

COMBINED MULTI-VIEW MATCHING ALGORITHM WITH LONG-STRIPS OF SATELLITE IMAGERY FROM DIFFERENT ORBITS

Jinxin Xiong^{a,*}, Yongjun Zhang^a

^a School of Remote Sensing and Information Engineering; Wuhan University, Wuhan 430049, China –
einbetter1995@hotmail.com

Commission III, WG III/1

KEY WORDS: Matching, Satellite Imagery, Long Strips, Multiple Orbits, Chinese Mapping Satellite-I

ABSTRACT:

Existing matching algorithms aim to match conjugate points among overlapping satellite scenes acquired from the same orbit and can generally achieve good matching performance. Unfortunately, no algorithm can avoid the difficulty of simultaneously processing the data sets of long-strip imagery acquired from different orbits. In this paper, the combined matching algorithm we propose introduces the LBP/C operator, which, when combined with existing feature detectors for the first time, can make possible the extraction of more stable interest points and candidates. At the same time, based on the typical characteristics of Chinese satellite imagery, we improved the filter method and achieved an effective combination of several image matching algorithms. A comparison among several kinds of matching transfer modes was presented; and to evaluate this algorithm, Chinese Mapping Satellite-I data are used as the reference data.

1. INTRODUCTION

Image matching was a key technology related to many problems in computer vision and photogrammetry, including object recognition, motion tracking, digital surface model (DSM) generation, etc. (Lowe, 2004). In the past 20 years, research related to processing satellite imagery has become a key focus. Due to the position and attitude information under fixed sampling intervals, the current matching algorithms are focused on area-based algorithms, especially the approximate epipolar geometric constraints (Heipke, 1996). These existing algorithms can achieve satisfactory matching results in the case of overlapping scenes acquired from the same orbit, but they cannot solve the following problems: 1) how to match with multiple long-strip satellite imagery simultaneously 2) how to eliminate the radiometric and geometric differences between adjacent orbits and achieve high-accuracy results. Moreover, the typical characteristics of imagery acquired from Chinese satellites (e.g., ambiguity in smooth areas, low direct georeferencing accuracy, and obvious differences in contrast) bring more challenges to image matching. Consequently, finding solutions for the above-mentioned problems in photogrammetry was an urgent issue.

In this study, we developed a practical matching algorithm based on the approaches in the literature. First, we divided the reference strip of imagery in each orbit into independently overlapping image patches; then, each patch was filtered to enhance the existing texture patterns and improve the noise-signal ratio (Baltasvias, 1991). After the generation of image pyramids, we combined the Harris-Z detector (Bellavia, 2008) with the LBP/C operator (Timo, 2002) to extract stable interest points. At the same time, the Förstner interest operator (Förstner, 1986) was used to refine these interest points. Then, through the global digital elevation model (Global DEM), the approximate geography scope covered by each image patch was computed. Moreover, the Geometrically Constrained Cross-Correlation (GC³) algorithm (Zhang, 2005 and 2006) was extended to match multiple long-strip imagery; and finally,

several strategies were proposed to detect and eliminate the mismatches. By comparing the results of several matching transfer modes, we determined the optimal manner in which to optimize the matching speed and reduce the number of mismatches.

In the second part, the new matching algorithm is described in detail, and the validity of the proposed matching algorithm is verified with Chinese Mapping Satellite-I data.

2. METHODOLOGY

In this paper, the proposed algorithm includes four parts: 1) Image Pre-Processing; 2) Image Matching Procedure; 3) Mismatch Detection and Elimination.

2.1 Image Pre-Processing

Image pre-processing method includes two parts: 1) Modified Wallis Filter; 2) Interest Points and Candidates Detection and Extraction

2.1.1 Modified Wallis Filter

Compared with aerial imagery and close-range imagery, the characteristics of satellite imagery generally include the following: 1) low spatial resolution, 2) unsymmetrical contrast of light and shade, 3) fuzzy texture structure. These factors seriously affect the reliability of feature extraction and image matching.

To solve these problems mentioned above, the Wallis filter is introduced. It can force the mean and contrast of an image to fit some given values effectively. Baltasvias (1991) used this filter to enhance the image texture pattern and eliminate radiometric difference. The general form of the Wallis filter is given by:

$$\mathbf{g}^w(x, y) = \mathbf{g}(x, y)r_l + r_0 \quad (1)$$

$$r_l = cs_f / (cs_g + s_f / c)$$

$$r_0 = bm_f + (1 - b - r_l)m_g \quad (2)$$

Where

In Equation 1, $\mathbf{g}^w(x, y)$ and $\mathbf{g}(x, y)$ are the filtered and original image respectively, r_o and r_l are the additive and multiplicative parameters, m_g and s_g are the mean and standard deviation of the original image, m_f and s_f are the target values for the mean and standard deviation, C was the contrast expansion constant, and b was the brightness forcing constant.

However, the existing Wallis filter has some problems: 1) the edges of roofs tend to expand, which causes deviations from their geometrical position; and 2) the image noise is partially enhanced, which hurts the positioning accuracy of the features extracted and has a detrimental effect on image matching.

We found that the additive parameter r_o causes the above problems. Owing to its restrictions on the gray level, unreasonable selection of r_o causes the excessive loss of gray information for the following reasons. First, r_o is mainly dependent on parameter b ; secondly, the multiplicative parameter r_l can change image contrast, whose value is dependent on parameter C ; and lastly, the high texture fidelity and enhance image contrast must be kept under the filter's capacity. In this study, we used the following method to improve the Wallis filter.

A highly adaptive parameter determination strategy: In the case of an image $f(x, y)$, we classified all pixels of this image into two sets (C_1, C_2).

$$\begin{aligned} \mathbf{c}_1 : \mathbf{f}(x, y) < T \\ \mathbf{c}_2 : \mathbf{f}(x, y) \geq T \end{aligned} \quad (3)$$

Where T was a classified threshold, \mathbf{c}_1 means the set of low-pass, and \mathbf{c}_2 means the set of high-pass. Then, b and c were calculated by the following equation:

$$\begin{aligned} b &= (255 - T + 1) / 255 \\ c &= ((255 - T + 1) / 255) + 0.3 \end{aligned} \quad (4)$$

In addition, the classified threshold T was a key point, which was critical to the validity of b and C :

$$T = \left(\frac{\sigma_B^2}{\sigma_\omega^2} \right)_{\max} \quad (5)$$

$$\begin{aligned} \sigma_B^2 &= \omega_1(m_1 - m)^2 + \omega_2(m_2 - m)^2 \\ \sigma_\omega^2 &= \omega_1\sigma_1^2 + \omega_2\sigma_2^2 \end{aligned} \quad (6)$$

In the above equation, ω_1 was the number of set \mathbf{c}_1 , m_1 and σ_1^2 were the mean value and variance of \mathbf{c}_1 , ω_2 were the number of set \mathbf{c}_2 , m_2 and σ_2^2 were the mean value and variance of \mathbf{c}_2 , and m was the mean value of the gray level of the image.

2.1.2 Interest Points and Candidates Detection and Extraction:

Strong relationships exist between the probability of the correct matching and the texture information content. Although detailed rules differ by surface variation, it has been shown that a successful rate of matching increases the probability of enriched texture information.

As a powerful tool for texture classification, the Local Binary Pattern (LBP) operator was an excellent measure of the spatial structure of local image texture. Ojala (2002) developed the LBP operator, which allows for detecting "uniform" patterns in circular neighborhoods of any quantization of the angular space and spatial resolution. Moreover, he enhanced the performance of the LBP operator through combining it with a rotation invariant local contrast (LC) measure that characterizes the contrast of local image texture.

In the case of an image, texture T was the joint distribution of the gray levels:

$$T = \mathbf{t}(\mathbf{g}_c, \mathbf{g}_0, \dots, \mathbf{g}_{P-1}) \quad (7)$$

Where P ($P > 1$) was the number of neighboring pixels, gray value \mathbf{g}_c was the center pixel of the local neighborhood, and \mathbf{g}_p was the neighboring pixel on a circle of radius R , as shown in Fig. 2. The coordinates of \mathbf{g}_p were calculated using Equation 8, and the gray values of neighbors which were not exactly in the center of the pixels were estimated by interpolation.

$$\begin{aligned} X_{g_p} &= X_{g_c} - R \cdot \sin(2\pi p / P) \\ Y_{g_p} &= Y_{g_c} + R \cdot \cos(2\pi p / P) \end{aligned} \quad (8)$$

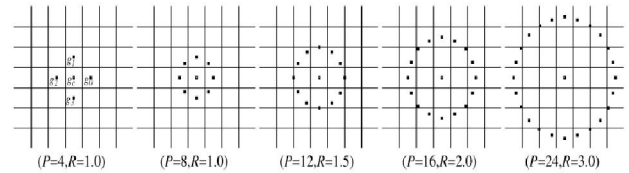


Figure 1: Circularly symmetric neighbor sets for different (P, R) (Ojala, 2002)

In fact, without losing information, the signed differences between the gray value of \mathbf{g}_c and \mathbf{g}_p ($P=0, \dots, P-1$) were not affected by variation in the mean luminance. Hence, the joint difference distribution was invariant against the grayscale shifts. By considering the signs of the differences, instead of their exact values, the operator could achieve the scale invariance:

$$T \approx \mathbf{t}(\mathbf{s}(\mathbf{g}_0 - \mathbf{g}_c), \mathbf{s}(\mathbf{g}_1 - \mathbf{g}_c), \dots, \mathbf{s}(\mathbf{g}_{P-1} - \mathbf{g}_c)) \quad (9)$$

$$\mathbf{s}(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (10)$$

By assigning a binomial factor 2^p for each sign $\mathbf{s}(\mathbf{g}_p - \mathbf{g}_c)$, we transformed Equation 9 into a unique LBP value. In this study, we set ($P=12, R=1.5$):

$$LBP_{P,R} = \sum_{p=0}^{P-1} \mathbf{s}(\mathbf{g}_p - \mathbf{g}_c) 2^p \quad (11)$$

Ojala designated patterns that have U values of at most 2 as "uniform" and proposed the following operator for grayscale and rotation invariant texture description instead of $LBP_{P,R}$:

$$LBP_{P,R}^{min2} = \begin{cases} \sum_{p=0}^{P-1} \mathbf{s}(\mathbf{g}_p - \mathbf{g}_c) & \text{if } U(LBP_{P,R}) \leq 2 \\ P+1 & \text{otherwise} \end{cases} \quad (12)$$

Where

$$U(LBP_{P,R}) = |\mathbf{s}(\mathbf{g}_{P-1} - \mathbf{g}_c) - \mathbf{s}(\mathbf{g}_0 - \mathbf{g}_c)| + \sum_{p=1}^{P-1} |\mathbf{s}(\mathbf{g}_p - \mathbf{g}_c) - \mathbf{s}(\mathbf{g}_{p-1} - \mathbf{g}_c)| \quad (13)$$

The $LBP_{P,R}^{riu2}$ operator was a grayscale invariant measure. It was an excellent measure of the spatial pattern, but it discarded contrast. So a rotation invariant measure was introduced to characterize the contrast of the local image texture:

$$LC_{P,R} = \frac{1}{P} \sum_{p=0}^{P-1} (g_p - \mu)^2 \quad (14)$$

Where

$$\mu = \frac{1}{P} \sum_{p=0}^{P-1} g_p$$

$LC_{P,R}$ was an invariant against shifts in grayscale. Since $LBP_{P,R}^{riu2}$ and $LC_{P,R}$ are complementary, their joint distribution $LBP_{P,R}^{riu2} / LC_{P,R}$, was expected to be a very powerful rotation invariant measure of local image texture.

When we compared feature detectors, Harris-Z was found to be more suitable to detect the interest points for large data sets of satellite imagery in our study. We also introduced the Förstner detector at this point and the surface fitting approach to get sub-pixel positioning. In this study, we combined Harris-Z with LBP/C to detect the stable interest points and candidates that optimize the validity of matching, which included the following five steps:

- 1) To ensure uniform distribution of the interest points, we used the grid of the feature extraction as a region unit, and a fixed number of features were extracted in each unit independently.
- 2) The interest strengths and LBP/C values in a region unit were calculated by the Harris-Z detector and the LBP/C operator.
- 3) Some points were saved when their LBP/C values were in a certain range. Then, we chose the point with the local maximum interest strength as the interest point that corresponded to the region unit. As a candidate, we chose the point with the maximum interest strength without consideration for its LBP/C value.
- 4) We used the Förstner detector and the surface fitting approach to obtain the sub-pixel positioning of the interest point and candidate.
- 5) Once the interest point failed to match or the accuracy of this match was low, the candidate was matched rather than the interest point; and the match with the higher accuracy was saved.

2.2 Image Matching Procedure

Due to linear pushbroom imaging mode, we can calculate the exterior orientation corresponding to every scanline through lagrangian interpolation. Thereby, in this paper, the proposed matching algorithm is based on the geometrically constrained cross-correlation. The algorithm includes the following several studies:

2.2.1 Automatic Computation for Approximate Elevation Range and Elevation Step through Global DEM

According to the position and attitude information of the scan lines, we use a hypothesis height to the frontal-project four corners of the corresponding reference image patch onto the ground. The approximate geography scope covered by the image patch was determined, and the maximum elevation H_{max} and the minimum elevation H_{min} are achieved through retrieval from Global DEM.

Then, we frontal-projected an interest point extracted onto the ground using H_{max} and H_{min} , and back-projected these object space points onto all of the search strips. The corresponding image coordinates (x_{max}^i, y_{max}^i) , (x_{min}^i, y_{min}^i) then could be determined, where i was the index of the different search strips. After that, we calculated the elevation step H_{pitch}^i with the following equation:

$$H_{pitch}^i = (H_{max} - H_{min}) / Length^i \quad (15)$$

$$\text{Where } Length^i = \sqrt{(x_{max}^i - x_{min}^i)^2 + (y_{max}^i - y_{min}^i)^2} \quad (16)$$

It was an iterative process: once H_{pitch}^i was an invalid value, we had to repeat the above procedure using another interest point.

2.2.2 Geometric and Radiometric Distortion Procedure

There were geometric and radiometric distortions in the case of strips acquired from different camera lenses. When the correlation window is defined in the reference image, this kind of distortion leads to an irregular and discontinuous shape of the corresponding window in the search images, which cannot be matched in a straightforward manner.

Gruen and Baltsavias (1985) extended the LSM method, which compensates this distortion and relaxes the strict geometric and radiometric assumption for normal cross-correlation methods. So, we introduced this method to eliminate the window warping:

- 1) In the reference image, we defined a correlated window Γ and a search window Γ' in which an interest point or candidate was central, then the corresponding pixel coordinates of the four corners and the center of Γ' were achieved by projection in the search image, which formed an irregular polygon Γ'' , as shown in Fig. 2. Accordingly, we calculated the parameters of the affine geometric transformation and linear radiometric transformation between these images.

$$x_2 = a_0 + a_1x + a_2y \quad (17)$$

$$y_2 = b_0 + b_1x + b_2y \quad (18)$$

$$g(x, y) = h_0 + h_1g_2(x_2, y_2) \quad (19)$$

Where $a_0, a_1, a_2, b_0, b_1, b_2$ were the parameters of the affine geometric transformation; h_0, h_1 were the parameters of the linear radiometric transformation.

- 2) Through Equations 17 and 18, the relationship of each pixel between Γ and Γ' was determined. The gray values in Γ' were interpolated by bilinear interpolation, and these gray values, recalculated by Equation 19, were assigned to the corresponding pixel in Γ .

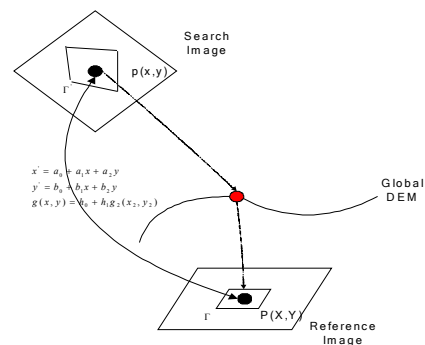


Figure 2. The principle of the affine geometric transformation and linear radiometric transformation between the reference image and search image

3) Once the angle θ was greater than 90 degrees on the very steep terrain surface, as shown in Fig. 3, W' became zero and made the procedure impossible. In this case, we found that when we iteratively computed the parameters of the transformation, the convergent speed was slow, and even the computation was divergent. So once the threshold of iteration times was crossed, we considered $\theta \geq 90$ and excluded it.

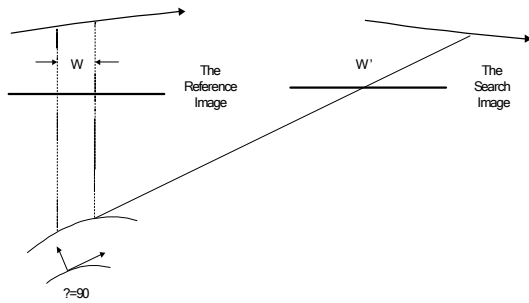


Figure 3. Image selection or exclusion based on the terrain slope and the sensor geometry. The size of the corresponding correlation window in the search image depends on the angle θ between the normal vector of the local object surface and the corresponding imaging ray.

4) The correlation window and search window are both square as shown in Fig. 4. Therefore, we could do a straightforward match and reduce the computational cost.

2.2.3 Extended Geometrically Constrained Cross-Correlation Strategy

Based on the area-based matching techniques, Zhang developed GC^3 (Geometrically Constrained Cross-Correlation) algorithm. With the GC^3 matching procedure, multiple images can be matched simultaneously and the epipolar constraint is integrated implicitly. As a similarity measure, SNCC (Sum of Normalized Cross-Correlation) greatly decreases mismatches caused by approximate texture distinction and occlusion problem. However, this algorithm causes high computational costs in determining the correct height value and has redundancy for its quality measuring procedure. Basically we extended this algorithm and save the computational costs:

- 1) For each reference image patch, the approximate elevation range and elevation step can be determined. After pre-processing the reference image patch, for each interest point or candidate, the corresponding approximate epipolar lines in the search strips can be further determined.
- 2) According to the position of approximate epipolar lines, we define the search image patches, and pre-process them.
- 3) Based on geometric and radiometric distortion procedure, we calculate the parameters of affine geometric transformation and linear radiometric transformation between the defined windows.
- 4) We combine a coarse-to-fine hierarchical approach with the efficient implementation of the epipolar geometry to match each point on the approximate epipolar lines. Once normalized cross-correlation coefficient in the certain level of pyramids is lower than threshold, we repeat the above procedures with different elevation in the certain range.
- 5) Until accomplishing the matching procedure, we compute the SNCC of all matching candidates.

- 6) Determine the sites of local maxima of the SNCC function and fit a smooth quadratic function for each local maximum with a local neighborhood of that maximum value.
- 7) If the second SNCC peak is less than half or one third of that of the first SNCC peak, the peak that have the largest SNCC value should represent the correct match. If not, we match inversely and if the difference between the two matches is less than 1.5 pixels in image space, the candidate should be the correct match.
- 8) Repeat 1)-7), until all interest points and candidate are matched.

2.3 Mismatches Detection and Elimination

After finishing the image matching procedure, there are inevitably some blunders that do not conform to the geometrical relationship in an object space, which we call mismatches. To maintain the matching accuracy, we must eliminate mismatches as accurately as possible. Thus, we classify all of the matches into two types, interior matches and adjacent matches. In this section, we present two strategies to eliminate mismatches

2.3.1 Mismatches Elimination Strategy within Orbit: In this study, we introduce the multi-rays forward intersection principle to eliminate mismatches and realize the object fusion for the matching results. The implementation strategy was as follows:

- 1) If an interest point or candidate finds multiple correspondences on the search strips, we assume that the exterior orientation elements corresponding to each image are known. We take the 3D coordinates of all matches in the object space as unknown; and based on the collinearity equations, we then list the error equations:

$$V_x = -a_{11}dX - a_{12}dY - a_{13}dZ - l_x \quad (20)$$

$$V_y = -a_{21}dX - a_{22}dY - a_{23}dZ - l_y \quad (21)$$

Where $a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}$ are the parameters related to the line elements of the exterior orientation; l_x, l_y are differences between observations and approximations.

- 2) If the number of the correspondences is n , we can list $(2n+2)$ error equations. To solve the iterative computation for unknown numbers, the least square adjustment was introduced. In the iterative procedure, we set the weights of each observation of the image points through a hypothesis test of the posterior variance component. If there are grosses in the candidates, the observation weights of the candidates become smaller and smaller until 0 was reached, which makes no sense in the adjustment process in order to eliminate the effect of a false match on the calculation of effectiveness.
- 3) After iterative convergence was achieved or the iterative times exceed the limited value, the root mean square (RMS) value was calculated. Once the RMS exceeds the given threshold, the corresponding matches are eliminated.

2.3.2 Mismatches Elimination Strategy between Adjacent Orbits

Once the position and attitude parameters acquired from different orbits exhibited obvious system shift, the matches between the adjacent orbits could not be intersected together in the object space. If we used the above strategy, we would face the problem of iterative computation convergence. Therefore,

we introduced an Object Space Shift (OSS) strategy to eliminate mismatches between the adjacent orbits:

1) The correspondences P1, P2, ..., Pn between the adjacent orbits were intersected within each orbit, and the shift of each point in the object space were calculated $(\Delta X_1, \Delta Y_1, \Delta Z_1)$, $(\Delta X_2, \Delta Y_2, \Delta Z_2)$, ..., $(\Delta X_n, \Delta Y_n, \Delta Z_n)$. Then, the mean shift A_x, A_y, A_z were calculated in Equation 22:

$$\begin{aligned} A_x &= (\Delta X_1 + \Delta X_2 + \Delta X_3 \dots + \Delta X_n) / n \\ A_y &= (\Delta Y_1 + \Delta Y_2 + \Delta Y_3 \dots + \Delta Y_n) / n \\ A_z &= (\Delta Z_1 + \Delta Z_2 + \Delta Z_3 \dots + \Delta Z_n) / n \end{aligned} \quad (22)$$

The mean square values in each direction m_x, m_y, m_z were calculated in Equation 23, and the RMS m_{error} between the adjacent orbits were calculated in Equation 24.

$$\begin{aligned} m_x &= \sqrt{((\Delta X_1 - A_x)^2 + (\Delta X_2 - A_x)^2 + (\Delta X_3 - A_x)^2 \dots + (\Delta X_n - A_x)^2) / n} \\ m_y &= \sqrt{((\Delta Y_1 - A_y)^2 + (\Delta Y_2 - A_y)^2 + (\Delta Y_3 - A_y)^2 \dots + (\Delta Y_n - A_y)^2) / n} \\ m_z &= \sqrt{((\Delta Z_1 - A_z)^2 + (\Delta Z_2 - A_z)^2 + (\Delta Z_3 - A_z)^2 \dots + (\Delta Z_n - A_z)^2) / n} \end{aligned} \quad (23)$$

$$m_{error} = \sqrt{(m_x \bullet m_x + m_y \bullet m_y + m_z \bullet m_z)} \quad (24)$$

2) Once the shift T_i was more than triple m_{error} , we deemed this correspondence as a mismatch.

$$\begin{aligned} T_i &= \sqrt{(\Delta X_i - A_x)^2 + (\Delta Y_i - A_y)^2 + (\Delta Z_i - A_z)^2} \\ T_i &\geq 3 \times m_{error} : \quad \text{Mismatch} \\ T_i &< 3 \times m_{error} : \quad \text{Correct Match} \end{aligned} \quad (25)$$

3. EXPERIMENTS AND ANALYSIS

3.1 Test Data Description

Figure 4 shows the real data for experiment, which consists of two long strips of imagery acquired by Chinese Mapping Satellite. These data are imaged on October, 2010, which covers Harbin in China, the coverage area is about 25000km². The satellite sensor works in pushbroom mode, and the stereo module consists of three lenses with one CCD line scanner for each, which provide a forward-, a nadir- and a backward-looking view. The degree of overlapping within orbit is about 98%, and the degree between adjacent orbits is about 10%. The parameters of CCD line sensors are shown in table 1.

Table 1. Camera parameters of the three linear array sensor for Chinese Mapping Satellite

Parameter	Forward-View	Nadir-View	Backward-View
Focal Length (mm)	717	650	717
GSD (m)	5.0	5.0	5.0
Intersection Angle (deg)	25.0	0.0	-25.0
Pixel Size (mm)	0.0065	0.0065	0.0065
Scan Line Width (pixel)	12000	12000	12000
Length of Strip (pixel)	142862	238100	142862
Geographical Scope (km ²)	42858.6	71430.0	42858.6
Data Volume (GB)	4.46	7.31	4.46
Base-Height Ratio		0.9323	

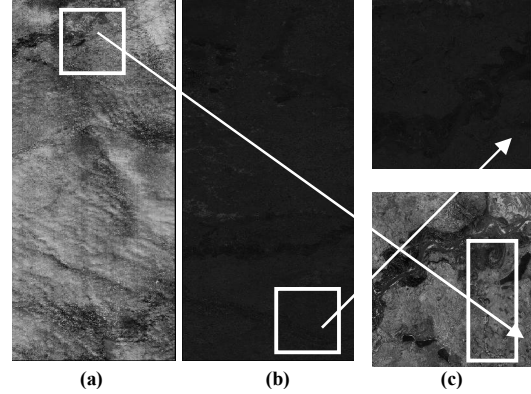


Figure 4. The Nadir-looking strips in the adjacent orbits: (a) The nadir-looking strip acquired from the first orbit; (b) The nadir-looking strip acquired from the second orbit; (c) The enlarged drawing of the overlapping area between the two strips

For testing and verifying the accuracy and reliability of the proposed algorithm, the geographical region covered by the test data includes not only farms and forests which have strong texture repeatability, but also residential quarters, mountain countries and smoothly open areas which are known as difficult terrain for matching. In addition, in order to analyze the matching accuracy quantitatively, the test data includes 33 planar control points and 15 elevation control points. After finishing the matching procedure, the ground control points are measured interactively on the test imagery, and then the bundle block adjustment is performed with matching points and control points.

3.2 Image Matching Testing

In order to improve the performance of the proposed matching algorithm, we compare different parameters in two ways: the number of pyramid levels and the size of the relative window at the top level.

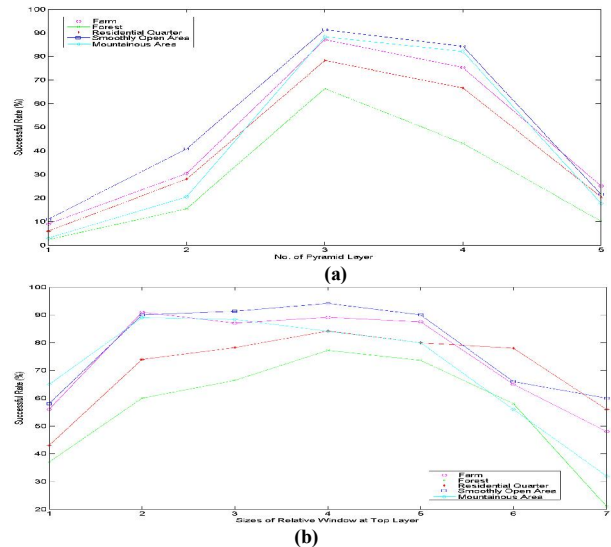


Figure 5. The successful rate of matching using the number of pyramid levels and the size of the relative window at the top level: (a) the effect of different number of pyramid levels; (b) the effect of different size of relative window at top level

Figure 5 shows the results in which the number of pyramid levels and the size of the relative window at the top level are varied. In graph (a), when the number is 3, the successful rate of matching is the highest toward the different terrain types. In graph (b), it is poor at the successful rate when the relative windows is set 7×7 pixels, but the results are not improved till 13×13 pixels. After that, a larger size of relative window can actually hurt the matching. Although the successful rate for 13×13 is a little lower than 11×11 and 9×9 on mountain countries and farms respectively, the overall successful rate is optimum. Therefore, through this experiment, we generate three pyramid levels and set the size of the relative window $R \times R$ as 13×13 . Moreover, the relative window is shrank to $(R-2) \times (R-2)$ level to level, and the corresponding search window is $(2 \times R-1) \times (2 \times R-1)$. The matching results are listed in table 3.

Table 3. Matching accuracy quantitative evaluation for different types of terrain

Terrain Type	Image Patch	Interest Points	Pyramid Level	Correct Matches/Successful Rate	RMS/Max Res (Pixel)
Farm	1_001	106	3	94 / 88.6%	0.27 / 1.76
	1_003	112	3	96 / 85.7%	0.34 / 2.12
	1_005	108	3	97 / 89.8%	0.29 / 1.41
	2_002	119	3	108 / 90.7%	0.41 / 1.02
	2_003	84	3	70 / 83.3%	0.44 / 1.37
Forest	1_007	65	3	50 / 76.9%	0.65 / 4.6
	2_001	52	3	41 / 78.8%	0.74 / 2.7
Residential Quarter	1_006	103	3	80 / 77.6%	0.32 / 2.3
	2_004	94	3	78 / 82.9%	0.47 / 2.15
Smoothly Open Area	1_002	106	3	97 / 91.5%	0.26 / 1.21
	1_004	106	3	93 / 87.7%	0.21 / 3.24
	2_005	96	3	91 / 94.7%	0.33 / 4.82
Mountain Country	2_006	114	3	95 / 84.1%	0.74 / 17.46

Table 3 illustrates the statistics of image matching for the test date on five types of terrain, through analysis for the matching results, we can see that:

- 1) For farms and forests which have strong texture repeatability, the range of the successful rate of matching is from 76.9% to 90.7%. Especially in areas covered by farms, the RMS error is within 0.5 pixel, which means that the proposed algorithm can avoid mismatching problem caused by texture repeatability.
- 2) For residential quarter and smoothly open areas, there are a few false matches, but most of matches are fine, the range of the successful rate is from 77.6% to 94.7%, the range of RMS error is from 0.21 to 0.47 pixel. Thus, the matching result is satisfactory.
- 3) For mountain countries, because of the rise or descend on terrain greatly, there are some obviously false matches, which greatly affect the precision of matching. But the RMS error is

still 0.74 pixel, which proves that the matching algorithm can be applied to this terrain.

4. CONCLUSION

In this study, we develop a practical algorithm for matching with large data sets of long-strips of satellite imagery. Through some experiments using real data set from Chinese Mapping Satellite, the proposed algorithm are proved to be valid, and the satisfactory accuracy is achieved under fully automatic computation.

At the present stage, the proposed methods mostly focus on multi-view matching with long strips of satellite imagery acquired from different orbits. In the future, we will keep on this research and realize the matching algorithm that support various satellite sensors, even various imaging modes.

REFERENCES

- DAVID, G. LOWE. 2004. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91-110
- Heipke, C. 1996. Overview of image matching techniques. In *Proceedings of OEEPE Workshop on the Application of Digital Photogrammetric Workstations*, Lausanne, Switzerland, Part III
- Baltsavias, E.P. 1991. Multiphoto geometrically constrained matching. Ph.D. dissertation, Institute of Geodesy and Photogrammetry, ETH Zurich, Report No. 49.
- F.Bellavia, D.Tegolo, C.Valenti. 2008. A non-parametric scale-based corner detector. *Pattern Recognition (ICPR)*, 19th International Conference, pp, 1-4
- Timo, Ojala and Matti, Pietikäinen. 2002. Multisolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 7, No. 24, pp. 971-987
- Foerstner, W. 1986. A Feature Based Correspondence Algorithm for Image Matching. *Proceedings of the Symposium from Analytical to Digital*, August 19-22, Rovaniemi, Finland, pp. 150-166.
- Zhang, Li. 2005. Auto Digital Surface Model (DSM) Generation from Linear Array Images. Zurich: Ph. D. Dissertation, ETH Zurich, Switzerland
- Zhang, L., Gruen, A. 2004. Automatic DSM Generation from Linear Array Imagery Data. *IAPRS*, Vol. 35, Part B3, pp. 128-133
- Q. Zhu., B. Wu., N. Wan. 2006. An interest point detect method to stereo images with good repeatability and information content. *Acta Electron. Sin.*, Vol. 34, No. 2, pp. 205-209
- Gruen, A and Baltsavias, E. P. 1985. Adaptive Least Squares Correlation with Geometrical Constraints. *SPIE*, Vol. 595, pp. 72-82
- Gruen, A. and Baltsavias, E. P. 1988. Geometrically Constrained Multiphoto Matching. *PE&RS*, Vol. 54, No. 5, pp. 633-641