

Article

An Image Matching Algorithm Integrating Global SRTM and Image Segmentation for Multi-Source Satellite Imagery

Xiao Ling ¹, Yongjun Zhang ^{1,*}, Jinxin Xiong ², Xu Huang ¹ and Zhipeng Chen ¹

¹ School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China; zyl_whu@163.com (X.L.); huangxu@whu.edu.cn (X.H.); chenpz@whu.edu.cn (Z.P.C.)

² 2012 Laboratory of HUAWEI Technology Co., Ltd., Shenzhen 518000, China; xalson2014@gmail.com

* Correspondence: zhangyj@whu.edu.cn; Tel.: +86-27-6877-1101

Academic Editors: Naser El-Sheimy, Zahra Lari, Adel Moussa, Xiaofeng Li and Prasad S. Thenkabail

Received: 7 May 2016; Accepted: 16 August 2016; Published: 19 August 2016

Abstract: This paper presents a novel image matching method for multi-source satellite images, which integrates global Shuttle Radar Topography Mission (SRTM) data and image segmentation to achieve robust and numerous correspondences. This method first generates the epipolar lines as a geometric constraint assisted by global SRTM data, after which the seed points are selected and matched. To produce more reliable matching results, a region segmentation-based matching propagation is proposed in this paper, whereby the region segmentations are extracted by image segmentation and are considered to be a spatial constraint. Moreover, a similarity measure integrating Distance, Angle and Normalized Cross-Correlation (DANCC), which considers geometric similarity and radiometric similarity, is introduced to find the optimal correspondences. Experiments using typical satellite images acquired from Resources Satellite-3 (ZY-3), Mapping Satellite-1, SPOT-5 and Google Earth demonstrated that the proposed method is able to produce reliable and accurate matching results.

Keywords: image matching; multi-source satellite imagery; epipolar line constraint; SRTM; image segmentation; matching propagation

1. Introduction

Image matching is probably the most important function in digital photogrammetry, as well as in automated modeling and mapping [1], which is used for finding corresponding pixels in a pair of images [2,3] and even in multiple images. The existing image matching techniques were developed over the past 50 years, and many empirical algorithms have been created. However, the performance of these matching algorithms is still not satisfactory and, in some cases, far from acceptable [1].

More and more satellites in orbit have become available around the world, and large datasets of multi-source satellite images need combined processing for various purposes. However, the traditional matching algorithms cannot meet this requirement effectively until the following several problems are resolved:

- Since multi-source satellite images are acquired from different sensors and perspectives, there are large geometric distortions between them (e.g., resolution difference, rotation, projection disparity, etc.). These distortions may decrease the correlation between correspondences and even may make the current matching algorithms invalid.
- To achieve numerous correspondences, the geometry constraint of triangulations is employed by many matching algorithms [3–5]. The key to this kind of constraint is that each triangle is

assumed to be a locally planar area. However, for areas with large topographic relief (e.g., urban areas and mountain areas), the assumption fails, which seriously affects the matching reliability.

- Image matching is a difficult and challenging problem when encountered with repetitive textural or untextured regions and occlusions [6]. A large number of human interactions are required to remove the errors in the matching results when dealing with poor textural images [7].

Figure 1 shows three typical image pairs acquired from different sensors. All except for the Google Earth image are Level 1B, the radiometric and internal geometric distortions of which have been corrected, and the relation between the images pixels coordinates and object space coordinates are formulated by the Rational Function Model (RFM) [8]. The expected location error of ZY-3 is 60 m; Mapping Satellite-1 is 80 m; SPOT-5 is less than 50 m; and the Google Earth is less than 20 m. In Figure 1a,b, the left image was acquired from the nadir CCD sensor of the ZY-3 satellite in 2012, and the right image was acquired from the forward CCD sensor of the Mapping Satellite-1 in 2010. The spatial resolution of the left image is 2.1 m, and the spatial resolution of the right image is 5 m; the resolution difference is more than two times. Moreover, there is a stereo angle of 25° between them; the radiometric changes caused by different imaging times, which pose a challenge to image matching, are obvious. Figure 1c,d was imaged in Guangdong, China. The spatial resolution of the left image is 3.5 m, and the spatial resolution of the right image is 10 m; the resolution difference is close to three times. Moreover, there is a stereo angle of 22° between them; this pair contains extensive mountainous areas, whose elevation range is 52~718 m. Figure 1e,f shows an image pair. The left image was downloaded from Google Earth [9], and the right one was acquired from the nadir CCD sensor of Mapping Satellite-1. The spatial resolution of the left image is 1 m, and the spatial resolution of the right image is 5 m; the resolution difference is five times, and significant perspective distortions exist, which means that the traditional matching algorithms may fail.

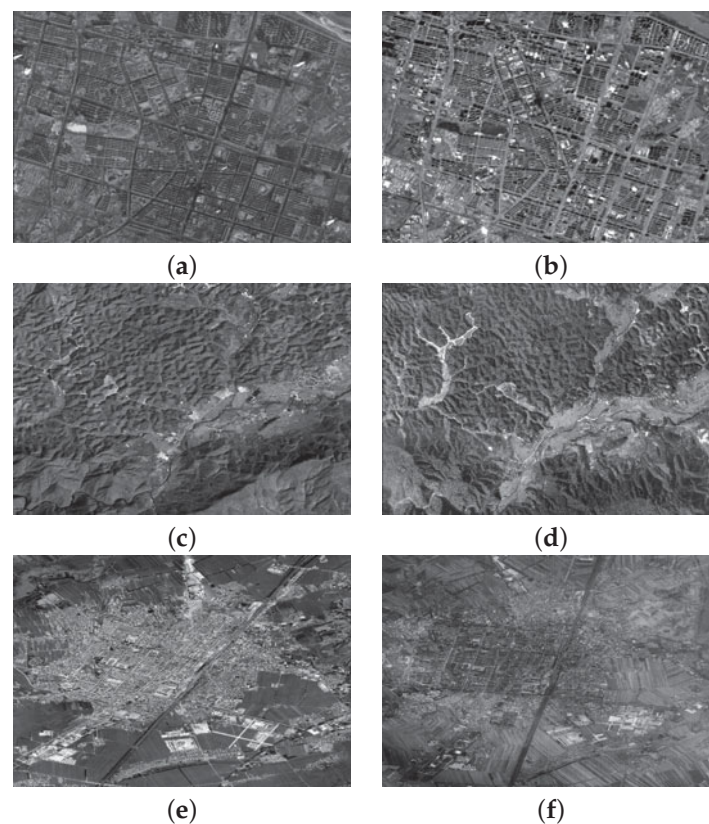


Figure 1. Examples of multi-source satellite images for image matching: (a) nadir-viewing image of the ZY-3 satellite; (b) forward-viewing image of Mapping Satellite-1; (c) backward-viewing image of the ZY-3 satellite; (d) panchromatic image of the SPOT-5 satellite; (e) Google Earth image under the geographic coordinate system; (f) nadir-viewing image of Mapping Satellite-1.

To solve the problems mentioned above, we propose an image matching algorithm in this paper that integrates global SRTM data and image segmentation. After presenting a literature review of the matching methods for satellite imagery in Section 1.1, the algorithmic outline is described in Section 1.2. Section 2 is dedicated to selecting and matching the seed points. A region segmentation-based matching propagation is presented in detail in Section 3. In Section 4, the pairs of satellite images illustrated in Section 1 are used for experimental analysis and quantitative evaluation; and our concluding remarks are presented in Section 5.

1.1. Related Studies

Over the last ten years, many important developments in image matching have been published. Some of these methods have been employed to process satellite imagery, and reliable and remarkable matching performances have been achieved. Zhang [8] developed the Geometrically-Constrained Cross-Correlation (GC³) algorithm to provide a dense pattern of mass points for Digital Surface Model (DSM) generation. This algorithm is based on the concept of multi-image matching guided from the object space and allows 3D reconstruction through matching all of the images at the same time. In this algorithm, a correlation function, Sum of Normalized Cross-Correlation (SNCC), was produced to replace Normalized Cross-Correlation (NCC) [10,11], which reduced the ambiguity caused by occlusions and surface discontinuities. This algorithm mainly depends on the accuracy of the epipolar lines, but an accurate DSM is difficult to obtain. In this case, the mismatch probability may increase substantially.

Scale-Invariant Feature Transform (SIFT) is a well-known detector in the computer vision field. Because the features extracted are highly distinctive and invariant to image scale and rotation, SIFT can provide robust matching across a substantial range of affine distortion, changes in 3D viewpoint, the addition of noise and changes in illumination [12]. Therefore, this detector is employed in photogrammetry [13]. Considering that SIFT can only detect blob-like feature points and produce relatively sparse matching results [14], many researchers have used it to provide reliable seed points. Long [15] and Silveira [16] proposed a hybrid matching method that combines a feature-based algorithm and an area-based algorithm. SIFT first is employed to provide a set of seed points; least square matching with region growing then is used in a subsequent step; and a dense and well-distributed cloud of corresponding points on a pair of satellite images is finally obtained. Although this procedure does not need human intervention, it has several problems: (a) satellite images are generally very large, which makes it quite time-consuming for the SIFT detector; (b) when this method is used for repetitive textural or homogeneous textural images, the region growing may not obtain dense and robust corresponding points; and (c) since the positioning and attitude parameters of satellite images are known in advance, which can simplify image matching and avoid mismatches, this auxiliary data should be used in the matching method.

Wu [4] and Zhu [5] presented a triangulation-based hierarchical image matching method. This method uses a few seed points to generate the initial Delaunay triangulations, and interest points are matched under the triangle constraint and the epipolar constraint. Once a pair of corresponding points is obtained, the triangulations are updated and extended dynamically. Based on this method, Wu et al. [3] developed a reliable and dense image matching on poor textural images, which takes advantage of the edge-constrained Delaunay triangulations and incorporates edge matching with point matching. Experimental analysis indicated that this method preserves the actual textural features and improves the image matching reliability, especially for poor textural images. However, the matching propagation strategy mainly depends on the hypothesis that each triangle is considered to be a locally planar area, so the sparse seed points therefore cannot ensure the smoothness within the initial Delaunay triangulations.

In addition to the above-mentioned algorithms, there are some recent and novel image matching algorithms. Xiong [17,18] developed a robust interest point matching algorithm that first detects the super points with the greatest interest strength. Then, a control network is constructed, and

each interest point is assigned a unique position and angle as its descriptor. Through a continuous process, the network is extended, and all interest points are matched. IAlruzouq and Habib [19] used Modified Iterated Hough Transform (MIHT) to estimate the corresponding linear features between stereo pairs. MIHT follows an optimal sequence, which takes into account the contribution of linear features with different orientations at various locations. Colerhodes et al. [20] introduced a registration algorithm that combined a simple yet powerful search strategy based on a stochastic gradient with two similarity measures, correlation and mutual information. Eugenio and Marques [21] developed two contour-matching techniques for satellite imagery, based on a general affine transformation, which modeled directly the corrections in the image domain without an explicit identification of the distortion sources. Chen and Shao [22] proposed a novel line-based matching method for high resolution remote sensing images. In this method, all of the extracted line segments are grouped into salient lines and general lines, and the hierarchical strategy employed in the method avoids a mass of iteration; a parameter adaptive Difference of Gaussians (DOG) method is used to detect keypoints. Besides, point matches on both sides of line segments are used to constrain the lines matching. Semi-Global Matching (SGM) [23,24] successfully combines the concepts of global and local stereo methods for accurate, pixel-wise matching at low runtime; the core algorithm considers pairs of images with known intrinsic and extrinsic orientation, and the method has been implemented for rectified and unrectified images [25]. Humenberger et al. [26] introduced a stereo matching approach consisting of a combination of census-based correlation, SGM disparity optimization, as well as segmentation-based plane fitting for enhancements on textureless and occluded areas.

1.2. The Proposed Approach

The algorithm proposed in this paper mainly focuses on those images acquired from different satellite sensors. Moreover, it tries to solve the matching ambiguity caused by repetitive and homogeneous textures. First, interest points are detected using the Harris detector [27]; the seed points are determined by the Local Contrast (LC) operator [28]; and an image pyramid (usually three levels) is generated. The global SRTM data are thereafter adopted, upon which an iterative method is used to achieve the true elevation range in the neighborhood of each seed point. Under the epipolar lines constraint, an SRTM-assisted image matching method then is developed to find the corresponding points of the seed points. In this process, the local geometric and radiometric distortion rectification is considered in the hierarchical matching strategy, and the epipolar lines are updated level by level in the pyramid. Finally, a region segmentation-based matching propagation method is presented, and numerous reliable matching results are obtained. The framework of this method is illustrated in Figure 2. This paper highlights the following two innovations:

- Based on the analysis of problems in the existing matching methods, considering the characteristics of multi-source satellite imagery, this paper develops a seed point matching method assisted by global SRTM. By the following research, the seed point selection, epipolar line generation, matching strategy optimization and mismatch detection, we can access reliable and high-precision corresponding points. Then, the corresponding points act as the seeds, which can provide reliable prior knowledge for the subsequent matching propagation.
- In order to obtain dense matching results, we propose a matching propagation method based on region segmentation. In this method, image segmentation technology is adopted innovatively in image matching algorithm in photogrammetry. Using the segmented regions as regional constraints, combined with the radiometric and geometric similarity, we constructed Distance, Angle and Normalized Cross-Correlation (DANCC), aiming at enhancing the validity of the texture-poor regions and repeat regions and to improve the accuracy rate and robustness of matching propagation.

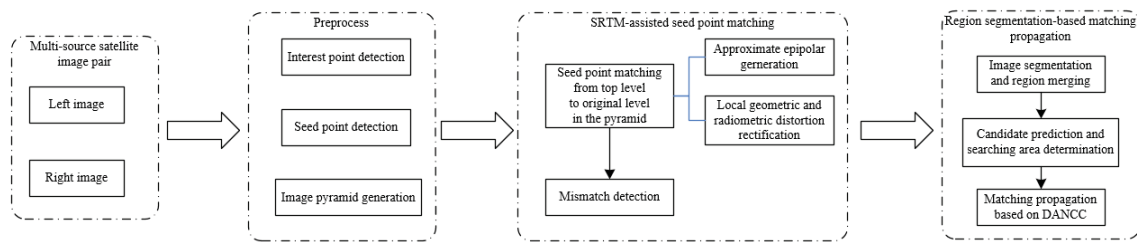


Figure 2. The framework of the proposed matching algorithm. DANCC, Distance, Angle and Normalized Cross-Correlation.

2. Global SRTM-Assisted Seed Point Matching

2.1. Seed Point Selection

If the neighborhood of a pixel includes high contrast textures, this pixel will be distinctly identifiable by the human eye. This case also may be suitable for gray-level matching. Therefore, the LC operator is applied in this paper. A powerful tool for measuring the contrast of local image texture, the LC operator is, by definition, invariant against shifts and rotation in grayscale [28]. In this paper, the local contrast values of all of the pixels are calculated, from which the distribution map is generated. Based on this map, the interest points with relatively high contrast are selected as the seed points. Figure 3b,d shows the distribution map; the brightness of a pixel increases as its local contrast value increases.

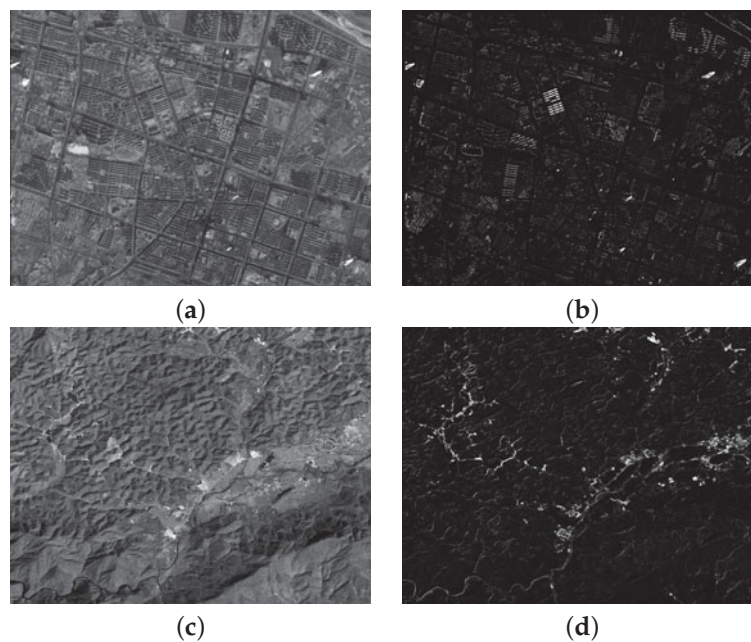


Figure 3. Examples of the distribution maps generated by the Local Contrast (LC) operator: (a) the satellite image of Test 1; (b) the distribution map of Test 1; (c) the satellite image of Test 2; (d) the distribution map of Test 2.

2.2. Epipolar Line Generation

There is no rigorous epipolar line for linear push broom satellite images (e.g., ZY-3, Mapping Satellite-1, SPOT-5, and so on) [29,30]. Only an approximate epipolar line can be generated. In this paper, the global SRTM data are employed, and an iterative method is used to provide a more accurate elevation. Through this method, the prediction error of the epipolar line decreases effectively, and a more reliable constraint serves for the subsequent image matching.

The principle of the iterative elevation determination is shown in Figure 4. A seed point is forward-projected onto the ground at an estimated height Z_0 , which in the paper is the height offset value of Rational Polynomial Coefficients (RPCs), and the initial plane coordinate (x_0, y_0) in the object space is obtained. According to this plane coordinate, the height Z_i can be acquired from the global SRTM data. Then, this point is forward-projected onto the ground at height Z_i again. According to the updated plane coordinate, height Z_{i+1} is acquired. It is an iterative process, step by step, until discrepancy $dz = Z_{i+1} - Z_i (i = 0, 1, 2, \dots)$ is lower than a given threshold, after which the accurate height Z is recorded and an epipolar line is established based on the projection track method [31–33]. Then, a series of correlation windows are established along the epipolar line, as shown in Figure 5.

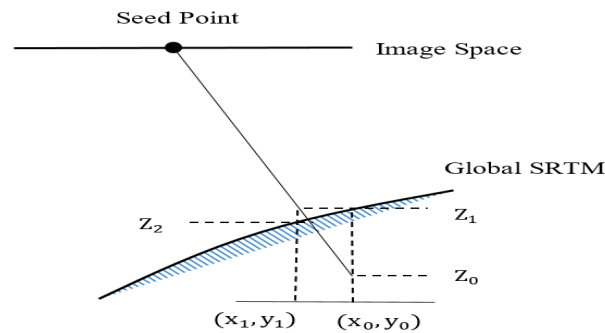


Figure 4. The principle of the iterative elevation determination.

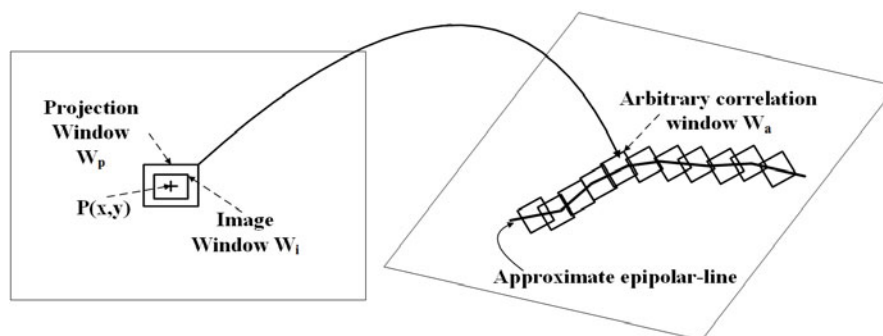


Figure 5. The principle of local geometric and radiometric distortion rectification.

2.3. Local Geometric and Radiometric Distortion Rectification

Local distortions may lead to low correlation if using the NCC method [10,11] and may even form an irregular or discontinuous correlation window. Therefore, as shown in Figure 5, for each seed point, each window is rectified along the approximate epipolar line to projection window; then, find the maximum correlation coefficient by using NCC with image window. More details are displayed below:

1. For a height Z_i , which has been mentioned in Section 2.2, forward-project and backward-project (Equation (1)) the projection window W_p , which includes this search area in the right image.

$$(x_c, y_c) = F_{RPCR}(F'_{RPCL}(x, y, Z_i)) \tag{1}$$

where $F'_{RPCL}(\cdot)$ denotes the RFM of the left image, $F_{RPCR}(\cdot)$ denotes the RFM of the right image and (x, y) here are image coordinates of an arbitrary node of projection window W_p .

2. Use Equation (2) to calculate the affine transformation and linear radiometric transformation parameters between projection window W_p and the correlation window W_a .

$$\begin{aligned}
 xc &= a_0 + a_1x + a_2y \\
 yc &= b_0 + b_1x + b_2y \\
 g'(xc, yc) &= h_0 + h_1g(x, y)
 \end{aligned}
 \tag{2}$$

where (x, y) and (xc, yc) are obtained from the step above, $a_0, a_1, a_2, b_0, b_1, b_2$ are variables of affine transformation, $g(\cdot)$ and $g'(\cdot)$ represent the left and right image gray value, respectively, and h_0, h_1 are the variables of linear radiometric transformation.

3. Apply the transformation parameters $a_0, a_1, a_2, b_0, b_1, b_2, h_0, h_1$ to resample the correlation window to the projection window.
4. For each pixel P in the projection window, compute the correlation coefficient between P and the seed point $P(x, y)$. Record the local maximum correlation coefficient and use affine transformation parameters to transform its coordinates back to the right image.
5. Repeat all steps above over $Z_i (i = 1, 2, 3, \dots, n)$; find the maximum correlation coefficient, and report it as the correspondence of the seed point.

2.4. Matching Strategy

Zhang [8] employed the GC^3 algorithm to process high-resolution satellite images and to provide robust matching points. This paper develops this algorithm and uses the global SRTM data and an iterative method to generate more accurate epipolar lines. When correspondences are obtained from one level in the image pyramid, the prediction error of the epipolar line in the image space is updated to the next lower level, which may increase the matching accuracy and stability in the hierarchical strategy.

Figure 6 shows the comparison of the prediction accuracy of the epipolar lines generated by GC^3 and the developed method. Figure 1a,b was selected as the test data. Eighty-six correspondences were measured manually as checkpoints; and the point marked using red crosses in Figure 6a is one of them. In Figure 6b, the red line is an epipolar line generated by the developed method on the top level, and the blue line is an epipolar line generated by GC^3 . Table 1 displays the prediction errors of the epipolar lines for all of the checkpoints. The Root Mean Square Error (RMSE) of the developed method is 22.8 pixels, and the RMSE of GC^3 is 29.3, which indicates that the prediction accuracy of the epipolar line is improved.

Table 1. Comparison of the prediction accuracy of epipolar lines. GC^3 , Geometrically-Constrained Cross-Correlation.

Method	Number of Checkpoints	Max Difference		RMSE (Pixel)
		Horizontal Direction (Pixel)	Vertical Direction (Pixel)	
Developed method	86	14.3	38.4	22.8
GC^3	86	16.5	47.2	29.3



Figure 6. Epipolar lines generated by GC^3 and the developed method: (a) a checkpoint on the test image; (b) zoomed view of epipolar lines; the red one is generated by the developed method, and the blue one is generated by GC^3 .

Figure 7 shows the matching result of the developed method. As seen in Figure 7, the correspondences are marked by red rectangles. Although large differences between the two images exist, the developed method still achieves visually accurate, albeit sparse, correspondences. Considering the potential mismatches, a subsequent process is provided.

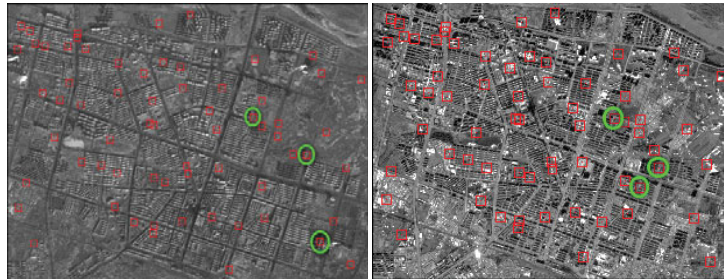


Figure 7. The matching result produced from the developed method, the correspondences are marked by red rectangles, and the outliers are marked by green circles.

In Figure 7, all of the correspondences are taken as two point sets, and the RANSAC algorithm is adopted to exclude the outliers [2,34]. Through the relationship between the two point sets, a model is generated to determine how many of the remaining correspondences fit this model, which is used as a criterion to determine the model that has the largest number of inliers. In this paper, a quadratic polynomial is used as the fitting model. The outliers are marked by green circles in Figure 7.

3. Region Segmentation-Based Matching Propagation

To achieve dense matching results, a matching propagation based on region segmentation is proposed. Through the matching propagation, the number of correspondences grows, and numerous matching results are produced. The details of this method are described in the following section.

3.1. Image Segmentation

To achieve reliable segmentation, the marker-based watershed algorithm is utilized to produce the initial partition, followed by a representation of the regions using a Region Adjacency Graph (RAG) [35]. Hierarchical region merging based on RAG then is applied to obtain the final segmentation result [36]. In Figure 8a, the initial segmentation result is shown. The over-segmented regions are extracted in order to obtain the fine details of the images, and the hierarchical merging process then is implemented. As shown in Figure 8b, the segmentation quality is obviously improved.

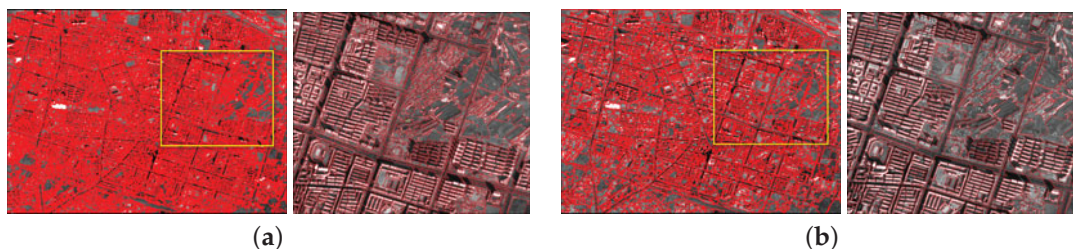


Figure 8. The results of image segmentation and region merging process: (a) the initial segmentation; (b) the segmentation after the merging process, the false segment is removed, and the segmentation quality is obviously improved.

3.2. Candidate Prediction

In this paper, each region extracted is considered to be a locally planar area. When a seed point lies within a region, several steps will be applied:

1. Apply Equation (1) to compute forecast coordinates (x_c, y_c) of seed point (x, y) on the right image with height value Z mentioned in Section 2.2.
2. Compute translation parameters (dx, dy) from:

$$\begin{aligned} dx &= x' - x_c \\ dy &= y' - y_c \end{aligned} \tag{3}$$

where (x', y') are the right image coordinates of the correspondence of seed point (x, y) .

3. All interest points (x_i, y_i) within the same segment will be predicted on the right image from:

$$(x'_i, y'_i) = F_{RPCR}(F'_{RPCL}(x_i, y_i, Z)) + (dx, dy) \tag{4}$$

If (x'_i, y'_i) is in the same segment with the correspondence of the seed point, it will become the remaining point, as seen in Figure 9.

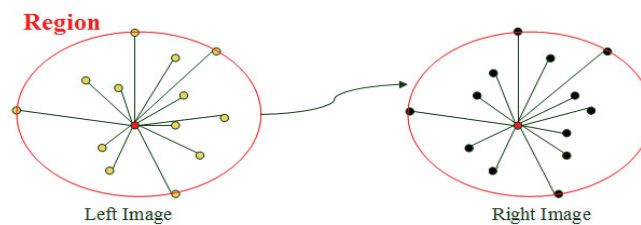


Figure 9. Candidate prediction in a region. The red point is a seed point; the remaining points marked in yellow are predicted on the right image. The points marked in black are the predicted points.

3.3. Searching Area Determination

Although the translation parameters are obtained, there is a tiny deviation between the predicted points and the real correspondences caused by the projection difference. Therefore, searching areas need to be determined to find the optimal correspondences. In Figure 10, the point marked in red is a seed point that lies within a region, which is considered a root node. Two other seed points nearest from the root node are selected as the starting nodes, the precondition for which is that the intersection angle θ must be larger than 0° and less than 180° . For the starting nodes, there is no need for them to lie within the same region with the root node. Once the starting nodes are determined, the local disparity $\Delta\lambda = \frac{\lambda' - \lambda}{\lambda}$ and $\Delta\theta = \theta' - \theta$ can be estimated; here, λ denotes the distance between P_2 and R , λ' denotes the distance between P'_2 and R' , θ denotes $\angle P_1 R P_2$ and θ' denotes $\angle P'_1 R' P'_2$. The candidate T' of remaining point T in the right image is taken as the center of a circle, and the radius r is estimated by Equation (5). The optimal correspondence of remaining point T will be found within the circular searching area.

$$r = \max(|L \cdot \Delta\lambda|, |L' \cdot \tan(\Delta\theta)|) \tag{5}$$

where L denotes the distance between T and R and L' denotes the distance between T' and R' .

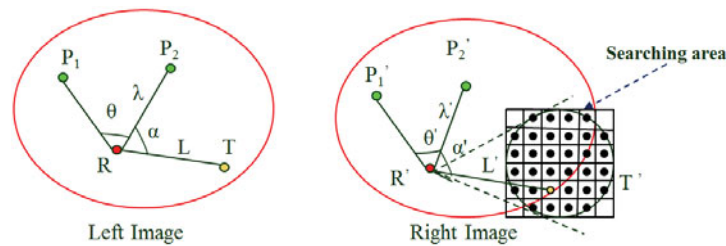


Figure 10. Illustration of the matching propagation. The root node marked in red and two starting nodes marked in green determine the local disparity $\Delta\lambda$ and $\Delta\theta$, then the circular searching area can be estimated.

3.4. Constraints for Matching Propagation

In this method, the remaining points are matched based on the following fundamental constraints.

- Segmentation constraint: As seen in Figure 10, the corresponding seed points R and R' lie within their respective regions in the left and right images, so both regions are assumed to be the corresponding regions (i.e., the correspondence of the remaining point T in left image should be first identified within the corresponding region in the right image). This constraint is helpful for pixels inside the corresponding region in order to raise the priority for being the correspondence.
- Spatial distance constraint: In a locally planar area, the spatial distance between the remaining points and the root node should be relatively consistent. Therefore, a weighted constraint based on spatial distance is employed. As seen in Figure 10, the closer the distance L' between root node R' and a candidate in a searching area is to a threshold, the higher weight the candidate is assigned. The threshold is defined by $L_{Threshold} = L + L \cdot \Delta\lambda$, where $\Delta\lambda$ is estimated as in Section 3.3.
- Spatial angle constraint: As seen in Figure 10, when point T' in the right image is assumed to be the correspondence of remaining point T in the left image, the intersection angle α should be consistent with angle α' in a locally planar area. Therefore, a weighted constraint based on spatial orientation is employed. When the intersection angle between a candidate and the two seed points in the right image is closer to a certain threshold, this candidate is assigned a higher weight, and the probability of correspondence is higher. The threshold is defined by $\alpha_{Threshold} = \alpha + \Delta\theta$, where $\Delta\theta$ is estimated as in Section 3.3.

3.5. An Integrated Similarity Measure (DANCC)

There have been several correlation methods presented for matching satellite imagery. NCC is often used in photogrammetry. NCC is simple for matching accurate correspondences, but it may be ineffective for matching in repeated or poor textural areas. Moreover, Xiong [17,18] presented a correlation method that evaluates the position and angle correlation for each remaining point, whereby the closest candidate becomes the correspondence. This method takes into account the geometric similarity, but radiometric similarity is not considered. If the geometric and radiometric similarities are integrated, the matching performance may be better than traditional measures. In this paper, an integrated similarity measure (DANCC) is presented for matching propagation, the steps for which follow:

1. Each region in the left image is considered to be an independently locally planar area, in which the seed point with the minimum error calculated by mismatch detection is selected as the root node. Another seed point nearest from the root node is selected as the starting node, after which the candidates for all remaining points that lie within the same region are predicted and the searching areas are determined.

- NCC is employed to calculate the correlation values for all of the sample points in the searching areas. Moreover, distance L' and intersection angle α' are assigned to them. Three element vectors therefore are used to describe each sample point, as shown in Figure 11.

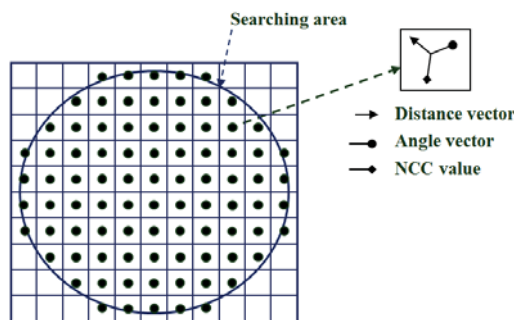


Figure 11. Illustration of the sample point description for DANCC.

- The best candidate for each remaining point is found by identifying its nearest neighbor in the database of sample points. The nearest neighbor is defined with the minimum Euclidean distance, which is described in Equation (6). $NCC_{Threshold}$ is the average correlation value in the searching area. $\alpha_{Threshold}$ and $L_{Threshold}$ are both calculated as in Section 3.4. In addition, a, b are the weight values of the distance and angle vectors ($0 < a < 1, 0 < b < 1, 0 < a + b < 1$), which are estimated in the following section.

$$E(x, y) = \begin{cases} \sqrt{a \cdot (L' - L_{Threshold})^2 + b \cdot (\alpha' - \alpha_{Threshold})^2 + (1 - a - b) \cdot \left(\frac{1}{NCC - NCC_{Threshold}}\right)^2}, & \text{if } NCC > NCC_{Threshold} \\ Invalid, & \text{otherwise.} \end{cases} \quad (6)$$

- For each region containing seed points, the correspondences of the remaining points that belong to this region are obtained by DANCC. In addition, when a remaining point is matched in this region, the new correspondence becomes a candidate for the starting node. If the angle θ assigned to a remaining point is close to zero, the starting node is changed. After the remaining points within this region are matched, they are grouped to expand to the regions that do not contain seed points. These regions select the joint correspondences or closest correspondences as root nodes and starting nodes, and the matching propagation is performed in these regions using Step 1–Step 3.

To compare the performance of DANCC with other similar measures, the image pair illustrated in Figure 1a,b was used as the test data. Figure 12a,b shows a zoomed view of a small part of the images after image segmentation. The correspondence marked with a red cross is a root node, and the correspondence marked with a green cross is a starting node. The point marked with a yellow cross in the left image is a remaining point, and its corresponding point needs to be found in the right image. Figure 12c–e shows the distributions of the similarity values using different methods. When the method presented by Xiong [17,18] was employed, there were several candidates with the minimum value in the searching area (Figure 12c), which caused a matching ambiguity. When NCC was selected, the candidate with the maximum NCC value was a false correspondence (Figure 12d). However, when DANCC was selected, the candidate with the minimum value was distinct and unique (Figure 12e). As displayed in Figure 12b, the correspondence determined by DANCC is accurate.

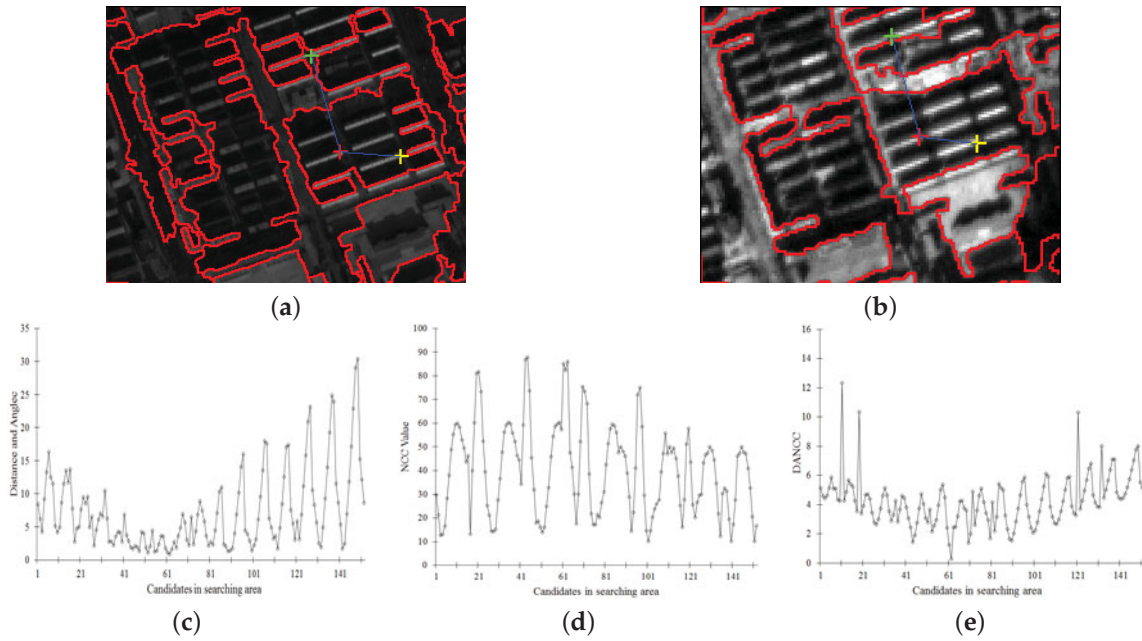


Figure 12. Matching performance comparison with different methods: (a) zoomed view of a small part of the left image; the root node is marked in red; a starting node is marked in green; and the remaining point is marked in yellow; (b) zoomed view of a small part of right image and the correspondence determined by DANCC is marked in yellow; (c) the distribution of similarity values in the searching area using the method presented by Xiong [17,18]; (d) the distribution of similarity values in the searching area using NCC; (e) the distribution of similarity values in the searching area using DANCC.

3.6. Weight Value Estimation

In Equation (6), there are two parameters that need to be estimated: weight a of the distance vector and weight b of the angle vector. As the weight values change, the performance of DANCC will be different.

Figure 13 shows experimental results in which weights a and b were varied. The three graphs were generated using three image pairs (Figure 1), which reflect the influence to the correct matching rate when a and b increase continuously with a fixed interval from 0–1. Considering the sensitivity of both weights, the fixed interval was set to 0.05. Although the three image pairs were acquired from different sensors and different areas, the similar performances are shown. The highest correct matching rate is obtained when weight a is set to the range of 0.15–0.2 and weight b is set to the range of 0.35–0.4. In this paper, we set weight a to 0.15 and set weight b to 0.35.

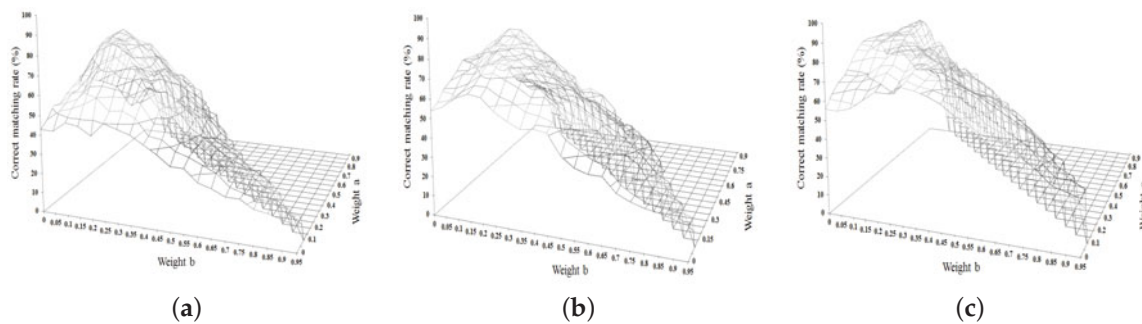


Figure 13. Experimental results using different weight values: (a) the statistics of the correct matching rate using an image pair (Figure 1a,b); (b) the statistics of the correct matching rate using an image pair (Figure 1c,d); (c) the statistics of the correct matching rate using an image pair (Figure 1e,f).

4. Experimental Result and Analysis

The three typical image pairs, as displayed in Figure 1, are used to evaluate the performance of the proposed method. Two other methods are used for comparison: (1) the GC³ method [8]. Although GC³ is a multi-image method, which is commonly used for dense matching, this method is also effective at providing sparse, but reliable corresponding points for block adjustment. Therefore, this method is widely introduced to interest point matching in the case of an image pair or multi-images. In this paper, the GC³ method is used to identify the corresponding points for the interest points extracted. (2) the Triangulation-based Affine-adaptive Cross Correlation (TAACC) method [4]. In this method, SIFT is used to match the key points and the triangulation constraint is developed to help find correct matches.

In order to quantitatively evaluate the performance of the matching methods, in Tests 1 and 2, the interest points are manually measured to achieve their corresponding points by using a commercial photogrammetric software system, Digital Photogrammetric Grid (DPGrid) workstation. The accurate RPCs are achieved after bundle block adjustment, which are used to calculate the 3D coordinates of matching results. Since no ground truth data are available for Tests 1 and 2, the 3D coordinates of the corresponding points manually measured are compared to the 3D coordinates of the matching points produced by the three matching methods, then their statistics of RMSE are obtained. In addition, the differences in image space are computed through comparing the coordinates of the corresponding points manually measured with the coordinates of the matching points; when the difference of an interest point is more than 1.2 pixels, the matching points of this interest point are taken as a mismatch. Therefore, RMSE, successful rate of matching and mismatching rate are used as indicators of the performance of image matching. In Test 3, a certain reference Digital Elevation Model (DEM) in the study area has been produced by DPGrid workstation through using aerial images. After interactive editing, the DEM, whose scale is 1:50,000, is generated as a reference DEM for comparison analysis. After image matching, the RPCs of Mapping Satellite-1 are updated by bundle block adjustment. Then, two elevation values of the point in each match pair can be obtained from the reference DEM. In the Google Earth image, the elevation value can be directly obtained, while that in Mapping Satellite-1 can be calculated by the iterative approach mentioned in Section 2.2. The two values should be very close if the match pair is correct and the point lies on the Earth's surface. For this criterion, some checkpoints on the Earth's surface are selected to evaluate the match result.

4.1. Experimental Results of Test 1

The image pair illustrated in Figure 1a,b is adopted as the test data in Test 1. The proposed method, GC³, and TAACC are used to produce the correspondences between both images. To ensure the comparisons under the same conditions, the parameter settings for the three methods were identical. For instance, the number of interest points is 3000, and the threshold of the cross-correlation coefficient in each pyramid level is the same.

As can be seen in Figure 14a, the matching results produced by TAACC show that a sparse matching distribution is obtained, and there are a number of mismatches. To determine the cause of the mismatches, Figure 15 shows the initial corresponding triangulations generated by the key points. Even though the SIFT algorithm is used to find the key corresponding points and RANSAC is used to exclude the outliers, the matching performance for this test data is not satisfactory, which caused the initial triangulations to be warped. TAACC mainly relies on accurate key points, and in this case, as shown in Figure 15, it provides false triangle constraints, which seriously increase the possibility of matching ineffectiveness.

From Figure 14b, it can be seen that the matching result produced by GC³ is accurate, but not dense. It also shows that integrating the epipolar geometry constraint is suitable for processing the satellite images acquired from different sensors. There are different imaging times, large local distortions, and similar building structures between the image pair, which may have led to low correlation and ambiguity for the correspondences if only NCC is used. As seen in Figure 14c, the

matching points derived from the proposed method are denser than the matching points derived from GC^3 , which reveals that the propagation method DANCC is valid and robust.

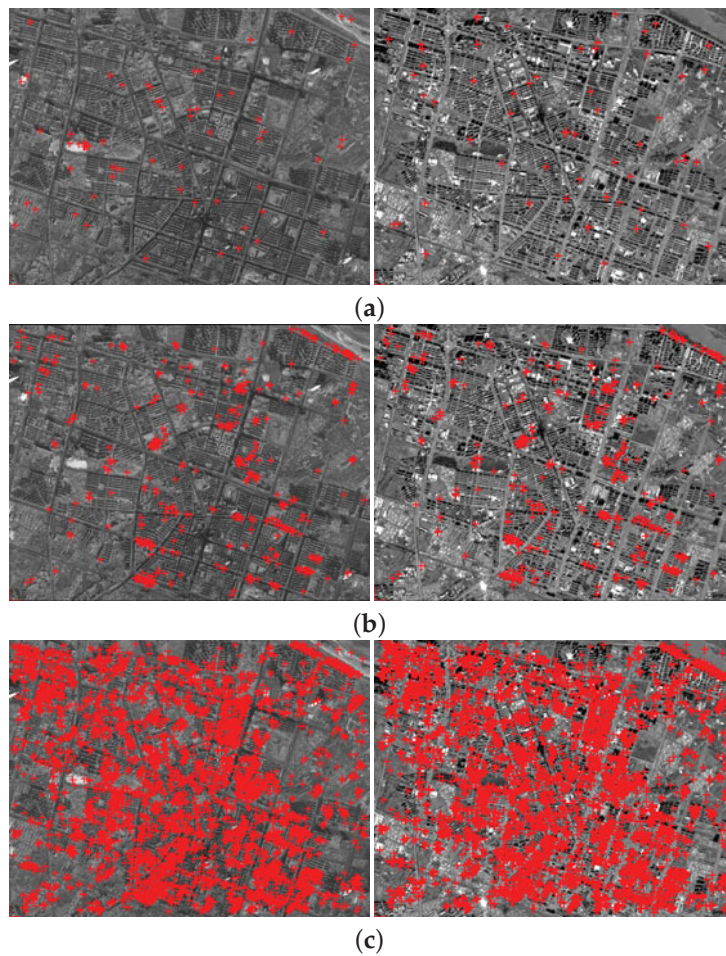


Figure 14. Comparison of the matching results produced by the three methods: (a) the matching result produced by TAACC; (b) the matching result produced by GC^3 ; (c) the matching result produced by the proposed method.

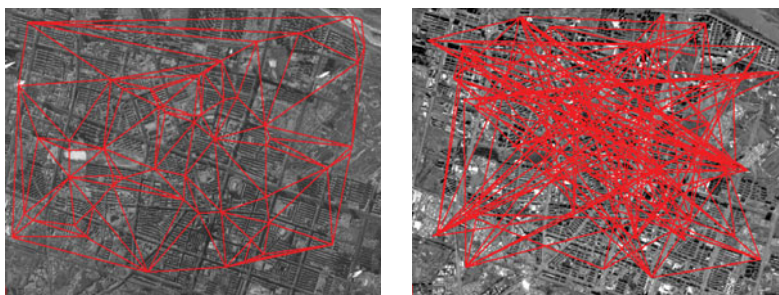


Figure 15. The initial corresponding triangulations produced by TAACC. The false triangle constraints cause mismatches as shown in Figure 14a.

In order to have a detailed comparison analysis, the indicators are listed in Table 2. For a total of 3000 interest points, TAACC produces only 79 correspondences, and the RMSE is more than 102 m. While the result generated by GC³ is improved, the number of correspondences is 691; the mismatching rate is 1.4%; and the RMSE reached 3.4 m. In contrast to the other two methods, the number of correspondences produced by the proposed method increases significantly; the mismatching rate reduces to 0.6%; and the RMSE is improved to 2.4 m. Although the change in the same region caused by different imaging times may affect the matching performance, the statistics still shows that the proposed method produced the best results.

Table 2. Experimental results produced by TAACC, GC³ and the proposed method.

Matching Method	Number of Interest Points	Number of Matching Points	Successful Rate of Matching	Mismatching Rate	RMSE (m)
TAACC	3000	79	2.6%	98.7%	102.3
GC ³	3000	691	23.0%	1.4%	3.4
Proposed method	3000	1946	64.8%	0.6%	2.4

4.2. Experimental Results of Test 2

The stereo pair, as shown in Figure 1c,d, is used in Test 2. The matching challenges for this pair are extensive mountainous areas with large topographic relief and homogeneous textures, which may cause ambiguity during the matching procedure.

In this test, a total of 26,044 interest points are extracted. As shown in Figure 16a, there are 11,672 correspondences matched by the TAACC method. Along with the generation of the accurate key points, the performance of the triangulation constraint works well. For the areas with good image texture, numerous correspondences are produced, but the correspondences for the mountainous areas are sparse. Figure 16b shows that 4985 correspondences are matched by GC³, which is a little worse than the performance of TAACC, indicating that the triangulation constraint may improve matching effectiveness better than the epipolar geometry constraint in local areas with large topographic relief. However, both GC³ and TAACC use NCC as their correlation method, which does not resist the ambiguity well in local homogeneous texture areas. Figure 16c shows the result produced by the proposed method, which produces 19,617 correspondences, as well as dense matching points distributed in the mountainous areas. Table 3 shows the detailed comparison results. The proposed method produces the best results, which demonstrates that this method can alleviate the ambiguity in homogeneous texture areas and achieve accurate matching results in mountainous areas with large topographic relief.

Table 3. Experimental results produced by TAACC, GC³ and the proposed method.

Matching Method	Number of Interest Points	Number of Matching Points	Successful Rate of Matching	Mismatching Rate	RMSE (m)
TAACC	26,044	11,672	44.8%	8.4%	10.3
GC ³	26,044	4985	19.1%	1.8%	6.2
Proposed method	26,044	19,617	75.3%	0.8%	5.0

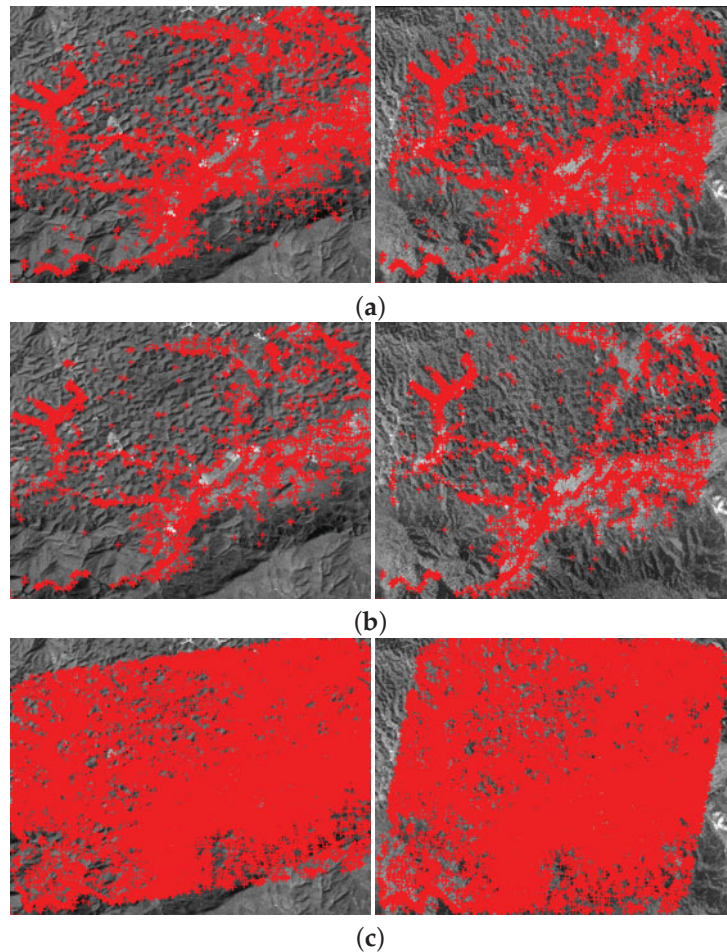


Figure 16. Comparison of the matching results produced by the three methods: (a) the matching result produced by TAACC; (b) the matching result produced by GC³; (c) the matching result produced by the proposed method.

4.3. Experimental Results of Test 3

To expand the application of the proposed method, the image pair shown in Figure 1e,f is used as the test data in Test 3. This pair consists of a Google Earth image downloaded from Google Earth [9] and a satellite image acquired from the nadir CCD sensor of Mapping Satellite-1.

Figure 17 shows that there are 5770 correspondences matched by the proposed method. The zoomed view of a local region, as marked with rectangles, is provided. It can be seen that this image pair is difficult to match, but the proposed method produces enough matches to declare its reliability as satisfactory.

To evaluate the performance of the proposed method, the RPCs of Mapping Satellite-1 are updated by bundle block adjustment, and the corresponding points are selected from the matching result as checkpoints, all of which lie on the Earth's surface. Their elevation values can be acquired from the reference DEM by two ways: one is directly obtained with its geographic coordinates from Google Earth; another is calculated by the iterative approach mentioned in Section 2.2 with its image coordinates in the Mapping Satellite-1 image. The differences were derived as shown in Figure 18. It can be seen that the two elevation profiles are very close, which indirectly indicates the high accuracy of the correspondences.

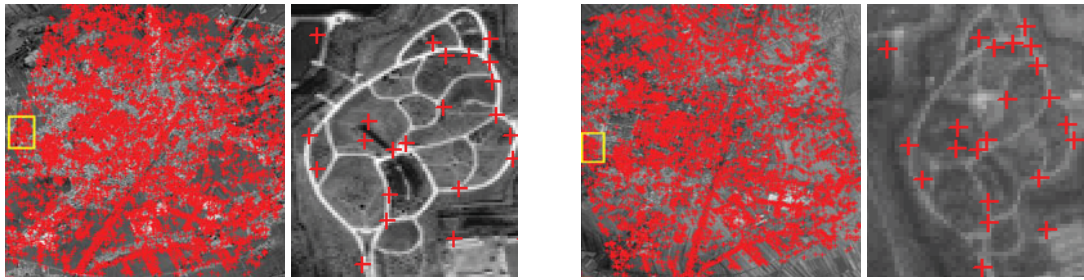


Figure 17. The matching result produced by the proposed method. The zoomed views of a local region, which is marked with a rectangle in each image, are provided to show the matching details.

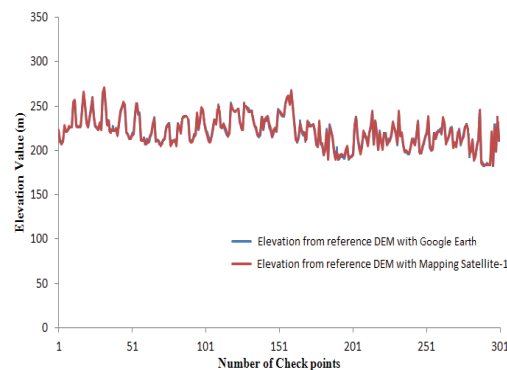


Figure 18. The elevation difference of checkpoints: the blue line is obtained from the reference DEM with their geographic coordinates from Google Earth, and the red one is from the reference DEM by the iterative approach mentioned in Section 2.2 with their image coordinates in the Mapping Satellite-1 image.

5. Conclusions

This paper presented an image matching algorithm that integrates global SRTM data and image segmentation. Using three typical image pairs acquired from different satellite sensors as test data, the experimental analysis resulted in the following conclusions:

1. This matching algorithm is fairly practical and effective for multi-source satellite images, which would be helpful in efforts to combine existing mass satellite data into a digital photogrammetry system.
2. For satellite images with large geometric distortions, this algorithm is capable of obtaining dense and reliable matching results. In addition, its region segmentation-based matching propagation performed quite well for image matching on repetitive or homogeneous textural images, for which achieving robust correspondences is difficult for existing methods.

Future research will focus on a better seed point extraction and a robust image segmentation algorithm. The robust seed points need to be provided without the known positioning and attitude parameters, which can develop the proposed algorithm for application to more image types, such as aerial and close-range images. Further effort will also involve the efficiency of the proposed algorithm and its effectiveness on different extremely difficult matching cases, such as images with dense tall buildings and forests.

Acknowledgments: This research was partially supported by the National Natural Science Foundation of China under Grants 41322010 and 41571434.

Author Contributions: Yongjun Zhang and Jinxin Xiong conceived of the study and designed the experiments. Xu Huang and Zhipeng Chen supervised the research. Xiao Ling wrote the main program and most of the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gruen, A. Development and status of image matching in photogrammetry. *Photogramm. Rec.* **2012**, *27*, 36–57.
2. Hartley, R.; Zisserman, A. *Multiple View Geometry in Computer Vision*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2003; pp. 672–674.
3. Wu, B.; Zhang, Y.S.; Zhu, Q. Integrated point and edge matching on poor textural images constrained by self-adaptive triangulations. *ISPRS J. Photogramm. Remote Sens.* **2012**, *68*, 40–55.
4. Wu, B.; Zhang, Y.S.; Zhu, Q. A Triangulation-based Hierarchical Image Matching Method for Wide-Baseline Images. *Photogramm. Eng. Remote Sens.* **2011**, *77*, 695–708.
5. Zhu, Q.; Wu, B.; Tian, Y.X. Propagation strategies for stereo image matching based on the dynamic triangle constraint. *ISPRS J. Photogramm. Remote Sens.* **2007**, *62*, 295–308.
6. Bleyer, M.; Gelautz, M. A layered stereo matching algorithm using image segmentation and global visibility constraints. *ISPRS J. Photogramm. Remote Sens.* **2005**, *59*, 128–150.
7. Heipke, C.; Oberst, J.; Albertz, J.; Attwenger, M.; Dorninger, P. Evaluating planetary digital terrain models-The HRSC DTM test. *Planet. Space Sci.* **2007**, *55*, 2173–2191.
8. Zhang, L. Automatic Digital Surface Model (DSM) Generation from Linear Array Images. Ph.D. Thesis, Swiss Federal Institute of Technology, Zurich, Switzerland, 2005.
9. Google Satellite Images. Available online: <http://www.google.com/earth/index.html> (accessed on 14 April 2013).
10. Helava, U. Digital correlation in photogrammetric instruments. *Photogrammetria* **1978**, *34*, 19–41.
11. Lhuillier, M.; Quan, L. Match propagation for image-based modeling and rendering. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 1140–1146.
12. Lowe, D.G. Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110.
13. Ai, M.; Hu, Q.W.; Li, J.Y.; Wang, M.; Yuan, H. A Robust Photogrammetric Processing Method of Low-Altitude UAV Images. *Remote Sens.* **2015**, *7*, 2302–2333.
14. Mikolajczyk, K.; Schmid, C. Scale & Affine Invariant Interest Point Detectors. *Int. J. Comput. Vis.* **2004**, *60*, 63–86.
15. Long, T.F.; Jiao, W.L.; He, G.J.; Zhang, Z.M. A Fast and Reliable Matching Method for Automated Georeferencing of Remotely-Sensed Imagery. *Remote Sens.* **2016**, *8*, 56.
16. Silveira, M.; Feitosa, R.; Jacobsen, K.; Brito, J.; Heckel, Y. A Hybrid Method for Stereo Image Matching. in Proceedings of the the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Beijing, China, 3–11 July 2008; pp. 895–901.
17. Xiong, Z. Technical Development for Automatic Aerial Triangulation of High Resolution Satellite Imagery. Ph.D. Thesis, University of New Brunswick, Fredericton, NB, Canada, 2009.
18. Xiong, Z.; Zhang, Y. A Novel Interest-Point-Matching Algorithm for High-Resolution Satellite Images. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 4189–4200.
19. IAlruzouq, R.; Habib, A. Semi-Automatic Registration of Multi-Source Satellite Imagery with Varying Geometric Resolutions. *Photogramm. Eng. Remote Sens.* **2005**, *71*, 325–332.
20. Colerhodes, A.; Johnson, K.; PLemoigne, J.; Zavorin, I. Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient. *IEEE Trans. Image Process.* **2003**, *12*, 1495–1511.
21. Eugenio, F.; Marques, F. Automatic satellite image georeferencing using a contour-matching approach. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 2869–2880.
22. Chen, M.; Shao, Z.F. Robust affine-invariant line matching for high resolution remote sensing images. *Photogramm. Eng. Remote Sens.* **2013**, *79*, 753–760.

23. Hirschmüller, H. Accurate and efficient stereo processing by semi-global matching and mutual information. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA, 20–25 June 2005; Volume 2, pp. 807–814.
24. Hirschmüller, H. Stereo Processing by Semiglobal Matching and Mutual Information. *IEEE Trans. Pattern Anal. Mach. Intell.* **2008**, *30*, 328–341.
25. Hirschmüller, H. Semi-global matching-motivation, developments and applications. In Proceedings of the Photogrammetric Week 11, Stuttgart, Germany, 5–9 September 2011.
26. Humenberger, M.; Engelke, T.; Kubinger, W. A census-based stereo vision algorithm using modified semi-global matching and plane fitting to improve matching quality. In Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), San Francisco, CA, USA, 13–18 June 2010; pp. 77–84.
27. Harris, C.; Stephens, M. A combined corner and edge detector. In Proceedings of the Alvey Vision Conference, Manchester, UK, 31 August–2 September 1988.
28. Ojala, T.; Pietikäinen, M.; Mäenpää, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 971–987.
29. Ji, S.P.; Yuan, X.X. Automatic Matching of High Resolution Satellite Images Based on RFM. *Acta Geod. Cartogr. Sin.* **2010**, *39*, 592–598.
30. Zhang, Y.J.; Wang, B.; Zhang, Z.X.; Duan Y.S.; Zhang, Y.; Sun, M.W.; Ji, S.P. Fully Automatic Generation of GeoInformation Products with Chinese ZY-3 Satellite Imagery. *Photogramm. Rec.* **2014**, *29*, 383–401.
31. Jiang, W.S.; Zhang, J.Q.; Zhang, Z.X. Simulation of Three-line CCD Satellite Images from Given Orthoimage and DEM. *Geomat. Inf. Sci. Wuhan Univ.* **2008**, *33*, 943–946.
32. Kratky, V. Rigorous photogrammetric processing of SPOT images at CCM Canada. *ISPRS J. Photogramm. Remote Sens.* **1989**, *44*, 53–71.
33. Duan, Y.S.; Huang, X.; Xiong J.X.; Zhang, Y.J.; Wang B. A Combined Image Matching Method for Chinese Optical Satellite Imagery. *Int. J. Digit. Earth* **2016**, *9*, 851–872.
34. Zhang, Y.J.; Wan, Y. DEM-assisted RFM Block Adjustment of Pushbroom Nadir Viewing HRS Imagery. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 1025–1034.
35. Haris, K.; Efstratiadis, S.N.; Maglaveras, N.; Katsaggelos, A.K. Hybrid image segmentation using watersheds and fast region merging. *IEEE Trans. Image Process.* **1998**, *7*, 1684–1699.
36. Beaulieu, J.M.; Goldberg, M. Hierarchy in picture segmentation: A stepwise optimization approach. *IEEE Trans. Pattern Anal. Mach. Intell.* **1989**, *11*, 150–163.



© 2016 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 (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).