remote sensing

MDPI

Article

Near Real-Time Automatic Sub-Pixel Registration of Panchromatic and Multispectral Images for Pan-Sharpening

Guangqi Xie [1], Mi Wang [1,*], Zhiqi Zhang [1,2], Shao Xiang [1] and Luxiao He [1]

[1] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China; xiegqrs@whu.edu.cn (G.X.); zzq540@whu.edu.cn (Z.Z.); xiangshao@whu.edu.cn (S.X.); heluxiao@whu.edu.cn (L.H.)
[2] School of Computer Science, Hubei University of Technology, Wuhan 430068, China
* Correspondence: wangmi@whu.edu.cn

Abstract: This paper presents a near real-time automatic sub-pixel registration method of high-resolution panchromatic (PAN) and multispectral (MS) images using a graphics processing unit (GPU). In the first step, the method uses differential geo-registration to enable accurate geographic registration of PAN and MS images. Differential geo-registration normalizes PAN and MS images to the same direction and scale. There are also some residual misalignments due to the geometrical configuration of the acquisition instruments. These residual misalignments mean the PAN and MS images still have deviations after differential geo-registration. The second step is to use differential rectification with tiny facet primitive to eliminate possible residual misalignments. Differential rectification corrects the relative internal geometric distortion between PAN and MS images. The computational burden of these two steps is large, and traditional central processing unit (CPU) processing takes a long time. Due to the natural parallelism of the differential methods, these two steps are very suitable for mapping to a GPU for processing, to achieve near real-time processing while ensuring processing accuracy. This paper used GaoFen-6, GaoFen-7, ZiYuan3-02 and SuperView-1 satellite data to conduct an experiment. The experiment showed that our method's processing accuracy is within 0.5 pixels. The automatic processing time of this method is about 2.5 s for 1 GB output data in the NVIDIA GeForce RTX 2080Ti, which can meet the near real-time processing requirements for most satellites. The method in this paper can quickly achieve high-precision registration of PAN and MS images. It is suitable for different scenes and different sensors. It is extremely robust to registration errors between PAN and MS.

Keywords: registration; pansharpening; near real-time; automatic

check for updates

Citation: Xie, G.; Wang, M.; Zhang, Z.; Xiang, S.; He, L. Near Real-Time Automatic Sub-Pixel Registration of Panchromatic and Multispectral Images for Pan-Sharpening. *Remote Sens.* **2021**, *13*, 3674. https://doi.org/10.3390/rs13183674

Received: 29 July 2021
Accepted: 12 September 2021
Published: 14 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 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 (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

From the perspective of how sensors work, remote sensing image data are split into PAN images and MS images [1]. The detector must receive enough energy to maintain the signal-to-noise ratio. The MS single-band receiving energy is weak, and the size of the detector can only be increased, resulting in the spatial resolution of an MS image being lower than that of the PAN image. Although the spatial resolution of PAN images is high, in order to increase the intensity of the energy received, its spectral range is wider and its spectral resolution is lower than those of MS images. In general, a PAN image has relatively high spatial resolution, but the spectral resolution is low; an MS image data adopts a high spectral resolution, and the spatial resolution is low. A pansharpening method combines the advantages of the two: it is a process of merging a high-resolution PAN image and a lower resolution MS image to create a single high-resolution color image [2–4].

Due to the complementarity of various datasets, pansharpening technology has become one of the main research topics in satellite image processing. Early methods such as hue–intensity–saturation (HIS) [5], principal component analysis (PCA) [6] and high pass filter (HPF) [7], used in Landsat TM and SPOT, have had achieved great success [8].

During the decades since, a large number of excellent algorithms have emerged. For example, smoothing filter-based intensity modulation (SFIM) [9,10], Brovey transformation (BT) [11,12], generalized Laplacian pyramid (GLP) [13], and Gram–Schmidt (GS) [14] are some classic pansharpening algorithms. According to the latest review paper [1,15], there are at least dozens of pansharpening methods. A useful benchmark consisting of recent advances in MS pansharpening was also presented in [16]. At the same time, there are many ways to take advantage of the depth of learning in pansharpening [17–22].

There is a certain difference in the resolution of PAN and MS images. For most satellites, PAN and MS image resolution differ four fold, and in other cases the resolution differs two-fold—e.g., the Landsat series. There also some other resolution differences. The NASA Earth Science program defined Standard Data Products in [23]. Most of the references related to fusion do not specify the level of the product, which assumes that PAN and MS images are accurately registered. However, this issue needs to be considered in actual production. General pansharpening can be processed on *Level 1A* (after radiometric correction) and *Level 2* (after geometric and radiometric correction) images [24]. For *Level 2*, the MS image can be directly up-sampled to register with the PAN image [25–28]. However, this method has some problems with high resolution images: (1) The amount of image data after geometric correction increases significantly, which significantly increases the computational burden for subsequent pansharpening. (2) The image after geometric correction has black edges, which is inconvenient for subsequent pansharpening. (3) The PAN and MS images are generally corrected separately in blocks. The blocks do not correctly account for actual size though, resulting in the corrected images having certain errors, especially in the mountainous areas. As described in [24,29], even with geometrically registered PAN and MS images, differences might appear between images.

Therefore, in practical applications, the first step of pansharpening must be high-precision registration of PAN and MS images [30,31]. Image registration is an important image processing procedure in remote sensing. Nevertheless, it is still difficult to find an accurate, robust, and automatic image registration method. Most of the traditional methods are based on the image side. They can be roughly divided into two categories. One is feature point matching, which includes SIFT [32–34] and its improved version SURF [35,36]. The second is template matching, which includes correlation coefficient [37], mutual information [38], and so on. These methods all rely on the relationship between the images, and it is difficult to complete the registration when the image difference is too large. In addition, matching based on feature points cannot eliminate random errors.

Therefore, in order to automatically obtain high-precision PAN and MS registration upscaled images, we fused *Level 1A* images and then corrected them to *Level 2*. In large-scale, automated production of pansharpening products, MS *Level 1A* images cannot be directly up-sampled, especially for high-resolution images, because they are not strictly projected on the same elevation surface. As we all know, the rational function model (RFM) is projected on the mean elevation surface. The PAN and MS images of linear push-broom satellites are sometimes not strictly the same, and the average elevation surface is often different. This is more obvious in places where the terrain is undulating, which has a great impact on the imaging conditions of low-orbit satellites with large roll angles and pitch angles. Enlarging the MS *Level 1A* image directly means that it is still projected on the average elevation surface of the MS image, which will have an obvious deviation from the PAN image. In this study, the method of differential geo-registration based on geographic location was used to complete the coarse-registration of PAN and MS images. Differential geo-registration normalized PAN and MS *Level 1A* images to the same direction and scale.

Although the PAN and MS images are taken almost simultaneously, there is still a certain delay between them, which causes the PAN and MS images to be incompletely registered. In addition, the tremors of any satellite platform [39,40] affect the accuracy of the PAN and MS images' registration. Although the relative relationship between PAN and MS will be calibrated in preprocessing, this can only eliminate systematic errors, and it is difficult to eliminate random errors [41,42]. Therefore, registration from the image

matching method is required. In this study, the method of rectification with tiny facet primitive [43] was used for high-precision sub-pixel registration.

Finally, due to the development of remote sensing satellite manufacturing technology in recent years, the amount of satellite imaging single scene data has increased sharply, and applications such as disaster relief and military reconnaissance have extremely high requirements for the efficiency of algorithms. Traditional CPU-based algorithms are very inefficient when processing tens of gigabytes or even hundreds of gigabytes, and often cannot meet real-world demands. There is also an urgent need for automated processing [44,45]. In order to achieve near real-time automatic sub-pixel registration of PAN and MS images, our method maps parallel algorithms to a GPU for processing. The method in this paper can greatly improve the registration efficiency while ensuring accuracy.

2. Methods

To achieve near real-time automatic registration of high resolution PAN and MS images, the algorithm in this paper was designed with full consideration of parallelism. The differentiation method is one of the best methods for parallel processing, in which the basic theory is to treat a complex model as consisting of many simple models. These simple models can be processed in parallel. They can be efficiently mapped to a GPU for processing. We used differential geo-registration to geometrically register PAN and MS images to obtain co-registered enlarged images, and eliminated potential errors through differential rectification with tiny facet primitive to achieve sub-pixel accurate registration. The entire process can be automated and processed in near real-time. The steps of the algorithm in this paper are as follows:

Step 1. Calculate the overlap range of PAN and MS images;
Step 2. Differential geo-registration for PAN and MS images;
Step 3. Differential rectification with tiny facet primitive for PAN and MS images.

The method in this paper was used as the preprocessing of PAN and MS images before fusion, which can accurately register them to improve the quality of fusion. We assumed that the bands of a given MS image were accurately registered, and only considered the misregistration between PAN and MS images. In addition, only the four bands of visible, near-infrared (VNIR) were considered in this experiment—that is, the bands of blue, green, red, and near-infrared. In fact, as long as the MS bands are accurately registered, no matter how many bands there are, the result should be the same. As the algorithm in this paper uses RGB to proportionally synthesize the pseudo-PAN image and the PAN image to match the control points and rectify all MS bands with the same parameters.

2.1. Calculate the Overlap Range

Although PAN and MS images are taken at the same time, in actual processing, one finds that the scene divisions of the two do not strictly correspond, especially in the case of different sources. Therefore, the overlap range of the two images needs to be calculated first.

As shown in Figure 1, the four corner points of PAN and MS were projected to the same elevation surface through the RFM model [46–48], and the overlap area on the object side was calculated. Then the overlap area was mapped back on their respective image coordinates to obtain overlay ranges for the PAN and MS images.

The RFM is a universal rational polynomial model that directly establishes the relationship between the pixel coordinates of an image point and the geographic coordinates of its counterpart. RFM hides the satellite sensor parameters and attitude orbit parameters, and has many advantages, such as versatility, high calculation efficiency, and coordinate back calculation without iteration. Therefore, it has been widely used and has become an international standard.

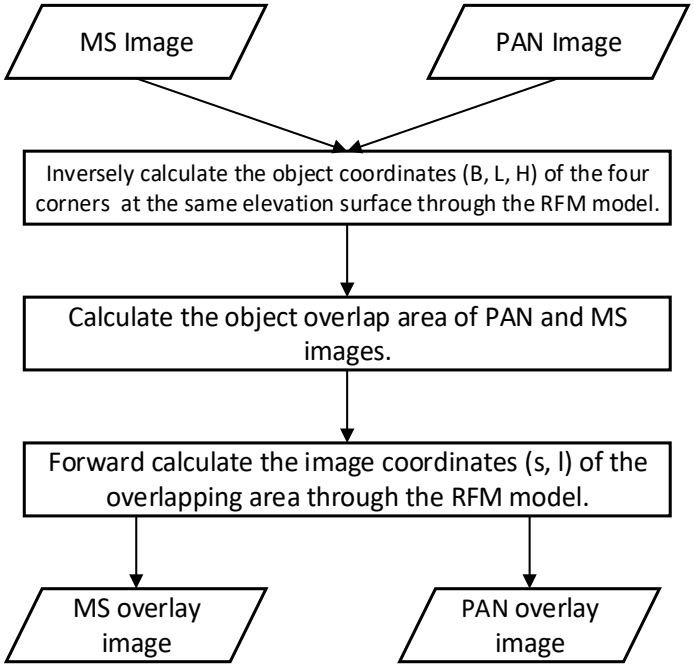


Figure 1. Calculate the overlap range. See Equation (1) for definitions of parameters.

In order to ensure the stability of the calculation, RFM regularizes the image coordinates (l, s), latitude and longitude coordinates (B, L), and the height of ellipsoid H to make the coordinate range between $[-1, 1]$. The formula for calculating the normalized image coordinates (l, s) corresponding to the image coordinates (l_n, s_n) of the image point is:

$$\begin{cases} l_n = \frac{l - LineOff}{LineScale} \\ s_n = \frac{s - SampleOff}{SampleScale} \end{cases}, \tag{1}$$

Among them, $LineOff$, $SampleOff$ are the translation values of the image-side coordinates, respectively, and $LineScale$, $SampleScale$ are the zoom values of the image-side coordinates, respectively.

The formula for calculating the normalized coordinates (U, V, W) of the object coordinates (B, L, H) is:

$$\begin{cases} U = \frac{B - LonOff}{LonScale} \\ V = \frac{L - LatOff}{LatScale} \\ W = \frac{H - HeiOff}{HeiScale} \end{cases}, \tag{2}$$

Among them, $LonOff$, $LatOff$, $HeiOff$ are the translation values of the object coordinates, and $LonScale$, $LatScale$, $HeiScale$ are the zoom values of the object coordinates.

For each scene image, the relationship between image coordinates and object coordinates can be expressed as a polynomial ratio as follows:

$$\begin{cases} l_n = \frac{Num_L(U,V,W)}{Den_L(U,V,W)} \\ s_n = \frac{Num_S(U,V,W)}{Den_S(U,V,W)} \end{cases}, \tag{3}$$

The numerator and denominator of the polynomial in the above formula are expressed as follows:

$$\begin{aligned}
Num_L(U,V,W) &= a_1 + a_2V + a_3U + a_4W + a_5VU + a_6VW + a_7UW + a_8V^2 \\
&\quad + a_9U^2 + a_{10}W^2 + a_{11}VUW + a_{12}V^3 + a_{13}VU^2 + a_{14}VW^2 \\
&\quad + a_{15}V^2U + a_{16}U^3 + a_{17}UW^2 + a_{18}V^2W + a_{19}U^2W + a_{20}W^3 \\
Den_L(U,V,W) &= b_1 + b_2V + b_3U + b_4W + b_5VU + b_6VW + b_7UW + b_8V^2 \\
&\quad + b_9U^2 + b_{10}W^2 + b_{11}VUW + b_{12}V^3 + b_{13}VU^2 + b_{14}VW^2 \\
&\quad + b_{15}V^2U + b_{16}U^3 + b_{17}UW^2 + b_{18}V^2W + b_{19}U^2W + b_{20}W^3 \\
Num_S(U,V,W) &= c_1 + c_2V + c_3U + c_4W + c_5VU + c_6VW + c_7UW + c_8V^2 \\
&\quad + c_9U^2 + c_{10}W^2 + c_{11}VUW + c_{12}V^3 + c_{13}VU^2 + c_{14}VW^2 \\
&\quad + c_{15}V^2U + c_{16}U^3 + c_{17}UW^2 + c_{18}V^2W + c_{19}U^2W + c_{20}W^3 \\
Den_S(U,V,W) &= d_1 + d_2V + d_3U + d_4W + d_5VU + d_6VW + d_7UW + d_8V^2 \\
&\quad + d_9U^2 + d_{10}W^2 + d_{11}VUW + d_{12}V^3 + d_{13}VU^2 + d_{14}VW^2 \\
&\quad + d_{15}V^2U + d_{16}U^3 + d_{17}UW^2 + d_{18}V^2W + d_{19}U^2W + d_{20}W^3
\end{aligned} \tag{4}$$

Among them, $a_i, b_i, c_i, d_i (i = 1, 2, \cdots, 20)$ are the rational polynomial coefficients (RPCs). In general, b_1, d_1 are both 1.

When both b_1, d_1 take the value 1, only 78 other rational polynomial coefficients need to be solved. The least square method can be used to solve RPCs through virtual control points. Thus Equation (1) can be written as follows:

$$\begin{aligned}
v_l &= Num_L(U,V,W) - l_n Den_L(U,V,W) \\
v_s &= Num_S(U,V,W) - s_n Den_S(U,V,W)
\end{aligned} \tag{5}$$

Then the error equation is:

$$\begin{aligned}
v_l &= \begin{bmatrix} 1 & V & U & \cdots & U^2W & W^3 & -l_nV & -l_nU & \cdots & -l_nU^2W & -l_nW^3 \end{bmatrix} \cdot \boldsymbol{X}_l - l_n \\
v_s &= \begin{bmatrix} 1 & V & U & \cdots & U^2W & W^3 & -s_nV & -s_nU & \cdots & -s_nU^2W & -s_nW^3 \end{bmatrix} \cdot \boldsymbol{X}_s - s_n
\end{aligned} \tag{6}$$

Among them, $\boldsymbol{X}_l = \begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_{19} & a_{20} & b_2 & b_3 & \cdots & b_{19} & b_{20} \end{pmatrix}^T$ and $\boldsymbol{X}_s = \begin{pmatrix} c_1 & c_2 & c_3 & \cdots & c_{19} & c_{20} & d_2 & d_3 & \cdots & d_{19} & d_{20} \end{pmatrix}^T$.

It is not difficult to find that the coefficients X_l and X_s to be solved are independent unknown parameters and can be solved separately. Take the first equation in Equation (6) as an example. If there are n virtual control points, the matrix form of the observation equation can be expressed as follows:

$$\boldsymbol{V}_l = \boldsymbol{A}_l X_l + \boldsymbol{L}_l \tag{7}$$

Among them,

$$\begin{aligned}
\boldsymbol{V}_l &= \begin{bmatrix} v_{l1} & v_{l2} & \cdots & v_{ln} \end{bmatrix}^T, \\
\boldsymbol{A}_l &= \begin{bmatrix} 1 & V_1 & U_1 & \cdots & U_1{}^2W_1 & W_1{}^3 & -l_{n1}V_1 & -l_{n1}U_1 & \cdots & -l_{n1}U_1{}^2W_1 & -l_{n1}U_1{}^3 \\ 1 & V_2 & U_2 & \cdots & U_2{}^2W_2 & W_2{}^3 & -l_{n2}V_2 & -l_{n2}U_2 & \cdots & -l_{n2}U_2{}^2W_2 & -l_{n2}U_2{}^3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & V_n & U_n & \cdots & U_n{}^2W_n & W_n{}^3 & -l_{nn}V_n & -l_{nn}U_n & \cdots & -l_{nn}U_n{}^2W_n & -l_{nn}U_n{}^3 \end{bmatrix}, \\
\boldsymbol{L}_l &= \begin{bmatrix} -l_{n1} & -l_{n2} & \cdots & -l_{nn} \end{bmatrix}^T,
\end{aligned}$$

In the same way, the adjustment equation of the second equation can be obtained:

$$\boldsymbol{V}_s = \boldsymbol{A}_s \boldsymbol{X}_s + \boldsymbol{L}_s \tag{8}$$

Among them,

$$\boldsymbol{V}_s = \begin{bmatrix} v_{s1} & v_{s2} & \cdots & v_{sn} \end{bmatrix}^T,$$

$$\boldsymbol{A}_s = \begin{bmatrix} 1 & V_1 & U_1 & \cdots & U_1{}^2W_1 & W_1{}^3 & -s_{n1}V_1 & -s_{n1}U_1 & \cdots & -s_{n1}U_1{}^2W_1 & -s_{n1}W_1{}^3 \\ 1 & V_2 & U_2 & \cdots & U_2{}^2W_2 & W_2{}^3 & -s_{n2}V_2 & -s_{n2}U_2 & \cdots & -s_{n2}U_2{}^2W_2 & -s_{n2}W_2{}^3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & V_n & U_n & \cdots & U_n{}^2W_n & W_n{}^3 & -s_{nn}V_n & -s_{nn}U_n & \cdots & -s_{nn}U_n{}^2W_n & -s_{nn}W_n{}^3 \end{bmatrix},$$

$$\boldsymbol{L}_s = \begin{bmatrix} -s_{n1} & -s_{n2} & \cdots & -s_{nn} \end{bmatrix}^T,$$

Estimate the optimal solution of the error equation system by least squares:

$$\begin{aligned} \boldsymbol{X}_l &= \left(\boldsymbol{A}_l{}^T\boldsymbol{A}_l\right)^{-1}\boldsymbol{A}_l{}^T\boldsymbol{L}_l \\ \boldsymbol{X}_s &= \left(\boldsymbol{A}_s{}^T\boldsymbol{A}_s\right)^{-1}\boldsymbol{A}_s{}^T\boldsymbol{L}_s \end{aligned} \tag{9}$$

2.2. Differential Geo-Registration

Different remote sensing images can be geometrically coarse-registered using geographic information, especially PAN and MS images, which are imaged at the same time; they are basically consistent geographically. Therefore, by setting a basic simulation plane, the PAN and MS images can be simulated on the same plane with geographic information, and then the PAN and MS geographic registration can be realized. In this study, the overlapping area of PAN and MS images was used as the simulation plane, which was calculated using geographic information.

As shown in Figure 2, the overlapping area obtained in the previous step was divided into small blocks by the method of differentiation. The smaller the block, the higher the accuracy. The corresponding points of the same name between each block of each PAN image and each block of each MS image were obtained by the RFM model. Additionally, we calculated the transformation model for each block through the points of the same name. The transformation model can adopt linear transformation, affine transformation, perspective transformation, etc. The computational burdens and accuracies of different transformation models vary. Generally speaking, the greater the computational burden, the higher the accuracy. This paper adopted perspective transformation as its transformation model to improve accuracy as a result of what is discussed in Section 3.4. Although different blocks had been rectified with different parameters, an overlapping area between them was used to maintain their continuity. The same design was also used in the differential rectification with tiny facet primitive.

When mapping to the GPU, we need some of the following principles. (1) The size of the thread block must be an integer multiple of the number of thread beams. (2) The thread block size should be at least 64 or more. (3) The thread block's maximum should exceed 1024; (4) Avoid using too much logical hype, such as a branch statement or a loop statement.

Perspective transformation [49] is a linear projection where three dimensional objects are projected on a picture plane, which can take three-dimensional information into consideration. As shown in the following formula, the perspective transformation first projects two-dimensional space *(s, l)* to three-dimensional space *(x, y, z)*.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} s \\ l \\ 1 \end{pmatrix}, \tag{10}$$

Among them, *(x, y, z)* are represented by *(s, l)* as follows:

$$\begin{aligned} x &= a_{11}s + a_{12}l + a_{13} \\ y &= a_{21}s + a_{22}l + a_{23} \\ z &= a_{31}s + a_{32}l + a_{33} \end{aligned} \tag{11}$$

Then project from three-dimensional space (x, y, z) to another two-dimensional space (s', l').

$$\begin{aligned} s' &= \frac{x}{z} = \frac{a_{11}s+a_{12}l+a_{13}}{a_{31}s+a_{32}l+a_{33}} \\ l' &= \frac{y}{z} = \frac{a_{21}s+a_{22}l+a_{23}}{a_{31}s+a_{32}l+a_{33}} \end{aligned} \tag{12}$$

In the above formula, a_{33} is generally equal to 1. There is a total of 8 unknowns, which can be calculated through 4 pairs of points with the same name. Additionally, when $a_{31} = 0$, and $a_{32} = 0$, the perspective transformation degenerates into an affine transformation.

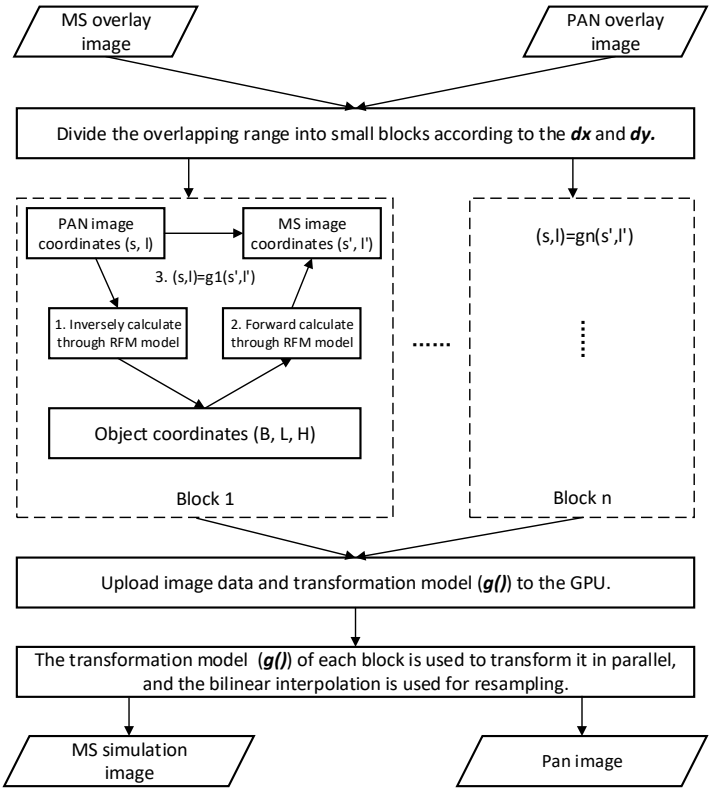


Figure 2. Differential geo-registration.

After the projection relationship is established, resampling is required. In this study, each block was transformed on the GPU according to its projection relationship in parallel. Interpolation methods for resampling include nearest neighbor, bilinear, and bicubic. Considering the efficiency and accuracy, we adopted the bilinear interpolation for resampling. Bilinear interpolation is performed using linear interpolation first in one direction, and then again in the other direction. Although each step is linear in terms of the sampled values and in the position, the interpolation as a whole is not linear but rather quadratic by the sample location. As shown in Equation (13), first interpolate in x-direction.

$$\begin{aligned} f(x,\, y_1) &\approx \frac{x_2-x}{x_2-x_q} f(Q_{11}) + \frac{x-x_1}{x_2-x_1} f(Q_{21}) \\ f(x,\, y_2) &\approx \frac{x_2-x}{x_2-x_q} f(Q_{12}) + \frac{x-x_1}{x_2-x_1} f(Q_{22}) \end{aligned} \tag{13}$$

In the above formula, $Q_{11} = (x_1, y_1)$, $Q_{12} = (x_1, y_2)$, $Q_{21} = (x_2, y_1)$, and $Q_{22} = (x_2, y_2)$, which are four points around the target point $p = (x, y)$. f is the value at the point.

Then it will be interpolated in y-direction to obtain the desired estimate.

$$f(x,\, y) \approx \frac{y_2 - y}{y_2 - y_1} f(x,\, y_1) + \frac{y - y_1}{y_2 - y_1} f(x,\, y_2) \tag{14}$$

PAN and MS images were normalized to the same direction and scale after differential geo-registration, which can be used directly for subsequent processing. However, there will be certain errors between them, such as the errors caused by tremor of the satellite platform [39,40] and random error by calibration [41,42], which needs to be further corrected.

2.3. Differential Rectification with Tiny Facet Primitive

PAN and MS images are taken almost at the same time, and they are also calibrated during preprocessing. In theory, the geometric relationship between the two images is strict. Its accuracy can reach the sub-pixel level after the differential geo-registration. However, during the actual processing, it was found that many data could not meet this standard, especially in the case of large roll, large pitch, and large terrain undulations. Therefore, this paper used differential rectification with tiny facet primitive to further correct the errors between PAN and MS images.

As can be seen in Figure 3, each geo-registered image was differentiated into many corresponding blocks. These blocks were uploaded to the GPU for parallel processing. The processing of each block included three steps: (1) matching to obtain the same name point; (2) calculating the transform model; (3) differential rectification with the transform model.

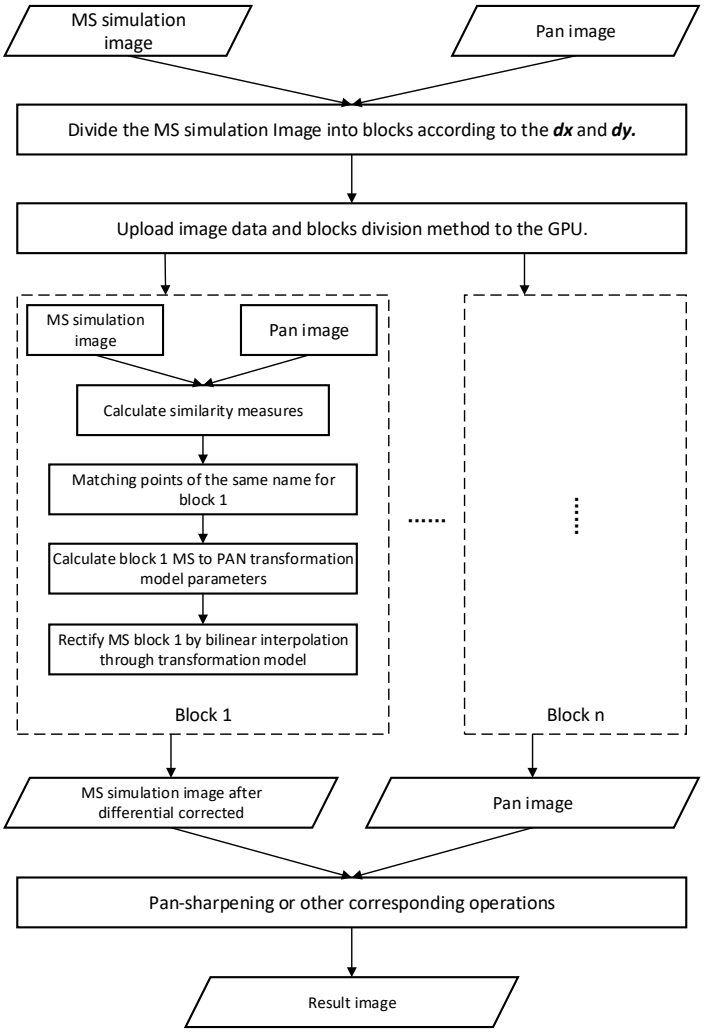


Figure 3. Differential rectification with tiny facet primitive.

Extracting points with the same name is based on template matching. This paper adopted the correlation coefficient as the method of similarity measurement in template matching. The correlation coefficient [50,51] is a standardized covariance function, which

can overcome the linear deformation of the gray level and maintain the invariance of the correlation coefficient. It is the most commonly used matching measure in image matching.

This paper assumed that the MS image was accurately registered, and only considered the misregistration between PAN and MS images. Therefore, we use the following approximate formula to use RGB to proportionally synthesize the pseudo-PAN image and PAN image to match the control points and rectify all MS bands with the same parameters.

$$\mathrm{PAN}_{\mathrm{pseudo}} = 0.299 * R + 0.587 * G + 0.144 * B \tag{15}$$

where $\mathrm{PAN}_{\mathrm{pseudo}}$ is defined as pseudo-PAN image.

For a certain point on the PAN image, a matching window of $m \times n$ is established with the point as the center, and a matching window of the same size and shape is established at the initial value point of the same name in the MS image, and the correlation coefficient is calculated as follows:

$$(c,r) = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} g_{i,j}g_{i+r,j+c} - \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} g_{i,j}\sum_{i=1}^{m}\sum_{j=1}^{n} g_{i+r,j+c}}{\sqrt{\left[\sum_{i=1}^{m}\sum_{j=1}^{n} {g_{i,j}}^2 - \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} g_{i,j}\right]^2}\left[\sum_{i=1}^{m}\sum_{j=1}^{n} {g_{i+r,j+c}}^2 - \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} g_{i+r,j+c}\right]^2} \tag{16}$$

where m,n represent the row and column size of the matching window; i,j represent the row number and column number in the matching window; $g_{i,j}$, $g_{i+r,j+c}$ represent the gray values in the PAN and MS image matching windows, respectively; c,r are the row and column of the MS image for the PAN image coordinate difference. For MS images that have undergone registration processing, take $c = r = 0$. For images that have not undergone band registration, the size of c,r is determined according to the camera design value.

After extracting the points with the same name, it is necessary to calculate the transform relationship between each block. The perspective transformation is also used at this stage; please refer to Section 2.2. Bilinear interpolation is also used in subsequent differential rectification; please refer to Section 2.2. Matching and rectification formulas were mapped to two different kernel functions for calculation.

2.4. Registration Quality Assessment Measures

This paper evaluated the effect of the algorithm from two angles: accuracy and efficiency. In terms of accuracy, the correlation coefficient was used to evaluate the registration accuracy of the algorithm. For the formula of the correlation coefficient, see Section 2.3. After the points with the same name were matched by the correlation coefficient, the root mean square error (RMSE) of these points were calculated as the final registration accuracy. For n correct checkpoints whose correct coefficient is greater than 0.9, the $\mathrm{RMSE}_{\mathrm{xy}}$ calculation formula is as follows:

$$\mathrm{RMSE}_{\mathrm{x}} = \sqrt{\frac{\sum_{i=1}^{n}\left(X_i' - X_i\right)^2}{n}}, \tag{17}$$

$$\mathrm{RMSE}_{\mathrm{y}} = \sqrt{\frac{\sum_{i=1}^{n}\left(Y_i' - Y_i\right)^2}{n}}, \tag{18}$$

$$\mathrm{RMSE}_{\mathrm{xy}} = \sqrt{\mathrm{RMSE}_{\mathrm{x}}^2 + \mathrm{RMSE}_{\mathrm{y}}^2}. \tag{19}$$

In the above formula, (X_i', Y_i'), $(X_i, Y_i)(i = 1, \ldots\ldots, n)$ are the image coordinates of the correct check point and its reference point. $\mathrm{RMSE}_{\mathrm{x}}$ and $\mathrm{RMSE}_{\mathrm{y}}$ are the root mean square error in x-direction and y-direction, respectively. In terms of efficiency, we used a pure CPU implementation method and a GPU method.

2.5. Pan-Sharpening Quality Assessment Measures

A pull resolution assessment infers the quality of the pansharpened image at the scale of the PAN image without resorting to a single reference image [16,52,53]. The quality with no reference (QNR) is one of the widely used full resolution assessments [16,52]. QNR consists of two parts: one is the spectral difference (D_λ), and the other is the spatial difference (D_s).

D_λ is calculated between the low-resolution MS image and the fused MS image. Calculate the universal image quality index (UIQI) [54] values between the two sets of bands at low resolution and high resolution. The difference in the corresponding UIQI values on the two scales produces the spectral distortion introduced by the pansharpening process.

D_s combines the UIQI values calculated between each MS band and PAN images downgraded to MS resolution, and between pan-sharpened MS and full-resolution PAN images. The absolute difference between the corresponding UIQI values averaged over all frequency bands produces D_s.

The effectiveness of the pansharpening method can be described from the two perspectives of spectrum and space. These two indicators can provide a unique quality index, referred to as QNR $\in [0, 1]$, with 1 being the best attainable value:

$$\mathrm{QNR} = (1 - D_\lambda)^\alpha (1 - D_S)^\beta \tag{20}$$

3. Results and Discussion

3.1. Data Introduction

The data used in this paper were from the GF-6 (GaoFen-6) and GF-7 (GaoFen-7) satellites, a series of high-resolution earth observation satellites planned by China's "China High-resolution Earth Observation System (CHEOS)"; ZY3-02 (ZiYuan3-02) of the National Land Satellite Remote Sensing Application Center of the Ministry of Natural Resources; and SV-1 (SuperView-1), China Aerospace Science and Technology Corporation's commercial remote sensing satellite. Among them, the resolution ratio of ZY3-02 PAN images to MS images is not the common 4 to 1, which allowed us to verify the versatility of the method in this paper. The selected images also include various topographies and landforms, such as mountains, plains, and coastal areas. The details of the data are shown in Table 1.

Table 1. The data.

Satellites	Parameters	Description		Satellites	Parameters	Description	
GF-6	Spectral range	PAN	0.45–0.90 μm	GF-7	Spectral range	PAN	0.45–0.90 μm
		MS	0.45–0.52 μm			MS	0.45–0.52 μm
			0.52–0.59 μm				0.52–0.59 μm
			0.63–0.69 μm				0.63–0.69 μm
			0.76–0.89 μm				0.77–0.89 μm
	Spatial resolution	PAN	2 m		Spatial resolution	PAN	0.8 m
		MS	8 m			MS	3.2 m
	landform	coastal			landform	plains	
ZY3-02	Spectral range	PAN	0.45–0.80 μm	SV-1	Spectral range	PAN	0.45–0.89 μm
		MS	0.45–0.52 μm			MS	0.45–0.52 μm
			0.52–0.59 μm				0.52–0.59 μm
			0.63–0.69 μm				0.63–0.69 μm
			0.77–0.89 μm				0.76–0.89 μm
	Spatial resolution	PAN	2.1 m		Spatial resolution	PAN	0.5 m
		MS	5.8 m			MS	2 m
	landform	mountains			landform	hills	

3.2. Precision Analysis

In order to verify the accuracy of the algorithm, we first registered the PAN and MS images described above. Figure 4 shows the registration results of these four pairs of images. It can be seen in the figure that the PAN and MS images were normalized to the same resolution and direction, and differential corrections were also performed. It is difficult to see the deviations visually, as this time the PAN and MS images were strictly registered. The corresponding details can be seen in Figure 5. Columns a and dare PAN images, columns b and e are MS images, and columns c and f are PAN and MS contrast images. Lines 1 and 2 are GF-6, lines 3 and 4 are GF-7, lines 5 and 6 are ZY3-02, and lines 7 and 8 are SV-1. These detailed images further demonstrate the registration accuracy of the method in this paper. Both PAN and MS images were registered. The detailed pictures were evenly cut out from each group of pictures, which shows that the overall deviation distribution is basically the same. For the ZY3-03 and SV-1 images, the accuracy of the registration results of the mountainous terrain with large undulations was also very high.

In this study, RMSE was used to quantitatively analyze the registration results and further analyze the accuracy of the registration method. As shown in Table 2, we compared the accuracy of the registration performed using different algorithms. SIFT with affine transformation was registered by affine transformation with matched SIFT points. The *Level 2* MS image was up-sampled to the scale of the *Level 2* PAN image. The *Level 1A* MS image was up-sampled to the scale of the *Level 1A* PAN image. The *Level 1A* image, after differential geo-registration, was registered by differential geo-registration described in Section 2.2. It can be seen from the table that among the four different geomorphological areas captured by the four satellites, the method in this paper had the best accuracy, and the accuracy was kept within 0.5 pixels. The registration with affine transformation by matched SIFT points cannot eliminate the relative geometric distortion of PAN and MS images. Therefore, the accuracy of registration using that method is closely related to internal distortion. The accuracy of the ZY3-02 image with a large amount of internal distortion was poor, and there were still residual errors in the registered images. The actual remote sensing image products cannot completely eliminate internal distortion. In addition, the method of registration by extracting feature points is not suitable for many images, because it is difficult to match the feature points in these cases, such as large differences in radiation response characteristics, severe cloud occlusion, and large differences in resolution. As shown in Figures A1 and A2 in Appendix A, the hyperspectral image of ZY1E (ZiYuan 1E—spatial resolution is 30 m) and the panchromatic image of GF-7 (Spatial resolution is 0.8 m) cannot match the SIFT feature points. However, the method in this paper could complete the rapid registration as long as the geometric measurements were accurate, which also shows its potential for image registration using images from different sources.

The directly enlarged *Level 1A* image is a result of a subset of the differential geo-registration, which is a special mode, and its accuracy will be lower than that of the differential geo-registered image. The accuracy of the differential geographic corrections of *Level 1A* image was better than that of the *Level 2* image, as each block was strictly geographically aligned when differential geographic correction. This accuracy depends on the stability of the satellite and the accuracy of camera calibration. Therefore, differential rectification with tiny facet primitive was required for further correction.

Level 2 PAN images and MS images which were enlarged based on geographic location were the least accurate. PAN and MS images are generally corrected separately in blocks, and the corresponding block is usually not consistent with the actual size, which results in the corrected image not strictly corresponding with the former, causing errors. After the high-precision registration in this study, the next step of pansharpening was carried out.

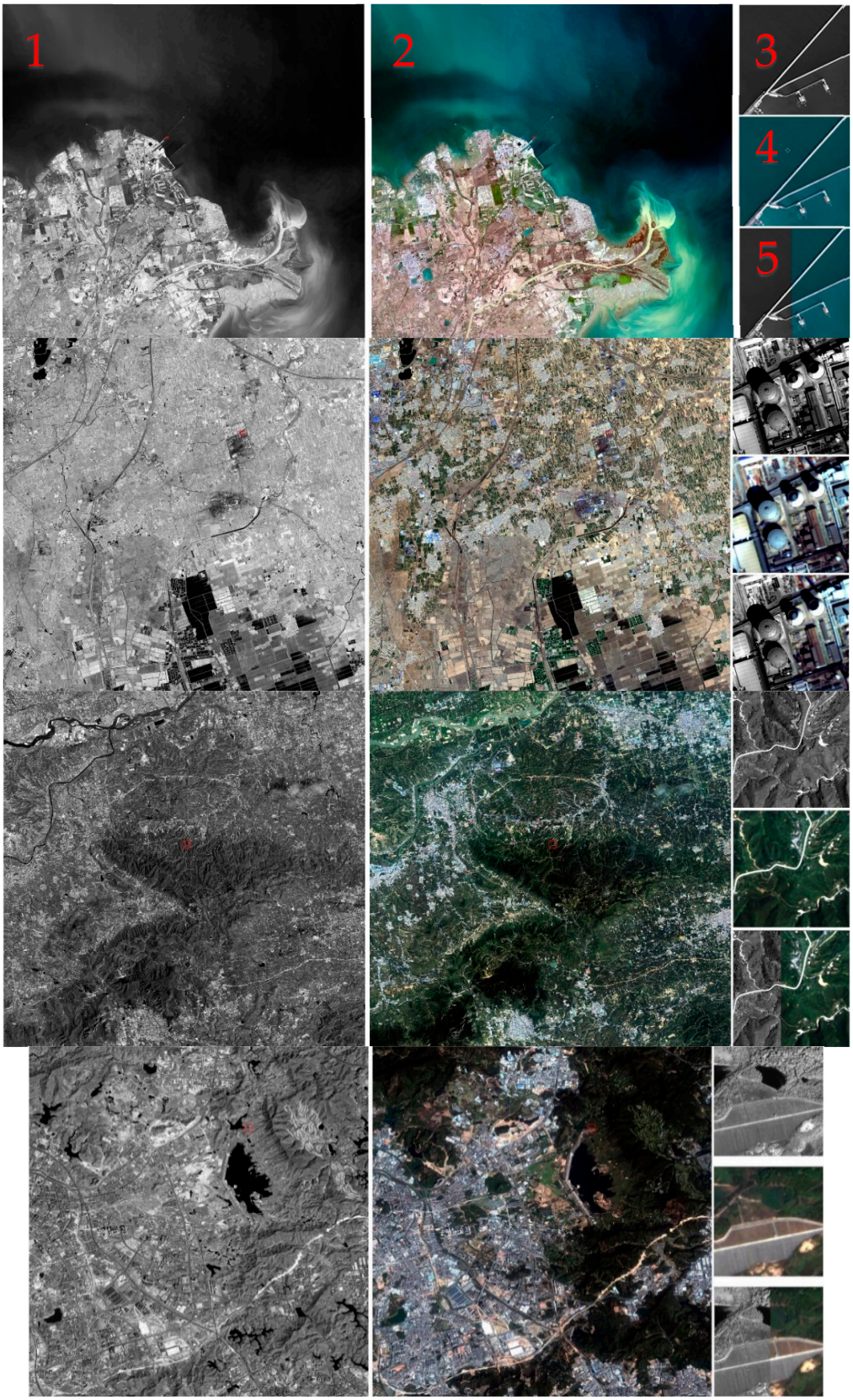


Figure 4. Registered images. (**a**) GF-6, (**b**) GF-7, (**c**) ZY3-02, (**d**) SV-1. (1) is a PAN image, (2) is an MS image up-sampled to PAN scale, (3) shows the details of (1), (4) shows the details of (2), and (5) is a comparison of details. Same in (**b**–**d**).

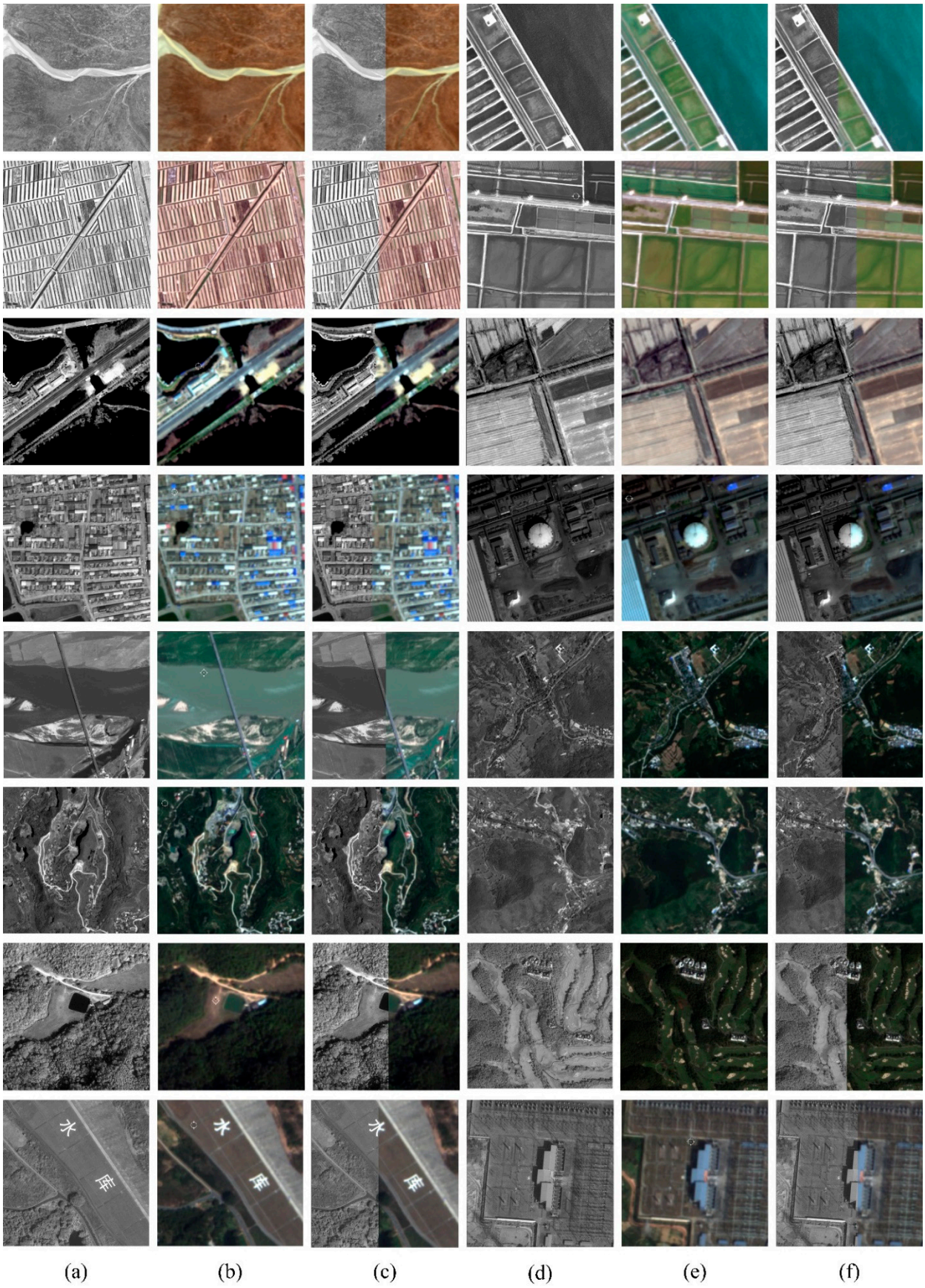
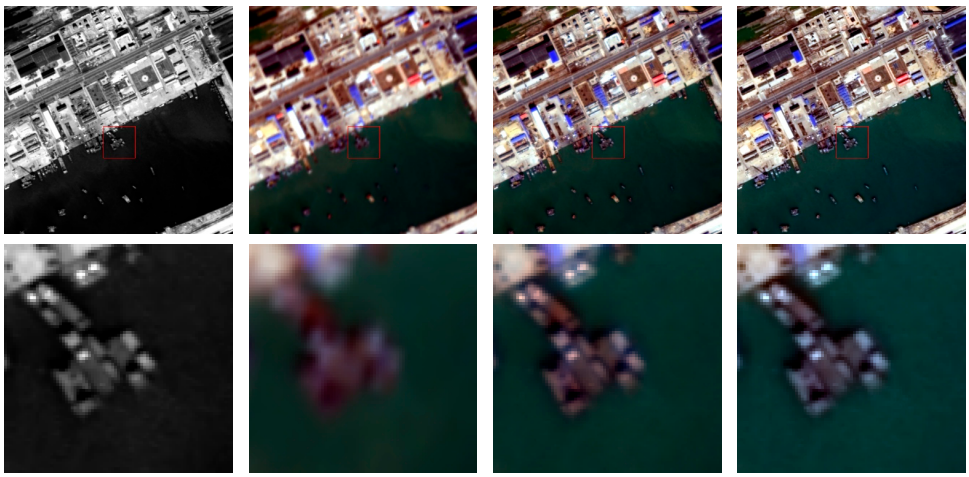


Figure 5. A detailed comparison of registered images. (**a**,**d**) PAN; (**b**,**e**) MS; (**c**,**f**) comparison. Lines 1 and 2 are GF-6, lines 3 and 4 are GF-7, lines 5 and 6 are ZY3-02, and lines 7 and 8 are SV-1.

Table 2. Registration accuracy.

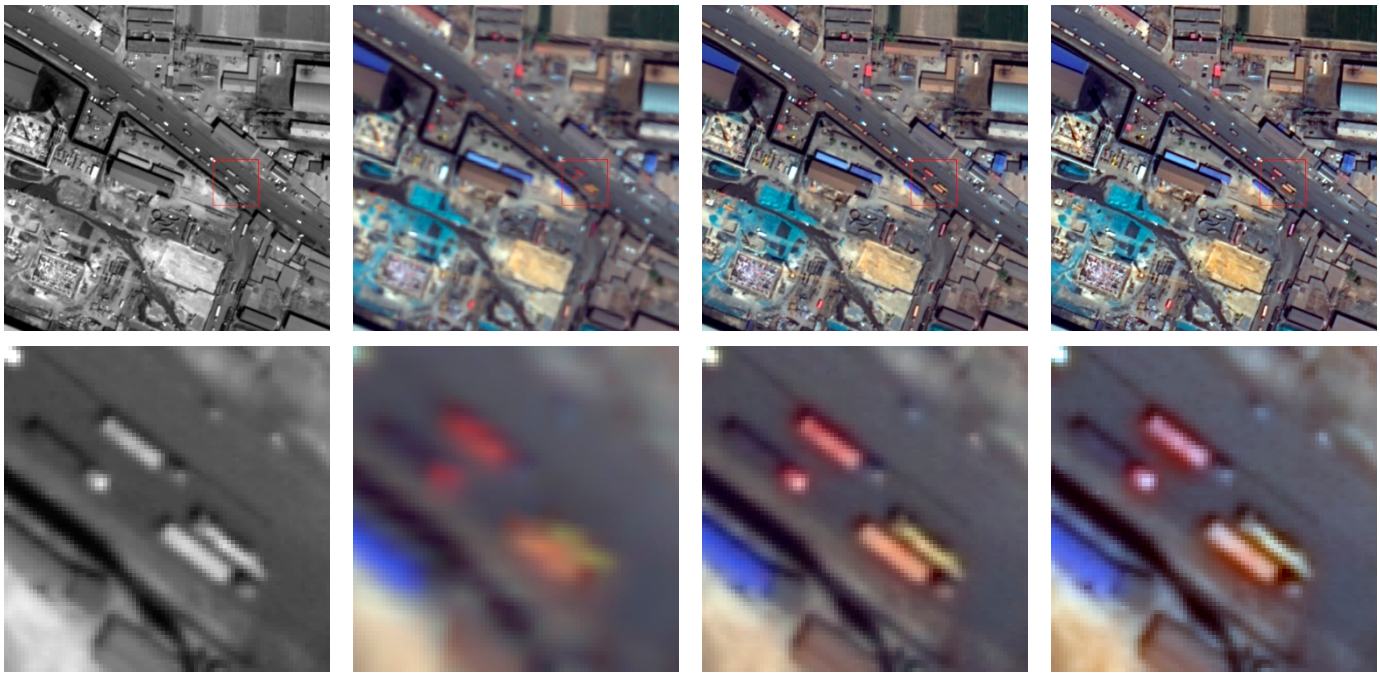
Satellites	Methods	$RMSE_x$/Pixel	$RMSE_y$/Pixel	$RMSE_{xy}$/Pixel
GF-6	SIFT with affine transformation	0.6134	0.5867	0.8488
	Level 2 image after enlarged	0.8047	1.7257	1.9041
	Level 1A image after enlarged	1.7595	1.6087	2.3841
	Level 1A image after differential geo-registration	*0.5510*	*0.5362*	*0.7688*
	Method of this paper	***0.1945***	***0.1761***	***0.2623***
GF-7	SIFT with affine transformation	0.4774	0.5379	0.7192
	Level 2 image after enlarged	0.5743	1.0138	1.1652
	Level 1A image after enlarged	*0.7035*	3.3993	3.4714
	Level 1A image after differential geo-registration	0.8669	*0.4634*	*0.9829*
	Method of this paper	***0.1863***	***0.2216***	***0.2895***
ZY3-02	SIFT with affine transformation	2.1153	0.9087	2.3022
	Level 2 image after enlarged	2.9053	1.4013	3.2256
	Level 1A image after enlarged	2.5494	3.4269	4.2712
	Level 1A image after differential geo-registration	*2.5064*	*1.8844*	*3.1358*
	Method of this paper	***0.1229***	***0.1096***	***0.1646***
SV-1	SIFT with affine transformation	0.6094	0.6687	0.9047
	Level 2 image after enlarged	2.0817	1.4188	2.5192
	Level 1A image after enlarged	*0.8166*	0.7857	1.1332
	Level 1A image after differential geo-registration	0.8531	*0.6298*	*1.0604*
	Method of this paper	***0.3140***	***0.2781***	***0.4195***

As shown in Figures 6–9, we used the Gram–Schmidt [14] method to fuse the images registered by the method in this paper and the *Level 2* enlarged image. The Gram–Schmidt method is sensitive to the registration accuracy. When the registration accuracy is poor, color deviation and ghosting will occur. It can be seen from the figures that the accurately registered images maintained good fusion effects in different scenes and different satellite images. The GS method exploits the intensity component as the first vector of the new orthogonal basis. The orthogonalization processes the MS vector. Then it finds its projection on the hyper-plane defined by the previously found orthogonal vectors and its orthogonal component such that the sum of the orthogonal and projection components is equal to the zero-mean version of the original vectorized band. Pansharpening is completed by substituting each intensity component with a histogram-matched P before performing the inverse transformation.



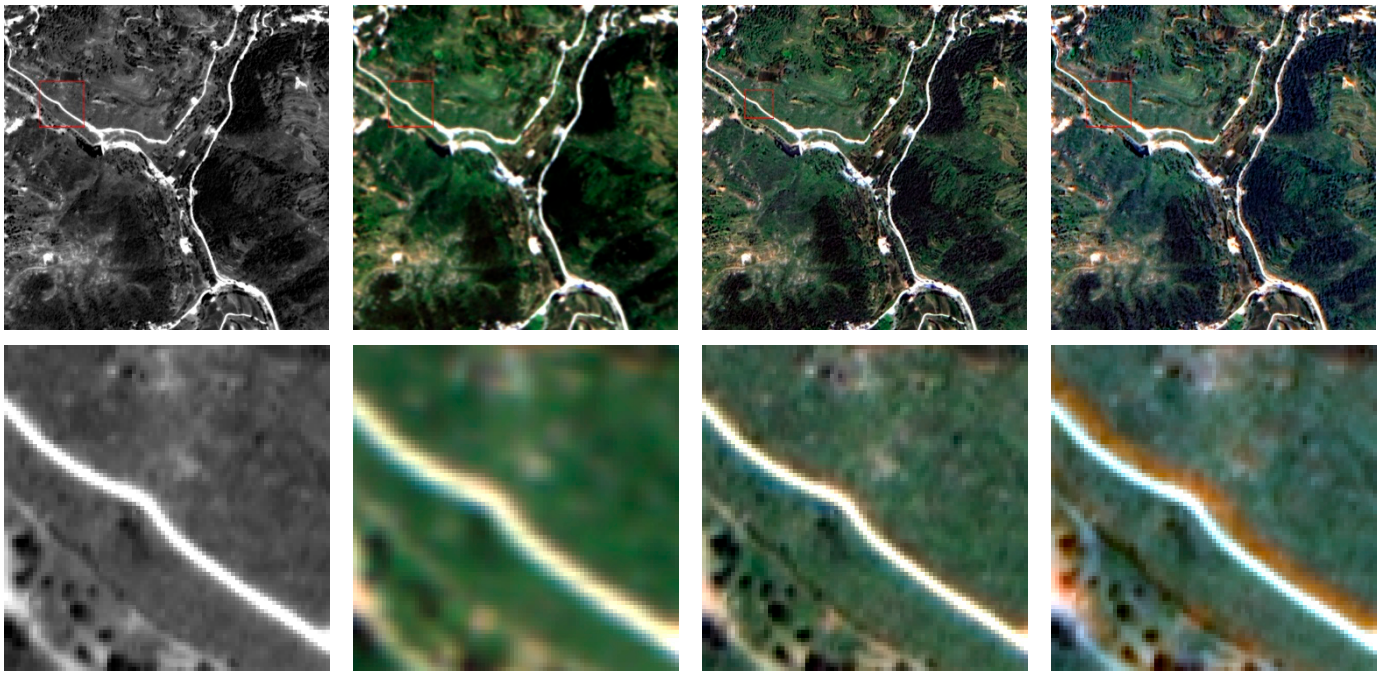
(**a**) GF-6 PAN (**b**) GF-6 MS (**c**) GF-6 Fusion by the method in this paper (**d**) GF-6 Fusion by *Level 2* enlarged image

Figure 6. Comparison of the GF-6 fusion image created using the method in this paper and the fusion image created using the *Level 2* PAN image and the MS enlarged image. (**a**) PAN image; (**b**) enlarged MS image; (**c**) fusion image created using the method in this paper; (**d**) fusion image created using the enlarged *Level 2* image.



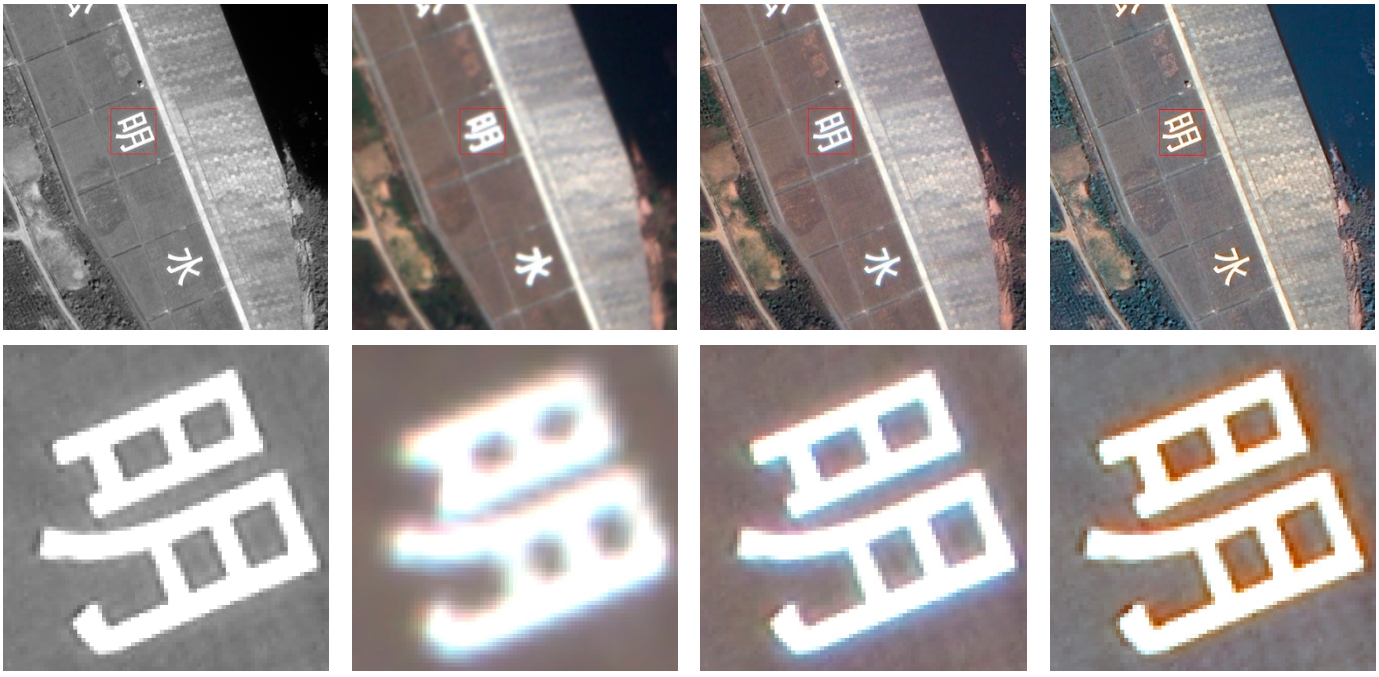
(**a**) GF-7 PAN (**b**) GF-7 MS (**c**) GF-7 Fusion by the method in this paper (**d**) GF-7 Fusion by *Level 2* enlarged image

Figure 7. Comparison of the GF-7 fusion image created using the method in this paper and the fusion image created using the *Level 2* PAN and MS enlarged image. (**a**) PAN image; (**b**) enlarged MS image; (**c**) fusion image created using the method in this paper; (**d**) fusion image created using the enlarged *Level 2* image.



(**a**) ZY3-02 PAN (**b**) ZY3-02 MS (**c**) ZY3-02 Fusion by the method in this paper (**d**) ZY3-02 Fusion by *Level 2* enlarged image

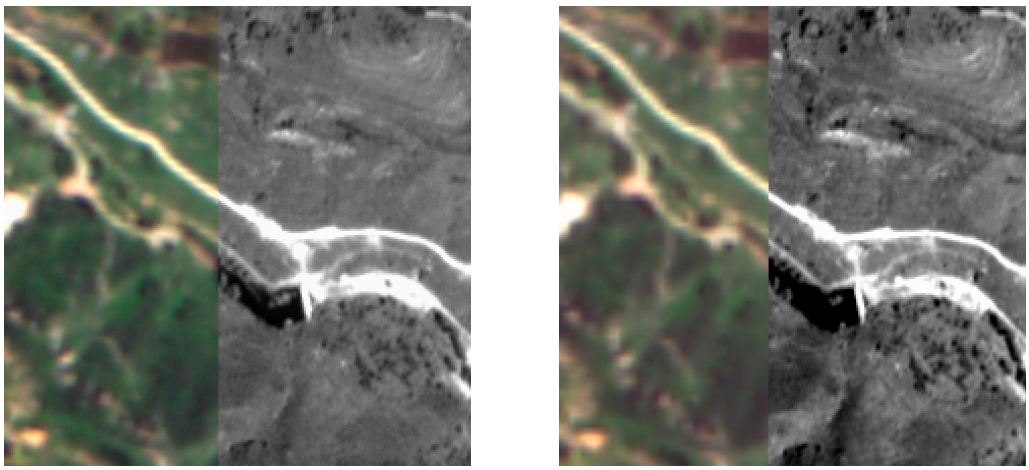
Figure 8. Comparison of the ZY3-02 fusion image created using the method in this paper and the fusion image created using the *Level 2* PAN and MS enlarged image. (**a**) PAN image; (**b**) enlarged MS image; (**c**) fusion image created using the method in this paper; (**d**) fusion image created using the enlarged *Level 2* image.



(**a**) SV-1 PAN (**b**) SV-1 MS (**c**) SV-1 Fusion by the method in this paper (**d**) SV-1 Fusion by *Level 2* enlarged image

Figure 9. Comparison of the SV-1 fusion image created using the method in this paper and the fusion image created using the *Level 2* PAN and MS enlarged image. (**a**) PAN image; (**b**) enlarged MS image; (**c**) fusion image created using the method in this paper; (**d**) fusion image created using the enlarged *Level 2* image.

Take ZY3-02 with obvious ghost phenomenon in Figure 8. As an example to show the effect of the algorithm in this paper. As shown in Figure 10, the *Level 2* PAN and up-sampled MS images are not fully registered, which also leads to obvious ghosting and color cast in the fusion result. The algorithm in this paper made the PAN and MS images basically completely registered, the effect of fusion was clear and no ghosting, and the color was maintained well.



(**a**) (**b**)

Figure 10. Comparison of the ZY3-02 registered image created using the method in this paper and the *Level 2* PAN and MS up-sampled image. (**a**) PAN and MS images registered using the algorithm in this paper; (**b**) *Level 2* PAN and MS up-sampled image.

The up-sampled *Level 2* images have particularly obvious color deviations and ghosting in the cases of ZY3-02 and SV-1, which is also consistent with the *Level 2* image error measured in the previous part of this paper. The *Level 2* image errors of GF-6 and GF-7 are relatively small, and the ghost phenomenon is not obvious, but there are serious color deviations.

As shown in Table 3, we used the QNR index at full resolution to further quantitatively demonstrate our method's superiority. It can be seen from the table that the algorithm in this paper could effectively improve the spectral and spatial consistency when other conditions were the same. The registration result will greatly affect the fusion effect. Therefore, the precise registration of PAN and MS images is of great significance for pansharpening.

Table 3. Global distortion/quality index integrated with fused data.

Satellites	Methods	D_λ (p = 1)	D_s (q = 1)	QNR (α = β = 1)
GF-6	Fusion by *Level 2* enlarged image	0.1178	0.1248	0.7721
	Fusion by method in this paper	***0.0234***	***0.0311***	***0.9463***
GF-7	Fusion by *Level 2* enlarged image	0.0587	0.0817	0.8645
	Fusion by method in this paper	***0.0433***	***0.0523***	***0.9067***
ZY3-02	Fusion by *Level 2* enlarged image	0.1392	0.2487	0.6467
	Fusion by method in this paper	***0.0393***	***0.0682***	***0.8951***
SV-1	Fusion by *Level 2* enlarged image	0.2684	0.305	0.5085
	Fusion by method in this paper	***0.1114***	***0.1271***	***0.7757***

In summary, the method in this paper could achieve high-precision registration of PAN and MS images, which is suitable for different scenes and different sensors. It is extremely robust to errors between PAN and MS images.

3.3. Efficiency Analysis

The CPU of the server used for efficiency analysis was an Intel Core i7-6700K @ 4.00 GHz, the GPU was an NVIDIA GeForce RTX 2080Ti, and the operating system was Ubuntu 18.04.4 LTS. In addition, we had 64 GB of RAM and a 256 GB high-speed Nvme SSD. The processing time discussed in this article does not include the time for reading and writing data. After the data were read in, they were simulated as the data flow in from the previous step. The specific processing time of each algorithm was analyzed through the log. The log printed out the time taken for each key step, accurate to the millisecond level.

It can be seen from Table 4 that the speedup ratio of the computational concepts which was mapped to the GPU was several hundred fold. However, the overall efficiency improvement was dozens of fold, because GPU processing requires additional processing, such as the CPU/GPU copying each other, which takes a long time and slows down the overall efficiency. In general, the processing time increased from ten minutes to seconds. The work that was completed in ten minutes can now be completed in tens of seconds. This is a great help to the completion of near real-time processing work. As can be seen in Figure 11, the processing efficiency of this method is almost proportional to the amount of data.

The reason why the improvement is so obvious is that the parallelism of the algorithms mapped to the GPU is very high. In differential geo-registration, the main time consuming factor is that each pixel is returned, and the pixels do not affect each other and can be handled simultaneously. In differential rectification, what is time consuming is the control point matching and block correction. Control points were obtained by calculating correlation coefficients. Not only can all control points be calculated at the same time, but each point can also have correlation coefficients calculated in parallel. The block rectification can also calculate all pixels at the same time, with high parallelism.

Take GF-7 as an example. Its PAN camera has a width of approximately 35,660 pixels, and it images approximately 7200 lines per second. The data volume per second is 490 MB.

The MS camera has a width of approximately 8915 pixels and images approximately 1800 lines per second. The amount of data per second is 122 MB. Therefore, for GF-7, the amount of PAN and MS data generated per second is about 612 MB; FUS (fusion) data totals about 1960 MB. It can be seen from Figure 11 that the algorithm in this paper ran almost as fast as said data could be collected, according to the conversion of the previous processing efficiency. The processing of one second of GF-7 data takes about 5.09 s, which is close to near real-time processing.

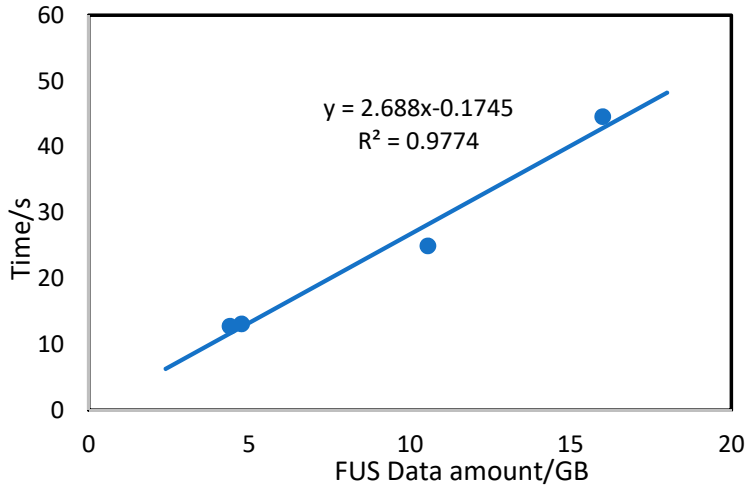


Figure 11. Time consumption with different amounts of FUS data.

Table 4. Efficiency analysis.

Satellites	Data Amount/GB	Processing Steps	CPU Time/s	CPU GPU Time/s	Speedup Ratio
GF-6	PAN:4.00 MS:1.00 FUS:16.00	Geo-registration	94.384	***0.214***	441.047
		Matching	891.895	***2.310***	386.102
		Rectification	141.609	***0.187***	757.270
		Other steps	***22.159***	41.848	0.530
		All times	1150.048	***44.559***	25.810
GF-7	PAN:2.64 MS:0.61 FUS:10.56	Geo-registration	59.747	***0.103***	578.944
		Matching	420.249	***1.336***	314.511
		Rectification	92.989	***0.108***	862.608
		Other steps	***15.535***	23.386	0.664
		All times	588.521	***24.933***	23.604
ZY3-02	PAN:1.10 MS:0.59 FUS:4.40	Geo-registration	25.913	***0.049***	528.837
		Matching	277.300	***0.630***	440.159
		Rectification	38.655	***0.051***	757.948
		Other steps	***6.319***	12.012	0.526
		All times	348.187	***12.742***	27.326
SV-1	PAN:1.19 MS:0.30 FUS:4.76	Geo-registration	30.307	***0.052***	582.827
		Matching	289.980	***0.662***	438.036
		Rectification	41.809	***0.055***	760.170
		Other steps	***6.862***	12.313	0.557
		All times	368.958	***13.082***	28.203

3.4. Discussion

In the method in this paper, the resampling method and the transformation model in block correction have large impacts on the results. The resampling method uses bilinear sampling. The transformation model for block correction uses perspective transformation. Here we discuss the reasons for choosing the bilinear resampling method and perspective transformation. The data used for this were the data introduced in Section 3.1. The analysis

of the transformation model used the four groups of PAN and MS images. The analysis of the resampling method used the GF-7 MS image.

The transformation model in block correction was applied in differential geo-registration efforts and differential rectification with tiny facet primitive. In order to illustrate the accuracy of the block transformation model in block correction, we tested affine transformation, linear transformation, and perspective transformation. It can be seen in Table 5 that the accuracy of perspective transformation is significantly better than those of affine transformation and linear transformation. The linear transformation only needs two points to fit, the affine transformation needs three points to fit, and the perspective transformation needs four points to fit. The four points of the perspective transformation correspond to the four corner points of a block, making it suitable for block correction. In the GF-6 image with relatively flat terrain (the Yellow River estuary), the difference in $RMSE_{xy}$ between these three models was not large, about 0.002. However, the $RMSE_{xy}$ of SV-1, which has large terrain undulations, was about 0.04. This shows that the more obvious is the terrain undulation, the more obvious the advantage of perspective transformation, a linear transformation or affine transformation can be used for flat areas. The areas with large terrain undulations require higher-precision perspective transformations.

After differential geographic registration, the image is normalized to the same scale and direction, and different interpolation algorithms have an impact on the similarity measure in the subsequent matching. We designed an experiment to demonstrate this effect. The experimental steps were as follows:

1. Select a scene image and down-sample it to different scales in a Gaussian pyramid.
2. Up-sample the down-sampling results to the original image size by nearest, bilinear, and bicubic methods.
3. Analyze the matching results of the grid points of different up-sampling methods. The correlation coefficient of the correct matching point is 0.9.
4. Analyze the peak signal-to-noise ratios (PSNRs) of the up-sampling methods. The higher the PSNR, the better.

Table 5. The impact of the transformation model for block correction on accuracy.

Satellites	Methods	$RMSE_x$/Pixel	$RMSE_y$/Pixel	$RMSE_{xy}$/Pixel
GF-6	affine transformation	0.5847	0.5137	0.7783
	linear transformation	0.5715	0.5165	0.7703
	perspective transformation	**0.5510**	0.5362	**0.7688**
GF-7	affine transformation	0.9077	0.4761	1.0250
	linear transformation	0.8891	0.4777	1.0093
	perspective transformation	**0.8669**	**0.4634**	**0.9830**
ZY3-02	affine transformation	2.5293	2.0146	3.2336
	linear transformation	**2.4338**	2.0050	3.1533
	perspective transformation	2.5064	**1.8844**	**3.1358**
SV-1	affine transformation	0.9126	0.6260	1.1067
	linear transformation	0.9212	0.6218	1.1114
	perspective transformation	**0.8531**	0.6298	**1.0604**

It can be seen in Table 6 that among the different sampling methods, bicubic had the best effect, followed by bilinear, and nearest was the worst. When the resolution difference reaches 16 times, the correlation coefficient is basically invalid. For the method in this paper, when the resolution difference is too large, the differential geographic registration is still applicable, but the differential correction based on the correlation coefficient is invalid. As shown in Figures A1 and A2 in Appendix A, it degenerated into differential geographic

registration, as differential geographic registration does not depend on the relationship between images.

Table 6. The impacts of the resample methods.

Ratio	Methods	Time	Matched/All Points	$RMSE_x$/Pixel	$RMSE_y$/Pixel	$RMSE_{xy}$/Pixel	PSNR
1	/	/	4896/4896	0	0	0	/
4	nearest	0.1342	591/4896	0.196568	0.187016	0.271319	29.4506
	bilinear	0.1478	658/4896	0.172961	0.183635	0.252265	29.4697
	bicubic	0.1778	846/4896	0.143572	0.151176	0.208487	29.7205
8	nearest	0.1199	68/4896	0.646531	0.573771	0.864416	27.6270
	bilinear	0.1216	87/4896	0.651209	0.659146	0.926578	27.6492
	bicubic	0.1408	109/4896	0.587586	0.587586	0.786102	27.8021
16	nearest	0.1328	5/4896	2.27818	1.20963	2.5794	26.4194
	bilinear	0.1445	6/4896	1.36651	1.48564	2.01853	26.4397
	bicubic	0.1363	6/4896	1.4527	2.54786	2.9329	26.5472
32	nearest	0.1235	0/4896	/	/	/	25.5210
	bilinear	0.1019	1/4896	7.49833	6.39505	9.85503	25.5403
	bicubic	0.1439	4/4896	3.73353	5.13013	6.34488	25.6268
64	nearest	0.1197	0/4896	/	/	/	24.7872
	bilinear	0.1092	0/4896	/	/	/	24.8015
	bicubic	0.1333	0/4896	/	/	/	24.8743

Unlike the method based on feature point matching, the method in this paper does not depend on the matching accuracy of feature points, and is also suitable for images with large radiation differences, cloud occlusion, and large-scale differences. However, the method in this paper relies on the initial geometric accuracy and requires accurate geometric calibration. The method in this paper needs to be carried out on the basis of the preprocessing of the originally captured image, and depends on the accuracy of the matching points.

4. Conclusions

We proposed a near-real-time PAN and MS sub-pixel registration method. Firstly, differential geo-registration is used to perform accurate geographical registration of PAN and MS images, which can normalize PAN and MS images to the same direction and scale. Compared with the direct enlargement of *Level 2* images, this method has higher registration accuracy. At this time, the main factors of the image registration accuracy are the relative calibration error of PAN and MS images, the error caused by the camera's high-frequency tremor, etc. These errors are difficult to eliminate on the object side. Therefore, this paper used sub-pixel differential rectification with tiny facet primitive on the image side to eliminate these potential errors. Through sub-pixel tiny facet element differential correction, the registration accuracy between PAN and MS images was basically stabilized within 0.5 pixels, so that PAN and MS images could be accurately registered.

Although the method in this paper could effectively improve the registration accuracy of PAN and MS images, the computational burden was very large—especially for the calculations of links, such as geographic registration, matching, and correction. Therefore, this article mapped these computationally large and highly parallel links to a GPU for processing. The processing efficiency of these links on the GPU was hundreds of times better that the processing efficiency on the CPU. Counting the extra operations brought about by GPU processing, the processing time of the GPU relative to the CPU was still multiples shorter. The processing time decreased from tens of minutes to tens of seconds. Additionally, taking GF-7 as an example, it took about 5.09 s to process the amount of data captured by GF-7 per second, which basically realizes near real-time processing.

In general, the algorithm in this paper could not only provide high-precision registration for PAN and MS fusion, but it has high processing efficiency and can basically achieve near real-time processing. In addition it is suitable for different scenes and different sensors. It is extremely robust to registration errors between PAN and MS images. It also has the potential for high-precision registration of sensors from different sources.

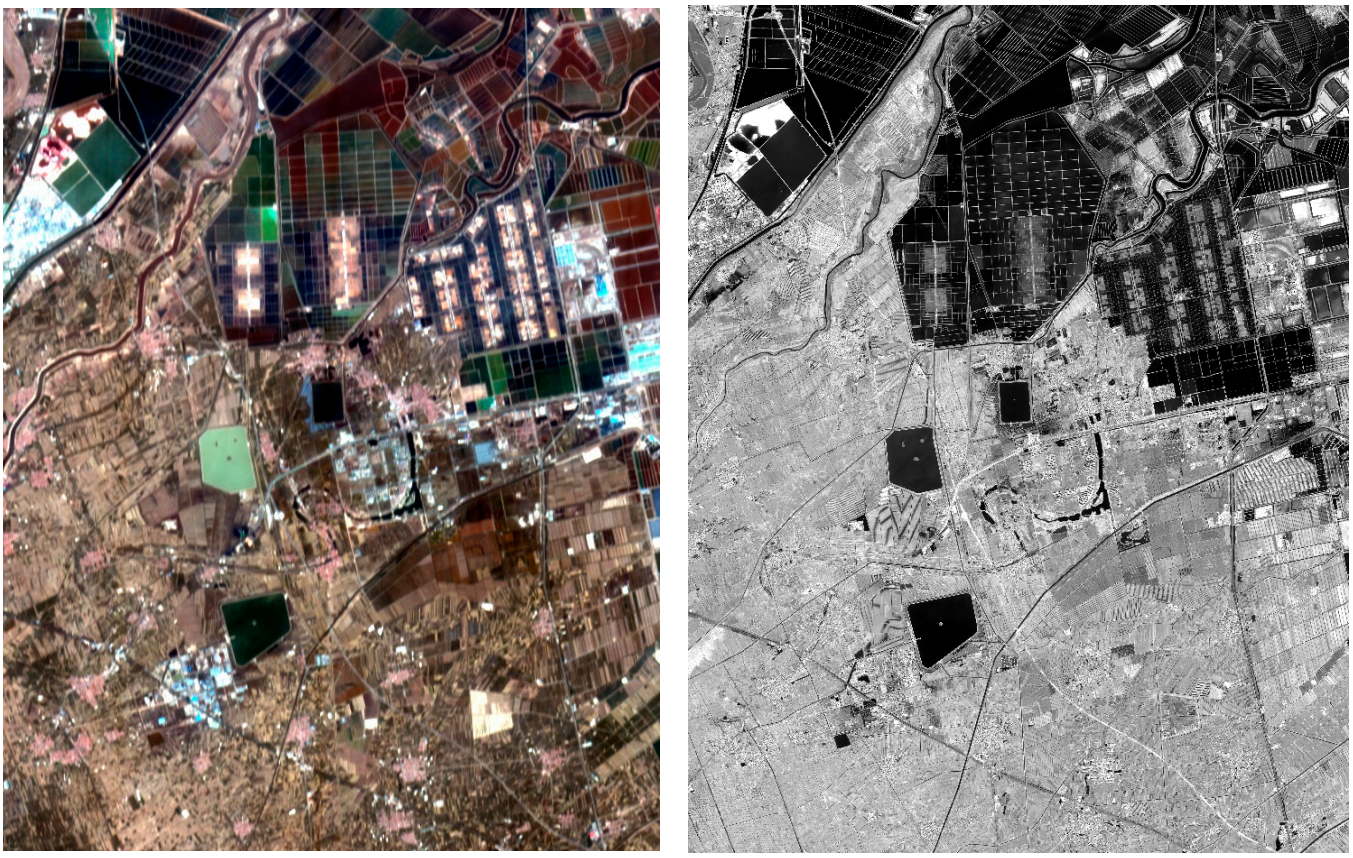
Author Contributions: Conceptualization, G.X. and M.W.; data curation, G.X.; formal analysis, G.X.; funding acquisition, M.W. and Z.Z.; investigation, S.X.; methodology, G.X., Z.Z. and L.H.; project administration, M.W.; resources, M.W.; software, G.X. and L.H.; supervision, M.W.; validation, S.X. and L.H.; visualization, G.X.; writing—original draft, G.X.; writing—review and editing, M.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China under projects 61825103, 91838303, 61901307, 91638301, and 91738302; the Open research fund of the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 20E01; the Key Research and Development Plan Project of Hubei Province under Grant 2020BIB006; and the Key Project of Hubei Provincial Natural Science Foundation under Grant 2020CFA001.

Acknowledgments: The authors would like to thank the China Centre for Resources Satellite Data and Application for providing the data used in this paper.

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

Appendix A



(**a**) (**b**)

Figure A1. ZY1E hyperspectral image (spatial resolution is 30 m) and GF-7 panchromatic image (spatial resolution is 0.8 m) registered by the method in this paper, which have a resolution difference of about 37.5 times. (**a**) A hyperspectral image of ZY1E. (**b**) A panchromatic image of GF-7.



Figure A2. The registration details of the ZY1E hyperspectral image and the GF-7 panchromatic image.

References

1. Dadrass Javan, F.; Samadzadegan, F.; Mehravar, S.; Toosi, A.; Khatami, R.; Stein, A. A Review of Image Fusion Techniques for Pan-Sharpening of High-Resolution Satellite Imagery. *ISPRS J. Photogramm. Remote Sens.* **2021**, *171*, 101–117. [CrossRef]
2. DadrasJavan, F.; Samadzadegan, F. An Object-Level Strategy for Pan-Sharpening Quality Assessment of High-Resolution Satellite Imagery. *Adv. Space Res.* **2014**, *54*, 2286–2295. [CrossRef]
3. Loncan, L.; De Almeida, L.B.; Bioucas-Dias, J.M.; Briottet, X.; Chanussot, J.; Dobigeon, N.; Fabre, S.; Liao, W.; Licciardi, G.A.; Simoes, M.; et al. Hyperspectral Pansharpening: A Review. *IEEE Geosci. Remote Sens. Mag.* **2015**, *3*, 27–46. [CrossRef]
4. Hasanlou, M.; Saradjian, M.R. Quality Assessment of Pan-Sharpening Methods in High-Resolution Satellite Images Using Radiometric and Geometric Index. *Arab. J. Geosci.* **2016**, *9*, 45. [CrossRef]
5. Haydn, R.; Dalke, G.W.; Henkel, J.; Bare, J. Application of the IHS Color Transform to the Processing of Multisensor Data and Image Enhancement. In Proceedings of the International Symposium on Remote Sensing of Arid and Semi-Arid Lands, Cairo, Egypt, 19–25 January 1982.
6. Kwarteng, P.; Chavez, A. Extracting Spectral Contrast in Landsat Thematic Mapper Image Data Using Selective Principal Component Analysis. *Photogramm. Eng. Remote Sens.* **1989**, *55*, 339–348.
7. Schowengerdt, R.A. Reconstruction of Multispatial, Multispectral Image Data Using Spatial Frequency Content. *Photogramm. Eng. Remote Sens.* **1980**, *46*, 1325–1334.
8. Chavez, P.; Sides, S.C.; Anderson, J.A. Comparison of Three Different Methods to Merge Multiresolution and Multispectral Data: Landsat TM and SPOT Panchromatic. *Photogramm. Eng. Remote Sens.* **1991**, *57*, 295–303.
9. Liu, J. Smoothing Filter-Based Intensity Modulation: A Spectral Preserve Image Fusion Technique for Improving Spatial Details. *Int. J. Remote Sens.* **2000**, *21*, 3461–3472. [CrossRef]
10. Wald, L.; Ranchin, T. Liu'Smoothing Filter-Based Intensity Modulation: A Spectral Preserve Image Fusion Technique for Improving Spatial Details'. *Int. J. Remote Sens.* **2002**, *23*, 593–597. [CrossRef]
11. Jiang, D.; Zhuang, D.; Huang, Y.; Fu, J. Survey of Multispectral Image Fusion Techniques in Remote Sensing Applications. *Image Fusion Its Appl.* **2011**, 1–23. [CrossRef]
12. Mandhare, R.A.; Upadhyay, P.; Gupta, S. Pixel-Level Image Fusion Using Brovey Transforme and Wavelet Transform. *Int. J. Adv. Res. Electr. Electron. Instrum. Eng.* **2013**, *2*, 2690–2695.
13. Hallabia, H.; Hamam, H.; Ben Hamida, A. A Context-Driven Pansharpening Method Using Superpixel Based Texture Analysis. *Int. J. Image Data Fusion* **2021**, *12*, 1–22. [CrossRef]
14. Klonus, S.; Ehlers, M. Performance of Evaluation Methods in Image Fusion. In Proceedings of the 2009 12th International Conference on Information Fusion, Seattle, WA, USA, 6–9 July 2009; IEEE: New York, NY, USA, 2009; pp. 1409–1416.
15. Vivone, G.; Dalla Mura, M.; Garzelli, A.; Pacifici, F. A Benchmarking Protocol for Pansharpening: Dataset, Preprocessing, and Quality Assessment. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 6102–6118. [CrossRef]
16. Vivone, G.; Dalla Mura, M.; Garzelli, A.; Restaino, R.; Scarpa, G.; Ulfarsson, M.O.; Alparone, L.; Chanussot, J. A New Benchmark Based on Recent Advances in Multispectral Pansharpening. *IEEE Geosci. Remote Sens. Mag.* **2021**, *9*, 53–81. [CrossRef]
17. Benzenati, T.; Kallel, A.; Kessentini, Y. Two Stages Pan-Sharpening Details Injection Approach Based on Very Deep Residual Networks. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 4984–4992. [CrossRef]
18. Zhang, Y.; Liu, C.; Sun, M.; Ou, Y. Pan-Sharpening Using an Efficient Bidirectional Pyramid Network. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 5549–5563. [CrossRef]
19. Dong, W.; Yang, Y.; Qu, J.; Xie, W.; Li, Y. Fusion of Hyperspectral and Panchromatic Images Using Generative Adversarial Network and Image Segmentation. *IEEE Trans. Geosci. Remote Sens.* **2021**, 1–13. [CrossRef]
20. Lei, D.; Chen, H.; Zhang, L.; Li, W. NLRNet: An Efficient Nonlocal Attention ResNet for Pansharpening. *IEEE Trans. Geosci. Remote Sens.* **2021**, 1–13. [CrossRef]
21. Qu, J.; Shi, Y.; Xie, W.; Li, Y.; Wu, X.; Du, Q. MSSL: Hyperspectral and Panchromatic Images Fusion via Multiresolution Spatial-Spectral Feature Learning Networks. *IEEE Trans. Geosci. Remote Sens.* **2021**, 1–13. [CrossRef]

22. Azarang, A.; Kehtarnavaz, N. Image Fusion in Remote Sensing: Conventional and Deep Learning Approaches. *Synth. Lect. Image Video Multimed. Process.* **2021**, *10*, 1–93. [CrossRef]
23. Parkinson, C.L.; Ward, A.; King, M.D. Earth Science Reference Handbook: A Guide to NASA's Earth Science Program and Earth Observing Satellite Missions. *Natl. Aeronaut. Space Adm.* **2006**, *277*.
24. Amro, I.; Mateos, J.; Vega, M.; Molina, R.; Katsaggelos, A.K. A Survey of Classical Methods and New Trends in Pansharpening of Multispectral Images. *EURASIP J. Adv. Signal Process.* **2011**, *2011*, 79. [CrossRef]
25. Pushparaj, J.; Hegde, A.V. Evaluation of Pan-Sharpening Methods for Spatial and Spectral Quality. *Appl. Geomat.* **2017**, *9*, 1–12. [CrossRef]
26. Liu, Q. Sharpening the Multispectral GF-2 Imagery Using the Modified Intensity-Hue-Saturation Approach: The Different Spectral Settings in Comparison. *IOP Conf. Ser. Mater. Sci. Eng.* **2020**, *768*, 062082. [CrossRef]
27. Liu, Q.; Zhou, H.; Xu, Q.; Liu, X.; Wang, Y. PSGAN: A Generative Adversarial Network for Remote Sensing Image Pan-Sharpening. *IEEE Trans. Geosci. Remote Sens.* **2020**, 1–16. [CrossRef]
28. Hu, J.; Hu, P.; Kang, X.; Zhang, H.; Fan, S. Pan-Sharpening via Multiscale Dynamic Convolutional Neural Network. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 2231–2244. [CrossRef]
29. Thomas, C.; Ranchin, T.; Wald, L.; Chanussot, J. Synthesis of Multispectral Images to High Spatial Resolution: A Critical Review of Fusion Methods Based on Remote Sensing Physics. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 1301–1312. [CrossRef]
30. Feng, R.; Du, Q.; Li, X.; Shen, H. Robust Registration for Remote Sensing Images by Combining and Localizing Feature- and Area-Based Methods. *ISPRS J. Photogramm. Remote Sens.* **2019**, *151*, 15–26. [CrossRef]
31. Hu, J.; He, Z.; Wu, J. Deep Self-Learning Network for Adaptive Pansharpening. *Remote Sens.* **2019**, *11*, 2395. [CrossRef]
32. Yi, Z.; Zhiguo, C.; Yang, X. Multi-Spectral Remote Image Registration Based on SIFT. *Electron. Lett.* **2008**, *44*, 107–108. [CrossRef]
33. Vural, M.F.; Yardimci, Y.; Temizel, A. Registration of Multispectral Satellite Images with Orientation-Restricted SIFT. In Proceedings of the 2009 IEEE International Geoscience and Remote Sensing Symposium, Cape Town, South Africa, 12–17 July 2009; Volume 3, pp. III-243–III-246.
34. Yu, L.; Zhang, D.; Holden, E.-J. A Fast and Fully Automatic Registration Approach Based on Point Features for Multi-Source Remote-Sensing Images. *Comput. Geosci.* **2008**, *34*, 838–848. [CrossRef]
35. Teke, M.; Temizel, A. Multi-Spectral Satellite Image Registration Using Scale-Restricted SURF. In Proceedings of the 2010 20th International Conference on Pattern Recognition, Istanbul, Turkey, 23–26 August 2010; IEEE: New York, NY, USA, 2020; pp. 2310–2313.
36. Yuan, Y.; Jiansheng, C.; Yu, M.; Anzhi, Y.; Yunlong, K.; Qingqing, H.; Xingchun, L. Registration of High Resolution Satellite Images Base on Scale-Orientation Restricted KAZE. *Sens. Lett.* **2014**, *12*, 802–807. [CrossRef]
37. Guo, J.; Yang, F.; Tan, H.; Wang, J.; Liu, Z. Image Matching Using Structural Similarity and Geometric Constraint Approaches on Remote Sensing Images. *J. Appl. Remote Sens.* **2016**, *10*, 045007. [CrossRef]
38. Kern, J.P.; Pattichis, M.S. Robust Multispectral Image Registration Using Mutual-Information Models. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 1494–1505. [CrossRef]
39. Wang, M.; Zhu, Y.; Jin, S.; Pan, J.; Zhu, Q. Correction of ZY-3 Image Distortion Caused by Satellite Jitter via Virtual Steady Reimaging Using Attitude Data. *ISPRS J. Photogramm. Remote Sens.* **2016**, *119*, 108–123. [CrossRef]
40. Wang, M.; Fan, C.; Pan, J.; Jin, S.; Chang, X. Image Jitter Detection and Compensation Using a High-Frequency Angular Displacement Method for Yaogan-26 Remote Sensing Satellite. *ISPRS J. Photogramm. Remote Sens.* **2017**, *130*, 32–43. [CrossRef]
41. Wang, M.; Cheng, Y.; Chang, X.; Jin, S.; Zhu, Y. On-Orbit Geometric Calibration and Geometric Quality Assessment for the High-Resolution Geostationary Optical Satellite GaoFen4. *ISPRS J. Photogramm. Remote Sens.* **2017**, *125*, 63–77. [CrossRef]
42. Wang, M.; Yang, B.; Hu, F.; Zang, X. On-Orbit Geometric Calibration Model and Its Applications for High-Resolution Optical Satellite Imagery. *Remote Sens.* **2014**, *6*, 4391–4408. [CrossRef]
43. Xing, S.; Tan, B.; Li, J.; Xu, Q.; Geng, Z. Approach of High Accurate Multisensor Remote Sensing Images Registration Based on Tiny Facet Primitive. *J. Pla Inst. Surv. Mapp.* **2003**, *2*.
44. Alcaras, E.; Parente, C.; Vallario, A. Automation of Pan-Sharpening Methods for Pléiades Images Using GIS Basic Functions. *Remote Sens.* **2021**, *13*, 1550. [CrossRef]
45. Wieland, M.; Martinis, S. A Modular Processing Chain for Automated Flood Monitoring from Multi-Spectral Satellite Data. *Remote Sens.* **2019**, *11*, 2330. [CrossRef]
46. Tao, C.V.; Hu, Y. A Comprehensive Study of the Rational Function Model for Photogrammetric Processing. *Photogramm. Eng. Remote Sens.* **2001**, *67*, 1347–1358.
47. Grodecki, J.; Dial, G. Block Adjustment of High-Resolution Satellite Images Described by Rational Polynomials. *Photogramm. Eng. Remote Sens.* **2003**, *69*, 59–68. [CrossRef]
48. Fraser, C.S.; Dial, G.; Grodecki, J. Sensor Orientation via RPCs. *ISPRS J. Photogramm. Remote Sens.* **2006**, *60*, 182–194. [CrossRef]
49. Mezirow, J. Perspective Transformation. *Adult Educ.* **1978**, *28*, 100–110. [CrossRef]
50. Benesty, J.; Chen, J.; Huang, Y.; Cohen, I. Pearson correlation coefficient. In *Noise Reduction in Speech Processing*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 1–4.
51. Dong, J.; Crow, W.T.; Tobin, K.J.; Cosh, M.H.; Bosch, D.D.; Starks, P.J.; Seyfried, M.; Collins, C.H. Comparison of Microwave Remote Sensing and Land Surface Modeling for Surface Soil Moisture Climatology Estimation. *Remote Sens. Environ.* **2020**, *242*, 111756. [CrossRef]

52. Alparone, L.; Aiazzi, B.; Baronti, S.; Garzelli, A.; Nencini, F.; Selva, M. Multispectral and Panchromatic Data Fusion Assessment without Reference. *Photogramm. Eng. Remote Sens.* **2008**, *74*, 193–200. [CrossRef]
53. Baronti, S.; Aiazzi, B.; Selva, M.; Garzelli, A.; Alparone, L. A Theoretical Analysis of the Effects of Aliasing and Misregistration on Pansharpened Imagery. *IEEE J. Sel. Top. Signal Process.* **2011**, *5*, 446–453. [CrossRef]
54. Wang, Z.; Bovik, A.C. A Universal Image Quality Index. *IEEE Signal Process. Lett.* **2002**, *9*, 81–84. [CrossRef]