Fast Double-Channel Aggregated Feature Transform for Matching Planetary Remote Sensing Images

Rong Huang [ID] , *Member, IEEE*, Genyi Wan [ID], Yingying Zhou [ID], Zhen Ye [ID], *Member, IEEE*, Huan Xie [ID], *Senior Member, IEEE*, Yusheng Xu [ID], *Member, IEEE*, and Xiaohua Tong [ID], *Senior Member, IEEE*

Abstract—**Matching planetary remote sensing images (PRSI) is essential for deep space exploration. Through matching optical images collected by different probes, the terrain of planets can be accurately mapped; however, PRSI lacks surface texture information and exhibits noticeable nonlinear radiation differences, e.g., illumination differences. It is not easy to achieve satisfactory results through traditional matching methods. In order to cope with the above problems, a new PRSI matching method has been proposed. It can extract low–high frequency features by building double-frequency scale space. In addition, we also use the nonmaximum suppression strategy for rejecting overlapped feature points, which reduces the time consumption and improves the matching accuracy. The experimental results show that the proposed method can effectively match PRSI and be superior to comparable methods.**

Index Terms—**Feature extraction, image matching, nonlinear radiation distortions, planetary remote sensing images (PRSI).**

I. INTRODUCTION

IMAGE matching is a fundamental task in 3-D reconstruction of planetary surfaces using multisource planetary remote sensing images (PRSI), serving as a prerequisite for many applications, such as crater detection [1], planetary navigation [2], and scientific exploration [3]. It aims to detect reliable geometric relationships in images from different times, sensors, and perspectives [4].

According to the utilized feature type, the matching of multisource remote sensing images (MRSI) can be divided into three types: 1) area-based, 2) feature-based, and 3) learning-based methods [5]. Area-based methods often use templates to search for the most similar parts of the whole image. These templates can be built from the spatial and frequency domain. The area-based methods are capable of achieving high matching accuracy. However, it is time-consuming and sensitive to scale

changes [6]. The feature-based methods match images using points [4] or geometric primitive, such as lines [7] and regions [8]. Compared with the area-based methods, the feature-based methods are more robust. Traditional feature matching methods, such as SIFT [9] and SURF [10] are widely used. However, for MRSI, apparent nonlinear radiation differences (NRDs) often occur, which leads to the invalidation of these features aiming at capturing grayscale changes. Therefore, researchers intend to achieve robust matching of MRSI by constructing descriptors over the frequency domain, such as RIFT [11] and HAPCG [12].

Moreover, the saliency structure of images is also utilized to match PRSI, e.g., CoFSM [13]. In recent years, in the field of matching, learning-based methods have also attracted attention [14], [15], [16], [17], [18]. Wu et al. [18] utilized both low-level geometric information and high-level context-aware information to improve matching accuracy. Zhang et al. [14] adopted the deep convolutional neural network (CNN) to learn the features of PRSI and achieved matching through the Delaunay triangulation strategy. Although learning-based methods have achieved promising results in some specific tasks, the type and number of datasets limit them and still lack generalization ability.

Although the methods mentioned above have achieved promising performance in image matching, there are still some limitations in the matching of PRSI. Compared with RS images acquired in Earth observation, the textures of PRSI are much weaker, and the differences in illumination are more serious. To solve the problems of NRDs and weak texture in PRSI, we propose a combined low–high frequency feature matching algorithm called fast double-channel aggregated feature transform (FDAFT) by capturing the information of low-frequency and high-frequency images. In addition, we match the descriptors from the double channels and integrate the matching results into the fast sample consensus (FSC) [19] to eliminate false matches to improve the matching accuracy. Fig. 1 illustrates the matching results of the FDAFT and the traditional method, such as PSO-SIFT [20], on a pair of planetary images. The significant contributions of this article are summarized as follows.

1) We propose a novel double-channel feature transform strategy for extracting low-frequency structure feature points and high-frequency phase feature points and aggregating them for feature matching.
2) A fast double-frequency scale space construction method is proposed to reduce the influence of PRSI scale changes
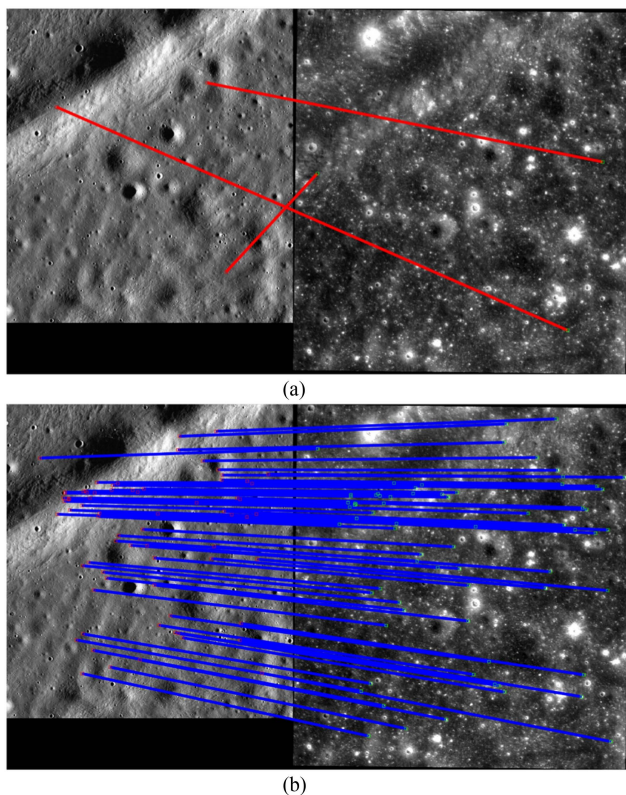
Fig. 1. Matching results of FDAFT and PSO-SIFT on sample planetary images. (a) Matching results of PSO-SIFT. (b) Matching results of FDAFT.

by using edge confidence maps (ECM) [21] and extracting phase features.

3) The nonmaximum suppression strategy and the FSC are involved in rejecting overlapped feature points and eliminating false matches, which improves the matching efficiency.

II. RELATED WORK

In this section, we review the development of the PRSI matching methods. The details are reviewed as follows.

A. Area-Based Methods

Area-based methods use a preset template to move within the window to be matched, thus comparing the similarity of features. In addition, in order to address the problems of rotation and scales in matching, the methods often use geographic reference coordinates for initial calibration to improve the matching accuracy [22]. This method of initial calibration based on geographic reference coordinates is widely used in matching optical-SAR images [8], [22], [23]. The key for area-based methods is establishing reliable similarity measures from the spatial or frequency domain instead of finding corresponding features from two images. Previous studies often used image grey information to match images, such as the sum of squared differences (SSD), normalized cross-correlation (NCC), and mutual information (MI). However, due to the existence of sensor differences and illumination differences in PRSIs, these

methods based on spatial domain features have proved difficult to apply to the images with the above problems [24]. Different from the grey change relationship of images captured from the spatial domain, similarity measurement built in the frequency domain can better reflect the structural features of images [25], [26], [27], [28]. Pedrosa et al. [25] constructs matching templates through fast Fourier transform (FFT) and calculates the probability volume between the templates and the images, thus realizing the detection of Mars impact craters. Although the method based on phase templates can deal with NRD and contrast between images well, it is less effective at rotated and scaled images and has certain limitations.

B. Feature-Based Methods

The feature-based methods extract similar features from two images, which can be point features, line features, surface features, etc. Traditional feature matching methods mainly include feature extraction [29], [30], [31], [32], construction of similarity measure [9], [10], [11], and removal of mismatched features [19], [33], [34]. Since the texture of PRSI is sparse and the structure is not apparent, most of the work focuses on the extraction and description of point features [24], [35]. Although there have been many matching methods in ground photogrammetry and computer vision, such as SIFT, SURF, ORB, etc., these methods based on image gray feature mainly focus on matching between homologous images. Because of the illumination differences of PRSI, the effect of using these methods is inferior. In order to solve the illumination variation existing in the image, Wu et al. [24] observed the double peaks in the dominant direction histogram of SIFT key points and proposed the use of adaptive suppression Gaussian function to level the histogram to reduce the illumination difference caused by the sun angle change. Zhang et al. [26] used the DEM of orbit image as an auxiliary image to simulate the change of solar illumination with the Lambert model. The most similar query image is obtained through the correspondence between the DEM image and the track image.

However, both Wu et al. [24] and Zhang et al. [26] 's methods are essentially weighted in the spatial domain to achieve the registration of PRSIs. Considering the sensitivity of spatial domain features (texture) in multimodal image matching, scholars use structure features to deal with NRDs. Li et al. [11] used to use the maximum phase index map (MIM) to address the NRDs between images. This method introduced phase features instead of image gradients, effectively solving the NRDs between heterogeneous images. Since then, the use of phase features to process MRSI has become mainstream due to the light invariance and NRD resistance of phase features [36]. For example, Xiang et al. [37] realized the registration of optical-SAR image through the global–local registration method and optimization of the phase congruency (PC) model. Yang et al. [38] designed a semiautomatic registration method for Chang'e-1 IIM images calculating NCC in the frequency domain. These methods promote the process of MRSI matching. Although the phase-based matching method has obtained satisfactory results in many applications, due to the weak texture characteristics of PRSI, it is difficult
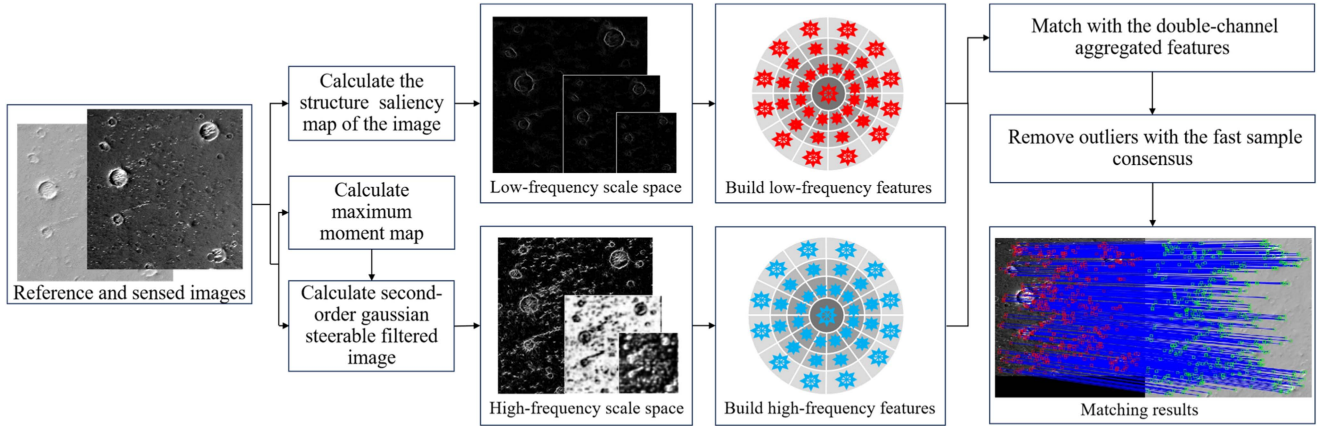
Fig. 2. Framework of the proposed FDAFT method.

to find the corresponding feature points by describing the local structure of the phase feature (the phase feature is mainly a description of the local structure of the image [39]). Recently, Yao et al. [13], Li et al. [4], and Wan et al. [40] have reduced the impact of NRD by calculating the global structure of images and using improved gradient features. This strategy of using global feature matching brings a new idea to solve the problem of weak texture in PRSI.

C. Deep Learning-Based Methods

In recent years, with the rapid development of deep learning, more and more learning-based methods have been used to match PRSI [2], [41], [42], [43]. Zhong et al. [41] proposed a feature detection and description method for planetary images, which obtained a sparse and reliable set of feature points by learning the depth features of images, called robust planetary features (RPFeat). Li et al. [42] used the CNN model (Net-model) to extract local features and global descriptors and use the nearest neighbor matching algorithm to match feature points. In addition, in order to accurately map the topography of Mars, Li et al. [2] used Superpoint [44] to extract basic shape features (triangles, cubes, checkerboard, and stars), and used SuperGlue [45] to complete the matching of narrow overlapping regions between adjacent CCD images. Compared with traditional algorithms, the deep learning method has the advantages of automatically optimizing parameters and constructing required descriptors [14]. However, because the deep learning method relies on a large amount of data [46] or time [47] for training, it is relatively difficult to construct PRSI datasets, and the application situation is complex and variable. As a result, learning-based approaches may be less transferable. In order to obtain stronger generalization ability and better applicability, the registration method based on deep learning still needs further research.

III. METHODOLOGY

Fig. 2 shows the framework of the proposed FDAFT method. The framework includes the construction of double-frequency scale space, feature point extraction and description, and false matches elimination.

A. Double-Frequency Scale Space

Although PRSI collected from different orbiters exists as obvious NRDs, all the structural features remain consistent. Therefore, we capture structural features (low-frequency features) from the images to match PRSI. In addition, scholars have found that phase features (high-frequency features) can better reflect the detail changes between images [11]. Thus, phase features can be enhanced by using phase information. We established low-frequency scale space and high-frequency scale space, respectively, from PRSI to capture the images' structural shape and phase details.

1) Low-Frequency Scale Space Construction: The ECM [21] of the image based on machine learning is used to calculate the structure features of each image. Besides, the Gaussian filter is applied to the ECM to construct low-frequency scale space. The low-frequency scale space is divided into N layers, and the scale of each layer is defined as follows:

$$\sigma_{s_n} = \sigma_{s_0} \cdot \sqrt[3]{2^n}, \ (n = 01, \ldots, N) \tag{1}$$

where σ_{s_n} represents the scale of the first layer, σ_{s_0} represents the scale of the nth layer, and N is the number of layers in the low-frequency scale space.

For low-frequency scale space, we need to define the filter window of each layer. The filter window of each layer can be calculated from the scale of the current layer and the filter window of the first layer. The definition is shown as

$$\begin{cases} \mathrm{FW}_n = \frac{\sigma_{s_n}^2 \cdot NO}{2} \\ L - \mathrm{FSpace} = \left\{ \mathrm{FW}_n \cdot J_p^n \right\}_{n=0}^{N} \end{cases} \tag{2}$$

where $L - \mathrm{FSpace}$ represents the low-frequency scale space, FW_n indicates the size of the filter window in the nth layer, NO is the initial size of the filter window, and J_p^n is the nth layer in $L - \mathrm{FSpace}$.

2) High-Frequency Scale Space Construction: Similar to the construction of low-frequency scale space, high-frequency scale space first extracts phase information from PRSI and then performs Gaussian blur. The maximum moment map ($M_{\max}$) of phase information shows the significance of image features and can reflect change details. FDAFT uses weighted PC (PC_w) [29]

to calculate $M_{\max}$

$$M_{\max} = \frac{1}{2} \left(c + a + \sqrt{b^2 + (a-c)^2} \right) \quad (3)$$

where $M_{\max}$ is the maximum moment map; a, b, and c are the three intermediate quantities, and the definition of a, b, and c are shown as

$$\begin{cases} a = \sum_{\theta} \left(pc_w\left(\theta\right) \cos\left(\theta\right) \right)^2 \\ b = 2 \sum_{\theta} \left(pc_w\left(\theta\right) \cos\left(\theta\right) \right) \left(pc_w\left(\theta\right) \sin\left(\theta\right) \right) \\ c = \sum_{\theta} \left(pc_w\left(\theta\right) \sin\left(\theta\right) \right)^2 . \end{cases} \quad (4)$$

By collecting texture information of images from different directions, the second-order steerable Gaussian filter can effectively process NRDs and enhance texture features in MRSI [22]. In this article, we use the second-order Gaussian steerable filter to further process $M_{\max}$ and construct $G_{2,\sigma}^{\theta} - M_{\max}$. The definition of the second-order Gaussian steerable filter is shown as

$$\begin{cases} G_{2,\sigma}^{0°} = G_{xx} = \left(-\frac{1}{2\pi\sigma^4}\right)\left(1 - \frac{x^2}{\sigma^2}\right) e^{\frac{-(x^2+y^2)}{2\sigma^2}} \\ G_{2,\sigma}^{90°} = G_{yy} = \left(-\frac{1}{2\pi\sigma^4}\right)\left(1 - \frac{y^2}{\sigma^2}\right) e^{\frac{-(x^2+y^2)}{2\sigma^2}} \\ G_{xy} = \frac{xy}{2\pi\sigma^6} e^{\frac{-(x^2+y^2)}{2\sigma^2}}, G_{2,\sigma}^{60°} = G_{yy} - G_{xy}, \ G_{2,\sigma}^{120°} = G_{yy} + G_{xy} \\ G_{2,\sigma}^{\theta} = \cos^2\left(\theta\right) G_{2,\sigma}^{0°} + \sin^2\left(\theta\right) G_{2,\sigma}^{60°} - 2\cos\left(\theta\right)\sin\left(\theta\right) G_{2,\sigma}^{120°} \end{cases} \quad (5)$$

where $G_{2,\sigma}^{\theta}$ filter is mainly composed of $G_{2,\sigma}^{0°}$, $G_{2,\sigma}^{60°}$, and $G_{2,\sigma}^{120°}$. σ represents the scale of the $G_{2,\sigma}^{\theta}$ filter. The final $G_{2,\sigma}^{\theta} - M_{\max}$ is achieved by using $G_{2,\sigma}^{\theta}$ filter to convolve with $M_{\max}$ in $(0, \frac{\pi}{6}, \frac{2\pi}{6}, \frac{3\pi}{6}, \frac{4\pi}{6}, \frac{5\pi}{6})$ and summing.

As $G_{2,\sigma}^{\theta} - M_{\max}$ has been constructed, we use the Gaussian filter to build high-frequency scale space. The construction process of the high-frequency scale space is the same as that of the low-frequency scale space, so we will not repeat the process.

B. Feature Point Extraction and Description

1) Feature Point Extraction and Nonmaximum Suppression Strategy: After constructing the double-frequency scale space, the feature points can be extracted. The feature points of the image include corner and blob points. For low-frequency images, we use FAST detector [48] to achieve corner points. For high-frequency images, we use the KAZE detector [49] to extract blob points. Besides, in order to eliminate unnecessary feature points and retain important feature points, an optimization strategy (OS) is adopted. Nonmaximum suppression was applied to blob points and corner points, respectively. We also calculate the significance score of the points set. The process is shown as

$$\begin{cases} f_{\text{score}}\left(\cdot\right) = \text{sortrows}(S_{\text{points}}|p_i^{\text{score}}) \\ \text{OS} = f_{\text{score}}\left(f_{\text{nms}}\left(S_{\text{points}}\right)\right) \end{cases} \quad (6)$$

where $f_{\text{score}}(\cdot)$ represents a set of feature points selected according to the significance score, S_{points} represents the feature points that need to be processed, p_i^{score} is the intensity value of each feature point got in the previous step; and the $f_{\text{nms}}(\cdot)$ represents the nonmaximum suppression function.

2) Log-Polar Coordinate Descriptor: The progress in constructing feature vectors is similar to the gradient location and orientation histogram (GLOH). We select a local block of $R \times R$ pixel size around each feature point. The direction histogram of the local block is evenly divided into 24 equal parts (bin) at 15° intervals, calculating the sum of the structural gradient amplitudes of each bin, and the bin with the highest score is selected as the main orientation. Furthermore, the bin scores greater than 0.8 times in the main orientation also be used as the auxiliary orientation. Different from the GLOH, we divide the orientation of each local block into 12 directions. Therefore, for each feature point, an eigenvector of $(2*16+1)*12 = 396$ dimensions will be generated. These feature vectors will be matched in their respective scale spaces, and the initial matching results will be achieved.

C. Matches Aggregation and Outlier Removal

The low-frequency feature descriptors and high-frequency feature descriptors of the two images are matched by using SSD metric, respectively. Then, we concatenate the matching results of low-frequency features and high-frequency features to obtain the aggregated matching results. The outliers are eliminated using FSC and the threshold of outliers is 3 pixels.

IV. EXPERIMENTAL RESULTS

This section evaluates the proposed method with four state-of-the-art matching methods. The details of comparison methods and the dataset are introduced as follows.

A. Datasets and Experimental Settings

The proposed method is compared with several baseline methods, including RIFT (2019) [11], LPSO (2022) [50], LNIFT (2022) [4], and HOWP (2023) [29]. The parameters of compared methods are tuned as the same. For example, we set feature points of each layer to 5000 for all comparison methods. In addition, we have established scale space with three layers for LNIFT and RIFT. Since RIFT is time consuming, we add parallel operations to it. The dataset we choose to evaluate our proposed method mainly collected from Mars and Moon. The dataset includes a total of four types of PRSI (illumination differences, weak textures, scales, and rotation), and each type contains 25 image pairs. Fig. 3 shows sample image pairs from the collected dataset.

Five metrics are used to evaluate our methods in experiments, which are the running time (RT), the ratio of corrected matches (RCM), the number of correct matches (NCM), the root mean square error (RMSE), and the success rate (SR). The RMSE reflects the precision of the matching, and the definition is shown in (7). If the RMSE > 3 pixels or NCM < 10 in a matching trial, it is considered failed for matching. In addition, we set the RMSE is set to 10 for failed matching trials

$$\text{RMSE} = \sqrt{\frac{1}{\text{NCM}} \sum_{i=1}^{\text{NCM}} \left[(x_i - x_{i'})^2 + (y_i - y_{i'})^2 \right]} \quad (7)$$
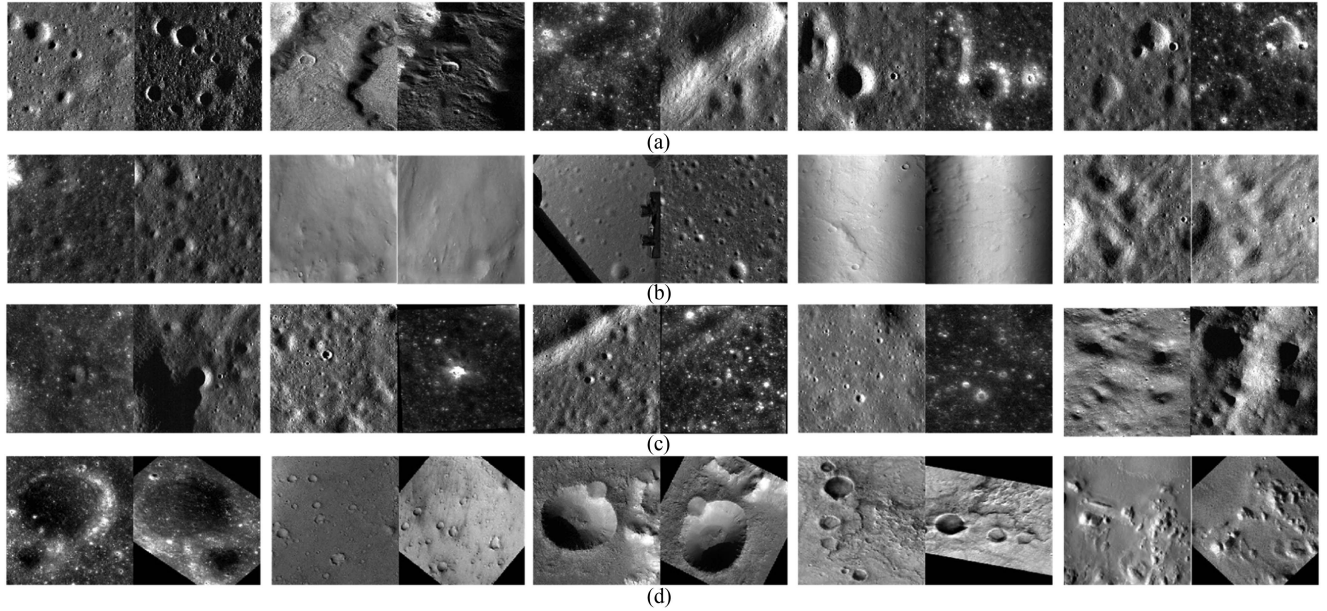
Fig. 3. Sample image pairs from the collected dataset. (a) Image pairs with illumination differences. (b) Image pairs with weak textures. (c) Image pairs with scale changes. (d) Image pairs with rotation changes.

TABLE I
DETAILS SETTINGS OF PARAMETERS

Experiments	Variable	Fixed Parameters
Parameter N	$N = [1,2,3,4,5]$	$R = 48$
Parameter R	$R = [32,40,48,56,64]$	$N = 3$

TABLE II
RESULTS OF PARAMETER R

Metric	$R, N = 3$				
	32	40	48	56	64
NCM	1343.9	1421.6	1540.2	1547.8	1550.3
SR/%	100	100	100	100	100

TABLE III
RESULTS OF PARAMETER N

Metric	$N, R = 48$				
	1	2	3	4	5
NCM	576.7	1121.3	1540.2	1610	1611.1
SR/%	100	100	100	100	100

where (x_i, y_i) is the coordinate for the correct match; (x_i', x_i') is the correct transformed coordinate of (x_i, y_i) through true transformation.

SR is the ratio of the successful trials between all matching trials, and RCM represents the RCM among all matches; the definition of RCM is shown as

$$\text{RCM} = \frac{\text{NCM}}{\text{NCM} + \text{NFM}} \qquad (8)$$

where NFM represents the number of false matches.

The proposed FDAFT mainly includes two parameters: 1) the number of layers in the low-frequency scale space and 2) the size of the descriptor, which are denoted as N and R. In order to seek appropriate values of parameters. Two independent experiments are designed and can be shown in Table I. The other parameters in this article are set according to the relevant research [23], [29], [50].

In general, the size of N is proportional to the number of feature points extracted as well as the time spent. In addition, the size of N is also related to the robustness of the method to the scale variation of the images. As the size of N increases, more feature points will be extracted. At the same time, as N becomes more significant, the method will become more robust to matching scale images. However, an increase in N will lead to a rise in computational overhead, resulting in a long RT of the method. R represents the radius of each feature block, and the larger R is, the more information the feature block contains, and

the easier it is to determine its location on the image. However, the increase of R will lead to an increase in the computation time. If the size of R is too small, the feature block cannot accurately describe the corresponding feature point which will lead to the failure of matching.

In order to achieve the appropriate parameters, we selected 40 image pairs, which covering all transform types, from the dataset for experiments. We design two independent experiments to explore the parameter settings of N and R. Only one parameter acts as a variable in each experiment, and the others are fixed. NCM and SR are used as evaluation metrics. Tables II and III show the results of two parameter experiments. Through Tables II and III, we can infer that 1) If the parameter N is too small, NCM will poorer. As N increases, NCM will also be improved. However, too large N will significantly increase computational complexity. Therefore, the parameter N is set to 3 considering both SR and NCM, as well as the computational complexity. 2) Table III shows the relationship between SR and
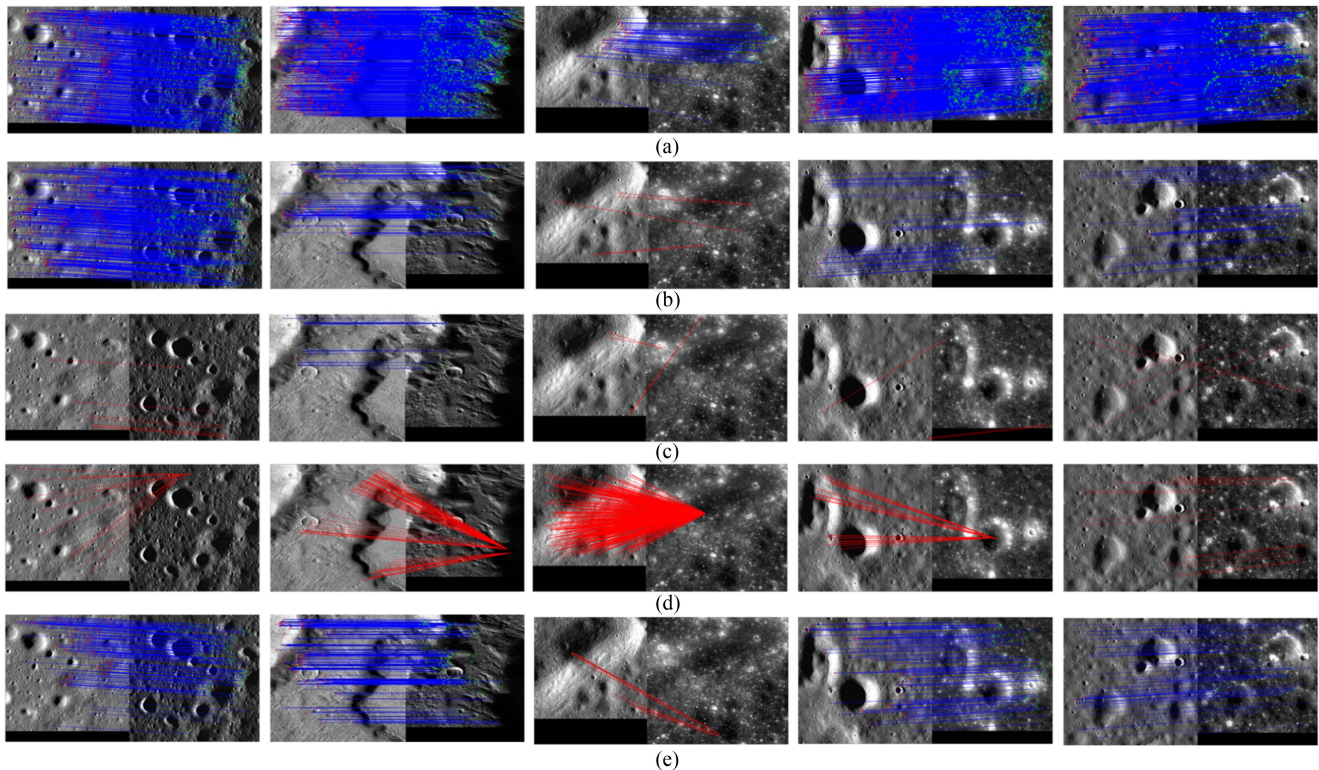
Fig. 4. Matching results of all methods on sample images with illumination differences. (a) FDAFT. (b) HOWP. (c) RIFT. (d) LNIFT. (e) LPSO.

TABLE IV
RESULTS OF ILLUMINATION DIFFERENCES ON FIVE EVALUATION METRICS

Method	RMSE	NCM	RCM	SR	RT
FDAFT	**2.168**	**1579.960**	**20.712**	**96.000**	**22.465**
HOWP	6.489	235.400	7.904	44.000	26.029
RIFT	8.760	119.360	2.526	16.000	50.068
LNIFT	9.686	9.080	0.298	4.000	36.236
LPSO	4.927	172.960	7.557	64.000	24.226

Noted that the lowest values in RMSE and RT and the highest values in NCM, RCM, and SR among all the methods are marked with bold texts.

NCM and the radius R of the feature descriptors. As R increases, the NCM becomes larger. To ensure the SR, NCM, and take runtime into count, set R to 48.

B. Experiment on Illumination Differences

The first experiment mainly evaluates the performance of the proposed method on illumination differences. These images mainly collected from lunar images. Table IV shows the results of all methods on five metrics. FDAFT achieves the best matching accuracy in all metrics compared to other compared methods. As seen, the RMSE of FDAFT on images with illumination differences is about 2.168 pixels, and the SR reaches 96%.

Between the comparison methods, HOWP, RIFT, and LPSO belong to the phase method, and LNIFT to the spatial strategy. Compared with HOWP, RIFT, and LPSO, the results of LNIFT are much worse. It can be seen from Table IV that the SR of LNIFT only reaches 4% on images with illumination differences. For image pairs with significant illumination differences,

the imaging transformation between them is nonlinear. Since the phase-based method realizes the matching by extracting the local structure of the images, the methods using phase are more suitable for nonlinear radiation problems than those using gradient.

However, although the above phase-based methods are more robust than the grayscale gradient-based methods, their results still need improvement in experiments. The SR of RIFT only has 16% and the SR of LPSO is 64%. Besides, the NCM of above phase-based methods are much lower than that of our proposed. Different from images collected on Earth, the NRDs on deep space images are more complex; Fig. 3(a) shows part image pairs with illumination differences and Fig. 4 shows the experimental results of different methods in Fig. 3(a). To be specific, although Wan et al. [51] demonstrated the illumination-invariant property of phase angle to the illumination changes resulting from different solar positions, via mathematical analysis and experiments. However, the local illumination of the lunar surface is more complex and variable. Thus, matching of deep space images is more difficult than that of the images of the Earth.

RIFT constructs the MIM to match image features. This method is less effective in planetary images because it may not find the correct maximum index layer due to the complexity and variability of local illumination. In addition, the weak texture environment in deep space images will also make the pixels near the feature points on the MIM share the same index. For HOWP, it designs a weighted bandwidth function model for the PC model, and this action helps curbing phase direction reversal and phase extreme value mutation. The results of Table IV shows

TABLE V
RESULTS OF WEAK TEXTURES ON FIVE EVALUATION METRICS

Method	RMSE	NCM	RCM	SR	RT
FDAFT	**1.832**	**2436.400**	**32.719**	**100.000**	**21.515**
HOWP	4.541	563.400	17.685	68.000	25.984
RIFT	6.111	314.880	8.410	48.000	45.131
LNIFT	8.804	21.960	0.735	16.000	35.100
LPSO	3.623	349.240	15.753	80.000	23.579

Noted that the lowest values in RMSE and RT and the highest values in NCM, RCM, and SR among all the methods are marked with bold texts.

TABLE VI
RESULTS OF SCALED IMAGES ON FIVE EVALUATION METRICS

Method	RMSE	NCM	RCM	SR	RT
FDAFT	**2.486**	**960.040**	**19.203**	**92.000**	**19.165**
HOWP	6.748	243.840	10.817	40.000	24.700
RIFT	6.835	145.520	6.273	40.000	44.246
LNIFT	9.432	11.160	0.599	8.000	33.543
LPSO	7.475	59.080	4.401	32.000	23.014

Noted that the lowest values in RMSE and RT and the highest values in NCM, RCM, and SR among all the methods are marked with bold texts.

this optimization method is effective, it can avoid shape changes caused by local illumination differences to a certain extent. Compared with RIFT, HOWP has a 28% improvement in SR. LPSO uses the local phase sharpness feature, the purpose of this method is to extract the local structural features of the image using the local phase coherence model. The experimental results of LPSO has proved that on the planetary images, this way is effective. The SR of LPSO reaches 64% and the NCM is about 172.960. Different from phase-based methods, LNIFT hopes to achieve matching by highlighting details. This method can enhance the detailed features of the image but at the same time magnify the effect of image noise, which makes the matching results of LNIFT unsatisfactory.

To address the problem of illumination differences in planetary images, we propose the FDAFT. We not only consider the local structural details of the image but also the overall contour features in the image. The ECM is used to obtain the contour features of the image, and the M_{max} is used to obtain the local structure features. Consider the various problems faced by matching planetary images, such as illumination differences in lunar images and weak textures in Mars images. We build descriptors in ECM and M_{max}, respectively, and perform the matching aggregation in the stage of outlier removal.

C. Experiment on Weak Textures

In order to verify the performance of the proposed method on weak texture images, 25 pairs of planetary images are chosen to evaluate the proposed method. Compared with the first dataset, the illumination differences of weak texture images are relatively small and the problem with weak textures is more serious. Table V shows the results of all methods on five metrics.

It can be seen from Table V that, compared to the first experiment, HOWP, RIFT, and LNIFT have improved their SRs by 24%, 32%, and 12%, respectively. These results show that matching images with weak textures is relatively easier than images with illumination differences in our dataset. Our method proposed in this article achieves the best matching performance on all metrics between all comparison methods, with the RMSE being about 1.832 pixels, the RCM reaching 32.719% and the RT about 21.515 s. In particular, on the NCM metric, we have more than twice as many correct matches as all other methods.

Fig. 3(b) shows part image pairs with weak textures, and Fig. 5 shows the experimental results of different methods in Fig. 3(b). Among all comparison methods, LPSO achieves the best matching performance, with the SR reaches 80%. Besides, we also find

the SR of all phase-based methods are more than 45%. To better investigate the performance of the phase-based methods on two different types of images, we present the matching results of RMSE in the first and the second experiment. By comparing Fig. 6(a) and (b), we can find that the phase-based methods perform significantly well on weak texture images. In the second experiment, the RMSE of HOWP is about 4 pixels, and the SR of HOWP reaches 68%. However, in the first experiment, HOWP only achieves a matching accuracy of about 2 pixels on only eleven pairs of images and the SR only reaches 44%. The matching results of HOWP in the first and second experiments show that the phase feature can extract the local structure of the image well, but the effect is relatively poor on images with complex local illumination changes. Experiments with RIFT have also confirmed this conclusion. Compared with the first experiment, the SR of RIFT in second experiment reached 3 times that in first experiment.

Fig. 5 and Table V show the results of different methods on weakly textured images. Among the methods for comparison, LPSO performs best. The SR of LPSO reaches 68%, and the RCM reaches 15.753%. Compared with the proposed method in this article, LPSO reduces RCM by 26.966%, and the matching time is 2.064 s longer. RIFT and LNIFT show the most tremendous variation in matching accuracy on two types of images. The SR of RIFT reaches 48%, and the SR of LNIFT reaches 16%. Although RIFT and LNIFT perform better in the second experiment than in the first, they still perform relatively poorly. RIFT builds feature descriptors on the MIM, which takes much time. In the second experiment, the RIFT takes an average time of about 45.131 s, which is slower than all other methods. Besides, the RCM of LNIFT only reaches 0.735%, and the average RMSE of LNIFT is about 9 pixels. The SR of LPSO reaches 80% in the second experiment and 64% in the first experiment. Since LPSO uses the local phase sharpness feature instead of the traditional gradient for matching, it can better capture the local structural features of the images. The idea of our method is similar to that of LPSO, both of which hope to obtain better matching results through enhancing structure features. The experiments prove that it is effective to match images by using the global structural features (contour features).

D. Sensitivity to Scale Changes

This part mainly evaluates the FDAFT's robustness to the scale changes. We selected 25 pairs of images with scale variations for evaluation. Table VI summarizes the results of all
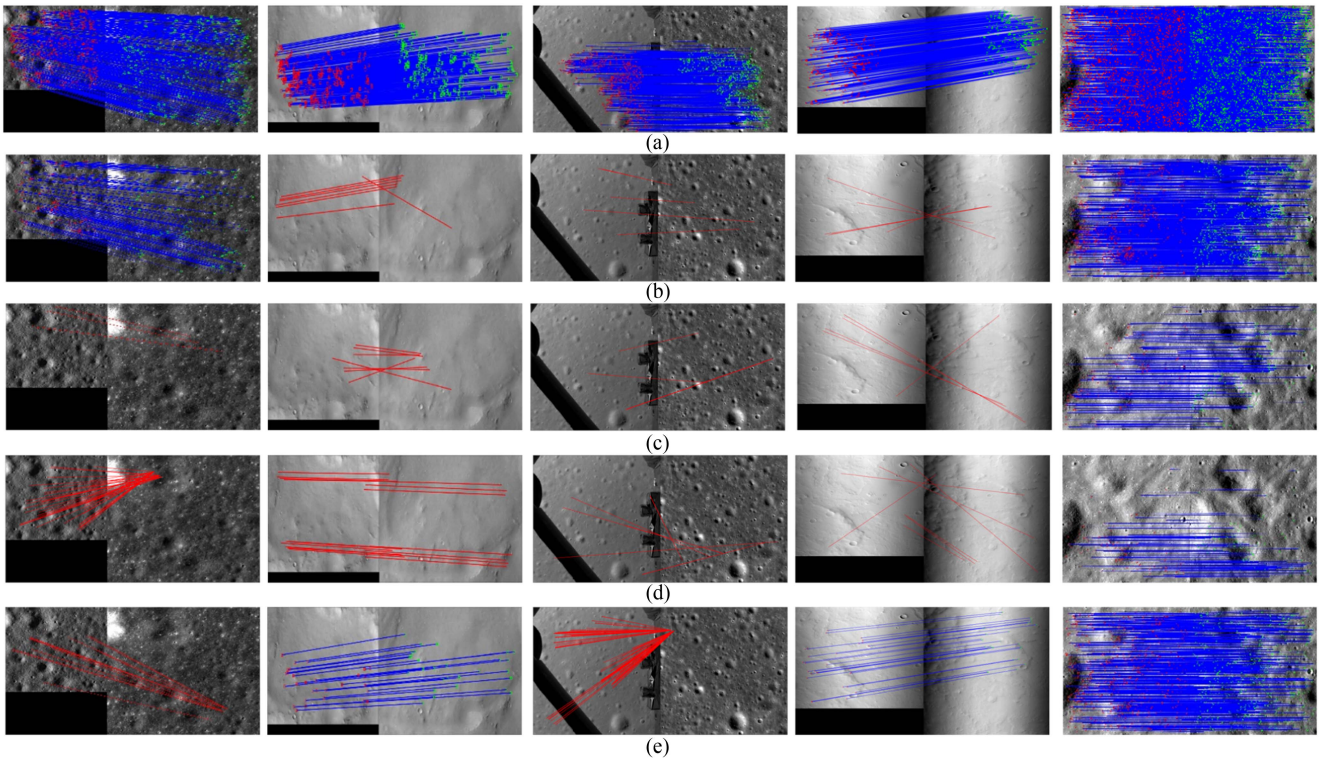
Fig. 5. Matching results of all methods on sample images with weak textures. (a) FDAFT. (b) HOWP. (c) RIFT. (d) LNIFT. (e) LPSO.
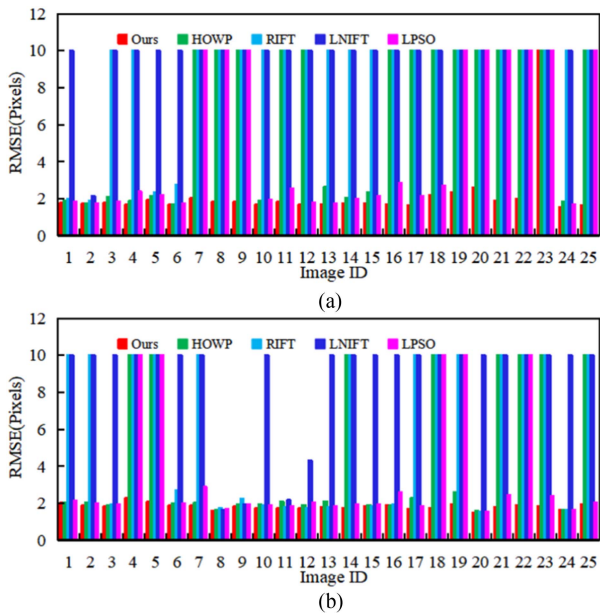

Fig. 6. RMSEs of different kinds of images. (a) Images with illumination differences. (b) Images with weak textures.

methods on five metrics. It can be found that LNIFT fail to match most test image pairs, and the SR of LNIFT only reaches 8%. RIFT builds feature descriptors on the MIM, which takes a lot of time. The experiments show that LNIFT performs poorly on scale changes. The NCM of LNIFT only reaches 11.16 and the

RCM only reaches 0.599. Since LNIFT unable to match images with illumination differences and weak textures robustly, the results are worse on planetary images with scale differences. HOWP and LPSO achieve better results than RIFT and LNIFT, the SR of HOWP and LPSO both reach than 30%. Compared to LPSO, HOWP improved by 6.416% in RCM metric and 8% in SR metric. It can be seen from Table VI that, FDAFT match almost all test images successfully except two image pairs. The RCM of FDAFT reaches 19.203% and the RMSE is about 2.486 pixels. Besides, our method consumes least time among all comparison methods, only 19.165 s. The results of experiments show that even if there are scale differences and, illumination differences or weak textures in the images, our method can still complete the image matching well. The average NCM of our method over 25 scale image pairs is about 960.040, far more than other methods, proving our method's reliability.

E. Sensitivity to Rotation Changes

This part mainly evaluates the FDAFT's robustness to the rotation changes. We select ten image pairs and perform different rotation transformations. Table VII summarizes the results of all methods on five metrics. It can be found that LPSO fails to match all test image pairs, and the SR of HOWP only reaches 8%. Since HOWP, and LPSO limit the main orientation to under 180°, they are sensitive to rotation changes. Although RIFT does not limit the main direction under 180°, RIFT is still sensitive to rotation. The RIFT's poor results may be due to its sensitivity to illumination differences and weak textures, as seen in Sections

TABLE VII
RESULTS OF ROTATED IMAGES ON FIVE EVALUATION METRICS

Method	RMSE	NCM	RCM	SR	RT
FDAFT	**2.765**	**518.320**	**8.862**	**92.000**	**22.233**
HOWP	9.395	51.120	1.687	8.000	26.740
RIFT	8.535	253.600	4.528	20.000	50.507
LNIFT	9.350	26.600	0.877	8.000	36.465
LPSO	9.090	20.640	1.396	12.000	24.128

Noted that the lowest values in RMSE and RT and the highest values in NCM, RCM, and SR among all the methods are marked with bold texts.

TABLE VIII
RESULTS OF ABLATION STUDY ON FOUR EVALUATION METRICS

Method	RMSE	NCM	RCM	SR
ECM	5.158	382.560	10.216	62.000
M_{max}	2.583	987.190	27.863	93.000
PC	5.562	191.940	12.813	55.000
MIM	8.970	59.570	2.852	13.000
ECM-PC	4.741	584.550	11.610	67.000
ECM-MIM	5.654	357.570	8.839	54.000
FDAFT	2.312	1370.510	20.070	95.000

IV-B and IV-C. Compared with the performance of HOWP, RIFT, and LPSO, the results of LNIFT are relatively better. However, since LNIFT mainly considers gradient changes in the spatial domain, LNIFT still performs worse on most test images, with the SR only reaching 8%. For rotation images, FDAFT achieves the best performances among all metrics, and these results prove the FDAFT's robustness to the rotation changes.

F. Ablation Study on FDAFT

To evaluate the effectiveness of the proposed FDAFT, we conduct ablation experiments on the test dataset. We replace the low–high frequency with the construction from M_{max}, ECM, PC, MIM, ECM-PC, and ECM-MIM, and keep other parts unchanged. Table VIII shows the comparison results in terms of NCM, RCM, SR, and RMSE.

Since FDAFT uses low–high frequency to build scale space and extract features, it can more effectively preserve blob and corner features from deep space images compared to other metrics. The SR of FDAFT reaches 95%, and the RMSE of FDAFT is about 2.312 pixels. Through using low–high frequency, our method achieves the best matching results among all metrics.

Besides, to further explore the effects of different metrics on matching results in FDAFT, the low-frequency features (ECM) and high-frequency features (M_{max}) in FDAFT are tested separately. The results of the ECM and M_{max} experiments are shown in Table VIII. We can find that two metrics proposed in this article are better than others: the SR of ECM reaches 62%, and the SR of M_{max} reaches 93%. Through building low–high frequency scale space and extracting features, FDAFT improves the accuracy to 95% compared to using only M_{max}. These results prove that our proposed double-channel aggregated features are effective.

To our surprise, MIM achieves the worst results among all metrics. The SR of MIM only reaches 13%, and the RMSE of MIM over 8 pixels. It is considering that MIM is mainly proposed to address the matching of the multimodal remote sensing image
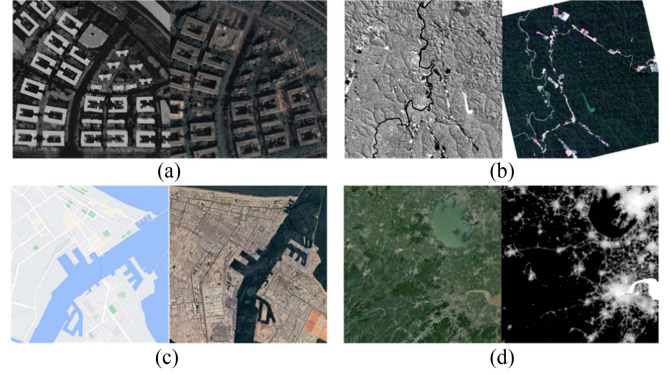

Fig. 7. Sample image pairs from the dataset [29]. (a) Optical-depth. (b) Optical-infrared. (c) Optical-map. (d) Day-night.

TABLE IX
RESULTS OF FOUR EVALUATION METRICS ON IMAGES OF THE EARTH BY FOUR METHODS

Method	RMSE	NCM	RCM	SR
ECM	2.398	115.400	6.416	94.000
M_{max}	2.986	83.720	5.558	86.000
PC	2.074	209.000	23.589	98.000
MIM	2.716	160.660	10.584	90.000

TABLE X
RESULTS OF ALL IMAGES ON FIVE EVALUATION METRICS

Method	RMSE	NCM	RCM	SR	RT
FDAFT	**2.312**	**1370.510**	**20.070**	**95.000**	**21.454**
HOWP	6.793	273.440	9.524	40.000	25.863
RIFT	7.560	208.340	5.434	31.000	47.488
LNIFT	9.318	17.200	0.627	9.000	35.336
LPSO	6.279	150.480	7.277	47.000	23.737

Noted that the lowest values in RMSE and RT and the highest values in NCM, RCM, and SR among all the methods are marked with bold texts.

pairs of the Earth. Thus, we add a set of images of Earth to test. These image pairs are collected by Zhang et al. [29] and Fig. 7 shows sample image pairs of the dataset. It can be seen from Table IX that MIM achieves great results among images of Earth, and the SR of MIM reaches 90%. The reason that MIM cannot achieve satisfactory results in deep space may due to the complexity and variability of local illumination, which can cause it to fail to find the correct maximum index layer. In addition, the weak texture environment in deep space images will also make the pixels near the feature points on the MIM share the same index.

The SR of PC reaches 55%, and the RMSE of PC is about 5.562 pixels. Since the local illumination of the lunar surface is more complex and variable, the impact of PC metrics on deep space images is less effective than on images of the Earth, as shown in Tables VIII and IX.

V. DISCUSSION

In this section, we evaluate the overall results of five methods on four types of images, as seen in Fig. 8 and Table X. During four kinds of images, FDAFT achieves the best performance on five metrics. The REMSE of FDAFT is about 2.312 pixels
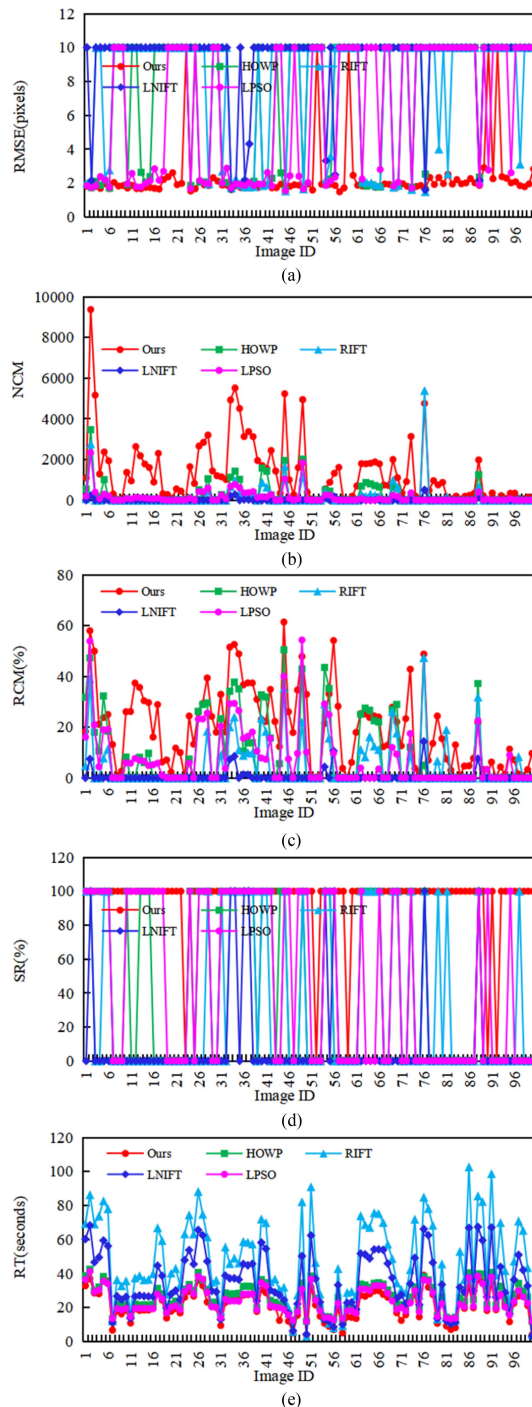
Fig. 8. Matching results of five methods on total images. (a) RMSE results. (b) NCM results. (c) RCM results. (d) SR results. (e) RT results.

and the SR reaches 95%. Besides, we have also found that our method is fast and has higher RCM than the other four comparison methods. The RCM of FDAFT reaches 20.070%. Four experiments show that our method is robust to planetary image matching. Compared with RIFT, LNIFT, and HOWP, the matching results of LPSO are relatively better. LPSO designs a local phase sharpness feature; the purpose of this method is to extract the local structural features of the image by using the local phase coherence model. This action helps curb phase direction reversal and phase extreme value mutation. However, LPSO

ignores the contribution of global features to image matching. Thus, the matching result of LPSO is worse than that of FDAFT. HOWP is higher than LPSO in the NCM metric, and is similar to LPSO in the RT metric. To obtain better matching results, HOWP and LPSO still need further improvement. RIFT constructs the MIM to match image features. Since the complexity and variability of local illumination, the RIFT does not match the PRSI or the images of the Earth. Besides, RIFT takes a long time to match and is, therefore, less effective in images of large scenes. Different from other compared methods, LNIFT computes image detail features in the spatial domain. Considering the NRDs between images, LNIFT is inferior to other methods in terms of matching effect. Effectively reducing the NRD of PRSI is an important problem when using spatial domain methods.

VI. CONCLUSION

In this article, we propose an aggregated low–high frequency feature matching method, called FDAFT. First, we construct a fast double-frequency scale space to capture contour and phase features. Subsequently, corner and blob points are extracted from high-frequency and low-frequency images. Finally, we use the nonmaximum suppression strategy to match the double channels' feature points. Different experiments show that FDATF is superior to other compared methods and can cope well with problems such as weak texture, scales, rotation, and illumination differences. Compared with other methods, our method extracts global and local features from images and proposes a two-channel aggregation strategy, which makes our method superior to others.

Although FDAFT achieves the best results among all comparison methods, it still needs to perfectly match in all experiments where the illumination difference, scale changes, and rotation difference are too large. Besides, the size of deep space images tends to be large, which often requires a larger radius to ensure reliable descriptors. Accordingly, in the future work, we will focus on improving the robustness of descriptors and reducing computational complexity by optimizing the framework of the method.

REFERENCES

[1] A. Tewari, V. Verma, P. Srivastava, V. Jain, and N. Khanna, "Automated Crater detection from co-registered optical images elevation maps and slope maps using deep learning," 2020, *arXiv:2012.15281*.

[2] Z. Li et al., "Photogrammetric processing of Tianwen-1 HiRIC imagery for precision topographic mapping on Mars," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4601916.

[3] K. Di et al., "Photogrammetric processing of rover imagery of the 2003 Mars exploration rover mission," *ISPRS J. Photogrammetry Remote Sens.*, vol. 63, no. 2, pp. 181–201, Mar. 2008.

[4] J. Li, W. Xu, P. Shi, Y. Zhang, and Q. Hu, "LNIFT: Locally normalized image for rotation invariant multimodal feature matching," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5621314.

[5] X. Xu, X. Li, X. Liu, H. Shen, and Q. Shi, "Multimodal registration of remotely sensed images based on Jeffrey's divergence," *ISPRS J. Photogrammetry Remote Sens.*, vol. 122, pp. 97–115, Dec. 2016.

[6] S. Suri and P. Reinartz, "Mutual-information-based registration of TerraSAR-X and Ikonos imagery in urban areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 2, pp. 939–949, Feb. 2010.

[7] B. Fang, K. Yu, J. Ma, and P. An, "EMCM: A novel binary edge-feature-based maximum clique framework for multispectral image matching," *Remote Sens.*, vol. 11, no. 24, Dec. 2019, Art. no. 3026.

[8] Y. Ye, L. Bruzzone, J. Shan, F. Bovolo, and Q. Zhu, "Fast and robust matching for multimodal remote sensing image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 11, pp. 9059–9070, Nov. 2019.

[9] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 187–198, Apr. 1997.

[10] H. Bay, T. Tuytelaars, and L. V. Gool, "SURF: Speeded up robust features," in *Proc. Eur. Conf. Comput. Vis.*, 2006, pp. 404–417.

[11] J. Li, Q. Hu, and M. Ai, "Rift: Multi-modal image matching based on radiation-variation insensitive feature transform," *IEEE Trans. Image Process.*, vol. 29, pp. 3296–3310, 2019.

[12] Y. Yao, Y. Zhang, Y. Wan, X. Liu, and H. Guo, "Heterologous images matching considering anisotropic weighted moment and absolute phase orientation," *Geomatics Inf. Sci. Wuhan Univ.*, vol. 46, no. 11, pp. 1727–1736, 2021.

[13] Y. Yao, Y. Zhang, Y. Wan, X. Liu, X. Yan, and J. Li, "Multi-modal remote sensing image matching considering co-occurrence filter," *IEEE Trans. Image Process.*, vol. 31, pp. 2584–2597, 2022.

[14] Y. Zhang, Y. Liu, H. Zhang, and G. Ma, "Multimodal remote sensing image matching combining learning features and Delaunay triangulation," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5635517.

[15] F. Ye, Y. Su, H. Xiao, X. Zhao, and W. Min, "Remote sensing image registration using convolutional neural network features," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 2, pp. 232–236, Feb. 2018.

[16] W. Ma, J. Zhang, Y. Wu, L. Jiao, H. Zhu, and W. Zhao, "A novel two-step registration method for remote sensing images based on deep and local features," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4834–4843, Jul. 2019.

[17] Y. Wu et al., "RORNet: Partial-to-partial registration network with reliable overlapping representations," *IEEE Trans. Neural Netw. Learn. Syst.*, to be published, doi: 10.1109/TNNLS.2023.3286943.

[18] Y. Wu, X. Hu, Y. Zhang, M. Gong, W. Ma, and Q. Miao, "SACF-Net: Skip-attention based correspondence filtering network for point cloud registration," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 33, no. 8, pp. 3585–3595, Aug. 2023.

[19] Y. Wu, W. Ma, M. Gong, L. Su, and L. Jiao, "A novel point-matching algorithm based on fast sample consensus for image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 1, pp. 43–47, Jan. 2015.

[20] W. Ma et al., "Remote sensing image registration with modified SIFT and enhanced feature matching," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 1, pp. 3–7, Jan. 2017.

[21] P. Dollár and C. L. Zitnick, "Fast edge detection using structured forests," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 8, pp. 1558–1570, Aug. 2015.

[22] Y. Ye, B. Zhu, T. Tang, C. Yang, Q. Xu, and G. Zhang, "A robust multimodal remote sensing image registration method and system using steerable filters with first- and second-order gradients," *ISPRS J. Photogrammetry Remote Sens.*, vol. 188, pp. 331–350, Jun. 2022.

[23] B. Zhu, J. Zhang, T. Tang, and Y. Ye, "SFOC: A novel multi-directional and multi-scale structural descriptor for multimodal remote sensing image matching," *Int. Arch. Photogrammetry, Remote Sens. Spatial Inf. Sci.*, vols. 127, pp. 113–120, May 2022.

[24] 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, 2018.

[25] M. M. Pedrosa, S. C. de Azevedo, E. A. da Silva, and M. A. Dias, "Improved automatic impact crater detection on Mars based on morphological image processing and template matching," *Geomatics Natural Hazards Risk*, vol. 8, no. 2, pp. 1306–1319, Dec. 2017.

[26] Y. Zhang et al., "Analysis of illumination conditions in the lunar south polar region using multi-temporal high-resolution orbital images," *Remote Sens.*, vol. 15, 2023, Art. no. 5691.

[27] B. Zhu, Y. Ye, L. Zhou, Z. Li, and G. Yin, "Robust registration of aerial images and LiDAR data using spatial constraints and Gabor structural features," *ISPRS J. Photogrammetry Remote Sens.*, vol. 181, pp. 129–147, Nov. 2021.

[28] Y. Xiang, R. Tao, L. Wan, F. Wang, and H. You, "OS-PC: Combining feature representation and 3-D phase correlation for subpixel optical and SAR image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 9, pp. 6451–6466, Sep. 2020.

[29] Y. Zhang et al., "Histogram of the orientation of the weighted phase descriptor for multi-modal remote sensing image matching," *ISPRS J. Photogrammetry Remote Sens.*, vol. 196, pp. 1–15, Feb. 2023.

[30] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. 4th Alvey Vis. Conf.*, 1988, pp. 147–151.

[31] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Proc. Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[32] J. Fan, Y. Ye, J. Li, G. Liu, and Y. Li, "A novel multiscale adaptive binning phase congruency feature for SAR and optical image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5235216.

[33] J. Bian, W.-Y. Lin, Y. Liu, L. Zhang, S.-K. Yeung, and M.-M. Cheng, "GMS: Grid-based motion statistics for fast ultra-robust feature correspondence," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 2828–2837.

[34] J. Zhao, J. Ma, J. Tian, J. Ma, and D. Zhang, "A robust method for vector field learning with application to mismatch removing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2011, pp. 2977–2984.

[35] W. C. Liu and B. Wu, "An integrated photogrammetric and photo clinometric approach for illumination-invariant pixel-resolution 3D mapping of the lunar surface," *ISPRS J. Photogrammetry Remote Sens.*, vol. 159, pp. 153–168, Jan. 2020.

[36] 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.

[37] X. W. Xiang, F. Wang, H. You, X. Qiu, and K. Fu, "A global-to-local algorithm for high-resolution optical and SAR image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5215320.

[38] Z. Yang, Z. Kang, and J. Yang, "A semiautomatic registration method for chang'E-1 IIM imagery based on globally geo-reference LROC-WAC mosaic imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 3, pp. 543–547, Mar. 2021.

[39] L. Yu, Y. Xu, and B. Zhang, "Single image super-resolution via phase congruency analysis," in *Proc. Vis. Commun. Image Process.*, 2013, pp. 1–6.

[40] 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.

[41] J. Zhong, J. Yan, M. Li, and J. P. Barriot, "A deep learning-based local feature extraction method for improved image matching and surface reconstruction from Yutu-2 PCAM images on the Moon," *ISPRS J. Photogrammetry Remote Sens.*, vol. 206, pp. 16–29, 2023.

[42] J. Li, P. Tu, J. Yu, and L. Liu, "Image feature matching of the Moon surface based on multi-task learning," in *Proc. Int. Conf. Electron. Inf. Eng. Comput. Sci.*, 2021, pp. 577–580.

[43] V. Franchi and E. Ntagiou, "Planetary rover localisation via surface and orbital image matching," in *Proc. IEEE Aerosp. Conf.*, 2022, pp. 1–14.

[44] D. DeTone, T. Malisiewicz, and A. Rabinovich, "Superpoint: Self-supervised interest point detection and description," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, 2018, pp. 224–236.

[45] P.-E. Sarlin, D. DeTone, T. Malisiewicz, and A. Rabinovich, "SuperGlue: Learning feature matching with graph neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 4938–4947.

[46] S. Wang, D. Quan, X. Liang, M. Ning, Y. Guo, and L. Jiao, "A deep learning framework for remote sensing image registration," *ISPRS J. Photogrammetry Remote Sens.*, vol. 145, pp. 148–164, 2018.

[47] Y. Ye, T. Tang, B. Zhu, C. Yang, B. Li, and S. Hao, "A multiscale framework with unsupervised learning for remote sensing image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5622215.

[48] E. Rosten, R. Porter, and T. Drummond, "Faster and better: A machine learning approach to corner detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 1, pp. 105–119, Jan. 2010.

[49] P. F. Alcantarilla, A. Bartoli, and A. J. Davison, "KAZE features," in *Proc. Eur. Conf. Comput. Vis.*, 2012, pp. 214–227.

[50] W. Yang, C. Xu, L. Mei, Y. Yao, and C. Liu, "LPSO: Multi-source image matching considering the description of local phase sharpness orientation," *IEEE Photon. J.*, vol. 14, no. 1, pp. 1–9, Feb. 2022.

[51] X. Wan, J. Liu, and H. Yan, "The illumination robustness of phase correlation for image alignment," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 10, pp. 5746–5759, Oct. 2015.

Rong Huang (Member, IEEE) received the Ph.D. degree in photogrammetry and remote sensing from 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.

Genyi Wan received the B.E. degree in geographical information sciences from 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 surveying and mapping with the College of Surveying and Geoinformatics, Tongji University, Shanghai, China.

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

Yingying Zhou received the B.E. degree in spatial information and digital technology from Shanghai Ocean University, Shanghai, China, in 2022. She is currently working toward the M.E. degree in surveying and mapping with the College of Surveying and Geoinformatics, Tongji University, Shanghai.

Her research interests include planetary remote sensing image matching.

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.

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.

Yusheng Xu (Member, IEEE) 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, Munich, Germany. Since 2022, he has been with Tongji University as a Full Professor. He has authored and coauthored more than 40 entries. His research interests include 3-D point cloud processing, image processing, and spaceborne photogrammetry.

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, Wuhan, 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 hyper-spectral images.