



Illumination invariant feature point matching for high-resolution planetary remote sensing images



Bo Wu^{*}, Hai Zeng, Han Hu

Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

ARTICLE INFO

Keywords:

Planetary imagery
SIFT
Matching
Co-registration

ABSTRACT

Despite its success with regular close-range and remote-sensing images, the scale-invariant feature transform (SIFT) algorithm is essentially not invariant to illumination differences due to the use of gradients for feature description. In planetary remote sensing imagery, which normally lacks sufficient textural information, salient regions are generally triggered by the shadow effects of keypoints, reducing the matching performance of classical SIFT. Based on the observation of dual peaks in a histogram of the dominant orientations of SIFT keypoints, this paper proposes an illumination-invariant SIFT matching method for high-resolution planetary remote sensing images. First, as the peaks in the orientation histogram are generally aligned closely with the sub-solar azimuth angle at the time of image collection, an adaptive suppression Gaussian function is tuned to level the histogram and thereby alleviate the differences in illumination caused by a changing solar angle. Next, the suppression function is incorporated into the original SIFT procedure for obtaining feature descriptors, which are used for initial image matching. Finally, as the distribution of feature descriptors changes after anisotropic suppression, and the ratio check used for matching and outlier removal in classical SIFT may produce inferior results, this paper proposes an improved matching procedure based on cross-checking and template image matching. The experimental results for several high-resolution remote sensing images from both the Moon and Mars, with illumination differences of 20°–180°, reveal that the proposed method retrieves about 40%–60% more matches than the classical SIFT method. The proposed method is of significance for matching or co-registration of planetary remote sensing images for their synergistic use in various applications. It also has the potential to be useful for flyby and rover images by integrating with the affine invariant feature detectors.

1. Introduction

High-resolution remote sensing images (Malin et al., 2007; McEwen et al., 2007; Robinson et al., 2010) are the major sources for the planetary spatial data infrastructure (PSDI) (Archinal et al., 2017), which contribute to all three types of products including the global reference system (Di et al., 2017), the 3D topographic mapping (Li et al., 2011; Di et al., 2014), and global orthoimage mosaics. The PSDI has set a foundation for further planetary applications and science, including the selection of landing sites (Kirk et al., 2008), the maneuvering of robotic rovers (Wu et al., 2013) and the detailed analysis of geological units such as craters (Watters et al., 2015) and rilles (Hurwitz et al., 2013). Accurate image matching is a key step in order to generate 3D topographic models and orthoimage mosaics from remote sensing images. Accurate image matching is also essential to remove inconsistencies arising between remote sensing images from different sources for co-registration and

synergistic use of them (Wu and Liu, 2017).

Two main types of feature point have been used in image matching in existing studies, namely corners (Harris and Stephens, 1988; Rosten et al., 2010) and salient blobs (Lowe, 2004; Alcantarilla et al., 2013). These features are generally located in regions with sufficient textural information, e.g. maximum autocorrelation for corners and strong textural patterns for salient blobs. In addition, feature descriptors are generally directly related to the intensity values around local features, via either grayscale values and gradients (Lowe, 2004) or the binary ordinal comparison of grayscale values (Calonder et al., 2012). And their descriptors are compared to find matches (Zhu et al., 2010).

The above strategies may be sufficient for use with remote sensing images of the Earth, due to their strong texture. The texture information generally comes from two sources, the variation of the materials (e.g. the surface albedo) and the reflectance ratio according to the bidirectional reflectance distribution function (BRDF). However, on the one hand,

^{*} Corresponding author.

E-mail address: bo.wu@polyu.edu.hk (B. Wu).

planetary bodies such as the Moon and Mars, show a relative lack of strong surface textural patterns. This is mainly because that the albedos of their surfaces are more homogenous. In fact, homogenous albedo assumption is commonly used in many applications, such as shape-from-shading for lunar surface reconstruction (Wu et al., 2018). On the other hand, the texture patterns caused by the variations of the surface normal and the radiance according to BRDF are generally in a larger scale, which may have less impact on local features. Therefore, the feature points selected from planetary images are usually located in regions around shadows cast by rocks or craters. Differences in the solar azimuth angle at image collection result in different grayscale values, and descriptors are also biased to the solar azimuth angle. This reduces the performance of descriptors and may even prevent them from retrieving enough matches.

To solve this problem, this paper proposes an illumination-invariant method of feature description with histogram adaption for point matching in high-resolution planetary images, based on the pioneering scale-invariant feature transform (SIFT) algorithm (Lowe, 2004). The key observation guiding the proposed method is that the gradients used in SIFT feature description are biased to the direction and the reverse direction of the solar azimuth angle, as shown in Fig. 1. In this histogram, the dominant direction, defined as the maximum magnitude of the gradients at a given feature location and scale, shows clear twin peaks corresponding to the solar azimuth angle. However, to enable generalization, the dominant direction should be distributed evenly. Concerning the possible anisotropic texture pattern on images either due to the albedo or the topography, as the most common features on planetary surfaces, such as craters and rocks are isotropic, and the SIFT features are detected on different scales, the anisotropic problem might be alleviated. In light of this observation, an automatically and adaptively tuned Gaussian-suppression function is first developed to level the histogram. Next, the gradients used to compute the SIFT descriptors are modified by the suppression factor. In addition, as the distribution of SIFT has changed, we find that the classical SIFT ratio check (Lowe, 2004) produces inferior matching results. In this paper, therefore, we compare the descriptors and then filter the putative matches using the normalized correlation coefficient (NCC).

The remainder of this paper is organized as follows. The next section presents a literature review on feature-point matching, with a focus on illumination problems. Section 3 first presents the overall workflow of the method, followed by its three major parts: tuning the suppression function, describing illumination-invariant features, and matching features. The experimental results obtained with five datasets from the Moon and Mars are presented in Section 4, followed by concluding remarks.

2. Related works

Successful feature-point matching faces two major obstacles: geometrical problems and radiometrical problems. The first is caused by

differences in viewpoints, spatial resolutions, and orbiter or flight directions, and is generally handled by invariant features in three aspects including scale, rotation and full affine deformations. Some studies of feature matching using corner detectors (Förstner and Gülch, 1987; Harris and Stephens, 1988) have neglected geometrical deformation, drastically reducing matching performance at even a moderate scale or with moderate rotation differences. In addition, geometric preprocessing is generally applied to reduce geometric distortions using known positional and altitudinal information on images.

To remove the prerequisite of known image geometry for robust image matching and thereby overcome the problems of scale and rotation deformation, theories of the scale space of images (Lindeberg, 1993, 1998) and the dominant orientation of feature points (Lowe, 2004) have been developed. Recent works have shown that scale and dominant direction can even be learned (Yi et al., 2016). Affine deformation, a superset of scale and rotational differences, has also been investigated (Matas et al., 2004; Mikolajczyk and Schmid, 2004; Wang et al., 2016). However, as the repeatability of fully affine-invariant features is inferior to that of their subset alternatives (Jazayeri and Fraser, 2010), affine-invariant feature description is not widely used in practice and has been superseded by the database-concatenation strategy (Morel and Yu, 2009; Calonder et al., 2012), which detects scale- and rotation-invariant features in all rectified images using synthesized affine parameters and concatenates them in the same database for feature matching.

As satellite images, either for the Earth or other planets, are collected in a well-controlled manner, their geometrical deformations are generally less severe than close range images, such as flyby or rover images. In practice, the difference of emission angles of satellite images most often does not exceed 30° and state-of-the-art invariant feature operator, such as SIFT, can handle this moderate amount of geometric deformations successfully. The main difficulties lie in the second set of problems mentioned above: radiometrical. Radiometrical problems are typically handled during feature description rather than detection (Hu and Wu, 2017). The most simple descriptors are grayscale values directly retrieved from local windows around the features of interest (Harris and Stephens, 1988; Brown et al., 2005), which are matched using the sum of squared differences or NCCs. Gradient information also provides useful feature descriptors (Mikolajczyk and Schmid, 2005), and gradients can be pooled into several regions (Brown et al., 2011) to allow minor off-setting, weighting and robust aggregation through histograms, creating forms such as a grid (Lowe, 2004; Bay et al., 2008), a concentric circle (Mikolajczyk and Schmid, 2005), a daisy (Tola et al., 2010) and even learned symmetrical patterns (Simonyan et al., 2014).

However, as descriptors based on gradient histograms are just linear combinations of grayscale values, they are inherently not invariant to complex illumination differences (Wang et al., 2016). The most common approach in cases of nonlinear and complex illumination differences involves taking advantage of phase information by moving the image to the frequency domain through the Fourier transform (Hu et al., 2016).

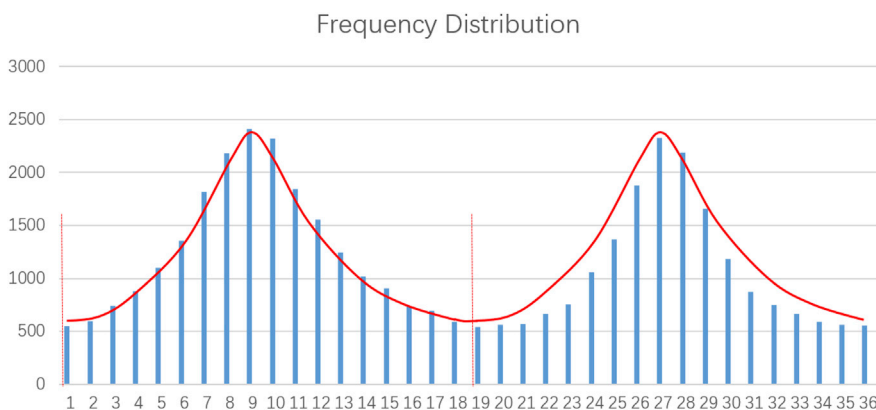


Fig. 1. Distribution of the dominant direction of SIFT features in 36 bins at intervals of 10° for the left image of the NAC-1 pair as shown in Fig. 5. The vertical axis shows the number of bins of each type. The solar azimuth angle is about 90°. The distribution clearly indicates that the gradients of SIFT descriptors are biased to the direction and the reversed direction of the solar azimuth angle.

This approach was first used for subpixel registration and translation (Foroosh et al., 2002) by measuring the cross-power spectrum in the frequency space, and then extended to general cases of scale and rotation (Morgan et al., 2010); it is also known as phase correlation. Another approach using the Fourier transform for feature matching is the phase-congruency model (Wong and Clausi, 2007), which is based on the assumption that features are located at points of maximum phase congruency in the Fourier components. This model has been empirically shown to be invariant to illumination and contrast conditions (Kovesi, 2003). The phase congruency moment values surrounding local features can be used directly as descriptors and matched (Wong and Clausi, 2007). This strategy has been improved by increasing its repeatability (Ye et al., 2012), using spatial and mutual information to enhance descriptors (Liang et al., 2014) and increasing the efficiency of computation through principal phase congruency (Li and Man, 2016). However, as a phase model is a compressed representation of an original image, significantly fewer reliable features can be extracted using this approach, which relies on sophisticated methods to recall matches and is thus less practical in real world applications.

In addition to the phase-congruency model, another widely adopted paradigm involves the detection and description of features in the original image space using rank or ordinal information on local patches (Verdie et al., 2015). As long as the relative order around local features remains relatively stable, the rank is invariant to complex illumination changes. The most simple strategy is to transform the 128-dimension vector of a SIFT descriptor into corresponding rank values given by the SIFT-Rank algorithm (Toews and Wells, 2009). Other recent works have used intensity order rather than descriptors (Wang et al., 2011), and sophisticated methods of region division and histogram aggregation have

been introduced to improve the results (Wang et al., 2016). In addition, a process of concatenating the results of binary comparison into a code string has been developed and used for efficient matching (Rublee et al., 2011; Calonder et al., 2012).

As described above, the feature descriptors based on gradients and their transformations are invariant to constant or linear changes of image intensities. The descriptors based on phase congruency and ordinal information may be invariant to more complex changes of images. However, these methods are built on the assumption that the changes are global or monotonic with the grayscale values of the images, but not for shadows. Shadows present abruptly local intensity changes on images caused by both the elevation and solar azimuth angle. In the case of shadows, the grayscale values of the images are inherently different and sometimes opposite. If the grayscale values or gradients used for feature description are not modified or compensated, the descriptors are absolutely biased. In addition, when matching high-resolution planetary images, the detected features are generally triggered by local variation in grayscale values and more specifically shadows caused by craters or rocks. As a result, the relative intensity order also changes locally and non-monotonically, which significantly degrades the performance of the methods described above. The method proposed in this paper handles local variations directly, by suppressing the strong gradient along the solar azimuth angle, and thereby alleviates illumination differences in local regions.

3. Histogram-adapted SIFT

The proposed method is based on the pioneering SIFT algorithm (Lowe, 2004), which still achieves state of the art performance relative to

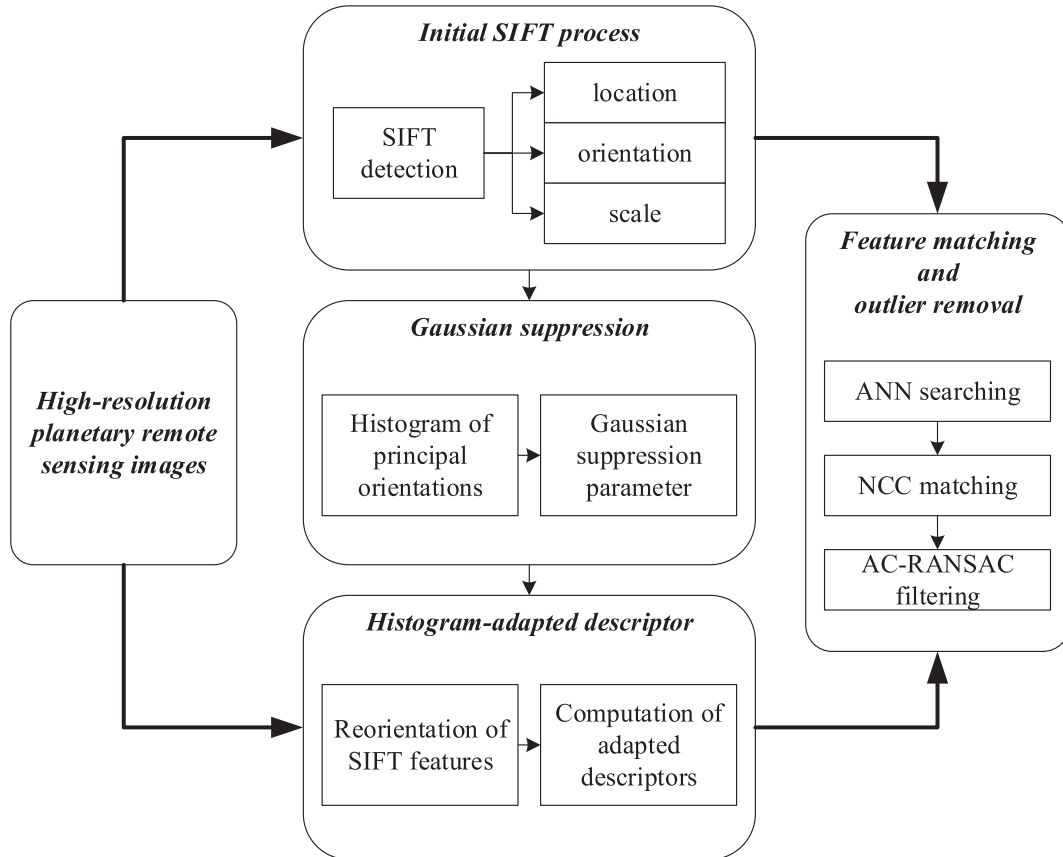


Fig. 2. Workflow of the proposed method. For a pair of planetary remote sensing images, the proposed method begins with the original SIFT (Scale Invariant Feature Transform) feature detection, then it is modified with the proposed Gaussian suppression and histogram-adapted description. Then the modified SIFT descriptors are matched by the ANN (Approximate Nearest Neighborhood) and NCC (Normalized Cross Correlation) and filtered by AC-RANSAC (A Contrario Random Sample Consensus).

many more recent variants (Schönberger et al., 2017). As shown in Fig. 2, the proposed method begins with the original SIFT process of feature detection, comprising the construction of image pyramids, the detection of features in each image pyramid, subpixel localization and the computation of dominant orientation. Next, as explained in Section 3.1, the dominant directions of orientation are aggregated into a histogram, which generally presents twin peaks aligned with the solar azimuth angle, as shown in Fig. 1. The histogram is then leveled using an adaptively tuned Gaussian suppression function. Next, the descriptors are computed at the specific scale and location of each feature; the novel approach here is that the dominant orientation is modified by the Gaussian suppression function and then extracted from the histogram of suppressed gradients, as detailed in Section 3.2. Improved methods of feature matching and outlier removal are also proposed, because the original SIFT ratio check produces inferior results due to the proposed modification of descriptor distribution.

3.1. Adaptive tuning of Gaussian suppression

After carrying out the initial process of SIFT detection shown in Fig. 2, information on scale and orientation is extracted along with information on the locations of the feature points. Orientation is defined as the direction with the maximal gradient magnitude. The principal orientations are then aggregated into a histogram, usually with bin intervals of 5° or 10° , with respect to image size and number of detected SIFT features, as shown in Fig. 1. Such histograms generally present twin peaks. Orientation at the peaks is commonly in accordance with the solar azimuth angle of the image, solar azimuth angle are usually available in the data archives for planetary images (ASU, 2017). Even for particular cases of terrain features with anisotropic patterns, the anisotropic problem might be alleviated by detecting features in different scales. However, as the peaks may also be slightly influenced by image geometry, the peak positions are detected using the proposed method to ensure accurate compensation.

To locate the peaks, the histogram is first smoothed using a symmetrical one-dimensional smoothing function similar to a cubic function, as follows:

$$b(i) = \frac{1}{16}((b(i-2) + b(i+2)) + (b(i-1) + b(i+1)) \times 4 + b(i) \times 6), \quad (1)$$

where $b(i)$ is the accumulation value for the i th component of the histogram. The smoothing function is used to remove possible outliers. Next, the twin peaks are detected jointly by selecting the maximum density of the two components with a stride of 180° , as follows:

$$\max_i b(i) + b(i + n/2), \quad (2)$$

where n is the total number of components in the histogram and $i = 0, 1, 2, \dots, n/2$ is the index of a given component. The peaks are further located at sub-bin level by linear interpolation, as below.

$$\bar{i} = i - 1 + 0.5 \times \frac{b(i-1) - b(i+1)}{b(i-1) - 2b(i) + b(i+1)} \quad (3)$$

After locating the twin peaks with the constraint of a 180° stride, the histogram is divided into two equal parts, each with a 90° interval centered around its peak. Next, the Gaussian suppression function denoted below is applied to each part of the histogram.

$$S = 1 - 2\delta \left(\frac{1}{d} \times e^{-\frac{(o-a)^2}{2\sigma^2}} - \frac{1}{2} \right) \quad (4)$$

S is the suppression factor and δ controls the magnitude of the suppression function and is tuned automatically, as described later. d is the normalization factor, which is the density of the histogram measured at its peak rather than through the accumulation of values shown in

Equation (1). o and a denote the angles of a specific component and the peak location in the histogram, respectively. σ is the variance of the normal distribution, related to the average orientation of SIFT features, which controls the sharpness of the suppression function. The effect of the suppression function S is illustrated in Fig. 3.

In Fig. 3 and Equation (4), the original Gaussian function is acquired by computing the peak and variance of all of the principal orientations of the SIFT features. The normalization function is a transition function used to normalize the original function to make 1 the peak value of the histogram, as represented by the blue line in Fig. 3. The suppression function is designed to scale and translate the normalized Gaussian function. When the value of δ is 0, no suppression is imposed and the proposed method of feature description is identical to the original SIFT method. For other values, the Gaussian suppression function S gives δ penalties to gradients of different magnitudes. In addition, S gives the same penalties to gradients at the tail of the Gaussian function, which is intended to represent the valley area, as shown in Fig. 1.

To select an appropriate value of δ , a series of potential values in the range $[0, 1]$ at intervals of 0.05 are tested at runtime. The value leading to the lowest standard deviation in all of the bins in both parts of the histogram is considered to have the best leveling effect on the orientation histogram. Fig. 4 shows the effects of different δ values. Notably, standard deviation decreases monotonically to reach a minimum point, which is considered the best leveling location. It then starts to increase at higher suppression values due to over-suppression, as shown in the bottom right of Fig. 4b.

3.2. Histogram-adapted feature description

The above suppression function is computed using a gradient histogram; however, as the suppression factor functions with gradient values, it can also be used in the feature-extraction stage, which involves the computation of principal orientation and feature descriptors. Fortunately, the SIFT framework already provides a weighting function for use before computing principal orientation. In the SIFT framework, for each pixel centered in a local patch, gradient magnitude ($m(x,y)$) and orientation ($o(x,y)$) are computed as follows:

$$\begin{aligned} m(x,y) &= \sqrt{(L(x,y+1) - L(x,y))^2 + (L(x+1,y) - L(x,y))^2} \\ o(x,y) &= \arctan(L(x,y+1) - L(x,y) / L(x+1,y) - L(x,y)), \end{aligned} \quad (5)$$

where $L(x,y)$ is the pixel at a specific scale.

The gradients are then pooled into 16 rectangular regions. In each

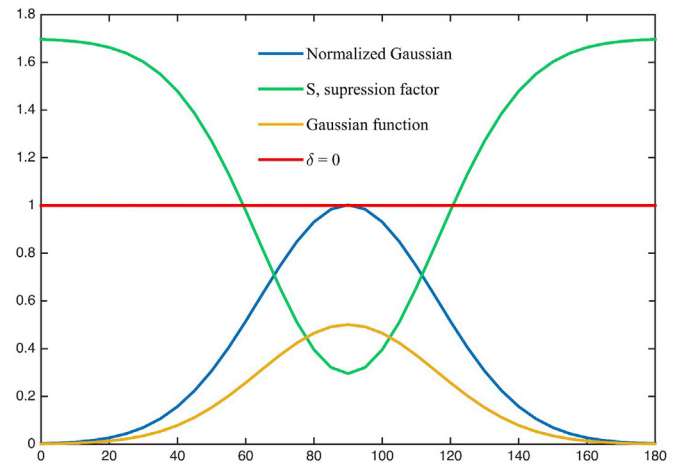


Fig. 3. Illustration of the suppression function S . The right axis demonstrates the effects of different scales of the tuned parameter δ . The peak position a is set to 90° . This suppression function takes effect in the histogram in the region of $[0^\circ, 180^\circ]$.

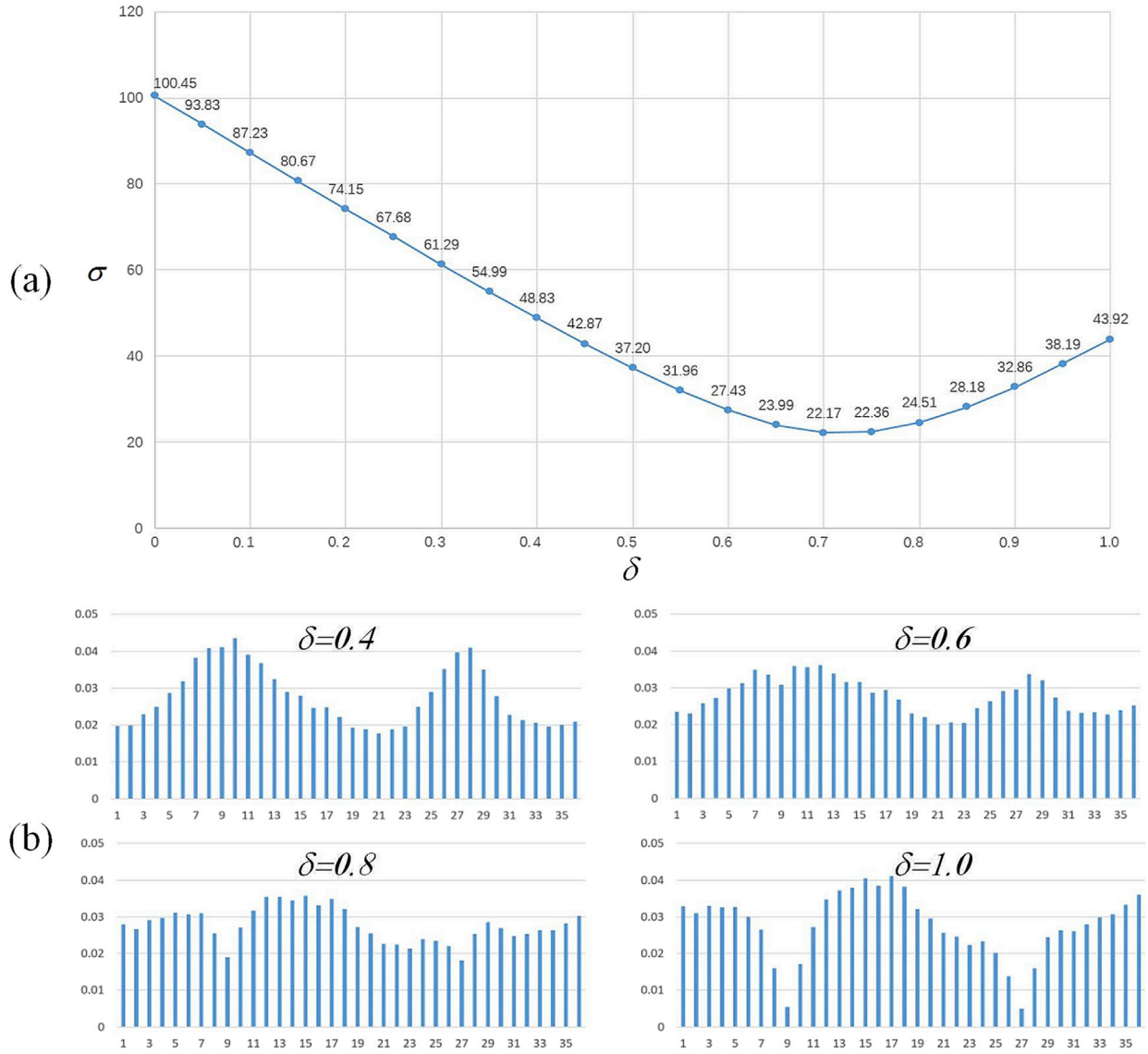


Fig. 4. Adaptive tuning of the parameter δ . (a) Standard deviations of the accumulated components with the specific suppression function and (b) corresponding histograms showing density function in four positions.

region, the values obtained for magnitude and orientation are aggregated into a local histogram, with a specific Gaussian weighting factor W , which gives higher weights to more central pixels and vice versa. Using the proposed method, this weighting function is directly multiplied by the suppression factor S for each pixel. Although the orientation parameter o in Equation (4) is designed for use with individual histogram components, it can also be used to estimate pixel orientation, as shown in Equation (5). The final magnitude used for histogram aggregation is obtained as shown below:

$$\bar{m}(x, y) = W(x, y) \times S(x, y) \times m(x, y), \quad (6)$$

where $S(x, y)$ is similar to the right-hand side of Equation (4) with orientation parameter $o(x, y)$ given in Equation (5) for each pixel (x, y) rather than the specific angle of the histogram. After this modification, the methods of estimating principal orientation and computing SIFT descriptors are identical to their counterparts in the original SIFT process.

3.3. Feature matching and outlier removal

In the original SIFT process (Lowe, 2004), descriptors are compared

using a binary partitioning tree to find the nearest and second-nearest neighbors, and the ratio of differences between the best and second-best matches are checked using a specific threshold. However, the performance of this ratio-check method is inferior to that of the proposed method, possibly because the distribution of the gradients has previously been modified. Therefore, a modified matching method is used in this paper, replacing a ratio-check threshold with an NCC threshold.

After extraction, the histogram-adapted SIFT descriptors are normalized using the Hellinger distance (Arandjelovic and Zisserman, 2012; Schönberger et al., 2017) rather than the L_2 -norm, as the former has been shown in a variety of studies to exhibit superior performance. Next, approximate nearest neighborhood (ANN) (Muja and Lowe, 2014) searching is conducted to find the best mutual matches, rather than the second-best matches as in the original SIFT process. As the putative matches suggested by descriptor comparison probably contain many outliers, it is suggested that a pre-filtering step similar to the ratio check be adopted before further outlier removal. In this paper, this is achieved by NCC matching in a square window of 21 pixels with a threshold of 0.6. It should be noted that, because this is just a pre-filtering step, the regular NCC matching in a fixed window with identical weights is adopted, rather than the more sophisticated methods considering both the SIFT

geometry and alternative weighting strategies (Barazzetti et al., 2010). Next, outlier removal using the random sample consensus (RANSAC) algorithm (Fischler and Bolles, 1981) is conducted, followed by verification of the homographic matrix. A recent variant of the RANSAC, the AC-RANSAC (Moisan et al., 2012) is used in the implementation, which features automatic threshold tuning.

4. Experimental evaluation

4.1. Description of datasets

The novel method proposed in this study is tested with two types of remote-sensing image acquired on the lunar surface and on Mars, respectively. The images comprise Narrow Angle Camera (NAC) images obtained on board the Lunar Reconnaissance Orbiter (LRO) and Context Camera (CTX) images obtained on board the Mars Reconnaissance Orbiter (MRO). Three pairs of lunar NAC images and two pairs of Mars CTX images are selected for experimental analysis. The lunar NAC images have a very high resolution of 0.5 m/pixel, while the Mars CTX images are coarser, with a resolution of about 6 m/pixel. As they are all pushbroom images, their geometrical deformations are mainly related to the emission angle of the camera and also may slightly be influenced by the swath coverage.

Details of the five image pairs chosen for experimental evaluation are listed in Table 1, such as location, solar azimuth angle at image collection, spatial resolution and image subset size. Although the images' differences in scale are not significant (up to 20%), each image pair shows conspicuous variation in solar azimuth angle, ranging from a moderate 20° to the ultimate 180°. The five image pairs are also shown in Fig. 5. Notably, most of the images in the experimental datasets feature large craters. The differences in position of the shadows inside the craters also reflect significant differences in solar azimuth angle. Indeed, differences are observed not only in the positions of shadows in large craters; small objects such as large rocks and small craters cast shadows with different orientations, reducing the performance of the SIFT descriptors by affecting their gradients.

The difference in illumination between the two lunar NAC-1 images is about 90°, and a larger crater is located at the center of each image; affine deformation at different incident angles is also noticeable. The lunar NAC-2 image pair features various small craters with small geometrical deformations, and the difference in solar azimuth angle is about 60°. For the lunar NAC-3 image pair, the difference in solar azimuth angle is about 180°. Notably, the majority of each crater is shadowed, and the area also contains a large number of rocks casting shadows in different directions. For the first Mars CTX image pair, the difference in solar azimuth angle is approximately 45°, and a large crater can be observed to the bottom left of each image, accompanied by some large rocks and small craters. And the second CTX image pair features wrinkled pattern inside the craters in the middle.

4.2. Evaluation of results

To evaluate the performance of the proposed method relative to that

of SIFT, the matching results for the five datasets obtained using the two methods are compared. As the same feature points are used with both the SIFT method and the proposed method, with differences only in the descriptors, the following three measures are also compared: matches retrieved by both methods, matches retrieved only using SIFT and matches retrieved only using the proposed method.

For the lunar NAC-1 pair, as shown in Fig. 6, the original SIFT algorithm retrieves 138 good descriptor matches. However, the proposed method yields 199 good matches. The proposed method replicates most of the matches obtained using the original SIFT algorithm. The matches missed by SIFT are mainly from the planar area, which shows no significant texture pattern. In addition, the matches not retrieved by SIFT and obtained only using the proposed method are located mainly near features with abrupt local shadow changes, such as small craters and rocks. Detailed statistics on these matching results are provided in Table 2.

Turning to the lunar NAC-2 pair, Fig. 7 shows that SIFT yielded 192 matches and the proposed method 271. The lunar NAC-1 and NAC-2 datasets are similar in terms of difference in scale (about 20%) and difference in solar azimuth angle (60°), but affine deformation is less severe in the NAC-2 pair. Detailed analysis of the matches also reveals effects similar to those for the first image pair: most of the original SIFT matches are recalled in the proposed method, but many additional matches near small craters and rocks are added.

The third image pair lunar NAC-3 shows a large difference in solar azimuth angle. Most of the large craters in these images are shadowed from different directions. The area depicted also contains many salient white regions of small rocks. As indicated in Fig. 8 and Table 2, 94% of the original SIFT matches are also retrieved using the proposed method. Approximately half of the matches using the proposed method are additions to the original set.

For the first Mars CTX image pair shown in Fig. 9, 75 and 117 matches are obtained using the SIFT and the proposed method, respectively. This image pair shows a conspicuous difference not only in illumination direction, but also in global contrast. Most of the matches missed by SIFT are located in places with no local changes of grayscale values, such as the bright slopes of large craters. However, most of the additional matches retrieved using the proposed method are located in places with local illumination changes, such as the cross-shaped area in the center of a crater.

For the second Mars CTX image pair shown in Fig. 10, 380 and 395 matches are obtained using the SIFT and the proposed method, respectively. This image pair shows wrinkled patterns both inside the crater located in the middle of the image and also outside the crater. The gradient direction of the pattern inside the crater is almost orthogonal to that of the shadow direction caused by the solar azimuth angle. And the gradient of the wrinkled pattern outside the crater is also different from the former. Even with the existence of this directional pattern, the peaks of the histogram of the SIFT dominant directions are also aligned with the solar azimuth angle centered at 130° and 310° for the left image of the CTX-2 pair in Fig. 5. This is because that, the scales of most SIFT features are in fact smaller than the wrinkled pattern and in local areas, the solar

Table 1

Specifications of the five image pairs. Notably, differences in solar azimuth angle within each pair range from 20° to the most extreme case of about 180° (in Lunar NAC-3).

Image Pair	Image ID	Image Center (°)	Solar Azimuth Angle (°)	Resolution (m)	Size (pixel)
Lunar NAC-1	M1154657182	(45.21, 295.7)	93.42	2.10	1982 × 2 279
	M1180556756	(44.18, 295.69)	22.29	1.95	2 138 × 3 386
Lunar NAC-2	M1103787673	(38.36, 325.29)	65.92	1.55	2042 × 2 507
	M07106240	(38.23, 325.1)	122.59	1.75	1812 × 2 221
Lunar NAC-3	M1105774283	(1.3, 23.18)	0.93	2.20	1 510 × 1 907
	M106726943	(1.03, 23.12)	176.00	2.50	1 329 × 1 678
Mars CTX-1	D19_034816_1921	(12.09, 313.88)	208.57	5.50	927 × 1 132
	P13_006176_1918	(11.83, 313.81)	163.68	5.51	924 × 1 129
Mars CTX-2	D13_032378_2 135	(33.56, 273.37)	137.47	5.82	553 × 457
	D15_033156_2 135	(33.53, 273.38)	153.61	6.01	564 × 548

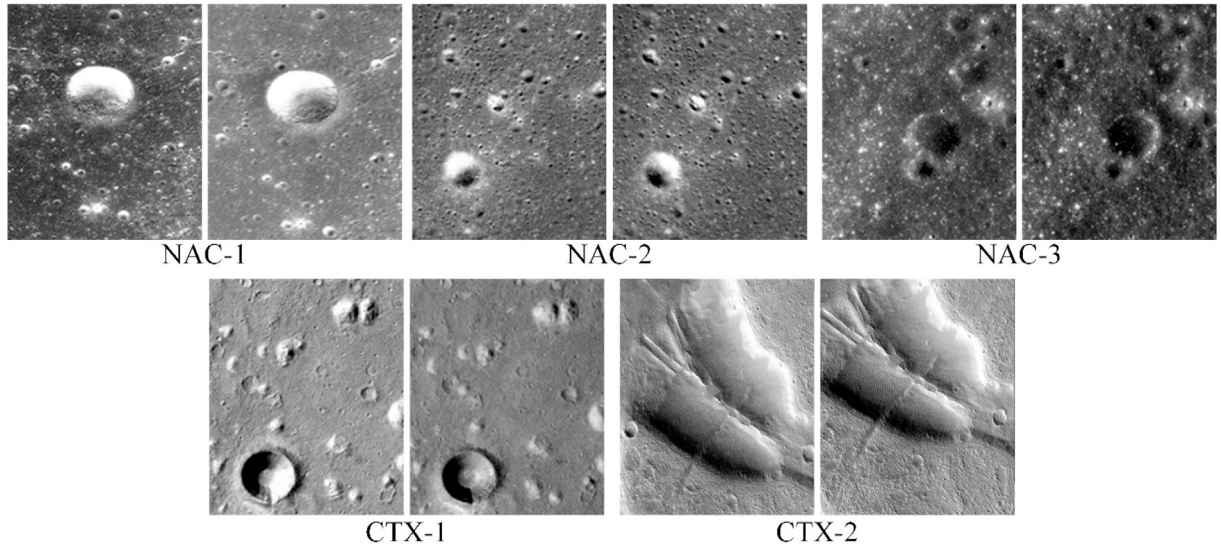


Fig. 5. Experimental datasets of the lunar NAC and the Mars CTX image pair. Notably, the positions of shadows in the craters are all different, reflecting differences in solar azimuth angle.

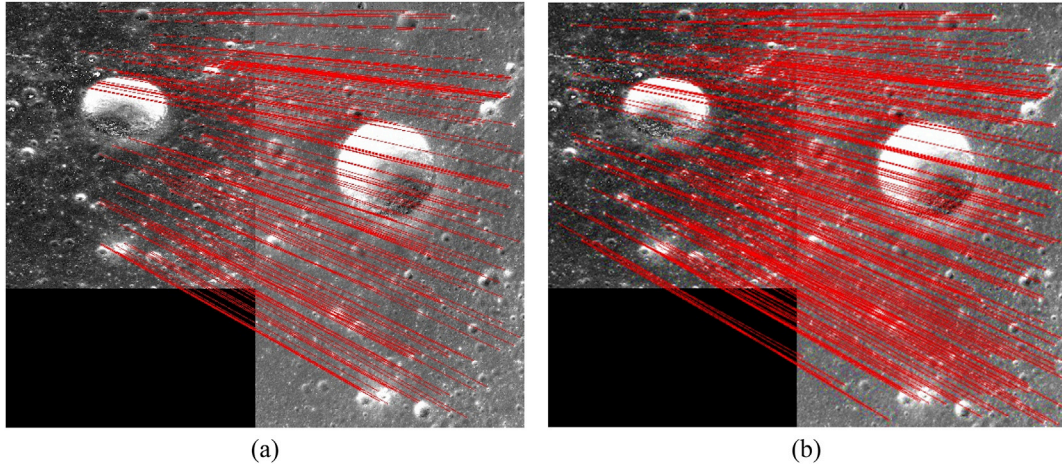


Fig. 6. Comparison of matching results for the lunar NAC-1 pair. (a) Matching results for the original SIFT and (b) matching results for the proposed method.

Table 2

Numbers of matches by SIFT, matches by proposed method, matches in common, matches by SIFT only and matches by proposed method only, shown from the second to the last column, respectively. The shaded row indicates the percentage of the original SIFT matches retrieved using the proposed method for each image pair.

Image Pair	#SIFT	#Proposed	#Both	#O-SIFT	#O-Proposed
Lunar NAC-1	138	199	114	24	75
		144%	83%	12%	66%
Lunar NAC-2	192	271	148	44	123
		141%	77%	23%	64%
Lunar NAC-3	125	177	118	7	59
		142%	94%	6%	47%
Mars CTX-1	75	117	43	32	74
		156%	57%	43%	99%
Mars CTX-2	380	395	368	12	27
		104%	97%	3%	7%

angle has the major impact on the gradients of the features. In addition, because the angle difference is smaller and this area is better textured, the original SIFT does not degrade much and the proposed method still outperforms SIFT slightly as shown in Table 2.

To further analyze the performance of the proposed method, we

investigate the additional matches by the proposed method in detail. The purpose is to determine whether the additional matches are directly influenced by the suppression of gradients around the twin peaks of the histogram. As shown in Table 3, the number of matches whose principal orientations are around the detected peaks of the histogram in a specific angle range as shown in Fig. 1, are recorded, namely those at 10°, 20° and 30° for the second, third and fourth columns, respectively. Notably, although the 30° region accounts for only a sixth of the total distribution, it contributes about two thirds of additional matches retrieved using the proposed method. This directly confirms that suppression increases the performance of the descriptors.

5. Conclusions

This paper proposes an improved SIFT matching method for high-resolution planetary remote sensing images that accommodates differences in illumination using histogram adaption. The guiding observation is that the principal orientation of SIFT features is highly biased toward the solar azimuth angle in textureless images of planetary surfaces with relatively homogeneous materials and albedos. The proposed method alleviates this effect by automatically detecting the peaks in a gradient histogram, and a Gaussian suppression function is adaptively tuned to

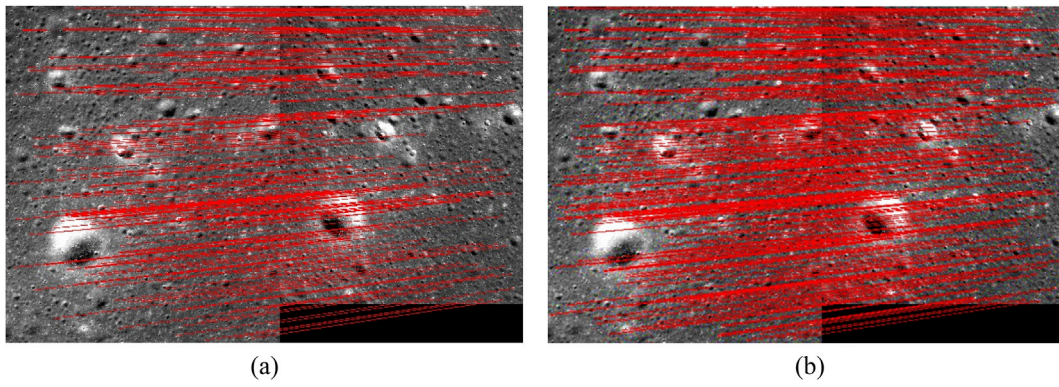


Fig. 7. Comparison of matching results for the lunar NAC-2 pair. (a) Matching results for the original SIFT and (b) matching results for the proposed method.

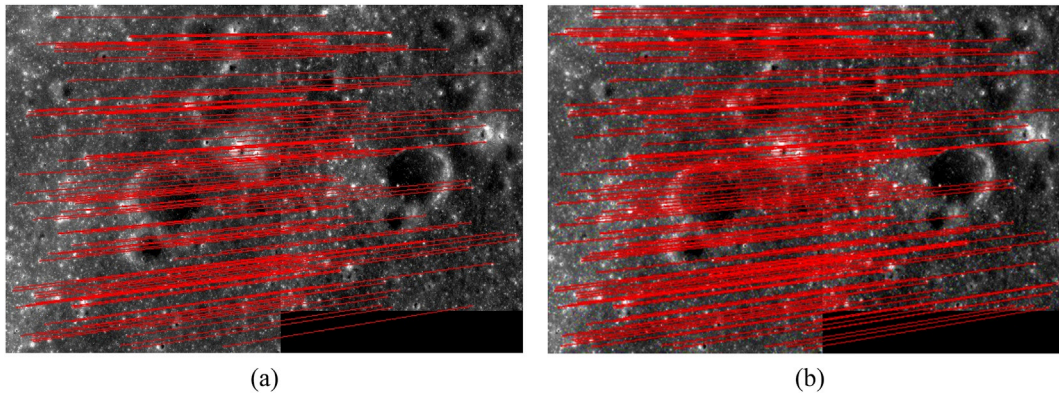


Fig. 8. Comparison of matching results for the lunar NAC-3 pair. (a) Matching results for the original SIFT and (b) matching results for the proposed method.

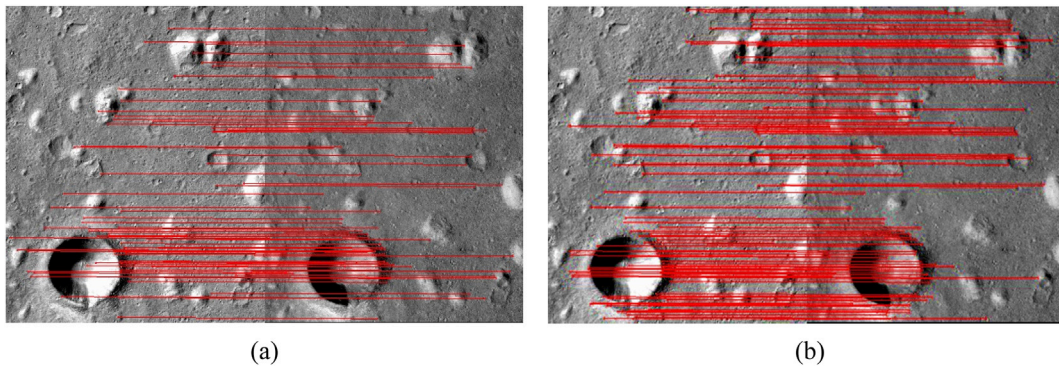


Fig. 9. Comparison of matching results for the Mars CTX-1 pair. (a) Matching results for the original SIFT and (b) matching results for the proposed method.

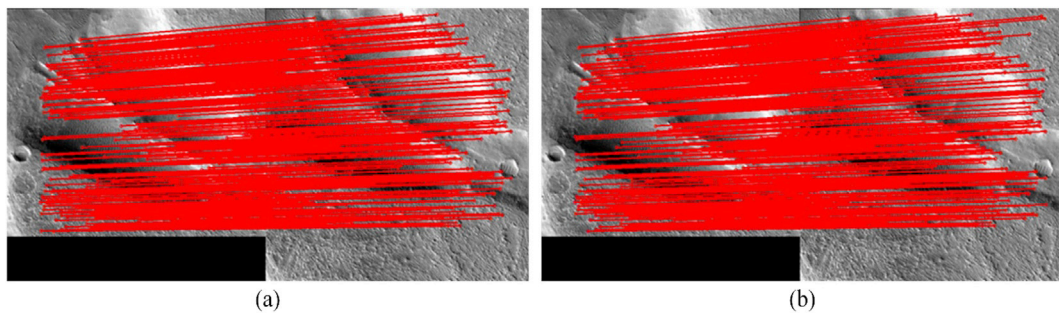


Fig. 10. Comparison of matching results for the Mars CTX-2 pair. (a) Matching results for the original SIFT and (b) matching results for the proposed method.

Table 3

Number of matches retrieved using the proposed method only whose principal orientations on the template images fall in a specific region around the detected peaks of the histogram, as shown in Fig. 1. The second, third and fourth columns show the results for the 10°, 20° and 30° regions, respectively. The shaded row shows the percentages with respect to the last column of Table 2.

Image Pair	#R-10	#R-20	#R-30
Lunar NAC-1	32	52	65
	43%	69%	87%
Lunar NAC-2	36	58	84
	29%	47%	68%
Lunar NAC-3	17	21	36
	29%	36%	61%
Mars CTX-1	17	35	48
	23%	47%	65%
Mars CTX-2	6	8	12
	22%	30%	44%

level differences in illumination. This suppression function is applied to the feature descriptors to improve principal-orientation estimation and robust histogram aggregation. The feature descriptors are matched using a modified method to account for changes in their distribution. Experimental evaluation reveals that the proposed method yields improvements of about 40%–60% as measured by the total number of correct matches. Further analysis confirms that the additional matches are retrieved predominantly due to the suppression of gradients.

The improved feature matching results from the proposed method can be used for the co-registration of multiple-source planetary remote sensing images for their synergistic use in various applications. The feature matching results can be used to constrain the subsequent dense matching, for example by triangulations generated from them (Wu et al., 2011). Similar strategy of accommodating illumination differences can also be applied for better dense matching using dense descriptors (Tola et al., 2010; Zbontar and LeCun, 2016), so that to generate 3D topographic models from planetary remote sensing images even with large illumination differences. In addition, for other types of planetary images with large geometric distortions, e.g. flyby and rover images, affine invariant feature detectors would be investigated and integrated with the illumination invariant descriptors.

Acknowledgments

The work described in this paper was funded by a grant from the Research Grants Council of Hong Kong (Project No.: PolyU 152086/15E) and grants from the National Natural Science Foundation of China (Project No.: 41471345 and Project No.: 41671426).

References

Alcantarilla, P.F., Nuevo, J., Bartoli, A., 2013. Fast explicit diffusion for accelerated features in nonlinear scale spaces. *IEEE Trans. Pattern Anal. Mach. Intell.* 34 (7), 1281–1298.

Arandjelovic, R., Zisserman, A., 2012. Three things everyone should know to improve object retrieval. In: *Proc. 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR2012)*. IEEE, Providence, RI, USA, pp. 2911–2918, 16–21 June.

Archinal, B.A., Laura, J., Kirk, R.L., Hare, T.M., Gaddis, L.R., Hagerty, J., 2017. Foundational data products needed to support planetary spatial data infrastructure. In: *Proc. Lunar and Planetary Science Conference*, The Woodlands, Texas, USA, March 20–24.

ASU, 2017. Lunar Reconnaissance Orbiter Camera. <http://roc.sese.asu.edu/>. (Accessed 17 August 2017).

Barazzetti, L., Scaioni, M., Remondino, F., 2010. Orientation and 3D modelling from markerless terrestrial images: combining accuracy with automation. *Photogramm. Rec.* 25 (132), 356–381.

Bay, H., Ess, A., Tuytelaars, T., Van Gool, L., 2008. Speeded-up robust features (SURF). *Comput. Vis. Image Understand.* 110 (3), 346–359.

Brown, M., Hua, G., Winder, S., 2011. Discriminative learning of local image descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.* 33 (1), 43–57.

Brown, M., Szeliski, R., Winder, S., 2005. Multi-image matching using multi-scale oriented patches. In: *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR2005)*. IEEE, San Diego, CA, USA, pp. 510–517, 20–25 June.

Calonder, M., Lepetit, V., Ozuysal, M., Trzcinski, T., Strecha, C., Fua, P., 2012. BRIEF: computing a local binary descriptor very fast. *IEEE Trans. Pattern Anal. Mach. Intell.* 34 (7), 1281–1298.

Di, K., Liu, B., Peng, M., Xin, X., Jia, M., Zuo, W., Ping, J., Wu, B., Oberst, J., 2017. An initiative for construction of new-generation lunar global control network using multi-mission data. *Int. Arch. Photogramm. Rem. Sens. Spatial Inf. Sci.* 42.

Di, K., Liu, Y., Liu, B., Peng, M., Hu, W., 2014. A self-calibration bundle adjustment method for photogrammetric processing of Chang'E-2 stereo lunar imagery. *IEEE Trans. Geosci. Rem. Sens.* 52 (9), 5432–5442.

Fischler, M.A., Bolles, R.C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* 24 (6), 381–395.

Foroosh, H., Zerubia, J.B., Berthod, M., 2002. Extension of phase correlation to subpixel registration. *IEEE Trans. Image Process.* 11 (3), 188–200.

Förstner, W., Gülch, E., 1987. A fast operator for detection and precise location of distinct points, corners and centres of circular features. In: *Proc. ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data*, Interlaken, Switzerland, pp. 281–305, 2–4 June.

Harris, C., Stephens, M., 1988. A combined corner and edge detector. In: *Proc. Alvey Vision Conference*, Manchester, UK, pp. 147–152, 31 August–2 September.

Hu, H., Ding, Y., Zhu, Q., Wu, B., Xie, L., Chen, M., 2016. Stable least-squares matching for oblique images using bound constrained optimization and a robust loss function. *ISPRS J. Photogramm. Rem. Sens.* 118, 53–67.

Hu, H., Wu, B., 2017. Bound-Constrained multiple-image least-squares matching for multiple-resolution images. *Photogramm. Eng. Rem. Sens.* 83 (10), 667–677.

Hurwitz, D.M., Head, J.W., Hiesinger, H., 2013. Lunar sinuous rilles: distribution, characteristics, and implications for their origin. *Planet. Space Sci.* 79, 1–38.

Jazayeri, I., Fraser, C.S., 2010. Interest operators for feature-based matching in close range photogrammetry. *Photogramm. Rec.* 25 (129), 24–41.

Kirk, R.L., Howington Kraus, E., Rosiek, M.R., Anderson, J.A., Archinal, B.A., Becker, K.J., Cook, D.A., Galuszka, D.M., Geissler, P.E., Hare, T.M., 2008. Ultrahigh resolution topographic mapping of Mars with MRO HiRISE stereo images: meter-scale slopes of candidate Phoenix landing sites. *J. Geophys. Res. Planets* 113 (E3), E00A24.

Kovesi, P., 2003. Phase congruency detects corners and edges. In: *Proc. 12th Digital Image Computing: Techniques and Applications*, Sydney, Dec 10–12.

Li, H., Man, Y., 2016. Robust Multi-source Image Registration for Optical Satellite Based on Phase Information, pp. 865–878.

Li, R., Hwangbo, J., Chen, Y., Di, K., 2011. Rigorous photogrammetric processing of HiRISE stereo imagery for Mars topographic mapping. *IEEE Trans. Geosci. Rem. Sens.* 49 (7), 2558–2572.

Liang, J., Liu, X., Huang, K., Li, X., Wang, D., Wang, X., 2014. Automatic registration of multisensor images using an integrated spatial and mutual information (SMI) metric. *IEEE Trans. Geosci. Rem. Sens.* 52 (1), 603–615.

Lindeberg, T., 1993. Detecting salient blob-like image structures and their scales with a scale-space primal sketch: a method for focus-of-attention. *Int. J. Comput. Vis.* 11 (3), 283–318.

Lindeberg, T., 1998. Feature detection with automatic scale selection. *Int. J. Comput. Vis.* 30 (2), 79–116.

Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* 60 (2), 91–110.

Malin, M.C., Bell, J.F., Cantor, B.A., Caplinger, M.A., Calvin, W.M., Clancy, R.T., Edgett, K.S., Edwards, L., Haberle, R.M., James, P.B., 2007. Context camera investigation on board the Mars reconnaissance orbiter. *J. Geophys. Res. Planets* 112 (E5).

McEwen, A.S., Eliason, E.M., Bergstrom, J.W., Bridges, N.T., Hansen, C.J., Delamere, W.A., Grant, J.A., Gulick, V.C., Herkenhoff, K.E., Keszthelyi, L., Kirk, R.L., Mellon, M.T., Squyres, S.W., Thomas, N., Weitz, C.M., 2007. Mars reconnaissance Orbiter's high resolution imaging science experiment (HiRISE). *J. Geophys. Res. Planets* 112 (E05S02).

Matas, J., Chum, O., Urban, M., Pajdla, T., 2004. Robust wide-baseline stereo from maximally stable extremal regions. *Image Vis. Comput.* 22 (10), 761–767.

Mikolajczyk, K., Schmid, C., 2004. Scale & affine invariant interest point detectors. *Int. J. Comput. Vis.* 60 (1), 63–86.

Mikolajczyk, K., Schmid, C., 2005. A performance evaluation of local descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.* 27 (10), 1615–1630.

Moisan, L., Moulon, P., Monasse, P., 2012. Automatic homographic registration of a pair of images, with a contrario elimination of outliers. *Image Process. On Line* 2, 56–73.

Morel, J., Yu, G., 2009. ASIFT: a new framework for fully affine invariant image comparison. *SIAM J. Imag. Sci.* 2 (2), 438–469.

Morgan, G.L.K., Liu, J.G., Yan, H., 2010. Precise subpixel disparity measurement from very narrow baseline stereo. *IEEE Trans. Geosci. Rem. Sens.* 48 (9), 3424–3433.

Muja, M., Lowe, D.G., 2014. Scalable nearest neighbor algorithms for high dimensional data. *IEEE Trans. Pattern Anal. Mach. Intell.* 36 (11), 2227–2240.

Robinson, M.S., Brylow, S.M., Tschimmel, M., Humm, D., Lawrence, S.J., Thomas, P.C., Denevi, B.W., Bowman-Cisneros, E., Zerr, J., Ravine, M.A., 2010. Lunar reconnaissance orbiter camera (LROC) instrument overview. *Space Sci. Rev.* 150 (1), 81–124.

Rosten, E., Porter, R., Drummond, T., 2010. Faster and better: a machine learning approach to corner detection. *IEEE Trans. Pattern Anal. Mach. Intell.* 32 (1), 105–119.

Rublee, E., Rabaud, V., Konolige, K., Bradski, G., 2011. ORB: an efficient alternative to SIFT or SURF. In: *Proc. IEEE International Conference on Computer Vision (ICCV2011)*. IEEE, Barcelona, Spain, pp. 2564–2571, 6–13 November.

Schönberger, J.L., Hardmeier, H., Sattler, T., Pollefeys, M., 2017. Comparative evaluation of hand-crafted and learned local features. In: *Proc. Conference on Computer Vision and Pattern Recognition (CVPR)*.

- Simonyan, K., Vedaldi, A., Zisserman, A., 2014. Learning local feature descriptors using convex optimisation. *IEEE Trans. Pattern Anal. Mach. Intell.* 36 (8), 1573–1585.
- Toews, M., Wells, W., 2009. SIFT-Rank: ordinal description for invariant feature correspondence. In: *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR2009)*. IEEE, Miami, FL, USA, pp. 172–177, 20–25 June.
- Tola, E., Lepetit, V., Fua, P., 2010. Daisy: an efficient dense descriptor applied to wide-baseline stereo. *IEEE Trans. Pattern Anal. Mach. Intell.* 32 (5), 815–830.
- Verdie, Y., Yi, K., Fua, P., Lepetit, V., 2015. TILDE: a temporally invariant learned detector. In: *Proc.*, 2015, pp. 5279–5288.
- Wang, Z., Fan, B., Wang, G., Wu, F.C., 2016. Exploring local and overall ordinal information for robust feature description. *IEEE Trans. Pattern Anal. Mach. Intell.* 38 (11), 2198–2211.
- Wang, Z., Fan, B., Wu, F., 2011. Local intensity order pattern for feature description. In: *Proc. IEEE International Conference on Computer Vision (ICCV2011)*. IEEE, Barcelona, Spain, pp. 603–610, 6–13 November.
- Watters, W.A., Geiger, L.M., Fendrock, M., Gibson, R., 2015. Morphometry of small recent impact craters on Mars: size and terrain dependence, short-term modification. *J. Geophys. Res. Planets* 120 (2), 226–254.
- Wong, A., Clausi, D.A., 2007. ARRSI: automatic registration of remote-sensing images. *IEEE Trans. Geosci. Rem. Sens.* 45 (5), 1483–1493.
- Wu, B., Zhang, Y., Zhu, Q., 2011. A triangulation-based hierarchical image matching method for wide-baseline images. *Photogramm. Eng. Rem. Sens.* 77 (7), 695–708.
- Wu, B., Guo, J., Hu, H., Li, Z., Chen, Y., 2013. Co-registration of lunar topographic models derived from Chang'E-1, SELENE, and LRO laser altimeter data based on a novel surface matching method. *Earth Planet Sci. Lett.* 364, 68–84.
- Wu, B., Liu, W.C., 2017. Calibration of boresight offset of LROC NAC imagery for precision lunar topographic mapping. *ISPRS J. Photogramm. Rem. Sens.* 128, 372–387.
- Wu, B., Liu, W.C., Grumpe, A., Wöhler, C., 2018. Construction of pixel-level resolution DEMs from monocular images by shape and albedo from shading constrained with low-resolution DEM. *ISPRS J. Photogramm. Rem. Sens.* <https://doi.org/10.1016/j.isprsjprs.2017.03.007>.
- Ye, Y., Xiong, L., Shan, J., 2012. Automated multi-source remote sensing image registration based on phase congruency. *ISPRS-Int. Arch. Photogramm. Rem. Sens. Spat. Info. Sci.* XXXIX (B6), 189–194.
- Yi, K.M., Trulls, E., Lepetit, V., Fua, P., 2016. LIFT: Learned Invariant Feature Transform. *arXiv preprint arXiv:1603.09114*.
- Zbontar, J., LeCun, Y., 2016. Stereo matching by training a convolutional neural network to compare image patches. *J. Mach. Learn. Res.* 17 (1–32), 2.
- Zhu, Q., Zhang, Y., Wu, B., Zhang, Y., 2010. Multiple close-range image matching based on a self-adaptive triangle constraint. *Photogramm. Rec.* 25 (132), 437–453.