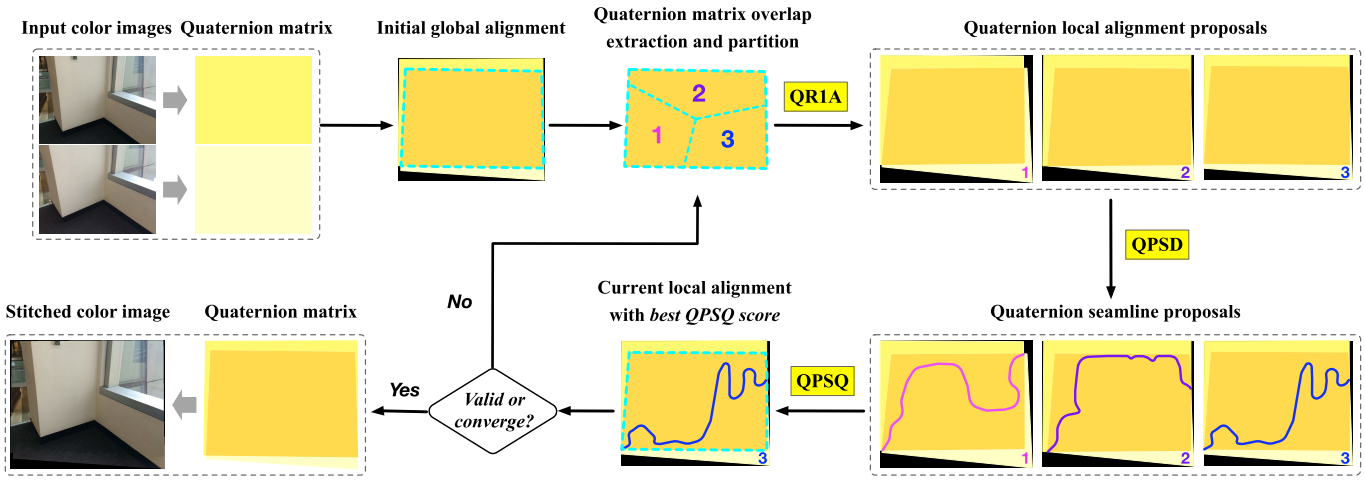


Fig. 3. The proposed AQCIS framework. A one-to-one mapping between the RGB color space and the quaternion space is first established. In the quaternion domain, the color information could be fully utilized to jointly learn the local alignment and seamline. The final color image stitching result is achieved automatically. See Section IV for details. The iterative details of the AQCIS framework are visualized in an additional material.

AQCIS adopts the joint optimization strategy to simultaneously learn the local alignment and seamline for the final color image stitching. In the process of joint optimization, the intermediately learned local alignment and seamline play complementary roles in advancing the color image stitching performance. Meanwhile, the intermediately learned color image stitching performance guides the subsequent optimization direction of local alignment and seamline. In each iteration, the current learned local alignment, seamline, and color image stitching performance only hold locally. That means, in the AQCIS framework, the quaternion computations of QR1A, QPSD, and QPSQ are iteratively reused for each other until the results achieve the final optimal or convergence conditions. This also enables the final stitching result to be obtained automatically.

C. AQCIS Framework

As illustrated in Fig. 3, input color images are first encoded into their quaternion representations and roughly aligned with a single homography for initial global alignment [36]. Then, we begin the iterative learning process for the joint optimization of local alignment and seamline.

In each iteration, we adaptively extract the current overlapped quaternion matrices and segment the global overlapping area into several local regions using the technique [37]. The number of partitions is based on the approximated number of sub-planes that compose the image scene in practice. We assume that each sub-plane predominates a local region. Empirically, we set the number of local regions at the range of $[3, 5]$ in this paper. In each local region, we learn the QR1A model to generate a quaternion local alignment proposal. On the basis of each quaternion local alignment proposal, we proceed to learn the QPSD model to obtain the corresponding quaternion seamline proposal. Subsequently, we utilize the QPSQ measure to evaluate the learned seamline qualities. The quaternion local alignment proposal with currently the best QPSQ score is selected as the underlying initial alignment for the next iteration. With this iterative learning process, we jointly optimize the QR1A and QPSD models under the guidance of the QPSQ metric. Accordingly, the selected best local alignment and seamline are iteratively updated in the quaternion domain towards the final optimal color image stitching performance.

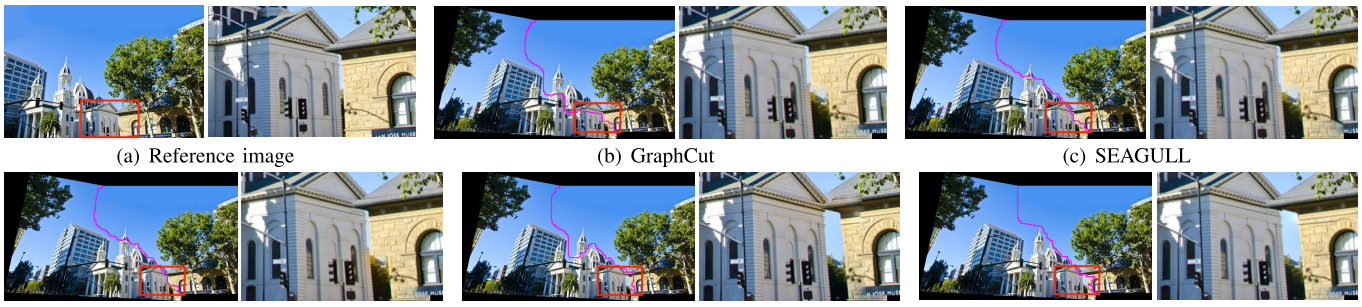
We record the best QPSQ score in each iteration. The iterative learning process is stopped when the relative difference of the QPSQ scores between two consecutive iterations is lower than a convergence threshold. Alternatively, it is considered to be optimum if the QPSQ score of currently learned the best seamline is lower than a valid threshold. Afterward, we recover the final stitched color image from its optimized quaternion representation.

V. Experiments

A. Experimental Settings

1) Parameters Setting: In the QPSD model, the parameters λ and ζ are both set to 1 by default. The threshold α is automatically selected using the technique [38] where the number of gray levels and the interval of the histogram are set to 100 and 0.01, respectively. Meanwhile, the threshold β is normally given by $\beta = \frac{4}{\kappa}$ where κ denotes the above-mentioned interval of histogram [20]. The parameter t of the quaternion local set equals 2 which shall be analyzed in the following ablation studies. In the AQCIS framework, we use the default parameters of the QR1A model [30] directly. For the iterative learning process, we set the convergence threshold and the valid threshold to 10^{-3} and 10^{-2}, respectively. The procedure is generally stopped within 10 iterations.

2) Datasets and Comparison Methods: We conduct experiments on public challenging datasets collected from [14], [24], and [25]. To save space, all the datasets and corresponding stitching results of different methods are provided in the supplementary material. The comparative methods include two widely used seamline detection algorithms in Graph-Cut [17] and SEAGULL [25], two recent seamline detection methods, i.e, PBSC [20] and QEIS [26], and two state-of-the-art overall color image stitching methods, i.e., LPC [5]



(a) Reference image (b) GraphCut (c) SEAGULL

(d) PBSC (e) QEIS (f) QPSD

Fig. 4. Comparison of seamline detention performance on the *PT-013* dataset [24]. (a) Reference image. Stitching seamlines of (b) GraphCut [17], (c) SEAGULL [25], (d) PBSC [20], (e) QEIS [26], and (f) QPSD. Stitching results of highlighted regions are magnified at the right side.

and ACIS-QR1A [30]. For a fair comparison, all seamline detection methods are performed on the same initial alignment, and the stitching order of all methods remains consistent on each dataset. Additionally, we adopt the Poisson blending technique [39] for all methods to improve their final aesthetic values.

3) Quantitative Evaluation Criteria: Despite the proposed QPSQ measure, we introduce four classical metrics, i.e., RMSE, SSIM [40], FSIMc [41], and ZNCC score [25], to assess seamline quality and color image stitching performance of different methods objectively. They are defined as follows:

$$\mathrm{RMSE}(l) = \frac{1}{MN}\sum_{i=1}^{N}\sum_{j=1}^{M}\|\mathbf{I}_0(x_{i,j}) - \mathbf{I}_1(x_{i,j})\|_{l_2}, \quad (15)$$

$$\mathrm{SSIM}(l) = \frac{1}{N}\sum_{i=1}^{N}\mathrm{SSIM}\left(\mathbf{I}_0(P_i), \mathbf{I}_1(P_i)\right), \quad (16)$$

$$\mathrm{FSIMc}(l) = \frac{1}{N}\sum_{i=1}^{N}\mathrm{FSIMc}\left(\mathbf{I}_0(P_i), \mathbf{I}_1(P_i)\right), \quad (17)$$

$$\mathrm{ZNCC\ score}(l) = \frac{1}{N}\sum_{i=1}^{N}\left(1 - \frac{\mathrm{ZNCC}\left(\mathbf{I}_0(P_i), \mathbf{I}_1(P_i)\right) + 1}{2}\right), \quad (18)$$

where N is the number of pixels along the seamline l, P_i denotes the local patch centered at the i^{th} pixel along the seamline, $\mathbf{I}_0(P_i)$ and $\mathbf{I}_1(P_i)$ are image patches extracted from the image $\mathbf{I}_0$ and $\mathbf{I}_1$, respectively. $x_{i,j}$ represents the j^{th} pixel in P_i, $\mathbf{I}_0(x_{i,j})$ and $\mathbf{I}_1(x_{i,j})$ are the corresponding values from image $\mathbf{I}_0$ and $\mathbf{I}_1$, respectively. Compared with QPSQ, these four metrics are all calculated in the real domain. In this paper, we uniformly set the patch size to 15×15 pixels, and hence $M = 255$. The reference image of the evaluation is consistent with that of the stitching order.

B. Performance of Seamline Detection

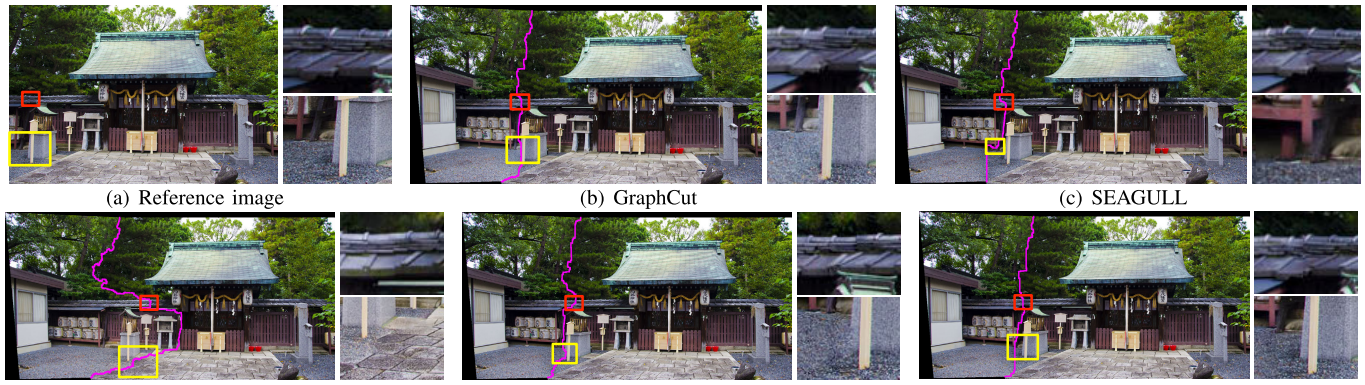
To verify the effectiveness of QPSD, we compare it with four peer seam-cutting methods, i.e., GraphCut [17], SEAGULL [25], PBSC [20], and QEIS [26]. The same initial alignment is given by a single homography. Two examples are displayed in Figs. 4 and 5. In the *PT-013* dataset [24], salient structures like buildings and lamps are misaligned severely due to large parallax. We use red rectangles to highlight specific local regions that seamlines pass through, and the corresponding stitching results are magnified at the right side. As can be seen, alignment errors are concealed by these seam-cutting methods to different extents. However, most seamlines cut through the right misaligned walls because they fail to characterize the differences of this local area sufficiently, resulting in obvious structural artifacts. Compared with them, QPSD generates a reasonable seamline achieving better stitching quality.

Another example is the *PT-053* dataset [24]. The repetitive textures like tiles and flagstones are misaligned severely due to large parallax. Specific local areas are highlighted in red and yellow rectangles and also magnified at the right side. These comparative algorithms optimize the seamline considering the alignment accuracy of only its two adjacent pixels. As shown in Fig. 5, their optimized seamlines are prone to cut through those misaligned areas with the same or similar textures since they cannot discriminate color differences effectively. In contrast, QPSD considers the local-area alignment accuracy along the seamline, and thus it is able to avoid generating such repetitive-textured artifacts to a certain extent.

C. Ablation Studies on QPSD

1) Quaternion Representation of Color Image: Characterizing differences between aligned color images is critical to optimizing a seamline, and the color difference is the most widely used. Therefore, an efficient color representation of color images is important to the sufficient exploration of color differences and is also beneficial to optimizing a seamless seamline. To verify the effectiveness of the quaternion representation, we substitute it with three common representation approaches, i.e., (1) just on gray images; (2) taking the mean values on color channels; and (3) considering color channels independently and then fusing the data, while keeping other parts unchanged in QPSD. We refer to the first two approaches as the gray representation and monochromatic model, while the third one is termed the concatenation model.

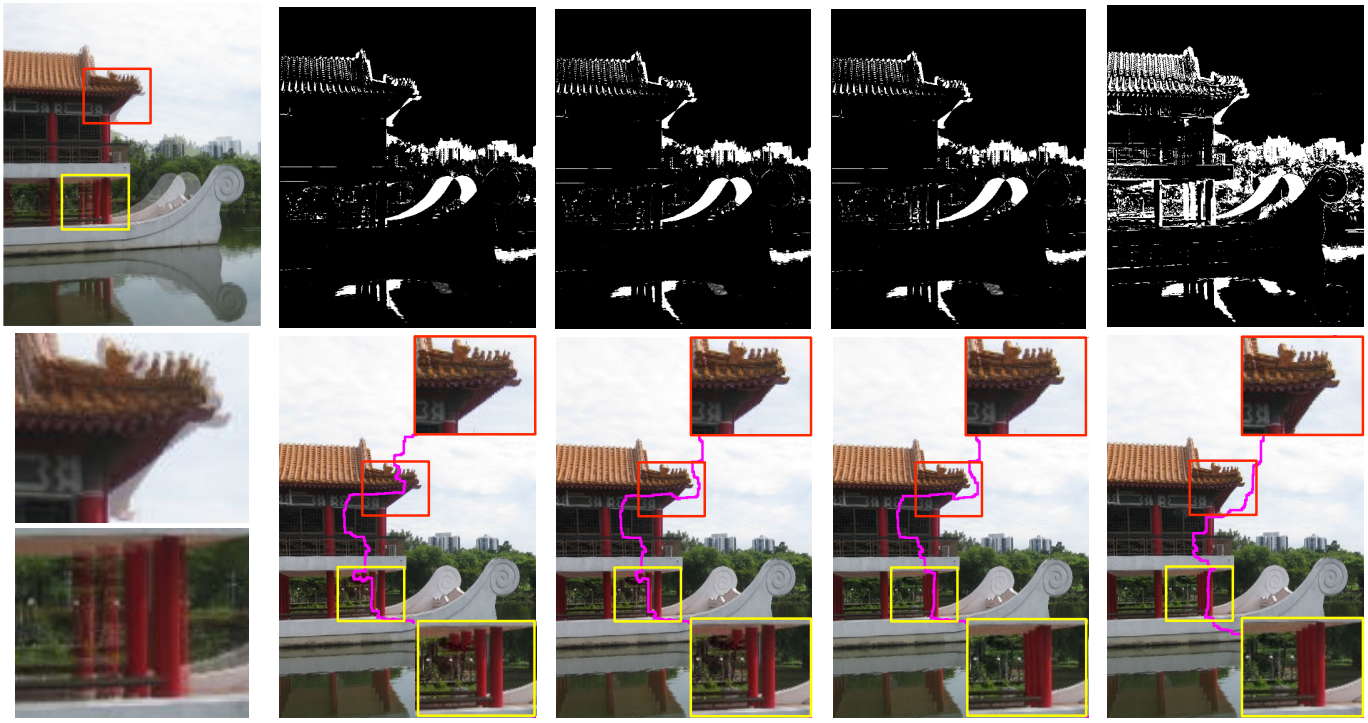
Fig. 6 illustrates an example on the *SEG-09* dataset [25]. Given the same image alignment, we use different representation approaches to calculate perceptual differences between aligned color images. The derived perceptual difference map can be visualized as a gray image to present the distribution and magnitude of perceived misalignment. Generally,



(a) Reference image (b) GraphCut (c) SEAGULL

(d) PBSC (e) QEIS (f) QPSD

Fig. 5. Comparison of seamline detention performance on the *PT-053* dataset [24]. (a) Reference image. Stitching seamlines of (b) GraphCut [17], (c) SEAGULL [25], (d) PBSC [20], (e) QEIS [26], and (f) QPSD. Stitching results of highlighted regions are magnified at the right side.



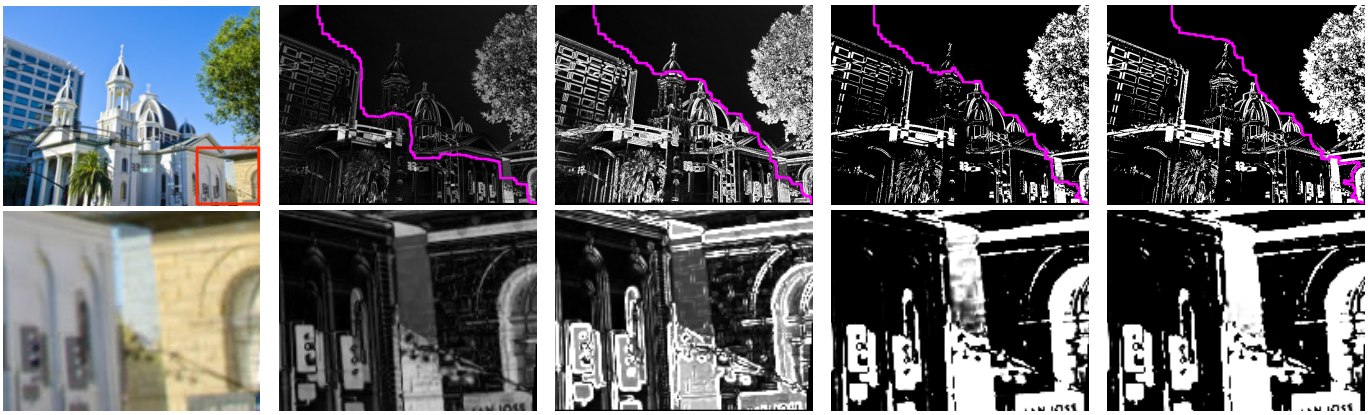
(a) Initial alignment (b) Gray representation (c) Monochromatic model (d) Concatenation model (e) Quaternion representation

Fig. 6. Ablation study on quaternion representation of color images on the *SEG-09* dataset [25]. (a) Initial alignment of overlapping area. Misaligned regions are magnified at the bottom. Exploration of color differences (top) and stitching seamlines (bottom) using the (b) gray representation, (c) monochromatic model, (d) concatenation model, and (e) quaternion representation. Stitching results of misaligned regions are magnified at right corners.

the brighter regions indicate larger differences, i.e., more alignment errors. Thus, an optimal seamline is supposed to cut through the black areas in a difference map, such that the overlapped color images have minimal difference cost. It can be observed from Fig. 6 (e) that the color differences of misaligned structures are fully characterized by the quaternion representation of color images. Correspondingly, the optimized seamline is able to avoid cutting through misaligned structures, generating a plausible stitching result. In contrast, the three common representation approaches do not work on the whole color space, and thus they cannot explore color differences sufficiently. Their optimized seamlines cut through misaligned structures, resulting in obvious artifacts to different extents.

2) Quaternion Perceptual Difference Map: QPSD utilizes the quaternion representation of color images to calculate the quaternion perceptual difference map. To validate the effectiveness of the quaternion perceptual difference map, we compare it with three common difference maps calculated in GraphCut, SEAGULL, and PBSC, i.e., (1) the color difference map; (2) the color and edge difference map; and (3) the perception-based difference map. Notably, these common difference maps are all defined in the real domain and we utilize the concatenation model to calculate their color differences uniformly. For fairness, we set $t = 1$ for $\dot{\mathbf{O}}$ in QPSD.

Fig. 7 shows the four difference maps on the *PT-013* dataset [24]. We extract the overlapping area to present the initial alignment in Fig. 7 (a). For illustration, we further magnify



(a) Initial alignment (b) Color difference (c) Color and edge difference (d) Perception difference (e) Quaternion perceptual difference

Fig. 7. Ablation study on quaternion perceptual difference map $\mathbf{D}_{percept}$ on the *PT-013* dataset [24]. (a) Initial alignment of overlapping area. The misaligned region is magnified at the bottom. Difference maps of (b) color difference [17], (c) color and edge difference [25], (d) perception-based difference [20], and (e) quaternion perceptual difference. Generated seamlines are delineated on the different difference maps. Example regions are magnified at the bottom.

(a) Initial alignment (b) $t = 1$ (c) $t = 2$ (d) $t = 3$ (e) $t = 4$

Fig. 8. Ablation study on quaternion local set $\dot{\mathbf{O}}$ on the *PT-053* dataset [24]. (a) Initial alignment of overlapping area. Generated seamlines are delineated on the quaternion perceptual difference maps when (b) $t = 1$, (c) $t = 2$, (d) $t = 3$, and (e) $t = 4$, respectively.

a misaligned region where the right yellow wall in one image is overlaid with the green plants and the blue sky in another image. As shown in Figs. 7 (b) and (c), alignment errors are characterized when the yellow wall is overlapped with the green plants, but cannot be effectively captured when it is overlapped with the blue sky. The perception-based difference map performs better in Fig. 7 (d); however, there are still some alignment errors not being captured. As a result, their generated seamlines all directly cut through the misaligned yellow wall, as shown in Figs. 4 (b)-(d). In comparison, all the above visual alignment errors are characterized sufficiently using the quaternion perceptual difference map in Fig. 7 (e). From Fig. 4 (f), we can observe it enables QPSD to generate a plausible seamline, avoiding such structural distortions.

3) Quaternion Local Set: To circumvent misaligned pixels in the low-textured and repetitive-textured areas, we optimize the seamline considering the alignment accuracy of its local area. To verify the efficacy of the quaternion local set $\dot{\mathbf{O}}$ on QPSD, we perform experiments with different values of the parameter t.

Fig. 8 shows the generated seamlines for the *PT-053* dataset [24]. We present the image alignment of a specific overlapping area and delineate seamlines on the quaternion perceptual difference map, respectively. As shown in Fig. 8 (a), the repetitive flagstones are misplaced severely. However, due to the same or similar textures, such differences are not distinguished sufficiently in the quaternion perceptual difference map. When $t = 1$, QPSD optimizes the seamline considering only its two adjacent color pixels, and thus the seamline cuts through misaligned flagstones easily in Fig. 8 (b). With the parameter value increasing, QPSD takes its local-area alignment accuracy into account to optimize the seamline. We can observe from Fig. 8 (c) that when $t = 2$, the seamline starts to circumvent those misaligned flagstones and generates a visually pleasing result, as shown in Fig. 5 (e). However, not the parameter value is larger the better. For instance, when $t = 3$ here, redundant alignment errors may mislead QPSD into optimizing the shortest path, i.e., the straight line.

TABLE I

INFLUENCE OF PARAMETER VARIATIONS ON SEAMLINE QUALITY MEASURED IN RMSE, SSIM, FSIMC, ZNCC SCORE, AND QPSQ

t	1	2	3	4	5	6
RMSE (%)	4.70	**3.36**	4.42	5.21	5.14	5.39
SSIM (%)	78.35	**86.19**	81.22	79.42	80.32	79.37
FSIMc (%)	93.96	**94.09**	93.18	90.55	82.77	82.29
ZNCC score	0.13	**0.10**	0.12	0.14	0.18	0.18
QPSQ (%)	5.38	**3.32**	7.04	8.58	8.44	8.76

To find an appropriate value for the parameter t, we give the empirical analysis. We collect a validation set containing five datasets that are randomly selected from [24] and [25]. We vary the value of parameter t from 1 to 6 to quantitatively evaluate the seamline quality. The average values with different choices are summarized in Table I. It can be clearly observed that the seamline qualities under five metrics are all improved when $2 \leq t < 3$; but generally decline if $t \geq 3$. The best seamline quality is obtained when $t = 2$. Thus, we set $t = 2$ by default in this paper.

D. Performance of Color Image Stitching

Although QPSD facilitates the generation of a plausible seamline in the presence of misalignment, the seamline quality is still founded on an accurate color image alignment.

(a) LPC (b) ACIS-QR1A (c) QPSD (d) AQCIS

Fig. 9. Comparison of color image stitching performance on the *PT-013* dataset [24] - the case with large parallax and repetitive textures. Stitching results of (a) LPC [5], (b) ACIS-QR1A [30], (c) QPSD, and (d) AQCIS. Sample regions are magnified at the right side.

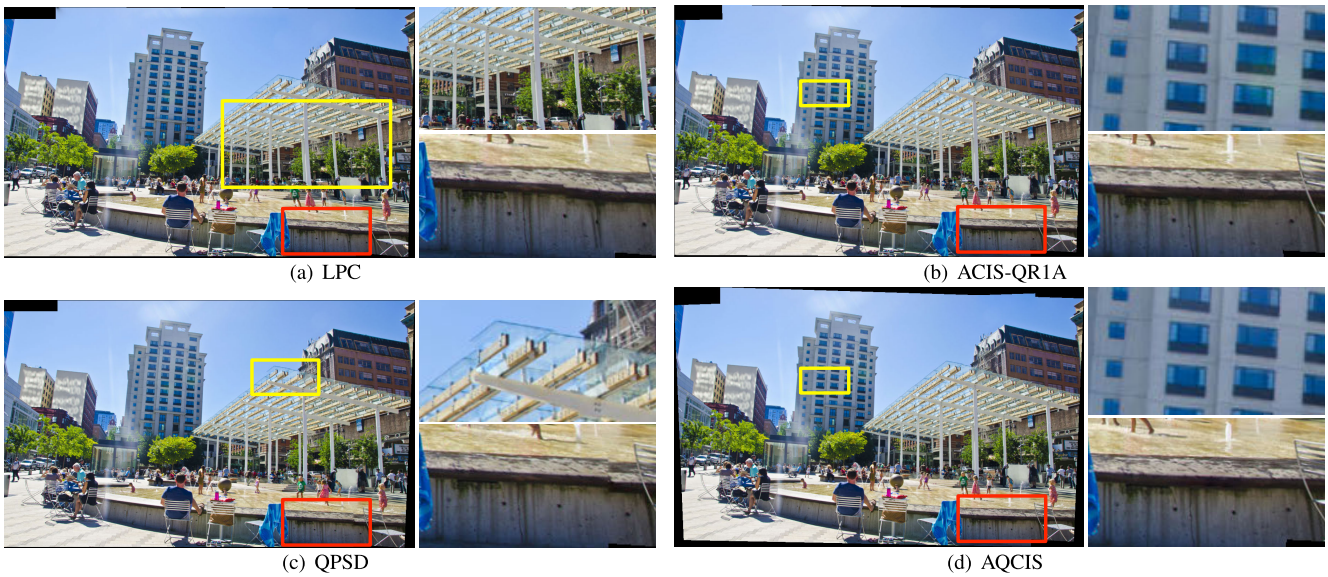
(a) LPC (b) ACIS-QR1A (c) QPSD (d) AQCIS

Fig. 10. Comparison of color image stitching performance on the *PT-053* dataset [24] - the case with large parallax and repetitive textures. Stitching results of (a) LPC [5], (b) ACIS-QR1A [30], (c) QPSD, and (d) AQCIS. Sample regions are magnified at the right side.

To evaluate the overall stitching ability of AQCIS, we compare it with QPSD and two state-of-the-art methods, i.e., LPC [5] and ACIS-QR1A [30] on various challenging datasets. Notably, both LPC and ACIS-QR1A adopt the PBSC [20] method for seamline detection.

In real-world applications, color image stitching often faces multiple difficulties simultaneously. As shown in Fig. 9, methods designed for large parallax are all able to generate visually plausible results for the *PT-013* dataset [24]. However, the repetitive bricks of the right yellow wall are misaligned by LPC, ACIS-QR1A, and QPSD, and their seamlines cut through these misaligned repetitive textures resulting in artifacts to different extents. By comparison, AQCIS aligns these repetitive bricks accurately and obtains higher seamline stitching quality clearly. For the *PT-053* dataset [24], there are more kinds of repetitive textures, including trees, tiles, flagstones, and the green roof. As shown in Fig. 10, seamlines of LPC and ACIS-QR1A are prone to pass through these misaligned repetitive-textured areas since they have the same or similar color values. Comparing AQCIS to QPSD, we can observe that its superior stitching performance is attributable to the improved local alignment accuracy along the seamline.

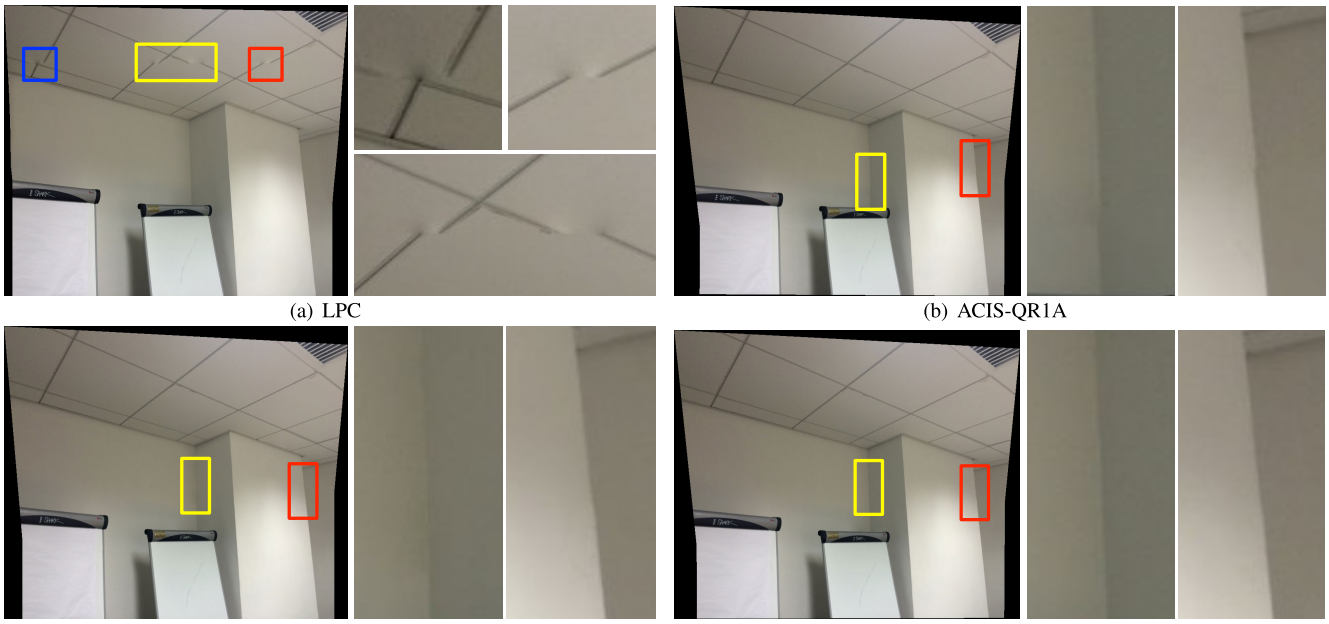
Examples of other challenging scenarios are further given. In the *PT-009* dataset [24], there are simultaneously existing large parallax, moving kids, and repetitive textures, as shown in Fig. 11. The presence of moving objects or occlusions may mislead feature detection and matching. LPC leverages feature points and lines trying to estimate the global alignment for the whole overlapping area. As a result, it distorts the front white pillars and also suffers from shearing artifacts. QPSD suffers from shearing artifacts as well. In contrast, ACIS-QR1A and AQCIS estimate the local alignment along the seamline, they are free from the above artifacts. However, ACIS-QR1A calculates the seamline considering only its two adjacent pixels. As a result, the stitched building of ACIS-QR1A in Fig. 11 (b) is not as straight as that of AQCIS in Fig. 11 (d). In the *DFW-roof* dataset [14], there is a combination of large parallax and low textures. As shown in Fig. 12, LPC generates apparent



(a) LPC (b) ACIS-QR1A

(c) QPSD (d) AQCIS

Fig. 11. Comparison of color image stitching performance on the *PT-009* dataset [24] - the case with large parallax, moving objects, and repetitive textures. Stitching results of (a) LPC [5], (b) ACIS-QR1A [30], (c) QPSD, and (d) AQCIS. Sample regions are magnified at the right side.



(a) LPC (b) ACIS-QR1A

(c) QPSD (d) AQCIS

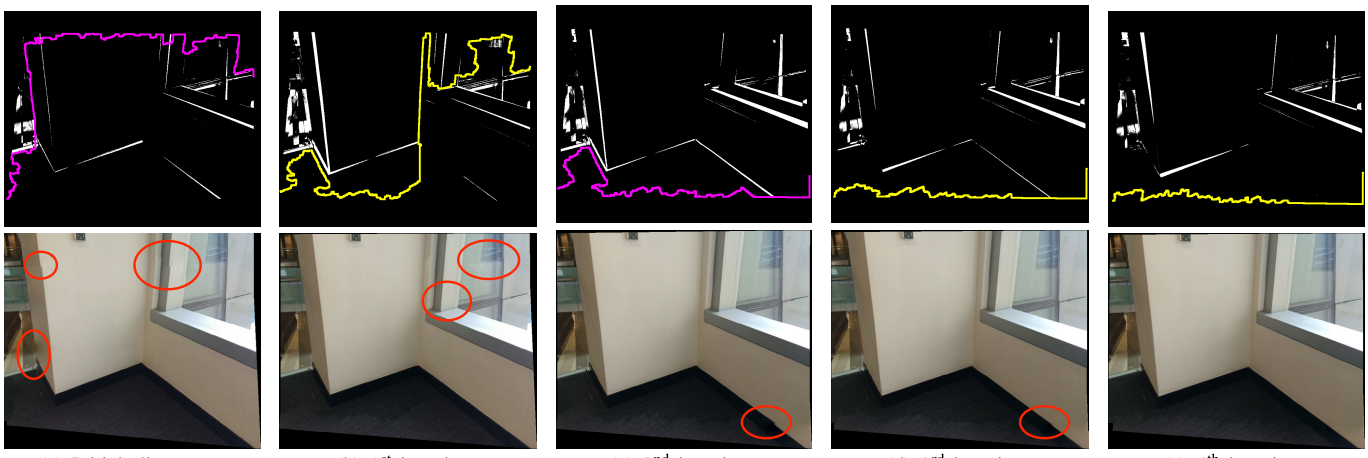
Fig. 12. Comparison of color image stitching performance on the *DFW-roof* dataset [14] - the case with large parallax and low textures. Stitching results of (a) LPC [5], (b) ACIS-QR1A [30], (c) QPSD, and (d) AQCIS. Sample regions are magnified at the right side.

artifacts due to insufficient feature correspondences. ACIS-QR1A and QPSD perform better but still suffer from slight artifacts at the connection of walls. In contrast, the result of AQCIS exhibits higher visually stitching quality. Overall, it may be inferred that AQCIS outperforms QPSD, LPC, and ACIS-QR1A in multiple kinds of challenging scenarios, demonstrating superiority in color image stitching.

E. Ablation Study on AQCIS

To provide an intuitive understanding of AQCIS, we conduct an ablation study on the *DFW-corner* dataset [14]. As shown in Fig. 13, we present the intermediate stitching results of AQCIS in each iteration, including the learned seamlines and stitched color images. To illustrate the optimized local alignment quality, we delineate seamlines on the quaternion perceptual difference maps. When the seamline cuts through the bright local areas in the difference map, artifacts will be generated in the stitched image accordingly. From Fig. 12 (a) to Fig. 12 (e), it can be clearly observed that the bright local areas, i.e., alignment errors, disappear along the seamline gradually. The corresponding artifacts are also eliminated little by little during the process of the joint optimization of our local alignment and seamline.

Table II records the intermediate seamline quality in each iteration. As can be seen, the RMSE and QPSQ scores



(a) Initial alignment (b) 1st iteration (c) 2nd iteration (d) 3rd iteration (e) 4th iteration

Fig. 13. Ablation study on AQCIS on the *DFW-corner* dataset [14]. Intermediate stitching results of the (a) Initial alignment, (b) 1st iteration, (c) 2nd iteration, (d) 3rd iteration, and (e) 4th iteration. Seamlines are delineated on the quaternion perceptual difference maps. Obvious artifacts are circled in red.

TABLE II

INTERMEDIATE SEAMLINE QUALITY MEASURED IN RMSE, SSIM, FSIMC, ZNCC SCORE, AND QPSQ

Iteration	*Initial*	1st	2nd	3rd	4th
RMSE (%)	4.78	4.05	**3.77**	3.83	3.83
SSIM (%)	**80.23**	77.59	71.70	69.40	70.13
FSIMc (%)	95.58	95.15	96.65	**97.50**	97.26
ZNCC score	**0.17**	0.26	0.28	0.24	0.26
QPSQ (%)	3.23	0.92	0.36	0.19	**0.001**

TABLE III

QUANTITATIVE EVALUATION MEASURED IN RMSE, SSIM, FSIMC, ZNCC SCORE, AND QPSQ

Dataset	Metric	PBSC [20]	QEIS [26]	LPC [5]	ACIS-QR1A [30]	QPSD	AQCIS
PT-009 [24]	RMSE(%)	6.29	5.87	5.05	**4.25**	4.81	4.34
	SSIM(%)	71.19	72.04	77.35	82.77	75.37	**83.93**
	FSIMc(%)	95.18	**95.93**	95.71	95.29	94.82	95.20
	ZNCC score	0.15	0.13	0.11	0.08	0.13	**0.07**
	QPSQ(%)	8.68	9.79	7.16	4.00	5.03	**2.83**
PT-013 [24]	RMSE(%)	6.43	5.81	**3.62**	5.34	7.90	5.63
	SSIM(%)	78.02	84.61	87.62	87.51	79.32	**94.97**
	FSIMc(%)	89.24	88.31	89.77	89.84	**91.27**	85.24
	ZNCC score	0.08	0.07	0.03	0.04	0.10	**0.02**
	QPSQ(%)	10.59	7.35	4.16	5.26	10.48	**2.27**
PT-053 [24]	RMSE(%)	10.24	11.35	10.14	**8.13**	10.16	8.50
	SSIM(%)	37.27	36.61	36.19	**51.95**	38.06	49.23
	FSIMc(%)	93.43	93.74	93.98	**94.67**	93.70	93.66
	ZNCC score	0.32	0.32	0.32	**0.26**	0.34	**0.26**
	QPSQ(%)	10.79	7.34	7.76	7.17	11.91	**4.42**
SEG-09 [25]	RMSE(%)	5.13	**2.40**	3.41	2.61	2.88	2.50
	SSIM(%)	68.75	87.95	84.22	**87.99**	86.54	87.39
	FSIMc(%)	95.37	95.65	**96.98**	96.19	95.35	96.26
	ZNCC score	0.23	0.26	**0.15**	0.17	0.18	0.17
	QPSQ(%)	4.46	2.08	2.53	1.63	2.22	**1.02**
DFW-roof [14]	RMSE(%)	1.78	–	2.28	**1.74**	2.33	1.96
	SSIM(%)	94.36	–	95.92	**96.81**	94.21	96.64
	FSIMc(%)	96.91	–	**98.65**	97.88	96.49	97.38
	ZNCC score	0.08	–	**0.01**	0.03	0.05.	0.04
	QPSQ(%)	0.37	–	0.63	0.35	0.55	**0.09**
DFW-corner [14]	RMSE(%)	5.54	5.89	4.96	4.73	4.78	**3.83**
	SSIM(%)	74.55	70.98	79.15	79.92	**80.23**	70.13
	FSIMc(%)	92.24	93.81	95.67	94.96	95.58	**97.26**
	ZNCC score	0.20	0.21	**0.10**	0.23	0.17	0.26
	QPSQ(%)	6.71	7.58	4.74	1.04	3.23	**0.001**

are iteratively improved with the joint optimization process. This reveals the fact that our learned local alignment and seamline are iteratively optimized, and also validates the effectiveness of the joint optimization strategy adopted in AQCIS. Furthermore, it is noteworthy that the RMSE metric cannot effectively discriminate the seamlines of the 3rd and 4th iterations, while the SSIM, FSIMc, and ZNCC scores even predict the learned seamline qualities with an opposite tendency. Compared with them, the QPSQ metric cannot only discriminate different seamlines effectively but also correlate well with human perception to evaluate the seamline quality.

F. Quantitative Evaluation

To comprehensively assess the stitching ability of QPSD and AQCIS, we carry out extensive quantitative experiments. Table III records the quantitative comparisons of the above qualitative samples. Additionally, we have expanded our comparative experiments on the entire datasets of PTIS [24] and SEAGULL [25] (see APPENDIX). The corresponding experimental results are provided in the supplementary file.

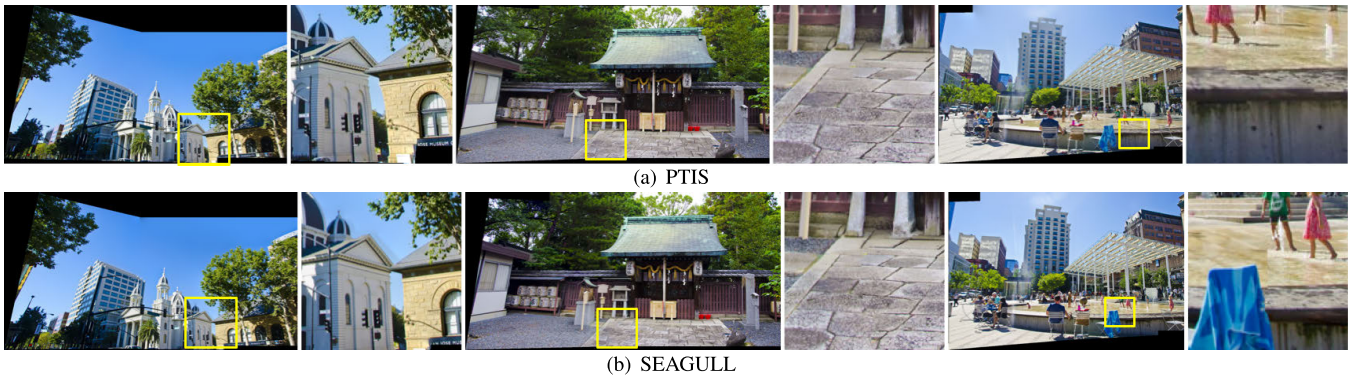
We highlight the best and the second-best scores under different metrics in bold and underlined, respectively; and "–" indicates the failure case. As reported in Table III, QPSD generally outperforms other seamline detection methods PBSC and QEIS, while AQCIS is superior to QPSD, achieving better color image stitching performance. Benefited from the improvement of alignment quality, LPC and ACIS-QR1A outperform PBSC in most cases. Additionally, ACIS-QR1A achieves comparable results with AQCIS, especially in low-textured scenarios. However, the seamline quality limits its effectiveness in the final color image stitching performance.

On the other hand, the RMSE and SSIM scores provide objective measures of geometric alignment accuracy and structural similarity, but may not be in accord with human beings' subjective evaluation. For example, AQCIS obtains the best RMSE and SSIM scores in the *PT-053* and *DFW-roof* datasets, while noticeable artifacts are still captured in Figs. 10 (b) and 12 (b). By comparison, the QPSQ score exhibits greater consistency with the qualitative results. In Table III, AQCIS achieves the best QPSQ scores constantly despite various challenging scenarios.

VI. DISCUSSION

A. Comparison With Seam-Driven Methods

The limitations of existing seam-driven methods have been discussed in the Introduction. Here, we present experimental



(a) PTIS

(b) SEAGULL

Fig. 14. Stitching results of (a) PTIS [24], (b) SEAGULL [25]. Example regions are magnified at the right side.

results of PTIS [24] and SEAGULL [25] for illustration. As shown in Figure 14, they put major efforts into optimizing a local alignment for seamline detection to tolerate large parallax. However, from the magnified local regions, we observe that stitching artifacts are still produced. Compared with them, AQCIS advances in four aspects: (1) the pixel-based QR1A model is utilized for accurate local alignment estimation; (2) the QPSD model is proposed for high-quality seamline detection; (3) the QPSQ measure is proposed to discriminate different seamlines to provide a specific optimization direction. (4) More importantly, AQCIS adopts the joint optimization strategy and the full quaternion-domain color image stitching strategy. It makes full use of the color information to jointly optimize the local alignment and seamline, achieving the final optimal color image stitching performance automatically.

B. Limitation and Future Work

Since input color images are usually partially overlapped, seamline detection is restricted to the overlapping area. To find a sensible local alignment enabling the optimal seamline stitching, AQCIS thus works on the overlapped color images as well. However, the estimated local alignment in the overlapping area may be biased in the non-overlapping area. As a result, projective distortions probably occur in the stitching results. In the future, the naturalness issue in color image stitching will be addressed.

VII. Conclusion

In this paper, we proposed an automatic quaternion-domain color image stitching method that was able to address multiple real-world difficulties with high-quality stitching results. Taking advantage of the quaternion representation of the color image, we proposed the QPSD model and the QPSQ measure to calculate and evaluate seamlines, respectively. Incorporating the pixel-based QR1A model with QPSD and QPSQ, we further developed a unified AQCIS framework. In this framework, the local alignment and seamline were jointly optimized in the quaternion domain to advance the final optimal color image stitching performance. Experimental results have shown that AQCIS outperforms the state-of-the-art methods in color image stitching, and has the capability of simultaneously addressing multiple difficulties, including large parallax, moving objects, occlusions, low textures, and repetitive textures.

References

[1] B. Zhuang and Q.-H. Tran, "Image stitching and rectification for hand-held cameras," in *Proc. ECCV*. Cham, Switzerland: Springer, 2020, pp. 243–260.

[2] X. Fan, J. Lei, Y. Fang, Q. Huang, N. Ling, and C. Hou, "Stereoscopic image stitching via disparity-constrained warping and blending," *IEEE Trans. Multimedia*, vol. 22, no. 3, pp. 655–665, Mar. 2020.

[3] C. Liu, S. Zhang, and A. Akbar, "Ground feature oriented path planning for unmanned aerial vehicle mapping," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 4, pp. 1175–1187, Apr. 2019.

[4] J. Li, Z. Wang, S. Lai, Y. Zhai, and M. Zhang, "Parallax-tolerant image stitching based on robust elastic warping," *IEEE Trans. Multimedia*, vol. 20, no. 7, pp. 1672–1687, Jul. 2018.

[5] Q. Jia et al., "Leveraging line-point consistence to preserve structures for wide parallax image stitching," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 12181–12190.

[6] P. Du, J. Ning, J. Cui, S. Huang, X. Wang, and J. Wang, "Geometric structure preserving warp for natural image stitching," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 3678–3686.

[7] K. Lin, N. Jiang, S. Liu, L.-F. Cheong, M. Do, and J. Lu, "Direct photometric alignment by mesh deformation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2701–2709.

[8] L. Nie, C. Lin, K. Liao, S. Liu, and Y. Zhao, "Unsupervised deep image stitching: Reconstructing stitched features to images," *IEEE Trans. Image Process.*, vol. 30, pp. 6184–6197, 2021.

[9] L. Nie, C. Lin, K. Liao, S. Liu, and Y. Zhao, "Deep rectangling for image stitching: A learning baseline," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 5740–5748.

[10] M. Afifi and M. S. Brown, "Deep white-balance editing," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 1394–1403.

[11] J. Zaragoza, T.-J. Chin, M. S. Brown, and D. Suter, "As-projective-as-possible image stitching with moving DLT," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 2339–2346.

[12] 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.

[13] J. Li, B. Deng, R. Tang, Z. Wang, and Y. Yan, "Local-adaptive image alignment based on triangular facet approximation," *IEEE Trans. Image Process.*, vol. 29, pp. 2356–2369, 2020.

[14] S. Li, L. Yuan, J. Sun, and L. Quan, "Dual-feature warping-based motion model estimation," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 4283–4291.

[15] Y.-S. Chen and Y.-Y. Chuang, "Natural image stitching with the global similarity prior," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2016, pp. 186–201.

[16] G. Zhang, Y. He, W. Chen, J. Jia, and H. Bao, "Multi-viewpoint Panorama construction with wide-baseline images," *IEEE Trans. Image Process.*, vol. 25, no. 7, pp. 3099–3111, Jul. 2016.

[17] V. Kwatra, A. Schödl, I. Essa, G. Turk, and A. Bobick, "Graphcut textures: Image and video synthesis using graph cuts," *ACM Trans. Graph.*, vol. 22, no. 3, pp. 277–286, 2003.
[18] A. Agarwala et al., "Interactive digital photomontage," in *Proc. ACM SIGGRAPH Papers*, Aug. 2004, pp. 294–302.
[19] J. Jia and C.-K. Tang, "Image stitching using structure deformation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 4, pp. 617–631, Apr. 2008.
[20] N. Li, T. Liao, and C. Wang, "Perception-based seam cutting for image stitching," *Signal, Image Video Process.*, vol. 12, no. 5, pp. 967–974, Feb. 2018.
[21] L. Li, J. Yao, X. Lu, J. Tu, and J. Shan, "Optimal seamline detection for multiple image mosaicking via graph cuts," *ISPRS J. Photogramm. Remote Sens.*, vol. 113, pp. 1–16, Mar. 2016.
[22] X. Chen, M. Yu, and Y. Song, "Optimized seam-driven image stitching method based on scene depth information," *Electronics*, vol. 11, no. 12, p. 1876, Jun. 2022.
[23] J. Gao, Y. Li, T.-J. Chin, and M. S. Brown, "Seam-driven image stitching," in *Proc. EG Short Papers*, 2013, pp. 45–48.
[24] F. Zhang and F. Liu, "Parallax-tolerant image stitching," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 3262–3269.
[25] K. Lin, N. Jiang, L.-F. Cheong, M. Do, and J. Lu, "Seagull: Seam-guided local alignment for parallax-tolerant image stitching," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2016, pp. 370–385.
[26] T. Liao, J. Chen, and Y. Xu, "Quality evaluation-based iterative seam estimation for image stitching," *Signal, Image Video Process.*, vol. 13, no. 6, pp. 1199–1206, Sep. 2019.
[27] C. Zou, K. I. Kou, and Y. Wang, "Quaternion collaborative and sparse representation with application to color face recognition," *IEEE Trans. Image Process.*, vol. 25, no. 7, pp. 3287–3302, Jul. 2016.
[28] X. Xiao and Y. Zhou, "Two-dimensional quaternion PCA and sparse PCA," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 7, pp. 2028–2042, Jul. 2019.
[29] X. Xiao, Y. Chen, Y.-J. Gong, and Y. Zhou, "2D quaternion sparse discriminant analysis," *IEEE Trans. Image Process.*, vol. 29, pp. 2271–2286, 2020.
[30] J. Li and Y. Zhou, "Automatic color image stitching using quaternion rank-1 alignment," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 19688–19697.
[31] W. R. Hamilton, "On quaternions; or on a new system of imaginaries in algebra," *Philos. Mag.*, vol. 33, no. 219, pp. 58–60, 1848.
[32] Y. Xu, L. Yu, H. Xu, H. Zhang, and T. Nguyen, "Vector sparse representation of color image using quaternion matrix analysis," *IEEE Trans. Image Process.*, vol. 24, no. 4, pp. 1315–1329, Apr. 2015.
[33] J. J. McAuley, T. S. Caetano, A. J. Smola, and M. O. Franz, "Learning high-order MRF priors of color images," in *Proc. 23rd Int. Conf. Mach. Learn.*, 2006, pp. 617–624.
[34] J. Zhang, S. Sclaroff, Z. Lin, X. Shen, B. Price, and R. Mech, "Minimum barrier salient object detection at 80 FPS," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 1404–1412.
[35] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 11, pp. 1222–1239, Nov. 2001.
[36] M. A. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.
[37] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.
[38] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst. Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, Jan. 1979.
[39] P. Pérez, M. Gangnet, and A. Blake, "Poisson image editing," *ACM Trans. Graph.*, vol. 22, no. 3, pp. 313–318, Jul. 2003.
[40] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[41] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.

Jiaxue Li is currently pursuing the Ph.D. degree with the Department of Computer and Information Science, University of Macau. Her research interests include color image alignment and stitching and low-rank and sparse matrix/tensor decomposition models.

Yicong Zhou (Senior Member, IEEE) received the B.S. degree in electrical engineering from Hunan University, Changsha, China, and the M.S. and Ph.D. degrees in electrical engineering from Tufts University, Medford, MA, USA.

He is a Professor with the Department of Computer and Information Science, University of Macau, Macau, China. His research interests include image processing, computer vision, machine learning, and multimedia security.

Dr. Zhou is a fellow of the Society of Photo-Optical Instrumentation Engineers (SPIE). He was recognized as one of "Highly Cited Researchers" in 2020, 2021, and 2023. He serves as an Associate Editor for IEEE TRANSACTIONS ON CYBERNETICS, IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, and IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.