

# *An ASIFT-based local registration method for satellite imagery*

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Wang, X., LI, Y., Wei, H. ORCID: <https://orcid.org/0000-0002-9664-5748> and Liu, F. (2015) An ASIFT-based local registration method for satellite imagery. Remote Sensing, 7 (6). pp. 7044-7061. ISSN 2072-4292 doi: <https://doi.org/10.3390/rs70607044> Available at <https://centaur.reading.ac.uk/41208/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

Published version at: <http://www.mdpi.com/journal/remotesensing>

To link to this article DOI: <http://dx.doi.org/10.3390/rs70607044>

Publisher: MDPI

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

Article

## An ASIFT-Based Local Registration Method for Satellite Imagery

Xiangjun Wang <sup>1</sup>, Yang Li <sup>1,\*</sup>, Hong Wei <sup>2</sup> and Feng Liu <sup>1</sup>

<sup>1</sup> State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, 92 Weijin Road, Tianjin 300072, China; E-Mails: xdocuxjw@vip.163.com (X.W.); tjuliufeng@tju.edu.cn (F.L.)

<sup>2</sup> Computer Vision Group, School of Systems Engineering, University of Reading, Berkshire RG6 6AY, UK; E-Mail: h.wei@reading.ac.uk

\* Author to whom correspondence should be addressed; E-Mail: y.li1987@hotmail.com; Tel./Fax: +86-22-2740-3395.

Academic Editors: Gonzalo Pajares Martinsanz and Prasad S. Thenkabail

Received: 20 March 2015 / Accepted: 22 May 2015 / Published: 29 May 2015

---

**Abstract:** Imagery registration is a fundamental step, which greatly affects later processes in image mosaic, multi-spectral image fusion, digital surface modelling, *etc.*, where the final solution needs blending of pixel information from more than one images. It is highly desired to find a way to identify registration regions among input stereo image pairs with high accuracy, particularly in remote sensing applications in which ground control points (GCPs) are not always available, such as in selecting a landing zone on an outer space planet. In this paper, a framework for localization in image registration is developed. It strengthened the local registration accuracy from two aspects: less reprojection error and better feature point distribution. Affine scale-invariant feature transform (ASIFT) was used for acquiring feature points and correspondences on the input images. Then, a homography matrix was estimated as the transformation model by an improved random sample consensus (IM-RANSAC) algorithm. In order to identify a registration region with a better spatial distribution of feature points, the Euclidean distance between the feature points is applied (named the S criterion). Finally, the parameters of the homography matrix were optimized by the Levenberg–Marquardt (LM) algorithm with selective feature points from the chosen registration region. In the experiment section, the Chang'E-2 satellite remote sensing imagery was used for evaluating the performance of the proposed method. The experiment result demonstrates that the proposed method can automatically locate a specific region with high registration accuracy between input images by achieving lower root mean square error (RMSE) and better distribution of feature points.

**Keywords:** satellite remote sensing; local image registration; image mosaic; ASIFT; RANSAC

---

## 1. Introduction

Image registration refers to aligning two image point sets that share the same scene in a common coordination system. It is widely used in remote sensing applications, including multi-spectral classification, environment surveillance, super-resolution reconstruction, change detection and outer space exploitation [1–4]. It is known that among those factors, the number of control points (CPs) and their spatial distribution across the registered region are of great importance to perform reliable registration [5]. Inadequate, inaccurate or ill-distributed CPs may lead to undesirable consequences [6,7]. In traditional registration approaches, CPs are obtained either manually selected by an experienced expert or pre-defined ground control points (GCP) [8]. Unfortunately, the accuracy of manually-selected CPs are dependent on the judgment of the specific expert, and this may vary from person to person, while it is expensive or even impossible to establish these GCPs in some areas, such as the Arctic or outer space planets. To solve the problem mentioned above, automatic registration algorithms that require no or little manual work have emerged in the last few decades. Automatic registration algorithms can be classified into two main categories, intensity based and feature based [2,9]. Intensity-based algorithms are more sensitive to variations in contrast, illumination, image noise and scale effects than feature-based ones [1,2]. Moreover, they are unsuitable for multispectral image registration, because the intensity value of multispectral image does not change linearly [10]. As a result, researchers gradually put their attention toward the feature-based local descriptors. Lowe [11,12] proposed a feature-based algorithm to extract scale-invariant key point and to find correspondence between image pairs, named scale-invariant feature transform (SIFT). SIFT locates key point by finding scale-space extreme in the difference-of-Gaussian (DoG) function; once a key point is detected, it will be described and matched based on its orientation, scale and coordination. SIFT and its variations are the most popular and successful feature detection algorithms in many remote sensing registration applications [2,9,13–16]. However, if the image pairs are obtained under a large difference of camera viewpoints, the correspondences generated by SIFT would be too few to perform a reliable registration. Affine-SIFT (ASIFT) is the anti-affine-enhanced edition of SIFT, and it outperformed many state-of-the-art methods in finding correspondences under affine distortion and large viewpoint changes [17]. Though ASIFT can acquire sufficient correspondences even under a severe viewpoint change, it suffers from low registration accuracy [18].

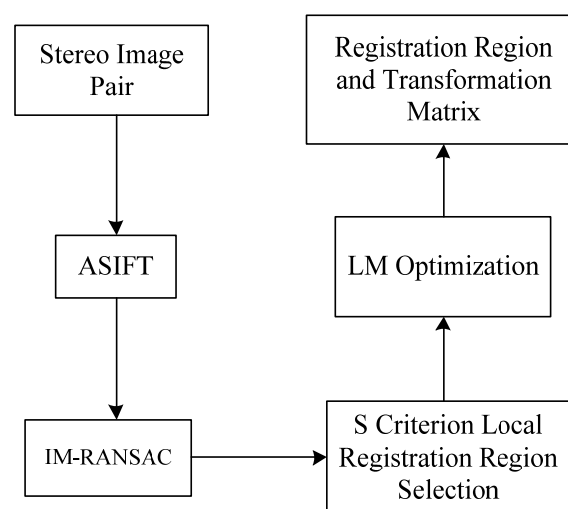
In many remote sensing applications, local registration accuracy is much more important than global registration accuracy. For instance, in the selection of a lunar landing site, in order to obtain the actual morphology of the lunar terrain, a high registration accuracy is essential for acquiring the three-dimensional information of the lunar surface. However, it is unnecessary to perform a global registration, since a small area with high registration accuracy within the preselected landing area is large enough for landing on the moon surface. Inspired by this practical demand, we developed a local registration method in this paper to identify a region between an input image pair with higher local registration accuracy than global registration has. Given a stereo image pair as the input reference and

sensed image, conventional feature-based registration methods take all of the feature points generated by a feature extraction algorithm as the CPs after the procedure of outlier removal, but pay little attention to the distribution of the CPs. The locations of the feature points actually depend on the context of input images, in other words the feature points are “randomly” distributed across the registration region. Toward achieving the maximum local registration accuracy between the reference and a sensed image in the proposed method, we firstly take advantage of the feature extraction capacity of ASIFT to obtain sufficient feature points and coarse correspondences (tentative correspondences), the mismatched and low accuracy correspondences among the tentative correspondences are then eliminated by the improved RANSAC (IM-RANSAC) algorithm, which is proposed in this paper to gain high accuracy inliers. Secondly, with the proposed S criterion, the standard deviation of Euclidean distance summation of each feature point to remaining feature points is calculated to select an image region in which the feature points are uniformly distributed. This ensures a reasonable distribution for the leaving correspondences. Finally, the non-linear parameter optimization algorithm Levenberg–Marquardt (LM) is employed to further minimize the matching error of correspondences. After the above three steps, it automatically secures a region in the image pair to be registered with a high registration accuracy.

The rest of paper is organized as follows. Section 2 describes the proposed method in detail. Section 3 gives the experimental results and a discussion, followed by the conclusions in Section 4.

## 2. Methodology

As shown in Figure 1, the developed local registration algorithm consists of four main steps. With the input of a stereo image pair, ASIFT is applied to generate a vast amount of correspondences; IM-RANSAC is then used for removing incorrect and low accuracy matching feature points to create a set of inliers; an automated region selection mechanism named the S criterion is further developed in the algorithm, combined with an optimization process to locate the image registration region and obtain the transformation model. The details of the method are discussed in the following sections.



**Figure 1.** Main steps of proposed method. LM, Levenberg–Marquardt.

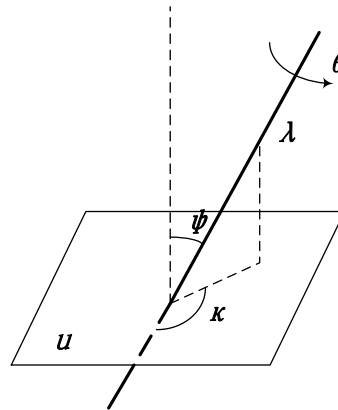
### 2.1. A Brief Introduction to ASIFT

The key idea behind ASIFT is to simulate different viewpoints of input images by changing the spatial position of camera optical axis, and then, the same procedure of SIFT is applied to extract correspondences from simulated images. It is known that any affine matrix  $\mathbf{A}$  can be decomposed as:

$$\mathbf{A} = \lambda \mathbf{R}_\theta T_t \mathbf{R}_\kappa = \lambda \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \kappa & -\sin \kappa \\ \sin \kappa & \cos \kappa \end{pmatrix} \quad (1)$$

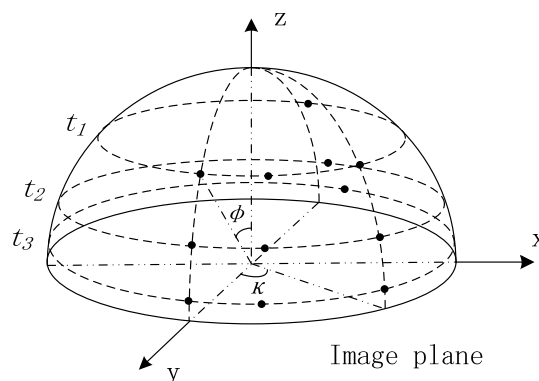
where, as shown in Figure 2,  $\lambda$  is a scale factor,  $\kappa$  is a rotation angle around image normal by the camera,  $\theta$  is the roll angle around the camera axis,  $T_t$  is a tilt defined as  $t = (1/\cos \phi)$ , where  $\phi$  is the angle between the camera optical axis and the image normal, and  $\mathbf{R}_\theta$  and  $\mathbf{R}_\kappa$  are rotation matrices. Because SIFT is robust to rotation and scale, thus  $\lambda$  and  $\theta$  can be neglected from Equation (1). As a result, the affine matrix  $\mathbf{A}$  becomes:

$$\mathbf{A} = \begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \kappa & -\sin \kappa \\ \sin \kappa & \cos \kappa \end{pmatrix} \quad (2)$$



**Figure 2.** Spatial relationship of image plane  $u$  and camera optical axis.

Therefore, ASIFT can achieve full affine invariance simulation by only changing the combination of the  $t$  and  $\kappa$  parameters. Some examples of the camera position according to different  $(t, \kappa)$  pairs are denoted by black dots in Figure 3. After locating the camera position, simulated affine images can be obtained from each of the positions. The simulated image pair with the largest number of correspondences is then used for registration. For details about ASIFT, please refer to [16].



**Figure 3.** Illustration of different camera positions.

## 2.2. Proposed Algorithm

### 2.2.1. Improved RANSAC

Random sample consensus (RANSAC) is an iterative algorithm used to estimate model parameters from a dataset, which was first introduced by Fischler and Bolles in 1981 [19]. The RANSAC algorithm is a typical method used in remote sensing image registration to eliminate incorrect correspondences (outliers) and to estimate the transformation matrix between stereo image pairs. However, the traditional RANSAC requires the setting of a predefined threshold, which often depends on specific problems. Unfortunately, it is usually difficult to give a proper threshold, because too large of a threshold will introduce big reprojection error to inliers, while too small of a threshold may lead to the failure of the algorithm due to insufficient inliers. Reprojection error refers to the image coordination difference between a point from the registered image and its matching point in the reference image. Here, in this paper, we develop an improved RANSAC algorithm, which is named IM-RANSAC, to give a new iteration strategy for the threshold selection. As the flowchart in Figure 4 shows, given a small positive number  $\Delta t$  as a threshold increment, let the initial threshold  $T = \Delta t$ , and a random sample set  $M$  with the size of  $m$  is selected from the ASIFT-generated tentative corresponding point set  $N$ . According to multi-view geometry [20], images of a plane captured at different viewpoints can be related by a homography transformation model. The terrain height variation within the range of the area where the experimental image pairs are taken can be ignored compared to an orbit height of 100 km of Chang'E-2, and the spatial resolution of a pixel in the horizontal is 7 meters. As a result, the experimental image pairs are considered as being captured from the same plane. Bearing these in mind, we use the eight-parameter homography matrix as the transformation model to estimate geometric distortions between stereo image pairs. The homography matrix  $\mathbf{H}$  is calculated by  $M$  through the least squares method; if the reprojection error between two correspondences is equal to or less than the threshold, they will be considered as inliers. Then, a consensus consisting of inliers is generated as the possible output of IM-RANSAC.  $P$  is the probability that the final homograph matrix is calculated by a random outlier containing  $M$ ; it is given by:

$$P = (1 - \omega^m)^k \quad (3)$$

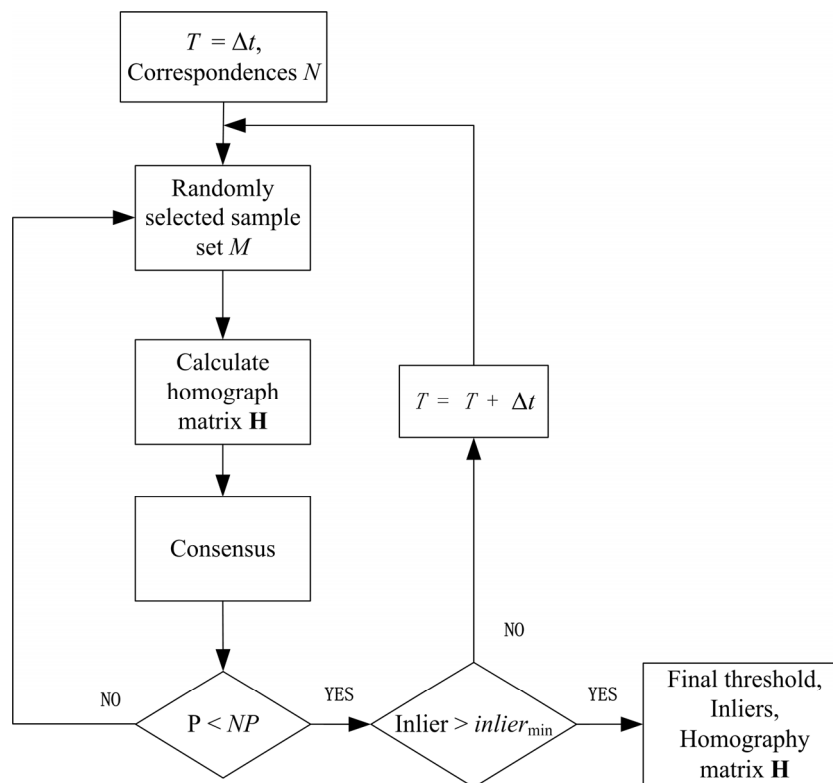
where  $\omega$  is the probability of choosing an inlier from the tentative corresponding point set  $N$ ;  $k$  is the number of times of sampling  $M$ . When  $P$  is less than a sufficiently small positive  $NP$ , the homography matrix  $\mathbf{H}$  is then considered to be acceptable. The number of inliers output by IM-RANSAC is constrained by the parameter  $inlier_{\min}$ , which represents for the minimum acceptable number of inliers, to ensure that there are enough inliers to go through the rest of the processes in the proposed algorithm. The  $inlier_{\min}$  is defined as [21]:

$$inlier_{\min} = \min_j \left\{ j : \sum_{i=j}^n Q_N(i) < \phi \right\} \quad (4)$$

where  $n$  is the total number of correspondences in  $N$ ; function  $Q_N(i)$  represents the probability that  $i$  out of  $n$  tentative correspondences support an arbitrary transformation model derived from an outlier-containing sample set  $M$ . When  $\sum_i^n Q_N(i)$  is smaller than a small positive  $\phi$ , the probability that an inlier set that has a number of inliers more than  $i$  is incorrect will be far less than  $\phi$ . Under the premise that the event of a correspondence belongs to inlier set is independent Bernoulli [22],  $Q_N(i)$  can be given by:

$$Q_N(i) = \beta^{i-m} (1-\beta)^{n-i+m} C_{n-m}^{i-m} \quad (5)$$

where  $\beta$  is the probability that an transformation model derived by outlier-containing sample set  $M$  is supported by one of the correspondences in  $N$ , which is not included in the outlier-containing sample set  $M$ ;  $i$  is in the value range of  $m < i \leq n$ ;  $m$  is the size of the  $M$ . If the number of inliers is less than  $inlier_{min}$ , the generated consensus will be rejected, meanwhile another IM-RANSAC evaluation begins with the updated threshold, which is increased by  $\Delta t$  on the previous one. With the increase of the threshold, the requirement of  $inlier_{min}$  needs to be satisfied, in other words, and the IM-RANSAC always has an output set of inliers with the minimum threshold unless the number of tentative correspondences is too few in  $N$ .



**Figure 4.** Flowchart of improved IM-RANSAC.

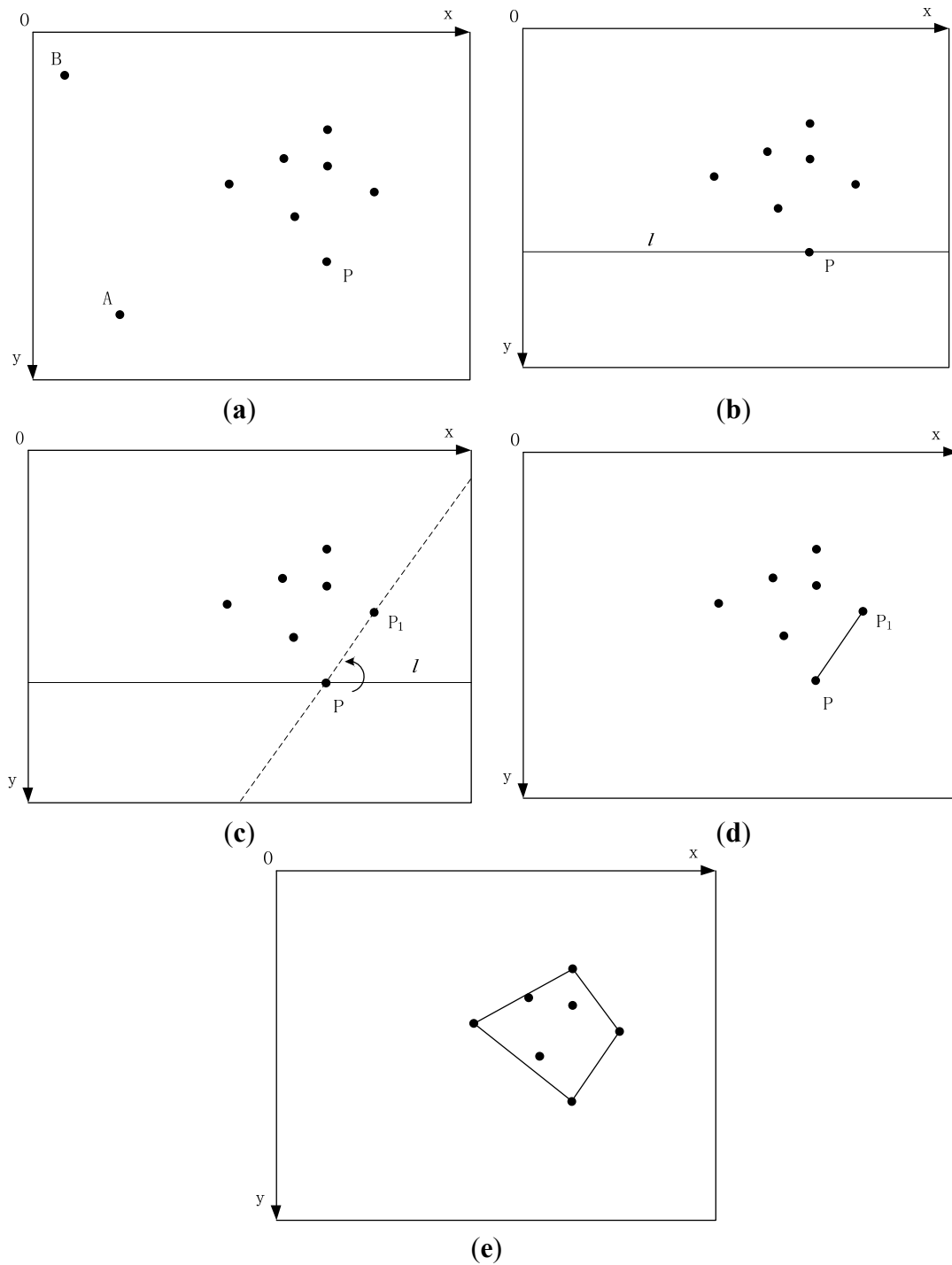
After the IM-RANSAC process, the final threshold, homography matrix  $\mathbf{H}$ , and inliers with the minimum final threshold are obtained simultaneously. The coarse correspondences that entered into IM-RANSAC are from ASIFT, so the quantity and reliability are promised unless the input stereo image pair is totally unmatchable. However, even when the input images are not matching at all, the algorithm will stop at ASIFT before it goes into IM-RANSAC. The stopping criterion of IM-RANSAC is controlled by  $inlier_{min}$ ; once the number of inliers in the consensus is more than  $inlier_{min}$ , IM-RANSAC will stop. Thus, the probability that the final threshold becomes too large to be reliable is very small. In IM-RANSAC, there is no need to pre-define the threshold for determining inliers, for it is automatically generated in the course of iteration by gradually increasing the initial input value. This dramatically improves the conventional RANSAC algorithm, which requires a pre-defined threshold in operation.

### 2.2.2. Automatic Selection of Registration Area

The value of final threshold from the IM-RANSAC output is dependent on the small increment; therefore, the increment should be set as small as possible to obtain inliers with small reprojection error. With the smaller threshold, the number of inliers will be also lesser. As a result, the distribution of some inliers will be scattered from others, leading to an unevenly spatial distribution, and the registration accuracy in the scattered portion will be low. Moreover, the derived homography matrix also will be less accurate even for the dense portion. In this paper, a novel algorithm named S criterion is applied to eliminate this kind of scattered point and automatically obtain a local registration region in which the feature points are distributed uniformly. Suppose that the Euclidean distance between point  $i$  and point  $j$  is denoted by  $d_{ij}$  and the standard deviation of Euclidean distance summation of point  $i$  to all of the other remaining feature points is denoted by  $S$ . If the point  $i$  has:

$$\left| \sum_{j=1}^n d_{ij} - \frac{\sum_{i=1}^n \sum_{j=1}^n d_{ij}}{n} \right| > S \quad (6)$$

the point  $i$ , as well as its matching point on the other image of the stereo image pair should be removed from the inliers. As shown in Figure 5a, the inliers on one of the images of the stereo image pair are illustrated in the o-x-y image coordinate system; round black dots represent the inliers. According to the S criterion, scattered points A and B should be removed. It can also use  $\sigma S$  ( $0 < \sigma \leq 1$ ) as the criterion to remove more feature points for a denser region, but with the consideration of the size of registration region; here, we chose  $\sigma=1$  in the following experiments. After the removal of bad distributed feature points, the automatic region selection mechanism of the S criterion can be summarized as follows: (1) the point that has the maximum coordinate in the y dimension among the remaining points of inliers is selected as the starting point. Draw a line through the starting point parallel to the x-axis; as shown in Figure 5b, point P is the starting point, and the derived line is denoted by  $l$ ; (2) rotate the line  $l$  counter clockwise (or clockwise) around the starting point until it reaches any one of the points of inliers; as shown in Figure 5c, the Point  $P_1$  is the reached point, and the dotted line is the stopping position of line  $l$ ; (3) connect the starting point and the reached point, as shown in Figure 5d; set the reached point as the new starting point, and continue to rotate the line  $l$  from its last stopping position; (4) repeat Steps (2) and (3) until the very starting point with the maximum y coordinate is reached by line  $l$  again; and (5) finally, the registration region as shown in Figure 5e can be automatically obtained. Both the distribution and the reprojection error of inliers optimized in this region result in better registration accuracy. The procedures above can locate a relative larger area with a profile of convex polygon, and they are computationally light algorithms with only a few steps. Because the coordinates of inlier points are not integers, we reassigned their locations to the nearest integer coordinates during the procedure.



**Figure 5.** Illustration of S criterion algorithm. (a) Distribution of IM-RANSAC outputted inliers. (b) Remaining inliers after the S criterion selection. (b–e) Demonstration of registration region locating by the S criterion.

### 2.2.3. LM Optimization

Although IM-RANSAC can obtain smaller reprojection error by means of automatically choosing a proper threshold, the homography transformation matrix  $\mathbf{H}$  is not yet in its optimal status, due to the fact that it is computed from all of the inliers outputted by IM-RANSAC rather than adapting to the S criterion selected region. Thus, we apply the nonlinear optimization algorithm LM [23] to optimize the

parameters of  $\mathbf{H}$  merely by the inliers in the S criterion selected area; by doing so, the homography transformation will be more precise with respect to the local registration region. We denote the parameters of  $\mathbf{H}$  by  $\mathbf{h} = [h_1, h_2, \dots, h_7, h_8]$  and the correspondences in the S criterion selected region by  $(x_i, y_i)$  and  $(X_i, Y_i)$  ( $i = 1, 2, 3, \dots, n$ ). Let  $(X'_i, Y'_i)$  be the homography matrix  $\mathbf{H}$  transformed coordinates of  $(x_i, y_i)$ ; thus, the transform equations were defined as:

$$X'_i = H_X(x_i, y_i; \mathbf{h}) = \frac{h_1 x_i + h_2 y_i + h_3}{h_7 x_i + h_8 y_i + 1} \quad (7)$$

$$Y'_i = H_Y(x_i, y_i; \mathbf{h}) = \frac{h_4 x_i + h_5 y_i + h_6}{h_7 x_i + h_8 y_i + 1} \quad (8)$$

The total reprojection error to be minimized between the correspondences can be formulated as:

$$E(\mathbf{h}) = \sum_{i=1}^n [(X'_i - X_i)^2 + (Y'_i - Y_i)^2] \quad (9)$$

The LM algorithm is particularly suited to a problem that is expressed by a sum of squared residuals. Therefore, the reduction of reprojection error is converted to a minimization problem in which optimal solution is given by the LM algorithm via minimizing over  $\mathbf{h}$ :

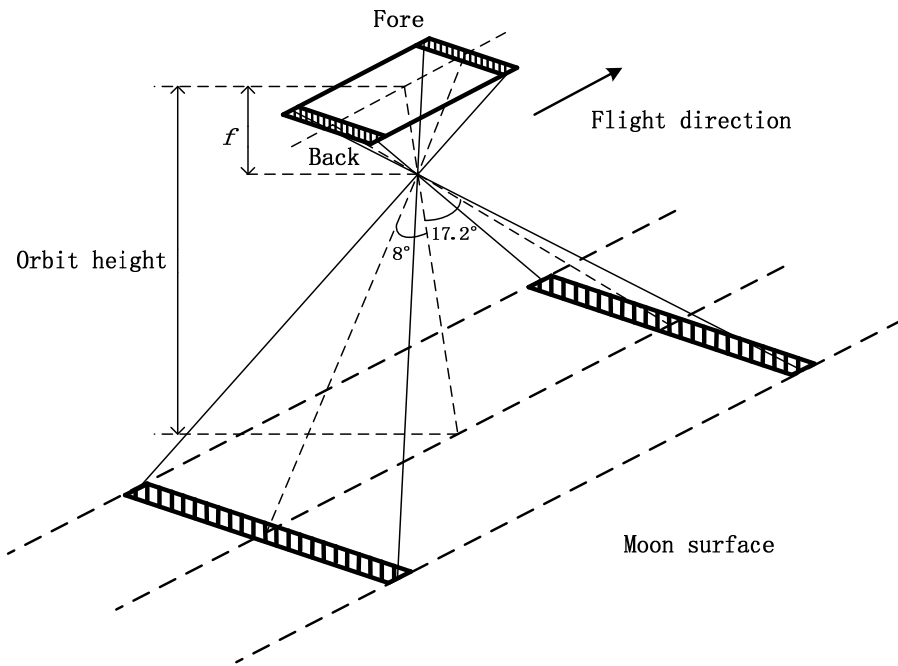
$$\hat{\mathbf{h}} = \arg \min_{\mathbf{h}} E(\mathbf{h}) \quad (10)$$

Usually, less than thirty iterations are sufficient for the convergence of the LM algorithm to the optimal solution. By now, the proposed algorithm (named AIRRL hereafter), which includes ASIFT, IM-RANSAC, the S criterion Region selection and LM optimization, is complete.

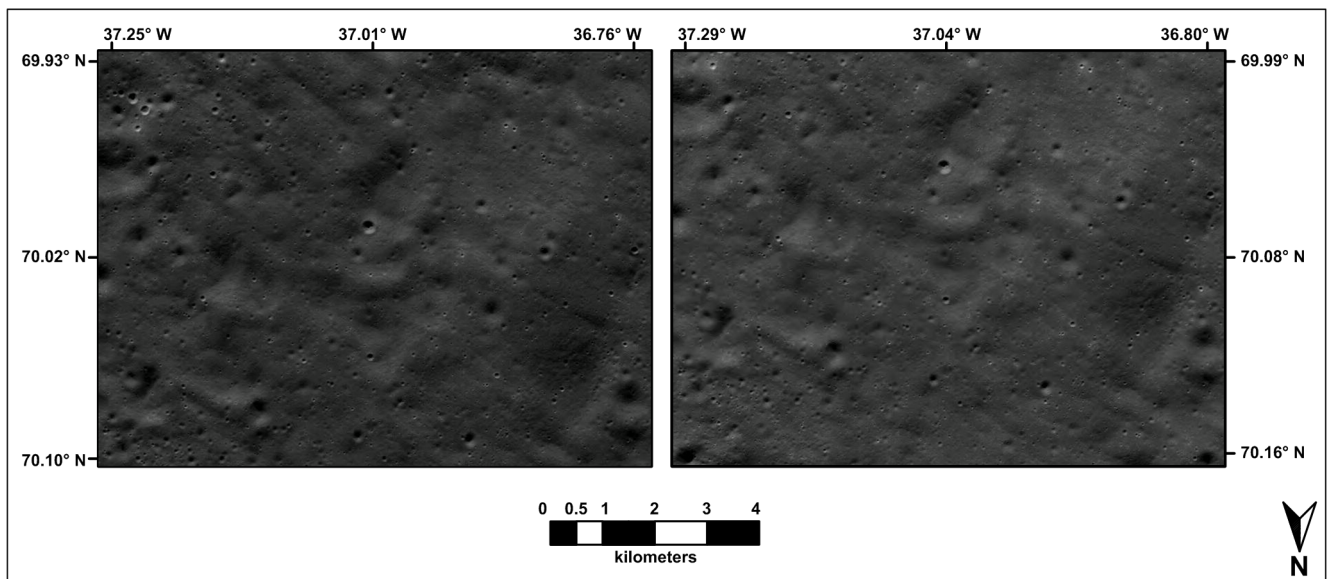
### 3. Experimental Result and Discussion

#### 3.1. Source Data Used in the Experiment

The experimental source data are the stereo image pairs captured by the linear push-broom imaging system of the Chang'E-2 satellite. It is a patch from No. 0581 orbit images with a spatial resolution of seven meters. The Chang'E-2 is the second lunar exploration satellite of China to acquire a three-dimensional morphological and tectonic model of the lunar surface, launched 1 October 2010. The CCD camera mounted on Chang'E-2 is a two-line push-broom sensor assembled on a focal plane separately, as shown in Figure 6. Two CCD arrays, which acquire forward and backward sight angle simultaneously, share the same optical axis with the same focal length of 144.2 mm. Each CCD line array has 6144 pixels, and the forward and backward sight angle is  $8^\circ$  and  $17.2^\circ$ , respectively; the orbit height is 100 km [24]. The original format of Chang'E-2 CCD images is in PDS (Planetary Data System) 2C-level. The PDS is created by NASA's Science Mission Directorate for planetary missions, and the 2C-level images represent the images after the radiometric calibration, geometric correction and photometric correction; thus, they can be directly used in image registration. In Figure 7, the stereo image pair used in the experiment is shown in the selenographic coordinates system. In order to reduce computational complexity, the resolution of the image pair is fixed to  $800 \times 600$  pixels.



**Figure 6.** Configuration of the Chang'E-2 Satellite.



**Figure 7.** Experimental stereo image pair shown in selenographic coordinates system. The backward view is on the left, and the forward view is on the right.

### 3.2. Discriminative Ability

The conventional RANSAC is applied to SIFT and ASIFT, forming SIFT-RANSAC (SIFT-R) and ASIFT-RANSAC (ASIFT-R), respectively, for performance comparison to the proposed local registration algorithm. Because AIRRL, SIFT-RANSAC and ASIFT-RANSAC are all SIFT-based algorithms, the initial matching results are obtained by comparing the ratio of the Euclidean distance of 128 dimensional feature vectors between the nearest neighbour and second nearest neighbour; thus, varying the threshold of the ratio (distance ratio) will affect the final registration result to a large extent. The smaller the value, the more accurate the matching result will be, but the inliers also will be

lesser. Lowe suggested 0.8 as the distance ratio [10]; thus, we chose a range of 0.6–0.8, with a 0.05 interval, to compare the performance of the three methods. The threshold of conventional RANSAC in SIFT-R and ASIFT-R is set to three, which is the empirical value typically used in remote sensing image registration [25], while the threshold increment of IM-RANSAC is set to 0.1 with respect to the compromise between time and accuracy. Because the transformation model is an eight-parameter homography, the size of  $M$  is set to four. The source codes of SIFT, RANSAC and ASIFT were written by Rob Hans [26] and Guoshen Yu *et al.* [27]. The value of parameters  $NP$ ,  $\beta$  in IM-RANSAC is set to 0.01, 0.01 according to the source code of RANSAC and [22], and the small positive  $\varphi$  in Equation (4) is set to a typical value of 0.05 [21].

In this experiment, we employ the RMSE of the Euclidean distance as the evaluation criteria to investigate the matching accuracy of the correspondences. This is defined as follows [28]:

$$RMSE = \sqrt{\frac{E(\mathbf{h})}{n}} \quad (11)$$

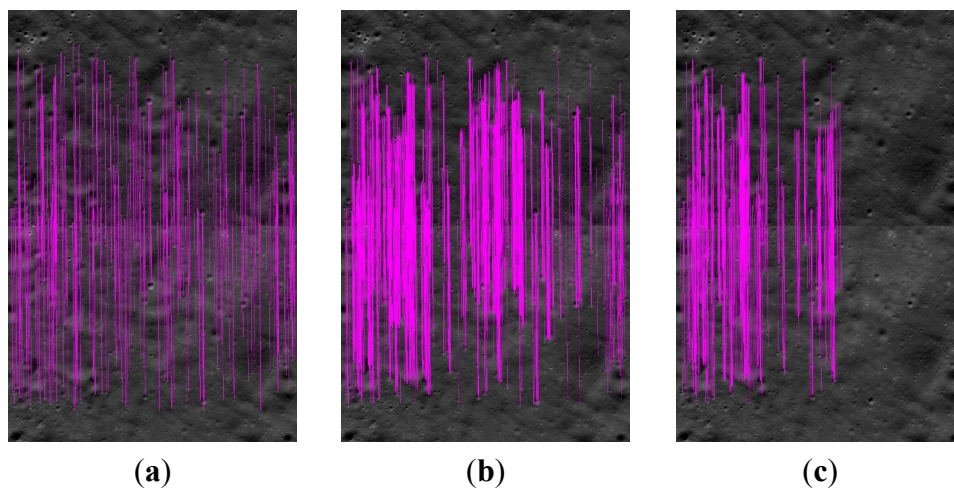
where  $E(\mathbf{h})$  is the total reprojection error in Equation (9). Additionally, the RMSE of intensity ( $I$ -RMSE) here represents the intensity difference at the pixel level; it is used for evaluating the quality of the image mosaic, as well as the quality of registration.  $I$ -RMSE is defined as follows:

$$I\text{-}RMSE = \sqrt{\frac{\sum_i^r \Delta_i^2}{r}} \quad (12)$$

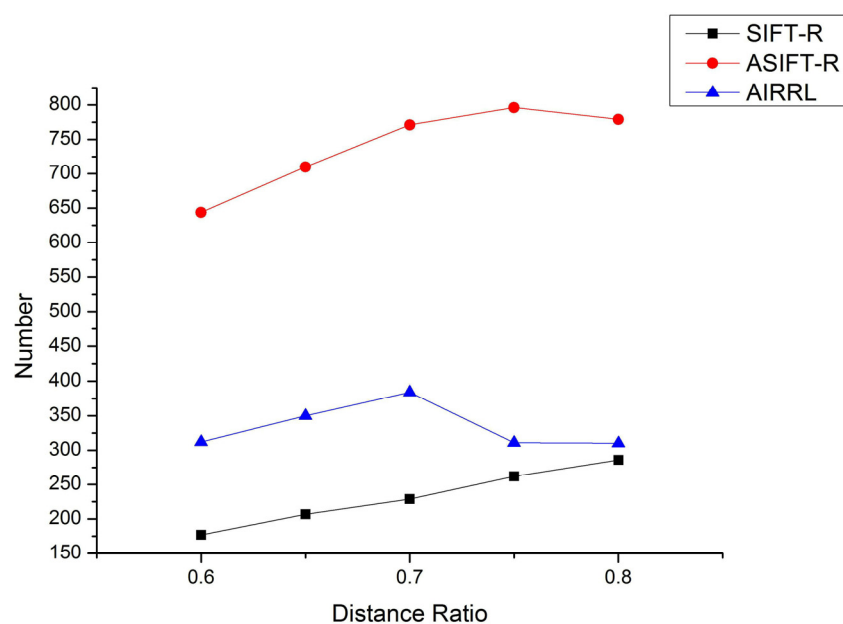
where  $r$  is the number of pixels in the registration region, and  $\Delta$  is the intensity difference between each pixel on the reference image from the local registration region and its corresponding pixel on the registered image.

The correspondences generated by the three methods are shown in Figure 8; from left to right are the matches acquired by SIFT-R, ASIFT-R and AIRRL. The final threshold automatically obtained by IM-RANSAC is 0.6. We can see the difference density and distribution of correspondences between the three methods. ASIFT-R has the densest correspondences, followed by AIRRL and SIFT-R. The distributions of inliers of ASIFT-R and SIFT-R are almost the same; their correspondences are spread all over the image, while AIRRL has no correspondences on the right side of the image. That is because the process of IM-RANSAC and region selection has filtered out many correspondences with relatively large reprojection error or undesired spatial distribution. Figure 8 shows registration results obtained at a distance ratio of 0.7, while this ratio varies between 0.6–0.8, and the number of correspondences generated by three methods is shown in Figure 9. As expected, ASIFT-R generated the greatest number of correspondences, at least two-times more than those of the other two methods. Although AIRRL eliminated many “unqualified” correspondences, it still has more correspondences than SIFT-R does. Figure 10 shows the RMSE value of reprojection error. AIRRL performed best among the three methods, and it is not sensitive to variation of the distance ratio. With the increase of the distance ratio, the RMSE value of ASIFT-R increases and reached its maximum 1.33 at a distance ratio of 0.8; while SIFT-R decreases after reaching a maximum value of 1.07 at a distance ratio of 0.7, but still about two-times larger than the RMSE value of AIRRL. The performances of ASIFT and SIFT are nearly same when the difference of imaging angle between the stereo image pair is small; however,

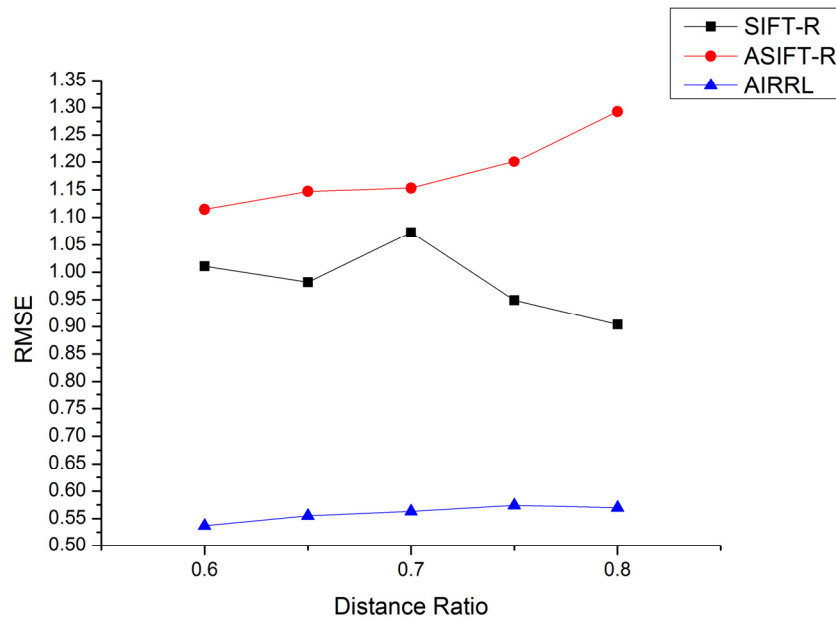
the result may change as the content of the stereo image pair varies; in this experiment, SIFT-R performed better than ASIFT-R with respect to the RMSE value. In Figures 9 and 10, the RMSE of ASIFT-R and SIFT-R are obtained at the empirical RANSAC threshold value of three, while the final thresholds generated by IM-RANSAC at each distance ratio are 0.6, 0.5, 0.6, 0.5 and 0.5, respectively. When the RANSAC threshold of SIFT-R is decreased to 0.8 manually, the correspondence is too little to perform a successful registration. On the other hand, though ASIFT-R can reach the same threshold and RMSE level of AIRRL, it needs manually attempting to gain an appropriate threshold, while AIRRL can automatically perform the process. The experiments were conducted on a laptop featuring a 2.6-GHz Intel core i5-3320 processor and 4 GB RAM. The average time consumption of the SIFT-R, ASIFT-R and AIRRL are 5.93 s, 7.05 s and 8.67 s, respectively.



**Figure 8.** Matching result of different methods. (a) Correspondences obtained by SIFT-R (R, RANSAC). (b) Correspondences obtained by affine SIFT (ASIFT-R). (c) Correspondences obtained by the proposed method.



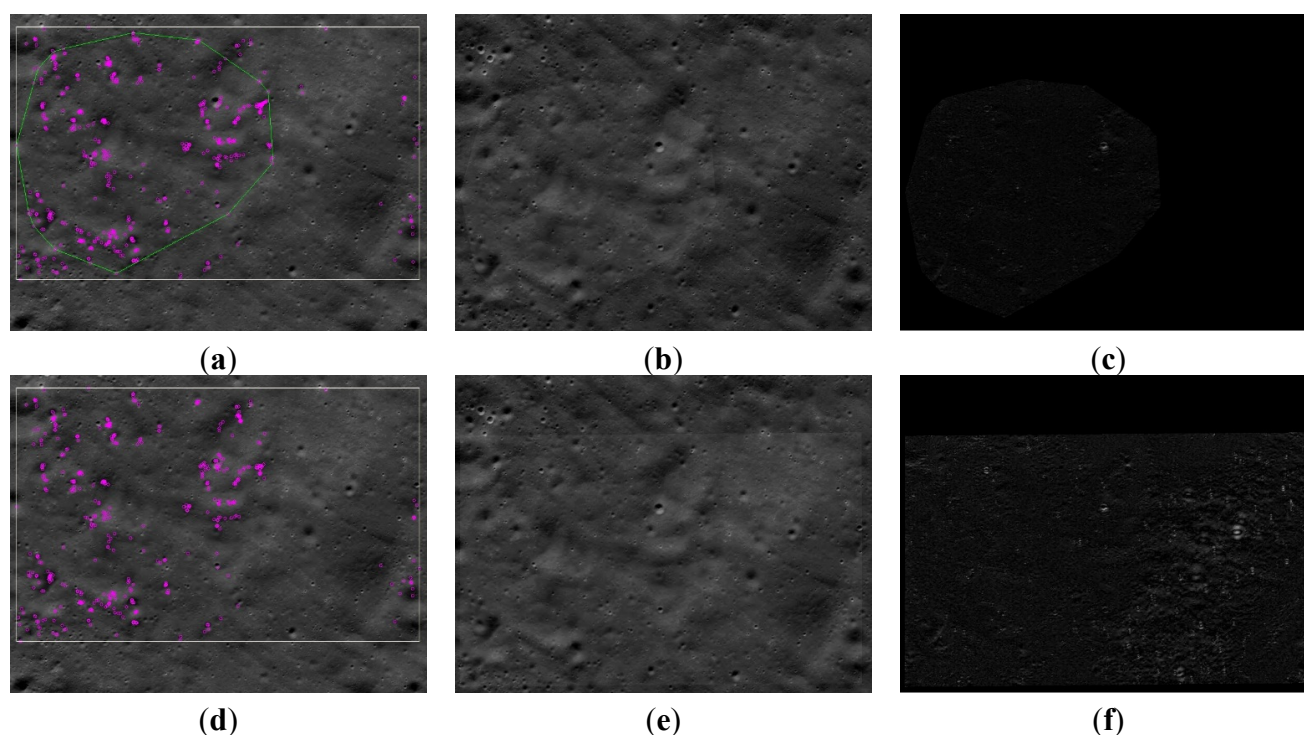
**Figure 9.** Number of correspondences of different methods.



**Figure 10.** RMSE between correspondences of different methods.

### 3.3. Local Registration Region Selection

In this section, the local registration region was determined by applying the S criterion, and then, the image mosaicking was performed to inspect the registration accuracy. According to the experiment results shown in Section 3.2, when the distance ratio is 0.7, AIRRL performed at its best, so we set the ratio to 0.7 in the later experiments. Figure 11a shows the result of registration region selection; tiny pink circles represent the locations of the feature point of inliers. The white rectangle is the area covered by all of the inliers from the output of IM-RANSAC, and the distribution of feature points is scattered in the right part of the region. If including these points in the registration process, the registration accuracy will be lowered. The region outlined by green line segments is the selection region of the S criterion; feature points in this region are obviously denser and evenner. Figure 11b shows the mosaic result of AIRRL; the backward quadratic interpolation is used to perform resampling. Figure 11c shows the intensity difference of image mosaic; the brighter the pixel point is, the larger the mosaic error will be. Figure 11d–f is the registration region, image mosaic and mosaic difference with all of the inliers from IM-RANSAC without the S criterion. In Figure 11f, we notice that the mosaic error of the right portion is much larger than that of the left portion, and the S criterion precisely excluded the right portion (as shown in Figure 11c). Due to that RANSAC-based outlier removal algorithms having a certain randomness, feature points in Figure 11a are different from those in Figure 11d, but the overall trend is constant. We conducted the comparison of the image mosaic between the S criterion and non-S criterion many times and found that the S criterion region selection could acquire lower *I*-RMSE, as shown in Table 1.

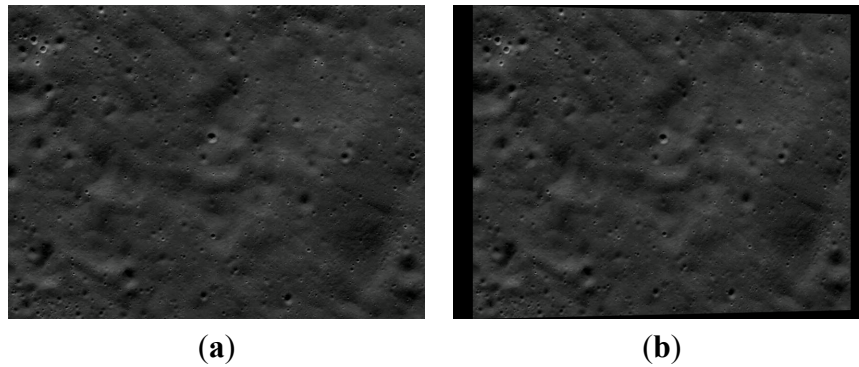


**Figure 11.** Demonstration of region selection and image mosaic. (a) S criterion region selection. (b) S criterion image mosaic. (c) Mosaic error of the S criterion. (d,e,f) The demonstration of the registration region, image mosaic and mosaic error without the S criterion.

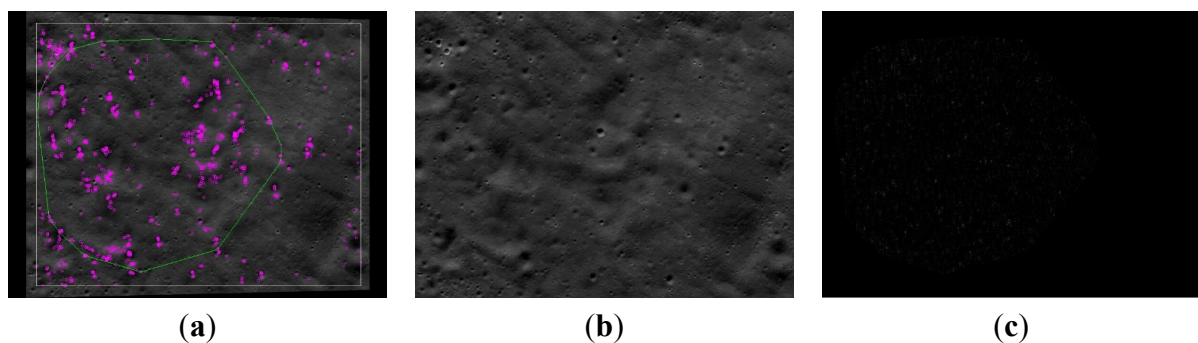
**Table 1.** The average of intensity ( $I$ )-RMSE with and without the S criterion.

Image	$I$ -RMSE
Figure 11c	7.380
Figure 11f	13.925

The result shown in Table 1 is just a relative comparison, and the absolute value of  $I$ -RMSE with the S criterion was high compared to the value of 4.50 in [5]. We notice that the difference of exposure level between the input image pair is too significant to be ignored, as the  $I$ -RMSE is based on the measurement of image intensity; this could be a major reason that affects the  $I$ -RMSE value. In order to avoid this problem, a synthetic image pair was developed for the experiment by choosing the backward view in Figure 7 along with its affine simulation as the new stereo image pair, as shown in Figure 12. The Figure 12b is the affine simulation of Figure 12a; the warp angle is  $25.2^\circ$ , which is the same intersection angle between the forward and the backward sight of Chang'E-2. The experimental result of affine image pair is shown in Figure 13; from left to right is the result of region selection, image mosaic and intensity difference. Comparing Figure 13c to Figure 11c, the exposure difference is corrected. The value of  $I$ -RMSE of the affine simulation with and without the S criterion in the selected region is 2.045 and 5.915, respectively.



**Figure 12.** Affine simulated stereo image pair. (a) Original experiment image. (b) Affine simulated image.



**Figure 13.** Region selection and mosaic result of the affine simulated image pair. (a) Region selection on the simulated image. (b) Image mosaic result. (c) Error of the image mosaic of the simulated image pair.

It should be pointed out that because of the large orbit height (100 km) of Chang'E-2, the captured moon surface images can be treated as from a plane at relatively planar terrains, so we use a homography matrix to estimate the transformation model between the input image pair in this experiment. However, from Figure 11c and Figure 11f, it can be seen that in the image portion with a large intensity gradient, for example in the registration area, which is hilly or involves craters, the mosaic error will be larger. A possible solution is to use a fundamental matrix, utilizing the epipolar constraint, to replace the homography matrix for better stereo correspondence. However, this approach may not be practical in dense matching, as shown in Figure 8, due to the epipolar constraint possibly resulting in one point matching many epipolar lines or *vice versa*, and it is hard to decide the exact match. Therefore, the homography transformation model is used in this research. Consequently, the registration accuracy of the proposed method will be lowered when the spatial resolution of image pair is low or the terrain relief (parallax) in the content of image pair is high.

#### 4. Conclusions

In this paper, we proposed a novel framework with various techniques to identify overlapping regions with high registration accuracy for a stereo image pair. The key contributes of the research can be summarized in four aspects. (1) The IM-RANSAC algorithm optimizes the inliers that are generated by ASIFT and dramatically improves the matching accuracy. In the case where ground

control points (GCPs) are not available, these matching points can be used as GCPs for image registration. (2) In the process of IM-RANSAC, there is no need to pre-define the threshold. The optimal threshold can be achieved automatically by an iteratively-incremental algorithm. It significantly simplifies the registration process. (3) The S criterion is developed to automatically obtain a better distribution region of correspondences between the image pairs according to the positions of the feature points. (4) The registration regions selected by the local image registration framework presented in this paper can be used as a seed portion for global registration. Future research will explore potential algorithms that can take the local registration region to grow into large-scale registered areas and also to narrow the reprojection error of the large topographic gradient portion to further improve the local registration accuracy.

### Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 60872097 and No. 61179043), the Tianjin Science & Technology Pillar Program (Grant No. 08ZCKFJC27900) and the National High Technology Research and Development Program 863 (Grant No. 2010AA122200).

### Author Contributions

All of the authors have made unique and important contributions. Xiangjun Wang gave the research plan and methodology. Yang Li performed image processing and data analysis; the manuscript was also written by him. Hong Wei helped to improve the algorithm and to revise the manuscript. Experiment data preparation and validation were accomplished by Feng Liu.

### Conflicts of Interest

The authors declare no conflict of interest.

### References

1. Zitová, B.; Flusser, J. Image registration methods: A survey. *Image Vis. Comput.* **2003**, *21*, 977–1000.
2. Lemmens, M.J.P.M. A survey on stereo matching techniques. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **1988**, *27*, 11–23.
3. Hui-Zhong, C.; Ning, J.; Jun, W.; Yong-Guang, C.; Luo, C. A novel saliency detection method for lunar remote sensing images. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 24–28.
4. Hussain, M.; Chen, D.M.; Cheng, A.; Wei, H.; Stanley, D. Change detection from remotely sensed images: From pixel-based to object-based approaches. *ISPRS J. Photogramm. Remote Sens.* **2013**, *80*, 91–106.
5. Lucchese, L.; Leorin, S.; Cortelazzo, G.M. Estimation of two-dimensional affine transformations through polar curve matching and its application to image mosaicking and remote-sensing data registration. *IEEE Trans. Image Process.* **2006**, *15*, 3008–3019.

6. Dai, X.; Siamak, K. The effects of image misregistration on the accuracy of remotely sensed change detection. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 1566–1577.
7. Townshend, J.R.G.; Justice, C.O.; Gurney, C.; McManus, J. The impact of misregistration on change detection. *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 1054–1060.
8. Ton, J.; Jain, A.K. Registering landsat images by point matching. *IEEE Trans. Geosci. Remote Sens.* **1989**, *27*, 642–651.
9. Goshtasby, A.A. Feature selection. In *2-D and 3-D Image Registration*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2005; pp. 43–61.
10. Yi, Z.; Zhiguo, C.; Yang, X. Multi-spectral remote image registration based on SIFT. *Electron. Lett.* **2008**, *44*, 107–108.
11. Lowe, D. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110.
12. Lowe, D.G. Object recognition from local scale-invariant features. In Proceedings of the 7th IEEE International Conference on Computer Vision, Kerkyra, Greece, 20–27 September 1999; Volume 2, pp. 1150–1157.
13. Goncalves, H.; Corte-Real, L.; Goncalves, J.A. Automatic image registration through image segmentation and SIFT. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 2589–2600.
14. Gong, M.; Zhao, S.; Jiao, L.; Tian, D.; Wang, S. A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 4328–4338.
15. Cai, G.-R.; Jodoin, P.-M.; Li, S.-Z.; Wu, Y.D.; Su, S.-Z.; Huang, Z.K. Perspective-SIFT: An efficient tool for low-altitude remote sensing image registration. *Signal Process.* **2013**, *93*, 3088–3110.
16. Zhu, Y.; Cheng, S.; Stanković, V.; Stanković, L. Image registration using BP-SIFT. *J. Vis. Commun. Image Represent.* **2013**, *24*, 448–457.
17. Morel, J.-M.; Yu, G. Asift: A new framework for fully affine invariant image comparison. *SIAM J. Imaging Sci.* **2009**, *2*, 438–469.
18. Ting, H.; Hongyan, Z.; Huanfeng, S.; Liangpei, Z. Robust registration by rank minimization for multiangle hyper/multispectral remotely sensed imagery. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2014**, *7*, 2443–2457.
19. Fischler, M.A.; Bolles, R.C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* **1981**, *24*, 381–395.
20. Hartley, R.; Zisserman, A. *Multiple View Geometry in Computer Vision*; Cambridge University Press: Cambridge, UK, 2003; p. 700.
21. Chum, O.; Matas, J. Matching with PROSAC—Progressive sample consensus. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), San Diego, CA, USA, 20–26 June 2005; Volume 1, pp. 220–226.
22. Brown, M.; Lowe, D.G. Recognising panoramas. In Proceedings of the 9th IEEE International Conference on Computer Vision (ICCV), Nice, France, 13–16 October 2003; pp. 1218–1227.
23. Marquardt, D.W. An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Ind. Appl. Math.* **1963**, *11*, 431–441.

24. Zhao, B.; Yang, J.; Wen, D.; Gao, W.; Chang, L.; Song, Z.; Xue, B.; Zhao, W. Overall scheme and on-orbit images of Chang'E-2 lunar satellite CCD stereo camera. *Sci. China Technol. Sci.* **2011**, *54*, 2237–2242.
25. Bin, F.; Chunlei, H.; Chunhong, P.; Qingqun, K. Registration of optical and sar satellite images by exploring the spatial relationship of the improved SIFT. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 657–661.
26. Hans, R. Opensift. Available online: <http://robwhess.github.io/opensift/> (accessed on 12 March 2015).
27. Yu, G.; Morel, J.-M. Asift: An Algorithm for Fully Affine Invariant Comparison. Available online: <http://www.ipol.im/pub/art/2011/my-asift/> (accessed on 12 March 2015).
28. Mikolajczyk, K.; Schmid, C. A performance evaluation of local descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 1615–1630.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).