A Mismatch Removal Method Based on Global Constraint and Local Geometry Preservation for Lunar Orbiter Images

Dayong Liu, Zhen Ye, *Member, IEEE*, Yusheng Xu, Rong Huang, *Member, IEEE*, Lin Xue, Hao Chen, Genyi Wan, Huan Xie, *Senior Member, IEEE*, and Xiaohua Tong, *Senior Member, IEEE*

***Abstract*—Mismatch removal is a crucial step in multiview matching of lunar orbiter images. This process involves complex challenges like terrain-induced distortion and nonunique geometric structures due to repetitive textures. Traditional methods, whether global constraints or local constraints, fall short in adequately addressing these issues in orbiter imagery. Therefore, this article proposes an effective method for mismatch removal of orbiter images based on global constraint and local geometry preservation combined with the imaging model. In this method, a clean neighborhood of each matching point based on the characteristic of centralized distribution of the back-projection residuals globally is constructed. In the local region, we define a local minimum geometric polygon consisting of the center feature point and its three neighbors, and combine the similarity of the back-projection difference vectors with the affine invariance of the polygon to distinguish the correct matches and mismatches by measuring the degree of local geometry preservation. A series of experiments encompassing parameter sensitivity analysis, comparison studies and ablation experiments were conducted on the lunar reconnaissance orbiter image datasets to demonstrate the effectiveness and reliability of the proposed method. The results indicate that our method exhibits a notable insensitivity to parameter variations, and outperforms other advanced methods in both qualitative and quantitative evaluations. Moreover, a large-scale orbiter images multiview matching tie points extraction framework is extended based on the proposed mismatch removal method, which can achieve better results than commonly used photogrammetric software in terms of the number and accuracy of tie points.**

***Index Terms*—Back-projection residual, local geometry preservation, mismatch removal, orbiter image matching.**

I. Introduction

MAPPING using orbiter images is a basic work of planetary exploration, providing key information for scientific research such as planetary geological and geomorphological analysis. Additionally, mapping products provide important data support for landing site selection, rover navigation and positioning in extraterrestrial object exploration missions [1]. Photogrammetric mapping using orbiter images commonly consists of the following steps: imaging model construction for orbiter images, sparse matching for tie points extraction [2], [3], [4], block adjustment [5], [6], dense matching and forward intersection [7], [8], [9] and DEM interpolation. Sparse matching constitutes a critical step in the mapping process [10], where the effectiveness of mismatch removal significantly impacts the matching results. This, in turn, exerts a substantial influence on subsequent processing stages. For example, during bundle adjustment, accurate matching points are required to iteratively solve for the correction values of exterior orientation parameters. Similarly, accurate matching points are also crucial in epipolar images generation, image registration and other tasks. Therefore, the performance of mismatch removal is pivotal in determining the overall quality and reliability of mapping outputs.

In recent years, a variety of methods have been developed to tackle the challenge of mismatch removal, with the most widely adopted one being the random sample consensus (RANSAC) [11] and its various adaptations. Other robust estimator-based methods include M-estimators [12], least-median-squares (LMedS) [13] and so on. M-estimators determines weights based on residuals to reduce mismatches [14]. However, this method fails with the high outlier ratio. LMedS is similar to RANSAC, and it selects the model with the smallest median residual while RANSAC focuses more on the number of inliers. These methods perform poorly when image transformations are complicated. Nonparametric models, exemplified by vector field consensus (VFC) [15], demonstrate superior proficiency in handling such scenarios compared with parametric models. This method removes mismatches from the vector field containing outliers through iterative optimization and obtain the optimal vector. Local linear transforming [16], as a variant of VFC, adds local geometric constraint to remove mismatches. However, these methods typically have cubic

Manuscript received 20 January 2024; revised 9 April 2024; accepted 7 May 2024. Date of publication 13 May 2024; date of current version 30 May 2024. This work was supported in part by the National Key R&D Program of China under Grant 2022YFF0504100, in part by the National Natural Science Foundation of China under Grant 42101447, Grant 42201478, and Grant 42221002, in part by the Shanghai Municipal Natural Science Foundation under Grant 21ZR1468100, and in part by the Shanghai Sailing Program under Grant 21YF1448800. *(Corresponding author: Zhen Ye.)*

The authors are with the College of Surveying and Geo-Informatics, Tongji University, Shanghai 200092, China, and also with the Shanghai Key Laboratory of Space Mapping and Remote Sensing for Planetary Exploration, Shanghai 200092, China (e-mail: liudy@tongji.edu.cn; 89_yezhen@tongji.edu.cn; yusheng_xu@tongji.edu.cn; rong_huang@tongji.edu.cn; 2011695@tongji.edu.cn; 1733511@tongji.edu.cn; genyi_wan@tongji.edu.cn; huanxie@tongji.edu.cn; xhtong@tongji.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2024.3400394

complexities, leading to high time costs and not being suitable for orbiter images with large sizes. Constraints based on local geometric or topological consistency have been implemented in many methods, such as grid-based motion statistics (GMS) [17], locality preserving matching (LPM) [18], and locality affine-invariant feature matching (LAM) [19]. These methods assume that the inliers have similar neighborhood structures and topology before and after transformation. Other methods, such as the point-to-line distance based method [20], combines epipolar line constraint with transformation model estimation. These methods usually achieve high processing efficiency. However, the influence of spatial local structure such as similar craters and rocks widely distributed on the lunar surface can significantly interfere with the performance of these methods.

Despite numerous methods previously mentioned, their performance is often suboptimal when applied to lunar orbiter images. There are several difficulties in mismatch removal for lunar orbiter images. First, most imaging sensors on orbiters are linear-array push broom cameras which are characterized by multicentric imaging [21]. In this case, linear-array images suffer from local deformations due to the rugged terrain relief or imaging viewpoint variations [22], [23], which limits the performance of the global methods based on robust estimation [24], [25]. Second, radiometric differences exist between different lunar orbiter images due to illumination variations [26]. Especially in the lunar south pole, there are many shadowed regions on image, and the quality of image matching is seriously affected. The initial matching result is relatively sparse and with a low correct ratio, making it difficult to remove outliers. Finally, lunar orbiter images are typically characterized by large dimensions, high aspect ratios and a large number of repetitive textures. The large-scale image size restricts some typical graph matching methods and nonparametric model methods from being suitable for lunar scenes due to their high complexity. The issue of repeated textures restricts the extraction of unique local structures, and increases the number of mismatches in the initial matching. This also affects the performance of local constraint-based mismatch removal methods [23]. Hence, it is still challenging for efficient orbiter image mismatch removal. Additionally, the imaging model of orbiter image including rigorous sensor model or rational function model is commonly provided by image vendors and is ignored in most existing methods.

In order to address the challenges mentioned above, this article proposes a novel method based on global constraint and local geometry preservation (GLGP), and combined with the imaging model. It is found that from the distribution of all putative matches on the whole image, the back-projection residuals of correct matches are concentrated within a certain range, while the residuals of mismatches are discretely distributed. Building upon this observation, a global constrained neighborhood construction strategy that aims to obtain clean neighbors around each matching point is developed. Subsequently, we introduce the back-projection difference vector (BDV) into the local geometry constraint to construct the local geometry preservation measure. After constructing the local minimum geometric polygon (LMGP) using each feature point and its clean neighbors, the mismatches are distinguished from correct matches by determining the degree of local geometry preservation in the LMGP. In addition, a multiview matching framework of orbiter image is extended based on the proposed method and achieves better results compared with Ames Stereo Pipeline (ASP) [27], which is a NASA's open source automated stereogrammetry software. Specifically, the main contributions of this article are summarized as follows.

1) We propose a mismatch removal method based on global constraint and local geometry preservation. This method can significantly enhance the performance of orbiter image matching.
2) We introduce a clean neighborhood construction strategy based on the characteristics of the back-projection residuals globally, which can effectively reduce the outlier rate of the neighborhood compared with the motion consistency method. A local geometry constraint measure is also developed based on local geometric structure preservation, which can more effectively distinguish between inliers and outliers.
3) Our method is extended to the multiview matching framework for orbiter images, which is capable of extracting tie points from thousands of lunar orbiter images, outperforming existing mainstream software in both quantity and accuracy.

II. Related Works

Mismatch removal is crucial for image matching. According to the different constraints used, the mismatch removal methods can be generally categorized into two types: global-constraint methods and local-constraint methods.

A. Global-Constraint Methods

The method based on global constraint, that is, by estimating the global transformation model and function or using some global consistency to achieve mismatch removal. The most commonly method is RANSAC [11], which estimates the global transformation parameters by using the smallest subset for model generation and adopting remaining data for geometric model verification. The model involving the most extensive set of correspondences will be chosen as optimal transformation model. Several optimized methods have been proposed, such as maximum likelihood estimation sample consensus [28], progressive sample consensus (PROSAC) [29], universal RANSAC (USAC) [30] and marginalizing sample consensus++ (MAGSAC++) [31]. However, RANSAC and its variants are sensitive to outlier ratios and the number of iterations. The most important is that RANSAC-based methods can achieve good performance only when the image pairs satisfy a certain global geometric transformation. Identifying correspondence function (ICF) [32] develops an iterative function based on diagnostic techniques and support vector machine for the correspondences estimation. By utilizing this function to assess the consistency of matches, it is possible to effectively identify and eliminate mismatches. VFC [15] reinterprets the challenge of feature matching as a problem of maximum a posteriori

estimation within a Bayesian framework, incorporating latent variables to exclude outliers from initial matches, thereby enhancing the performance of feature matching. In recent years, learning-based mismatch removal methods have also attracted much attention. Learning to find good correspondences [33] is characterized by its deep learning architecture using a multilayer perceptron, focusing on analyzing and processing pixel coordinates rather than image data, while embedding global information to learn to find good correspondences. Learning for mismatch removal (LMR) [34] trains a classifier and converts mismatch removal into a classification problem for processing. Although learning-based methods can obtain better performance, their generalization capabilities are greatly affected by the training data set, that is, the performance decreases when processing models that are not in the training set [35]. Alternatively, graph matching [36] can also solve the correspondence problem, such as mode-seeking [37], [38], dual decomposition [39], and graph shift [40]. However, graph matching falls into the category of NP-hard problems. It is very complex to solve and is not usually used to handle the orbiter images with larger dimensions [41], [42]. These methods nearly use the information of all putative matches to obtain the global transformation model or consistency. Therefore, the performance of these methods will drastically decrease in cases of a high proportion of outliers or independent motion structures presented in the putative matches.

B. Local-Constraint Methods

Compared with global-constraint methods, local-constraint methods exhibit a superior robustness in handling local distortions and diverse transformations [43]. GMS [17] is a method that encapsulates motion smoothness into a certain number of matches in the grid and uses statistical likelihood to remove mismatches. LPM [18] proposes a method that focuses on maintaining the local geometric structures of the inliers. The core idea is that for a correct match, the geometry structures within the neighborhood is invariant after the transformation. LPM, categorized under graph matching, effectively sidesteps the substantial computational complexities commonly inherent in traditional graph matching approaches. On this basis, guided locality preserving matching (GLPM) [44] designs a guided matching strategy, which uses satisfactory matching results obtained from a smaller set, characterized by a higher proportion of inliers, to guide the matching process in a larger set of putative matches. Frame-based locality preservation matching [45] incorporates the frame knowledge into the LPM, demonstrating robustness against extreme zoom and high outlier ratio. Local graph structure consensus [46] combines integer quadratic programming with local graph structure consistency compensation terms to establish a comprehensive and efficient optimization framework using graph matching. Top K rank preservation (TopKRP) [47] designs a metric, the weighted Spearman's footrule distance, to describe the similarity of the top K rank of matching points. To extract the correct match from the putative set, multiscale top K rank preservation (mTopKRP) [48] introduces a strategy for building neighborhoods at multiple scales, which is robust to different types of degradation compared with TopKRP. LAM [19] uses the area ratio invariance of affine transformation to construct the local barycentric coordinates to remove mismatches. Then, LAM constructs local matching coordinate matrices (MCMs) which are used to determine the accuracy of the remaining matches through the minimization of the MCMs' ranks. Neighborhood manifold preserving matching (NMP) [49] uses motion consistency to construct neighborhoods and exploits the property that correlated neighborhood point pairs in a high-dimensional manifold preserve neighborhood structure in a low-dimensional manifold to remove mismatches. Local affine preservation (LAP) [43] selects neighbors with the highest motion consistency to construct neighborhood and defines the minimum topological unit (MTU). LAP calculates the cost of MTU based on the area ratio invariance of affine transformation to filter out mismatches. Although these methods fully mined local constraints information, the performance degrades in scenarios where lunar orbiter images present local geometric distortions, similar texture information, or when initial putative matches are sparse within a local area.

Moreover, there are also several methods that combine global and local constraints. Wu et al. [23] proposed a two-step mismatch removal method based on local and global constraints. The core of this method lies in utilizing the invariance of local neighborhood topological structures and the similarity of descriptors of correct matching points to obtain the initial set of inliers. Subsequently, the inliers acquired are employed to compute the transformation model between the images, selecting points that meet the criteria as the final results. However, for images with large local deformations, the global constraints of this method are invalid and the performance will be greatly affected. The locality-guided global-preserving optimization (LOGO) [50] combines affine transformation consistency with global-preserving affinity matrix, and proposes a local-to-global mismatch removal method. This method employs local topological consistency to determine a small initial set of high inlier ratio matches. Combining the topological structure of the affinity matrix, it leverages reliable geometric information to get additional inliers, thereby achieving the elimination of erroneous matches. However, when the putative matches set is sparse, the performance of this method will decrease. Different from the above methods, our method introduces sensor imaging model into the global constraint, and uses it to enhance the local constraint so as to remove mismatches.

III. Methodology

The workflow of the proposed GLGP mismatch removal method is illustrated in Fig. 1, which includes two parts. In the first part, the global constraint neighborhood (GCN) is constructed based on the centralized distribution characteristics of the back-projection residual. The second part is divided into the following three steps. First, the local BDV is constructed, and the similarity between the BDVs of the matching point and its neighbors is determined. Then, the local affine invariance of the LMGP formed by the matching points and neighbors is calculated. Finally, the degree of local geometry preservation (LGP) of LMGP is obtained combining the similarity of the

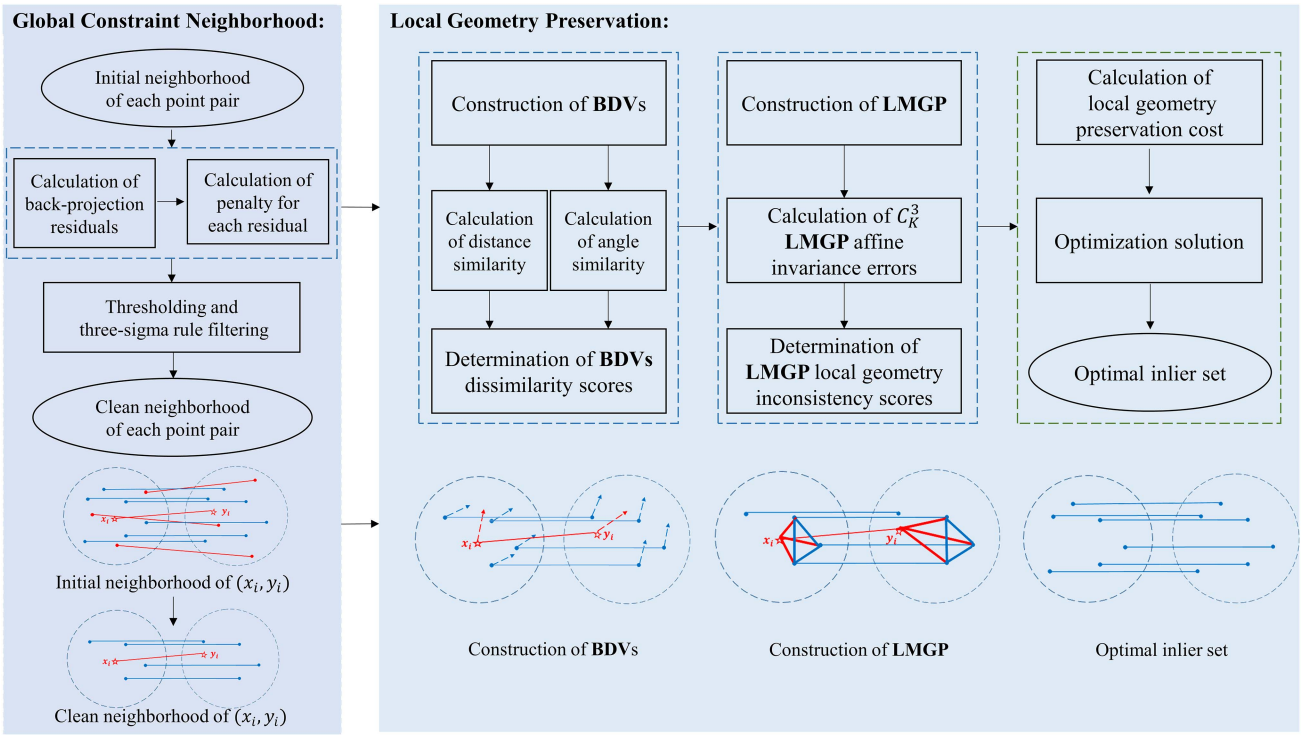

Fig. 1. Workflow of the proposed method and the schematic illustration of key steps. (x_i, y_i) is the matching point pair currently being processed. The red lines indicate the outliers, while the blue ones represent the inliers.

BDVs with the local affine invariance, which is employed to remove the mismatches.

A. Neighborhood Construction Based on Back-Projection Residual Constraint

Local-constraint methods generally require the participation of neighbors. It is crucial to construct a clean neighborhood with low outlier rates because the successful judgment of local geometric properties depends on the selected neighborhood. Previous methods such as LPM and GLPM usually adopt K-nearest neighbors (KNN) to build neighborhood. The mTopKRP uses KNN ranking similarity to construct neighborhood. However, the methods based on KNN are not robust when processing the putative set with a high outlier ratio. Some methods such as NMP and LAP employ motion consistency to obtain neighborhoods. Affected by orbital variations, inconsistent resolution or geometry distortion of orbiter images, motion consistency cannot always obtain the clean neighborhood. Therefore, a global constraint based on the distribution characteristic of back-projection residuals is developed to construct a clean neighborhood for matching points of orbiter images.

Given a set of putative set $S = \{(x_i, y_i)\}_{i=1}^{N}$ obtained from two orbiter images containing N features, where x_i and y_i represent the pixel position vectors of the matching points on the left and right images, respectively. The triangulation technique [51] is adopted to calculate the three-dimensional (3-D) space coordinates of the matching points. Then, the back-projection coordinates x_i' and y_i' are calculated from the calculated 3-D coordinates based on the collinearity equation. The schematic diagram is shown in Fig. 2(a).

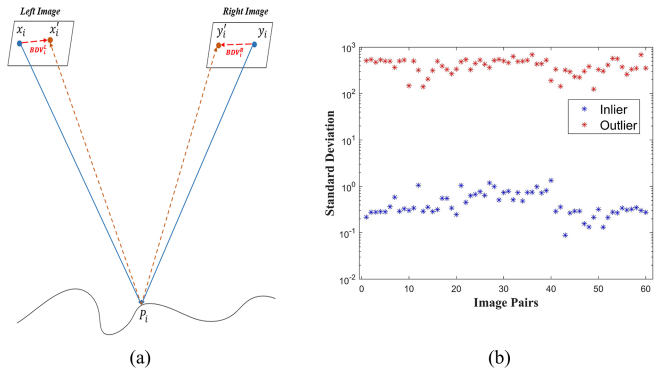

Fig. 2. Back-projection residual of matching point pair. (a) Red dotted line represents the BDV and the vector magnitude is the back-projection residual on the left and right images. The final back-projection residual value is the mean of these two values. (b) Standard deviation of the back-projection residuals for inliers(blue) and outliers(red). The logarithmic scale is utilized for the Y-axis.

Based on the original image point and the back-projected point, the mean residual bpj_res$_i$ of two images, that is the back-projection residual of the matching point pair, can be obtained as follows:

$$\text{bpj\_res}_i = \frac{(|v_{xi}| + |v_{yi}|)}{2} \tag{1}$$

$$v_{xi} = x_i - x_i'$$

$$v_{yi} = y_i - y_i' \tag{2}$$

where v_{xi} and v_{yi} represent the back-projection difference vector of the *i*th matching point pair on the left and right images, respectively.

Whether using the rigorous sensor model or rational function model, the imaging model errors are inevitable, such as sensor calibration errors, orbiter positional and attitude errors and instrument installation errors [52], [53]. These errors are mainly presented as systematic errors, leading to deviations between the real imaging model and the theoretical model, manifesting as back-projection residuals on the image plane. Systematic errors in the imaging model, compared with the random errors from mismatches, take a leading role in the back-projection residuals [53]. Fig. 2(b) shows the statistical results of the standard deviation of the back-projection residuals for inliers and outliers of sixty lunar reconnaissance orbiter (LRO) narrow angle cameras (NAC) image pairs. Due to the large differences in variance, the logarithmic scale is utilized for the *Y*-axis. The statistical results show that the standard deviation corresponding to the residuals of the inlier is small, which indicates that residuals are distributed within a certain range, while the residuals of the outlier are randomly distributed at different values. In other words, the back-projection residuals of inliers on the lunar orbiter images exhibit a characteristic of concentrated distribution.

In order to quantify the distribution of back-projection residuals, we use a predefined back-projection residuals cutoff to limit the range of residuals, then sort the remaining back-projection residuals and calculate the median, which is approximated as the centrally distributed values $\text{bpj\_res}_{\text{cen}}$ of back-projection residuals. We define the penalty value to normalize the error as follows:

$$k\ (\text{err}_i) = 1 - e^{-\frac{1}{2}\cdot\left(\frac{\text{err}_i}{\tau}\right)^2} \tag{3}$$

where err_i denotes the difference between two measures, τ is a predefined parameter, and $k(\text{err}_i)$ denotes the penalty value of error. The penalty value of back-projection residual could be obtained through (3) as follows:

$$k\ (\text{bpj\_res}_i) = 1 - e^{-\frac{1}{2}\cdot\left(\frac{\text{bpj\_res}_i - \text{bpj\_res}_{\text{cen}}}{\tau_0}\right)^2} \tag{4}$$

where τ_0 is the predefined parameter; $k(\text{bpj\_res}_i)$ denotes the penalty value of the back-projection residual, ranging from 0 to 1. A smaller value indicates that the putative match is more likely to be correct. Therefore, the matches with penalty values greater than an empirical threshold are discarded. The mean value μ and standard deviation σ of the penalty value are calculated in the remaining set of matching points. The matches with penalty value in the range of $\mu \pm 3\sigma$ are used to construct a clean putative match set N_c as follows:

$$N_c = \{x_i | \, i = 1,\ 2, \ldots, N,\ k\,(\text{bpj\_res}_i) \in (\mu - 3\sigma, \mu + 3\sigma)\}. \tag{5}$$

Finally, we use KNN to obtain the neighbors of each matching point in N_c as its global constrained neighborhood $N^i_{c\_k}$.

B. Similarity of Local Back-Projection Difference Vectors

Local back-projection difference vector is the vector pointing from a matching point to its back-projection point. Each putative match has a corresponding back-projection difference vector in both left and right images. The back-projection residual of the correct matching point in the orbiter image not only exhibits global consistency but also shows similarity in local region. Therefore, the deviation between the BDVs of the *i*th and *j*th feature points is defined as follows:

$$\begin{aligned} \text{dis\_}err^L_{ij} &= ||v_{xi}| - |v_{xj}|| \\ \text{dis\_}err^R_{ij} &= ||v_{yi}| - |v_{yj}|| \end{aligned} \tag{6}$$

$$\text{ang\_}err_{ij} = \left| \frac{(v_{xi}, v_{xj})}{|v_{xi}| \cdot |v_{xj}|} - \frac{(v_{yi}, v_{yj})}{|v_{yi}| \cdot |v_{yj}|} \right| \tag{7}$$

where $\text{dis\_}err^L_{ij}$ represents the distance deviation between v_{xi} and v_{xj} on the left image, and $\text{dis\_}err^R_{ij}$ represents the distance similarity error between v_{yi} and v_{yj} on the right image. $\text{ang\_}err_{ij}$ represents the angle deviation between BDVs on the left and right images. Given that distance and angle are of different magnitudes, in order to integrate them more rationally, we convert the deviation of BDVs into the corresponding penalty values according to (3) as follows:

$$\begin{aligned} k\,(\text{dis\_}err_{ij}) &= \frac{k\left(\text{dis\_}err^L_{ij}\right) + k\left(\text{dis\_}err^R_{ij}\right)}{2} \\ &= 1 - \frac{e^{-\frac{1}{2}\cdot\left(\frac{\text{dis\_}err^L_{ij}}{\tau_1}\right)^2} + e^{-\frac{1}{2}\cdot\left(\frac{\text{dis\_}err^R_{ij}}{\tau_1}\right)^2}}{2} \end{aligned} \tag{8}$$

$$k\,(\text{ang\_}err_{ij}) = 1 - e^{-\frac{1}{2}\cdot\left(\frac{\text{ang\_}err_{ij}}{\tau_2}\right)^2} \tag{9}$$

$$bdv\_sco_{ij} = k\,(\text{dis\_}err_{ij}) + k\,(\text{ang\_}err_{ij}) \tag{10}$$

where τ_1 and τ_2 are the predefined parameters, $k(\text{dis\_}err_{ij})$ represents the distance similarity penalty, $k(\text{ang\_}err_{ij})$ represents the angle similarity penalty, and bdv_sco_{ij} represents the BDVs' dissimilarity score of *i*th and *j*th feature points, ranging from 0 to 2. A smaller value indicates more similar of the BDVs, which also indicates more similar of the correctness between two corresponding putative matches.

C. Local Affine Invariance and Local Geometry Preservation

It is inadequately robust if merely using the similarity of BDVs to distinguish between inliers and outliers. For example, when the feature point on one image being matched to the vicinity of its corresponding correct position in another image, the above judgment criteria will make it difficult to accurately remove this mismatch. In order to strengthen the local geometric constraints, three neighboring points from the neighborhood $N^i_{c\_k}$ are selected to construct the local minimum geometric polygon (LMGP). A total of V ($\text{V} = C^3_K$) polygons can be constructed from the neighborhood of x_i. Although the orbiter image does not satisfy a global transformation for the entire image region, the local region approximately satisfies affine transformation. In other words, the geometric structure of inliers in LMGP remains unchanged after affine transformation, while outliers do not exhibit this characteristic.

As shown in Fig. 3, (x_i, y_i) is a putative matching point, (x_A, y_A), (x_B, y_B) and(x_C, y_C) are three neighbors of (x_i, y_i).

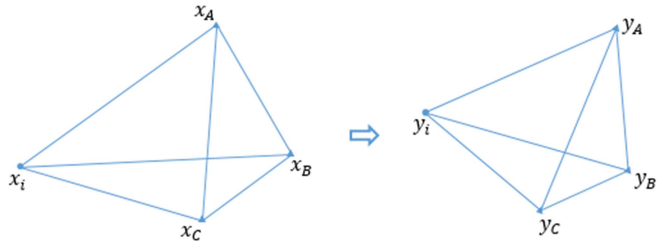

Fig. 3. LMGP of x_i (Left) and y_i (Right). Points x_A, x_B and x_C are three clean neighbors of x_i, and points y_A, y_B, y_C, and y_i are the corresponding matching points on the right image.

x_A, x_B, x_C and x_i construct a LMGP on the left image, while y_A, y_B, y_C and y_i construct the right LMGP.

Based on the local affine invariance, the triangles with common side selected from the LMGP should have area ratio consistency and satisfy the following equation:

$$\begin{aligned}
\frac{S_{x_i x_B x_C}}{S_{x_A x_B x_C}} &= \frac{S_{y_i y_B y_C}}{S_{y_A y_B y_C}} = \frac{h^x_{iBC}}{h^x_{ABC}} = \frac{h^y_{iBC}}{h^y_{ABC}} \\
\frac{S_{x_i x_A x_C}}{S_{x_B x_A x_C}} &= \frac{S_{y_i y_A y_C}}{S_{y_B y_A y_C}} = \frac{h^x_{iAC}}{h^x_{BAC}} = \frac{h^y_{iAC}}{h^y_{BAC}} \\
\frac{S_{x_i x_A x_B}}{S_{x_C x_A x_B}} &= \frac{S_{y_i y_A y_B}}{S_{y_C y_A y_B}} = \frac{h^x_{iAB}}{h^x_{CAB}} = \frac{h^y_{iAB}}{h^y_{CAB}}
\end{aligned} \tag{11}$$

where $S_{x_i x_B x_C}$ represents the area of triangle formed by three points x_i, x_B, x_C. h^x_{iBC}, h^x_{iAC}, and h^x_{iAB} represent the vertical distance from x_i to the side of $\overline{x_B x_c}$, $\overline{x_A x_c}$ and $\overline{x_A x_B}$, respectively, and h^y_{iBC}, h^y_{iAC}, and h^y_{iAB} represent the vertical distance from y_i to the side of $\overline{y_B y_c}$, $\overline{y_A y_c}$, and $\overline{y_A y_B}$ respectively. Due to the existence of matching errors, the local affine invariance error of LMGP can be expressed as follows:

$$\begin{aligned}
\text{loc_}err_{iA} &= \left| \frac{S_{x_i x_B x_C}}{S_{x_A x_B x_C}} - \frac{S_{y_i y_B y_C}}{S_{y_A y_B y_C}} \right| = \left| \frac{h^y_{ABC}}{h^x_{ABC}} h^x_{iBC} - h^y_{iBC} \right| \\
\text{loc_}err_{iB} &= \left| \frac{S_{x_i x_A x_C}}{S_{x_B x_A x_C}} - \frac{S_{y_i y_A y_C}}{S_{y_B y_A y_C}} \right| = \left| \frac{h^y_{BAC}}{h^x_{BAC}} h^x_{iAC} - h^y_{iAC} \right| \\
\text{loc_}err_{iC} &= \left| \frac{S_{x_i x_A x_B}}{S_{x_C x_A x_B}} - \frac{S_{y_i y_A y_B}}{S_{y_C y_A y_B}} \right| = \left| \frac{h^y_{CAB}}{h^x_{CAB}} h^x_{iAB} - h^y_{iAB} \right|.
\end{aligned} \tag{12}$$

Assuming that the neighbors of matching points are correct, the local affine invariance error can represent the discrepancy between the ratios of distances from the feature point and its neighbor to the common side in the left and right images. This equates to an error of local geometric consistency between the feature point and its neighbor. So loc_err_{iA}, loc_err_{iB} and loc_err_{iC} represent the local geometric consistency error between(x_i, y_i) and (x_A, y_A), (x_i, y_i) and (x_B, y_B), (x_i, y_i) and(x_C, y_C), respectively. It should be noted that the angle of triangles involved in the calculation of the local affine invariance error cannot be extremely small, such as less than 1°.

In order to normalize the errors, we convert the local geometric consistency error to the corresponding penalty value by (3) as follows:

$$\begin{aligned}
\text{loc_}sco_{iA} &= k\ (\text{loc_}err_{iA}) = 1 - e^{-\frac{1}{2}\cdot\left(\frac{\text{loc_}err_{iA}}{\tau_3}\right)^2} \\
\text{loc_}sco_{iB} &= k\ (\text{loc_}err_{iB}) = 1 - e^{-\frac{1}{2}\cdot\left(\frac{\text{loc_}err_{iB}}{\tau_3}\right)^2} \\
\text{loc_}sco_{iC} &= k\ (\text{loc_}err_{iC}) = 1 - e^{-\frac{1}{2}\cdot\left(\frac{\text{loc_}err_{iC}}{\tau_3}\right)^2}
\end{aligned} \tag{13}$$

where the τ_3 is the predefined parameter, loc_sco_{iA}, loc_sco_{iB} and loc_sco_{iA} represent the local geometric inconsistency score of between(x_i, y_i) and (x_A, y_A), (x_i, y_i) and (x_B, y_B), (x_i, y_i), and(x_C, y_C), respectively.

Combining the dissimilarity score of BDVs with local geometric inconsistency score, the cost value of local geometry preservation is defined as follows:

$$\begin{aligned}
\text{cost}_u\,(i) \ = \text{bdv_}sco_{iA} \ \cdot \text{loc_}sco_{iA} + \text{bdv_}sco_{iB} \cdot \text{loc_}sco_{iB} \\
+ \text{bdv_}sco_{iC} \cdot \text{loc_}sco_{iC}
\end{aligned} \tag{14}$$

where bdv_sco and loc_sco are calculated form (10) and (13), respectively, $\text{cost}_u(i)$ represents the cost of local geometry preservation of the uth LMGP. The larger the cost is, the worse the degree of local geometry preservation is, indicating a higher probability of mismatch. Although a relatively reliable neighborhood can be obtained using back-projection residual constraints, there are still some outliers in the neighborhood that will perturb the local geometry structure. In order to alleviate this problem, the cost values are sorted in ascending order and the average of the top ξV costs are calculated to represent the cost value $c(x_i, y_i)$ for local geometry preservation of matching point as follows:

$$c\,(x_i, y_i) = \frac{1}{\xi \text{V}} \sum_{u=1}^{\xi \text{V}} \text{cost}_u(i). \tag{15}$$

In general, the scaling factor ξ is empirically set to 0.3.

D. Objective Function and Solution

Due to the unavoidable outliers in the set of putative set S, in order to maximize the retention of correct matches, the mismatch removal is formulated into an optimization problem as follows:

$$I^* = \arg\min_I C\,(I; S; \lambda) \tag{16}$$

with the cost function C defined as follows:

$$C\,(I; S; \lambda) = \sum_{i \epsilon I} c\,(x_i, y_i) + \lambda\,(N - |I|) \tag{17}$$

where $|I|$ means the cardinality of I, $c(x_i, y_i)$ is devoted to penalizing the wrong correspondences, i.e., a larger value indicates that the putative match is more likely to be wrong, and vice versa. The second term serves the purpose of discouraging the presence of outliers. Parameter $\lambda > 0$ is used to control the tradeoff between these two terms.

In order to solve the optimization function, a N-dimensional binary vector $\mathbf{p}$ is introduced to associate the putative set S, where $p_i = 0$ and $p_i = 1$ represent the ith match to be wrong

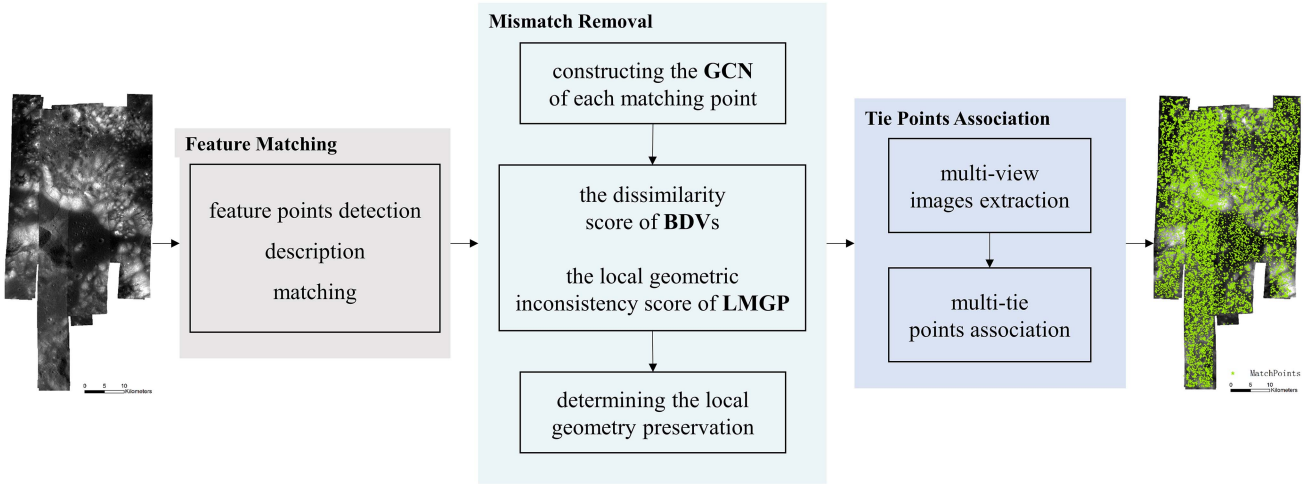

Fig. 4. Overview of the large-scale multiview orbiter images mapping framework. The leftmost is the multiview image, after feature matching, mismatch removal and tie points association, ultimately yields the rightmost tie points result. The green points represent the multiview tie point.

Algorithm 1: GLGP Algorithm.

Input: putative correspondences $S = \{(x_i, y_i)\}_{i=1}^{N}$, parameters $\mathrm{T} = \{\tau_j\}_{j=0}^{3}$, K, λ;
Output: inlier set I^*;
1 Calculate the back-projection residual bpj_res_i of the (x_i, y_i) and the concentrated distribution value bpj_res_{cen}, calculate the back-projection residual penalty value using (3) and (4);
2 Calculate the mean value and variance of the penalty value, construct the clean putative match sets N_c using (5);
3 Search K nearest neighbors in N_c for each feature point x_i, then obtain neighborhood $N_{c\_k}^{i}$;
4 Calculate bdv_sco_{ij}, loc_sco_{ij} and $cost_u(i)$ for each feature point using (10), (13) and (14);
5 Calculate the cost $\{c(x_i, y_i)\}_{i=1}^{N}$ using (15);
6 Determine I^* using (20) and (21).

and correct, respectively. Accordingly, (17) is equivalent to

$$C\left(\mathbf{p}; S; \lambda; \tau\right) = \sum_{i=1}^{N} p_i c\left(x_i, y_i\right) + \lambda \left(N - \sum_{i=1}^{N} p_i\right). \quad (18)$$

After merging and simplifying, we can obtain the following:

$$C\left(\mathbf{p}; S; \lambda; \tau\right) = \sum_{i=1}^{N} p_i \left(c_i - \lambda\right) + \lambda N \quad (19)$$

where $c_i = c(x_i, y_i)$. The costs c_i of geometric preservation between feature points is fixed and can be precalculated for a given putative match. Consequently, any matching point with a c_i value smaller than λ will result in a negative value, and vice versa. Therefore, the analytical solution of the minimized cost function C is as follows:

$$p_i = \begin{cases} 1, c_i \leq \lambda \\ 0, c_i > \lambda \end{cases} \quad i = 1, 2, \ldots, N. \quad (20)$$

Thus, the optimal set of inliers can be obtained as follows:

$$I^* = \{i | p_i = 1, i = 1, 2, \ldots, N\}. \quad (21)$$

The procedure of the proposed GLGP method is concluded in Algorithm 1.

E. Multiview Tie Points Extraction Framework

Based on the proposed method, a large-scale multiview orbiter images tie points extraction framework is extended. Fig. 4 shows the overview of this framework. An optimal set of weighted box filters which approximates the Laplacian of Gaussian operator is used to extract feature points on the orbiter image [54]. Gradient descriptors are constructed and utilized to acquire putative matches through nearest neighbor matching. Subsequently, the proposed method is used to remove mismatches. Finally, the extraction of tie points is extended from stereo to multiview imagery, and a refined strategy for tie points association is employed for the extraction of multitie points. Specifically, each image is sequentially set as a base image, and images forming stereo pairs with the base image are selected to construct an image set according to the image pair relationships. The tie points within this image set are then associated using the union-find algorithm. Finally, by removing the redundant results, the obtain of tie points across multiview imagery is achieved.

IV. Experimental Data and Details

In order to qualitatively and quantitatively evaluate the performance of GLGP, the feature matching method in the proposed framework is employed to obtain putative matches, and the correctness of each putative correspondence in all image pairs is manually checked. Besides, a large-scale lunar orbiter image set is applied to validate the effectiveness of the proposed framework. Several experiments are designed to conduct detailed testing and analysis of the proposed method.

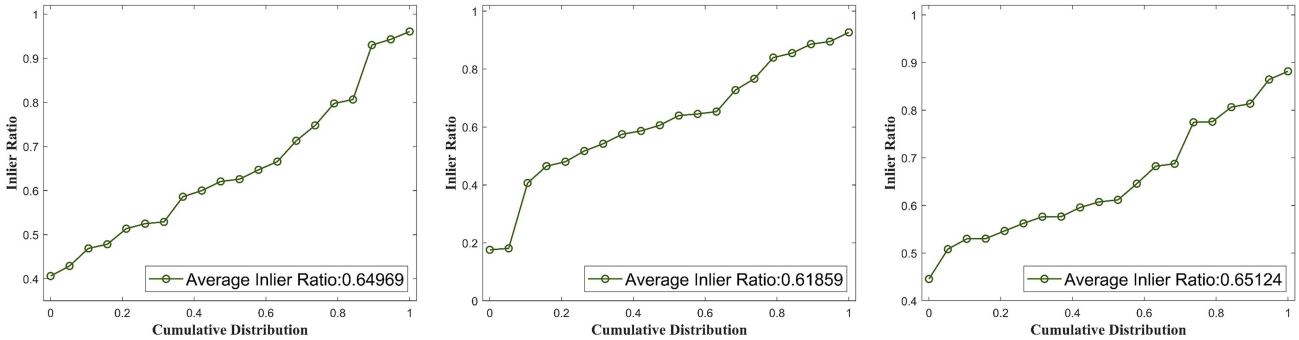

Fig. 5. Inlier ratio of image pairs. From the left to right: the inlier ratio of dataset 1, dataset 2 and dataset 3.

A. Datasets

The datasets are taken from the LRO NAC imagery [55] which are publicly available in the link provided.[1] NAC images are able to deliver panchromatic images at meter-scale resolution across a 5 km wide area from an orbit of 50 km, and are characterized by their expansive dimensions and high aspect ratios, typically manifesting as long strips. The task of mismatch removal in NAC images is notably challenging. Due to the influence of illumination, the texture of identical lunar surfaces, particularly in regions with craters or rocks, exhibits significant variations across different images. Moreover, due to the terrain relief or imaging viewpoint variations, geometric distortions exist in the orbiter images. The above factors lead to a low inlier ratio in lunar orbiter images.

In this study, 60 image pairs were selected in three regions, including Apollo-17 landing site, Chang 'e-4 landing site, and lunar South Pole region, which are named Dataset 1, Dataset 2, and Dataset 3, respectively. The image pairs in the dataset are representative, encompassing a range of challenges such as repeated textures, significant local distortions, and sparse putative matches. The quantity of average putative matches for three datasets are respectively 452, 255, and 174, while their corresponding average inlier ratios are 0.650, 0.619, and 0.651. We plot the cumulative distributions of the inlier ratio and Fig. 5 shows the inlier ratio of each image pair in three datasets. There are several image pairs with low inlier ratio in the dataset, and the lowest ratio is 0.176. The image size of Dataset 1 and Dataset 2 is 52 224 × 5064 pixels, and the Dataset 3 is 52 224 × 2532 pixels. Additionally, in order to evaluate the performance of the proposed framework, we collected 1144 NAC images of the Apollo 17 region, constituting 2599 valid image pairs. The area corresponding to these images is approximately 100 × 300 km. The image pairs contain an average of 362 putative matches. Similarly, the image size is 52 224 × 5064 pixels. Details of the dataset are listed in Table I.

TABLE I
DETAILS OF DATASET USED IN THE EXPERIMENTS

Name	Quantity of Images, Pairs and Average Putative Matches	Average Inlier Ratio	Central Position	Average Pixel Size (m/Pixel)	Purpose
Dataset 1	26, 20, 452	0.650	20.19°N 30.77°E	0.487	Quantitative Evaluation
Dataset 2	32, 20, 255	0.619	45.46°S 177.59°E	0.831	
Dataset 3	18, 20, 174	0.651	89.47°S 137.41°W	1.093	
Apollo17 Region	1144, 2599, 362	\	20.19°N 30.77°E	0.921	Large-Scale Experiments

B. Experimental Designs and Evaluation Metrics

In order to test and analyze the performance of the proposed method, a series of experiments have been designed as follows.

The qualitative comparisons are conducted with several state-of-the-art methods, among which the global-based methods are RANSAC [11] and BpjRes. The latter uses the three-sigma rule based on global back-projection residuals to remove mismatches. The local-based methods are GMS [17], LPM [18], mTopKRP [48], LOGO [50], and LAP [43]. These seven methods are compared with GLGP to demonstrate the effectiveness of the proposed method.

In the quantitative experiments, the proposed method is compared with the global-based methods and the local-based methods. In addition to the methods mentioned above, the global method is supplemented with the optimized variant MAGSAC++ [31] of RANSAC, and the local-based method is supplemented with NMP [49]. In addition, the learning-based method LMR [34] is put together with the global method for comparison. Three metrics are used for quantitative evaluation, i.e., Precision (P), Recall (R), and F-score(F). And Runtime is used to evaluate method efficiency

$$P = \frac{N_{\mathrm{TP}}}{N_{\mathrm{TP}} + \mathrm{N_{FP}}}, \; R = \frac{N_{\mathrm{TP}}}{N_{\mathrm{TP}} + N_{\mathrm{FN}}}, \; F = \frac{2\mathrm{PR}}{\mathrm{P} + \mathrm{R}} \tag{22}$$

where N_{TP} and N_{FP} represent the number of correct matches and mismatches found by each matching method, N_{FN} represents the number of the correct matches not distinguished.

Ablation experiments are conducted on 20 randomly selected image pairs to demonstrate and analyze the reliability of each component of the proposed method, which are divided into two parts. The first part tests various neighborhood construction

[1] [Online]. Available: https://ode.rsl.wustl.edu/ moon/index.aspx

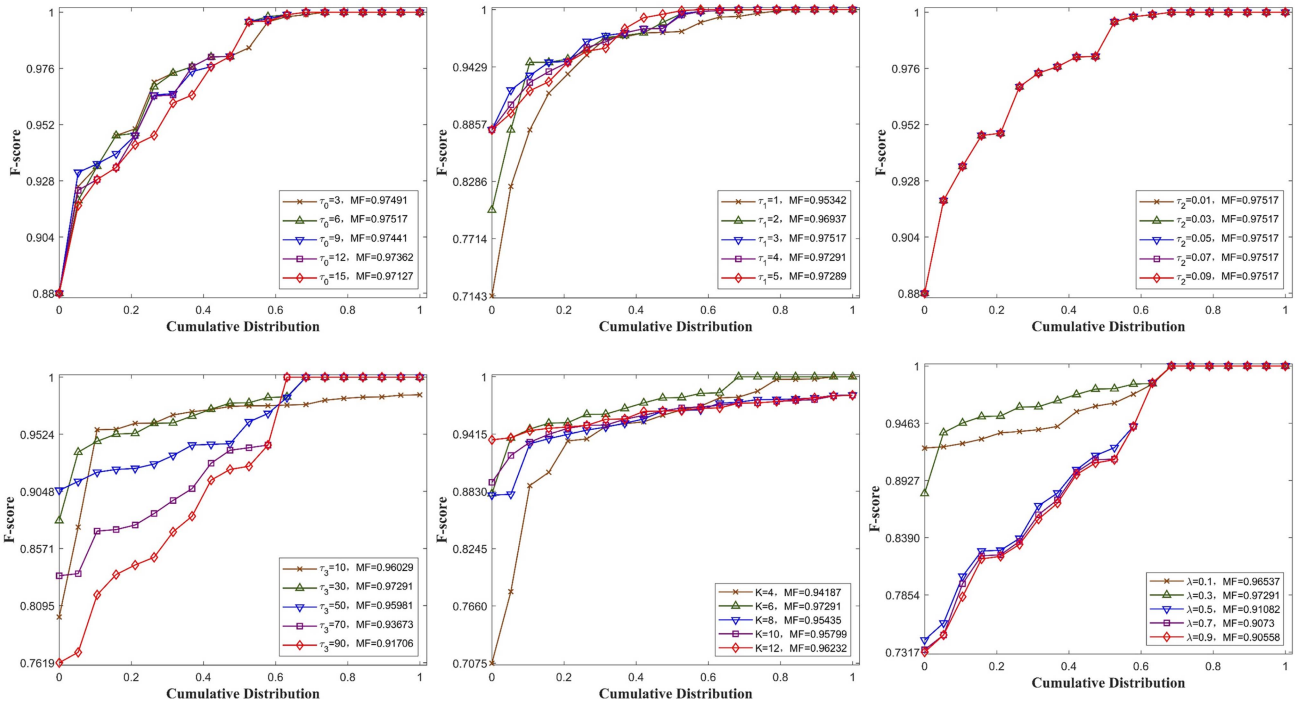

Fig. 6. F-score relative to the cumulative distribution under varying parameters configurations. MF represents the mean value of F-scores.

methods. Specifically, four methods are selected to construct neighborhoods, the k-nearest neighbor neighborhood (KNN) [43] and the motion consistency neighborhood (MCN) [43]. The neighborhood constructed using three-sigma rule of global back-projection residuals, is also named BpjRes in the ablation experiments. The global-constraint neighborhood in our GLGP is named GCN. The performance is evaluated using the same local geometry preservation metric with these four different neighborhood construction methods. In the second part, the same neighborhood construction method is used with different local geometry or topology metrics to evaluate the performance of these variants in detail. The method of neighborhood manifold consensus is named NMC [49], the consensus of neighborhood topology used by LPM [18] is named CNT, the local affine invariance with filtering mechanism is named LAI [43] and the local geometry preservation proposed in this paper is named LGP.

The performance of the developed tie points extraction framework is tested using large-scale orbiter images and compared with the results generated by ASP software. The number and accuracy of tie points are used to quantitatively evaluate the performance of the proposed framework, where the accuracy is the back-projection residual calculated by using the exterior orientation parameters corrected through bundle adjustment [56].

C. Parameter Setting and Sensitivity Analysis

The methods mentioned above for qualitative and quantitative comparisons are all implemented based on publicly available codes released by the original authors. In order to achieve the best performance of these methods, extensive experiments were conducted to select the most suitable parameters. It should be noted that in the comparative experiments, the LMR method employs the training model provided by the authors due to the lack of sufficient data for retraining.

On the other hand, the parameters in our GLGP method that need to be set include: $\{\tau_j\}_{j=0}^{3}$ used to calculate the penalty value, the number of the neighbors K, and the threshold λ for differentiating between inliers and outliers. In order to analyze the sensitivity of GLGP for parameter setting, 20 image pairs are randomly chosen from the mentioned three datasets and tested with different parameters.

The results are illustrated in Fig. 6. The larger parameter τ indicates the higher tolerance of error. As the parameter K increases, a larger number of LMGPs are involved in the cost calculation. An excessive or insufficient number of LMGPs will lead to a reduction in the F-score. The increase in the number of neighbors causes many points that are far away to be selected as neighbors, the local geometric structure will become unstable, and the inaccurate judgment of the corresponding relationship between feature points will lead to a decrease in F-score. The results reveal a trend that the scores are initially increasing and subsequently stable for the six parameters. It can be concluded that the GLGP method is insensitive to parameters τ_0, τ_1, and τ_2, while changes in τ_3 relatively affect the performance of GLGP, yet the overall average F-score remains satisfactory. The results of parameter K indicates that selecting either too many or too few neighbors will affect the performance of GLGP, highlighting the importance of choosing an optimal number of neighbors. Regarding the final parameter, it is evident that when λ is set below 0.5, its variation has a minimal effect on the performance of GLGP. However, when λ is not less

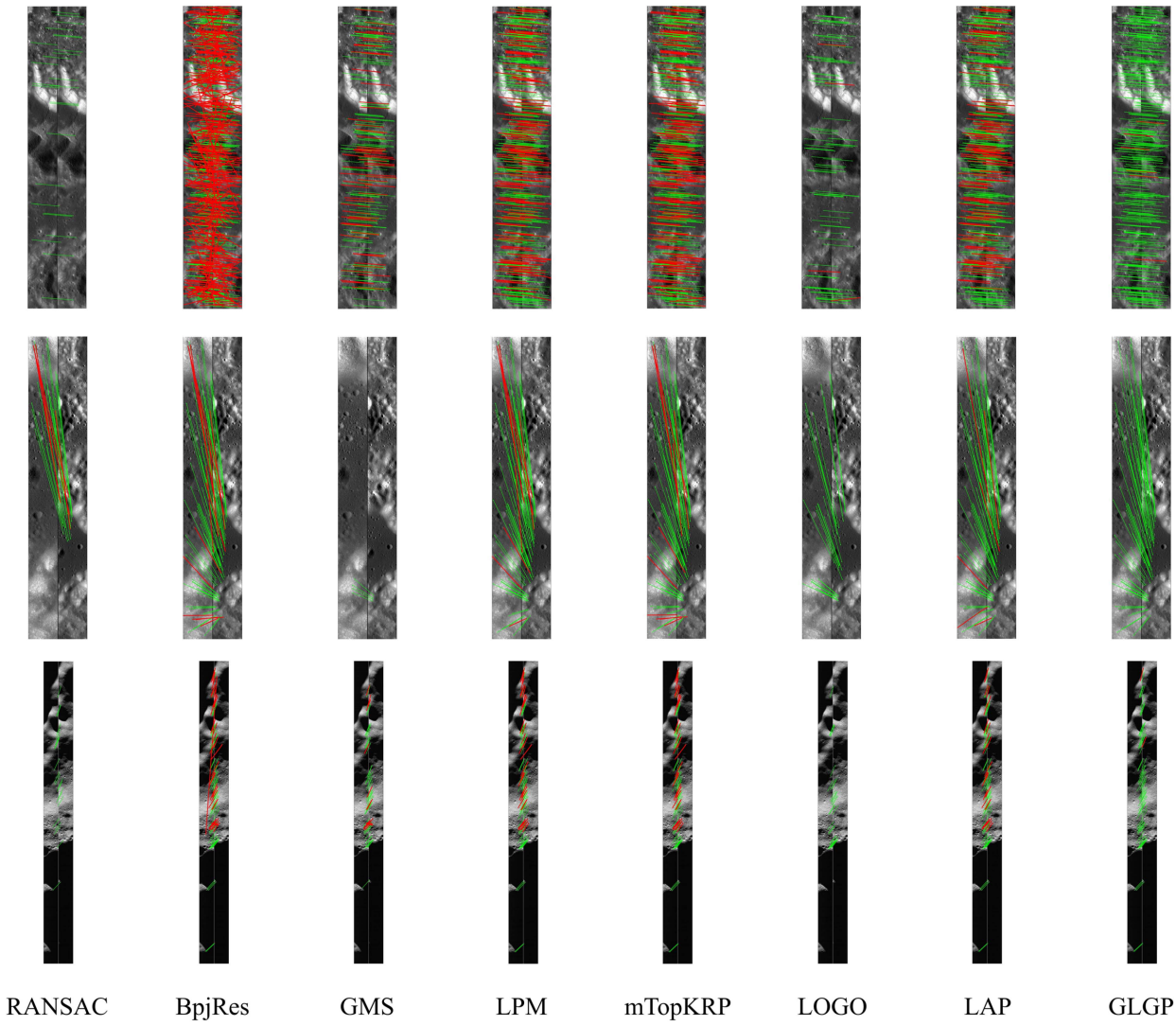

Fig. 7. Qualitative results of orbiter image pairs from three datasets. The red and green lines represent outliers and inliers, respectively.

than 0.5, the performance of GLGP decreases significantly, attributed to the relaxed threshold allowing the retention of mismatches with costs below the threshold. Nevertheless, GLGP still achieves a reasonable mean F-score. Overall, the variations of six parameters influence the performance of GLGP method to some extent, but the lowest F-score still exceeds 0.7, and all average F-scores surpass 0.9. These results indicate that GLGP exhibits a sound robustness to parameter variations. In the subsequent evaluation, the default values are empirically set as $(\tau_0,\ \tau_1,\ \tau_2,\ \tau_3) = (6,\ 3,\ 0.05,\ 30)$, $K = 6$, $\lambda = 0.3$. In addition, when quantifying the distribution of back-projection residuals, we set 200 pixels as the back-projection residual cutoff value empirically according to the positioning accuracy of LRO NAC imagery.

V. Experimental Results and Discussion

In this section, we comprehensively study the performance of the proposed GLGP method through qualitative and quantitative experiments, and discuss the experimental results in detail. Moreover, the effectiveness of the neighborhood construction and local geometric metrics in the GLGP is verified through ablation experiments. Finally, the feasibility and superiority of our developed framework for processing large-scale orbiter images are validated by the large-scale tie points extraction experiment.

A. Qualitative Experiments

To visually illustrate the effectiveness of our method, three sample orbiter image pairs from three datasets were selected for qualitative comparison. All these image pairs exhibit repetitive textures, and their putative matches are relatively sparse. The last two sets of image pairs also have obvious local distortions. As can be seen from the results in Fig. 7, in the first image pair, RANSAC [11] can effectively remove mismatches, but exhibits poor performance in retaining correct matches. BpjRes only considers the back-projection residuals, resulting in the worst performance. Other methods are unable to simultaneously remove outliers with similar motion orientations to inliers

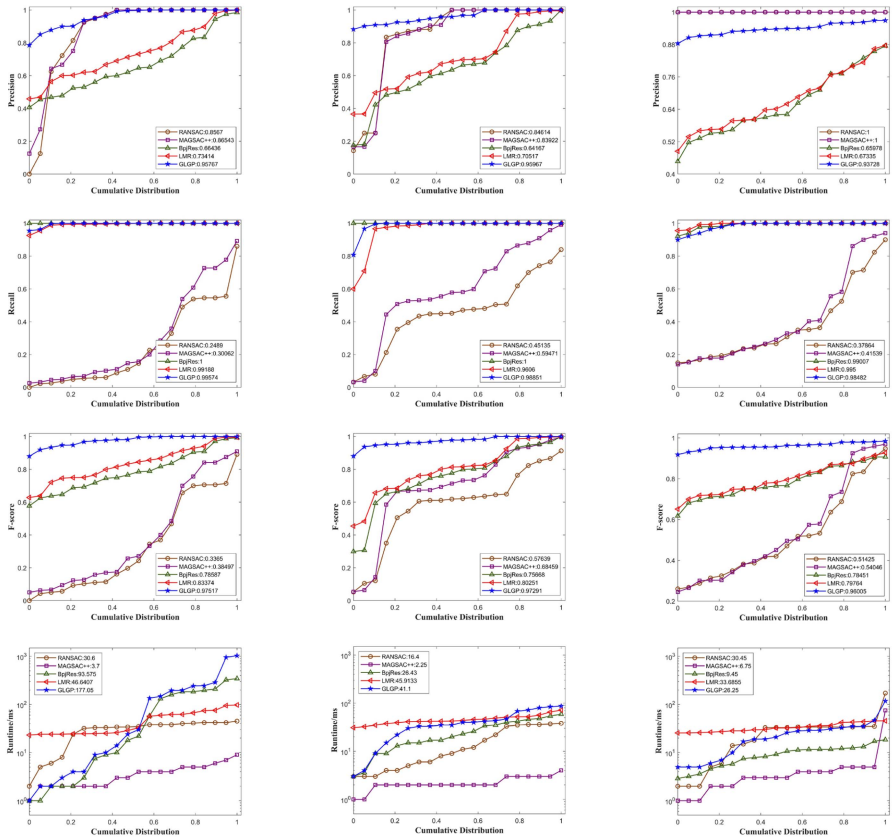

Fig. 8. Quantitative results of RANSAC, MAGSAC++, BpjRes, LMR, and GLGP on the three datasets. From (top) to (bottom), precision, recall, F-score and runtime (in milliseconds) with respect to the cumulative distribution.

while maintaining a high recall. In contrast, the GLGP method effectively preserves most of inliers, while only a very small number of outliers are not removed, achieving satisfactory result. The second image pair contains local geometric distortions. Compared to other low precision and low recall methods, GLGP removes all mismatches and preserve as many correct matches as possible. In the third image pair, the images contain areas without texture due to the influence of illumination, which results in sparse putative matches. GLGP can also preserve correct matches in this case. Other methods either preserve too many outliers or fail to preserve any inliers locally, leading to unsatisfactory results. The results demonstrate that GLGP is capable of effectively handling various kinds of complex orbiter image matching tasks, obtaining the results with high precision and recall.

B. Quantitative Evaluation

Fig. 8 shows the quantitative comparison results on three datasets of the proposed method with other global-based methods and learning-based method. The first to the last rows show the statistical results about precision, recall, F-score, and runtime, respectively. RANSAC [11] and MAGSAC++ [31] are methods based on resampling and global constraint. These global-based methods perform poorly when handling image pairs with large local geometric distortions. BpjRes approximately regards the back-projection residual as satisfying Gaussian distribution, and uses the three-sigma rule to remove mismatches. However, due to the imaging model errors and other random errors the mismatches cannot be accurately distinguished only by the back-projection residuals. LMR trains a classifier to remove mismatches, and like other learning-based methods, the training model is a primary factor limiting the performance of the method.

The experimental results indicate that there is little difference in precision between RANSAC and MAGSAC++. The latter improves the performance of RANSAC in terms of recall and F-score, and exhibits the shortest execution time. BpjRes method obtains the worst results compared with others in terms of precision, but are significantly better than other two global-based methods in terms of recall and even better than GLGP in some cases. This indicates that the BpjRes can retain as many correct matches as possible, but at the same time a large number of mismatches are also retained, leading to a poor F-score result. LMR performs comparably to BpjRes on three datasets, with slightly better precision but still falling short of requirements. It exhibits better recall, indicating its ability to identify more correct matches. However, due to the retention of many mismatches, the overall performance of this method is relatively poor. The reason for this is that there are differences in lunar and earth scenes, and the pretrained model lacks specificity for lunar orbiter images, leading to the degraded performance of this

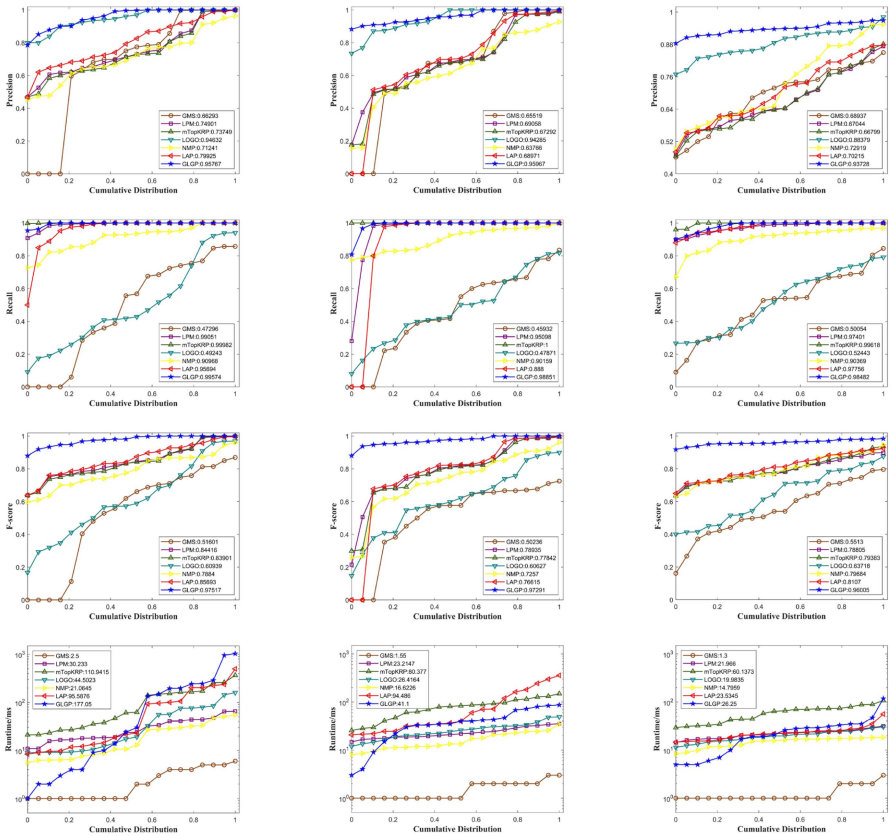

Fig. 9. Quantitative results of GMS, LPM, mTopKRP, LOGO, NMP, LAP, and GLGP on the three datasets. From (top) to (bottom), precision, recall, F-score and runtime (in milliseconds) with respect to the cumulative distribution.

method. The experimental result is consistent with the limitation analysis mentioned above. Since GLGP not only adopts back-projected residuals, but also appends local geometric constraints, it exhibits better performance than global-based methods in terms of precision and recall.

Fig. 9 shows the quantitative comparison results on three datasets of the proposed method with other local-based methods, including GMS [17], LPM [18], mTopKRP [48], LOGO [50], NMP [49], and LAP [43]. It can be seen from the results that GMS has the worst performance. LPM, mTopKRP, NMP, and LAP are locality consistency assumption-based methods and can achieve better performance than GMS. In terms of precision, there is little difference between these four methods. In terms of recall, compared with GMS, the performance of these four methods is improved, with mTopKRP even achieving the best results among all the comparison methods. However, since global geometric information is not considered, the performance is still worse compared with GLGP. LOGO adopts a method that combines the locality-guided strategy and global-preserving affinity matrix. The results of LOGO exhibit a high precision but low recall leading to unsatisfactory performance. The possible reason is LOGO still relies on the method of estimating the local affine transformation matrix in local constraints. When there is obvious local distortion in the image pair, the affine matrix cannot accurately distinguish the correct matches resulting in a lower recall of this method. GLGP imposes global constraints as neighbor screening criteria instead of defining the correctness of matching points. Locally, GLGP fully capitalizes on the characteristics of local geometric constraints, and therefore can obtains the satisfactory results. It is clear that GLGP ranks first in term of precision and F-score on all three datasets and only worse than mTopKRP in terms of recall on dataset 2. In terms of runtime, for the image size of 52 224 $\times$ 5064 pixels, it takes less than 180 ms to process each image pair, which is acceptable in practical applications. In the future, we will focus on improving the efficiency of the proposed method through parallel optimization of locally constrained steps.

C. Ablation Experiment and Discussion

Fig. 10 shows the statistical results about F-score of GLGP variants with different neighborhoods and local geometry or topology metrics.

From the results in Fig. 10(a), on one hand, it can be seen that using the same local geometry preservation metric, the variants with BpjRes and KNN exhibit the worst performance. These two methods construct neighborhoods using all outliers from the putative matches, leading to poor performance. Both variants have a mean F-score less than 0.5. The variant with MCN improves performance and its mean F-score is 0.75. It can achieve better results on some image pairs, but worse on others, especially image pairs with a high outlier ratio, which

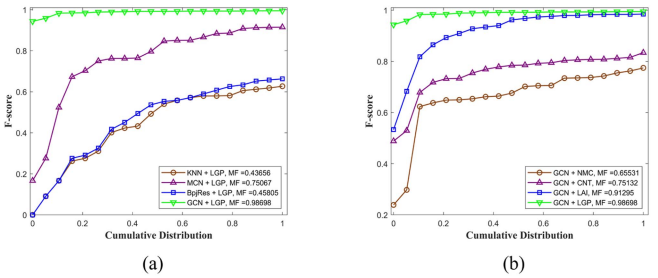

Fig. 10. Statistical results about F-score of GLGP variants with different neighborhoods and local geometry or topology metrics. (a) Results of various neighborhood construction methods combined with the same local geometry preservation metric. (b) Results of same neighborhood construction method combined with different local geometry or topology metrics.

indicates the performance of this variant is unstable. Compared with the results of the above three variants, the performance of the variant with GCN is significantly improved with the mean F-score better than 0.98. This clean neighborhood construction can supply a more stable and reliable structure for subsequent judgment of local geometry preservation.

On the other hand, the performance of variants with NMC and CNT are relatively poor based on the same neighborhood. From the results in Fig. 10(b), the mean F-score of variants with NMC and CNT are 0.66 and 0.75, respectively, which are not satisfactory. NMC constructs the metric based on the neighborhood consensus distance but does not take the directionality of edges into account, while the constraints imposed by CNT are relatively weaker. Due to the strong local affine invariance constraint, the LAI metric performs better, with a mean F-score of 0.91. In contrast, LGP combines the similarity constraint of the back-projection difference vectors with the local affine invariance constraint, leading to the best performance. In summary, the neighborhood construction method of GLGP provides a reliable neighborhood structure, and the local geometry preservation metric is able to distinguish inliers and outliers more accurately. Both aspects make significant contributions to the overall effectiveness of the GLGP method.

D. Large-Scale Multiview Matching Tie Points Extraction

Fig. 11 shows the multiview matching and tie points extraction results of ASP software and our framework on large-scale LRO NAC images within a region of about 100 km × 300 km near the Apollo-17 landing site. It is obvious from the comparison results that our framework can obtain more tie points for multiview images. ASP extracted 258 360 tie points, while our developed framework successfully extracted 940 883 tie points. The significant increase in the quantity demonstrates the enhanced capability of our framework in tie points extraction.

In terms of the reliability of tie points, 60 sets of image pairs were randomly selected and the back-projection residuals of the tie points were calculated using the adjustment-corrected external orientation parameters to evaluate the accuracy. As shown in Fig. 12, our results exhibit an average back-projection residual of less than 2 pixels, with the root mean square of the residuals also being less than 2 pixels. In contrast, the ASP

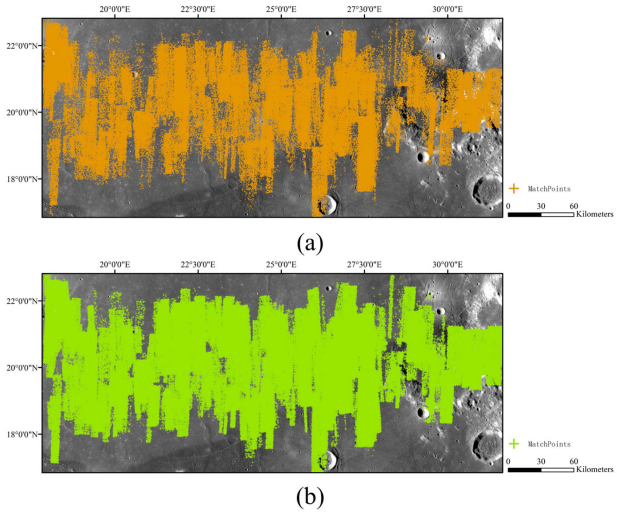

Fig. 11. Tie points extracted by (a) ASP and (b) our framework.

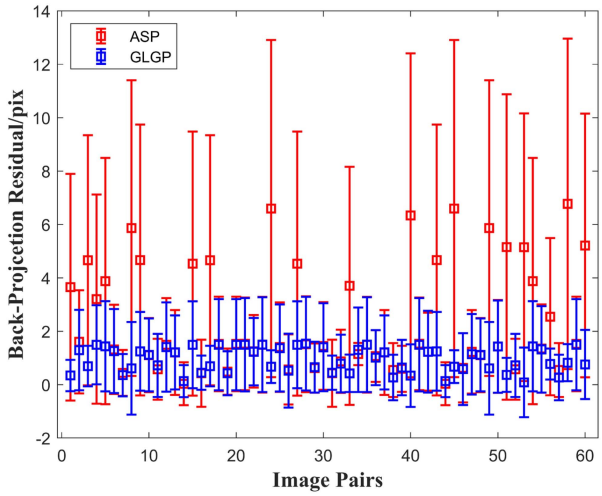

Fig. 12. Back-projection residuals after bundle adjustment calculated by sixty sample image pairs.

results significantly lag behind ours in accuracy. Therefore, it is concluded that the quantity and accuracy of tie points derived from our developed framework are superior to those obtained from the commonly used planetary mapping software, which can significantly benefit large-scale orbiter mapping endeavors.

VI. Conclusion

In this article, we propose a mismatch removal method based on global and local geometric constraints for orbiter images. Based on the consistency of global geometry and well-preserved local geometry of feature points, we formulate the mismatch removal into an optimization problem and derive the closed-form optimal solution. The global constraint in the model uses the back-projection residuals to construct a clean neighborhood. The local geometric constraints integrate the local affine invariance with the similarity of BDVs, to distinguish the inliers and outliers. Experiments on lunar orbiter image datasets in different

regions demonstrate that our method outperforms the recent advanced methods. Meanwhile, we extend a large-scale orbiter images multiview matching framework to obtain high-precision tie points. According to the result of large-scale data processing, our framework has demonstrated its ability to surpasses the performance of current mainstream planetary mapping software especially in terms of the quantity and accuracy of tie points.

This article primarily validates the proposed method through lunar orbiter images, yet the methodology is also capable of processing images from other orbiters. Nevertheless, the exterior orientation parameters' errors of the orbiter images may impose a limitation on GLGP performance. In future, the proposed geometric constraints could be improved to reduce the impact of imaging model errors, and more validations on other different orbiter images will be explored.

ACKNOWLEDGMENT

The authors would like to thank the authors of the related comparison methods mentioned in this study for providing the implementation codes.

REFERENCES

[1] H. Chen et al., "CNN-based large area pixel-resolution topography retrieval from single-view LROC NAC images constrained with SLDEM," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 9398–9416, Oct. 2022, doi: 10.1109/JSTARS.2022.3214926.

[2] G. Wan et al., "Multimodal remote sensing image matching based on weighted structure saliency feature," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 4700816, doi: 10.1109/TGRS.2023.3347259.

[3] H. Li, X. Zheng, M. Dong, G. S. Xia, and H. Xiong, "Locally nonlinear affine verification for multisensor image matching," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5617016, doi: 10.1109/TGRS.2021.3139429.

[4] S. P. Ji, C. Zeng, Y. J. Zhang, and Y. L. Duan, "An evaluation of conventional and deep learning-based image-matching methods on diverse datasets," *Photogrammetric Rec.*, vol. 38, no. 182, pp. 137–159, Jun. 2023, doi: 10.1111/phor.12445.

[5] Q. Fu et al., "A GPU-accelerated PCG method for the block adjustment of large-scale high-resolution optical satellite imagery without GCPs," *Photogrammetric Eng. Remote Sens.*, vol. 89, no. 4, pp. 211–220, Apr. 2023, doi: 10.14358/pers.22-00051r2.

[6] B. Wu, H. Hu, and J. Guo, "Integration of chang'E-2 imagery and LRO laser altimeter data with a combined block adjustment for precision lunar topographic modeling," *Earth Planet. Sci. Lett.*, vol. 391, pp. 1–15, Apr. 2014, doi: 10.1016/j.epsl.2014.01.023.

[7] Z. Ye, Y. S. Xu, H. Chen, J. W. Zhu, X. H. Tong, and U. Stilla, "Area-based dense image matching with subpixel accuracy for remote sensing applications: Practical analysis and comparative study," *Remote Sens.*, vol. 12, no. 4, Feb. 2020, Art. no. 696, doi: 10.3390/rs12040696.

[8] H. Albanwan and R. Qin, "A comparative study on deep-learning methods for dense image matching of multi-angle and multi-date remote sensing stereo-images," *Photogrammetric Rec.*, vol. 37, no. 180, pp. 385–409, 2022, doi: 10.1111/phor.12430.

[9] X. Tong et al., "Image registration with fourier-based image correlation: A comprehensive review of developments and applications," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 10, pp. 4062–4081, Oct. 2019, doi: 10.1109/JSTARS.2019.2937690.

[10] M. Chen, R. J. Qin, H. Q. He, Q. Zhu, and X. Wang, "A local distinctive features matching method for remote sensing images with repetitive patterns," *Photogrammetric Eng. Remote Sens.*, vol. 84, no. 8, pp. 513–524, Aug. 2018, doi: 10.14358/pers.84.8.513.

[11] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," in *Readings in Computer Vision*, M. A. Fischler and O. Firschein, Eds. San Mateo, CA, USA: Morgan Kaufmann, 1987, pp. 726–740.

[12] P. H. S. Torr and D. W. Murray, "The development and comparison of robust methods for estimating the fundamental matrix," *Int. J. Comput. Vis.*, vol. 24, no. 3, pp. 271–300, 1997, doi: 10.1023/A:1007927408552.

[13] J. Salvi, X. Armangue, and J. Pages, "A survey addressing the fundamental matrix estimation problem," in *Proc. Int. Conf. Image Process.*, 2001, vol. 2, pp. 209–212, doi: 10.1109/ICIP.2001.958461.

[14] J. Li, Q. Hu, M. Ai, and S. Wang, "A geometric estimation technique based on adaptive M-estimators: Algorithm and applications," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 5613–5626, May 2021, doi: 10.1109/JSTARS.2021.3078516.

[15] J. Ma, J. Zhao, J. Tian, A. L. Yuille, and Z. Tu, "Robust point matching via vector field consensus," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1706–1721, Apr. 2014, doi: 10.1109/tip.2014.2307478.

[16] J. Ma, H. Zhou, J. Zhao, Y. Gao, J. Jiang, and J. Tian, "Robust feature matching for remote sensing image registration via locally linear transforming," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 12, pp. 6469–6481, Dec. 2015, doi: 10.1109/tgrs.2015.2441954.

[17] J. Bian, W.-Y. Lin, Y. Matsushita, S.-K. Yeung, T.-D. Nguyen, and M.-M. Cheng, "GMS: Grid-based motion statistics for fast, ultra-robust feature correspondence," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 2828–2837, doi: 10.1109/cvpr.2017.302.

[18] J. Ma, J. Zhao, J. Jiang, H. Zhou, and X. Guo, "Locality preserving matching," *Int. J. Comput. Vis.*, vol. 127, no. 5, pp. 512–531, 2019, doi: 10.1007/s11263-018-1117-z.

[19] J. Y. Li, Q. W. Hu, and M. Y. Ai, "LAM: Locality affine-invariant feature matching," *ISPRS J. Photogrammetry Remote Sens.*, vol. 154, pp. 28–40, Aug. 2019, doi: 10.1016/j.isprsjprs.2019.05.006.

[20] Y. Wan and Y. J. Zhang, "The P2L method of mismatch detection for push broom high-resolution satellite images," *ISPRS J. Photogrammetry Remote Sens.*, vol. 130, pp. 317–328, Aug. 2017, doi: 10.1016/j.isprsjprs.2017.06.009.

[21] R. Gupta and R. I. Hartley, "Linear pushbroom cameras," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 9, pp. 963–975, Sep. 1997, doi: 10.1109/34.615446.

[22] A. Sedaghat, M. Mokhtarzade, and H. Ebadi, "Uniform robust scale-invariant feature matching for optical remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4516–4527, Nov. 2011, doi: 10.1109/TGRS.2011.2144607.

[23] Y. Wu et al., "A two-step method for remote sensing images registration based on local and global constraints," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 5194–5206, May 2021, doi: 10.1109/jstars.2021.3079103.

[24] T.-A. Teo, L.-C. Chen, C.-L. Liu, Y.-C. Tung, and W.-Y. Wu, "DEM-aided block adjustment for satellite images with weak convergence geometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 4, pp. 1907–1918, Apr. 2010, doi: 10.1109/TGRS.2009.2033935.

[25] Z. Ye et al., "Illumination-robust subpixel Fourier-based image correlation methods based on phase congruency," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 1995–2008, Apr. 2019, doi: 10.1109/tgrs.2018.2870422.

[26] B. Wu, H. Zeng, and H. Hu, "Illumination invariant feature point matching for high-resolution planetary remote sensing images," *Planet. Space Sci.*, vol. 152, pp. 45–54, Mar. 2018, doi: 10.1016/j.pss.2018.01.007.

[27] R. A. Beyer, O. Alexandrov, and S. McMichael, "The ames stereo pipeline: NASA's open source software for deriving and processing terrain data," *Earth Space Sci.*, vol. 5, no. 9, pp. 537–548, Sep. 2018, doi: 10.1029/2018ea000409.

[28] P. H. S. Torr and A. Zisserman, "MLESAC: A new robust estimator with application to estimating image geometry," *Comput. Vis. Image Understanding*, vol. 78, no. 1, pp. 138–156, Apr. 2000, doi: 10.1006/cviu.1999.0832.

[29] O. Chum and J. Matas, "Matching with PROSAC - progressive sample consensus," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2005, vol. 1, pp. 220–226, doi: 10.1109/CVPR.2005.221.

[30] R. Raguram, O. Chum, M. Pollefeys, J. Matas, and J.-M. Frahm, "USAC: A universal framework for random sample consensus," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 2022–2038, Aug. 2013, doi: 10.1109/tpami.2012.257.

[31] D. Baráth, J. Noskova, M. Ivashechkin, and J. Matas, "MAGSAC++, a fast, reliable and accurate robust estimator," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 1301–1309, doi: 10.1109/CVPR42600.2020.00138.

[32] X. R. Li and Z. Y. Hu, "Rejecting mismatches by correspondence function," *Int. J. Comput. Vis.*, vol. 89, no. 1, pp. 1–17, Aug. 2010, doi: 10.1007/s11263-010-0318-x.

[33] K. M. Yi, E. Trulls, Y. Ono, V. Lepetit, M. Salzmann, and P. Fua, "Learning to find good correspondences," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 2666–2674, doi: 10.1109/cvpr.2018.00282.
[34] J. Ma, X. Jiang, J. Jiang, J. Zhao, and X. Guo, "LMR: Learning a two-class classifier for mismatch removal," *IEEE Trans. Image Process.*, vol. 28, no. 8, pp. 4045–4059, Aug. 2019, doi: 10.1109/tip.2019.2906490.
[35] J. Chen, S. Chen, Y. Liu, X. Chen, Y. Yang, and Y. Zhang, "Robust local structure visualization for remote sensing image registration," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 1895–1908, Jan. 2021, doi: 10.1109/JSTARS.2021.3050459.
[36] N. Mohammadi, A. Sedaghat, and M. Jodeiri Rad, "Rotation-invariant self-similarity descriptor for multi-temporal remote sensing image registration," *Photogrammetric Rec.*, vol. 37, no. 177, pp. 6–34, 2022, doi: 10.1111/phor.12402.
[37] M. Cho and K. M. Lee, "Mode-seeking on graphs via random walks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2012, pp. 606–613, doi: 10.1109/CVPR.2012.6247727.
[38] C. Wang, L. Wang, and L. Q. Liu, "Progressive mode-seeking on graphs for sparse feature matching," in *Proc. Eur. Conf. Comput. Vis.*, 2014, vol. 8690, pp. 788–802.
[39] L. Torresani, V. Kolmogorov, and C. Rother, "Feature correspondence via graph matching: Models and global optimization," in *Proc. Eur. Conf. Comput. Vis.*, 2008, vol. 5303, Art. no. 596.
[40] H. Liu and S. Yan, "Common visual pattern discovery via spatially coherent correspondences," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2010, pp. 1609–1616, doi: 10.1109/CVPR.2010.5539780.
[41] J. Y. Li, Q. W. Hu, and M. Y. Ai, "4FP-structure: A robust local region feature descriptor," *Photogrammetric Eng. Remote Sens.*, vol. 83, no. 12, pp. 813–826, Dec. 2017, doi: 10.14358/pers.83.12.813.
[42] J. Li, Q. Hu, M. Ai, and R. Zhong, "Robust feature matching via support-line voting and affine-invariant ratios," *ISPRS J. Photogrammetry Remote Sens.*, vol. 132, pp. 61–76, 2017, doi: 10.1016/j.isprsjprs.2017.08.009.
[43] X. Y. Ye, J. Y. Ma, and H. L. Xiong, "Local affine preservation with motion consistency for feature matching of remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5613912, doi: 10.1109/tgrs.2021.3128292.
[44] J. Ma, J. Jiang, H. Zhou, J. Zhao, and X. Guo, "Guided locality preserving feature matching for remote sensing image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4435–4447, Aug. 2018, doi: 10.1109/tgrs.2018.2820040.
[45] L. Shen, Q. Xin, J. H. Zhu, X. T. Huang, and T. Jin, "Frame-based locality preservation matching for images involving large-scale transformations," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Jul. 2022, Art. no. 5625813, doi: 10.1109/tgrs.2022.3187842.
[46] X. Y. Jiang, Y. F. Xia, X. P. Zhang, and J. Y. Ma, "Robust image matching via local graph structure consensus," *Pattern Recognit.*, vol. 126, Jun. 2022, Art. no. 108588, doi: 10.1016/j.patcog.2022.108588.
[47] J. Jiang, Q. Ma, T. Lu, Z. Wang, and J. Ma, "Feature matching based on top K rank similarity," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2018, pp. 2316–2320, doi: 10.1109/ICASSP.2018.8461504.
[48] X. Jiang, J. Jiang, A. Fan, Z. Wang, and J. Ma, "Multiscale locality and rank preservation for robust feature matching of remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6462–6472, Sep. 2019, doi: 10.1109/tgrs.2019.2906183.
[49] X. Ye and J. Ma, "Neighborhood manifold preserving matching for visual place recognition," *IEEE Trans. Ind. Inform.*, vol. 19, no. 7, pp. 8127–8136, Jul. 2023, doi: 10.1109/tii.2022.3216619.
[50] Y. F. Xia and J. Y. Ma, "Locality-guided global-preserving optimization for robust feature matching," *IEEE Trans. Image Process.*, vol. 31, pp. 5093–5108, Jul. 2022, doi: 10.1109/tip.2022.3192993.
[51] G. Slabaugh, R. Schafer, and M. Livingston, "Optimal ray intersection for computing 3D points from n-view correspondences," *Deliverable Rep.*, pp. 1–11, Feb. 2001.
[52] Z. Ye et al., "Estimation and analysis of along-track attitude jitter of ZiYuan-3 satellite based on relative residuals of tri-band multispectral imagery," *ISPRS J. Photogrammetry Remote Sens.*, vol. 158, pp. 188–200, 2019, doi: 10.1016/j.isprsjprs.2019.10.012.
[53] M. Wang, B. Yang, F. Hu, and X. Zang, "On-orbit geometric calibration model and its applications for high-resolution optical satellite imagery," *Remote Sens.*, vol. 6, no. 5, pp. 4391–4408, May 2014, doi: 10.3390/rs6054391.
[54] V. R. Jakkula, "Efficient feature detection using OBAloG: Optimized box approximation of Laplacian of Gaussian," Ph.D. dissertation, Dept. Elect. Comput. Eng., Kansas State Univ., Manhattan, KS, USA, 2010.
[55] M. S. Robinson et al., "Lunar reconnaissance orbiter camera (LROC) instrument overview," *Space Sci. Rev.*, vol. 150, no. 1–4, pp. 81–124, Jan. 2010, doi: 10.1007/s11214-010-9634-2.
[56] C. Chen et al., "Large-scale block bundle adjustment of LROC NAC images for lunar south pole mapping based on topographic constraint," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 17, pp. 2731–2746, 2024, doi: 10.1109/JSTARS.2023.3346199.

Dayong Liu received the B.E. degree in surveying and mapping engineering from the China University of Mining and Technology, Xuzhou, China, in 2017. He is currently working toward the Ph.D. degree in the College of Surveying and Geoinformatics, Tongji University, Shanghai, China.

His research interests include photogrammetry and planetary 3-D mapping.

Zhen Ye (Member, IEEE) received the Ph.D. degree in cartography and geoinformation from Tongji University, Shanghai, China, in 2018.

He was a Postdoctoral Researcher with the Chair of Photogrammetry and Remote Sensing, Technical University of Munich, Munich, Germany, from 2018 to 2020. He is currently an Associate Professor with the College of Surveying and Geoinformatics, Tongji University. His research interests include photogrammetry and remote sensing, high-resolution satellite image processing and planetary 3-D mapping.

Yusheng Xu was born in 1989. He received the B.S. and M.E. degrees in surveying and mapping from Tongji University, Shanghai, China, in 2011 and 2014, respectively, and the Ph.D. (Dr.-Ing., *summa cum laude*) degree in photogrammetry and the habilitation (*venia legendi*) in point cloud analysis from the Technical University of Munich (TUM), Munich, Germany, in 2019 and 2022, respectively.

From 2019 to 2022, he was a Lecturer with TUM. Since May 2022, he joined Tongji University as a Full Professor. He has authored and coauthored more than 60 entries. His research interests include 3-D point cloud processing, image processing, and spaceborne photogrammetry.

Rong Huang (Member, IEEE) received the Ph.D. degree in photogrammetry and remote sensing from the Technical University of Munich, Munich, Germany, in 2021.

She was a Postdoctoral Research Fellow with the Hong Kong Polytechnic University, Hong Kong, from 2021 to 2022. She is currently an Assistant Professor with the College of Surveying and Geoinformatics, Tongji University, Shanghai, China. Her research interests include photogrammetry and remote sensing, point cloud processing, and 3-D mapping.

Lin Xue received the B.S. degree in geographic information science from Southwest Jiaotong University, Chengdu, China, in 2017, and the M.E. degree in surveying and mapping engineering from Wuhan University, Wuhan, China, in 2019. He is currently working toward the Ph.D. degree in the College of Surveying and Geoinformatics, Tongji University, Shanghai, China.

His research interests include photogrammetry and planetary 3-D mapping.

Hao Chen received the M.Sc. degree in geographic information system from Tongji University, Shanghai, China, in 2020.

From 2020 to 2021, he was a Research Assistant with the College of Surveying and Geoinformatics, Tongji University. His research interests include spatial data processing, planetary mapping (mainly for lunar surface and small bodies), and deep learning.

Genyi Wan received the B.E. degree in geographical information sciences from the Suzhou University of Science and Technology, Suzhou, China, in 2020, and the M.E. degree in surveying and mapping engineering from Capital Normal University, Beijing, China, in 2023. He is currently working toward the Ph.D. degree in the College of Surveying and Geoinformatics, Tongji University, Shanghai, China.

His research interests include planetary remote sensing image matching and 3-D mapping.

Huan Xie (Senior Member, IEEE) received the B.S. degree in surveying engineering and the M.S. and Ph.D. degrees in cartography and geoinformation from Tongji University, Shanghai, China, in 2003, 2006, and 2009, respectively.

From 2007 to 2008, she was a Visiting Scholar with the Institute of Photogrammetry and Geoinformation, Leibniz Universität Hannover, Hannover, Germany. She is currently a Professor with the College of Surveying and Geo-Informatics, Tongji University. Her research interests include satellite laser altimetry and hyperspectral remote sensing.

Xiaohua Tong (Senior Member, IEEE) received the Ph.D. degree in geographic information system from Tongji University, Shanghai, China, in 1999.

Between 2001 and 2003, he was a Postdoctoral Researcher with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, China. He was a Research Fellow with Hong Kong Polytechnic University, Hong Kong, in 2006, and a Visiting Scholar with the University of California, Santa Barbara, CA, USA, between 2008 and 2009. He is currently a Professor with the College of Surveying and Geoinformatics, Tongji University. His research interests include remote sensing, geographic information system, uncertainty and spatial data quality, and image processing for high-resolution and hyperspectral images.