



# Article Integrated Preprocessing of Multitemporal Very-High-Resolution Satellite Images via Conjugate Points-Based Pseudo-Invariant Feature Extraction

Taeheon Kim 🗅 and Youkyung Han \*🕩

Department of Civil Engineering, Seoul National University of Science and Technology, Seoul 01811, Korea; honey25@seoultech.ac.kr

\* Correspondence: han602@seoultech.ac.kr; Tel.: +82-2-970-6566

Abstract: Multitemporal very-high-resolution (VHR) satellite images are used as core data in the field of remote sensing because they express the topography and features of the region of interest in detail. However, geometric misalignment and radiometric dissimilarity occur when acquiring multitemporal VHR satellite images owing to external environmental factors, and these errors cause various inaccuracies, thereby hindering the effective use of multitemporal VHR satellite images. Such errors can be minimized by applying preprocessing methods such as image registration and relative radiometric normalization (RRN). However, as the data used in image registration and RRN differ, data consistency and computational efficiency are impaired, particularly when processing large amounts of data, such as a large volume of multitemporal VHR satellite images. To resolve these issues, we proposed an integrated preprocessing method by extracting pseudo-invariant features (PIFs), used for RRN, based on the conjugate points (CPs) extracted for image registration. To this end, the image registration was performed using CPs extracted using the speeded-up robust feature algorithm. Then, PIFs were extracted based on the CPs by removing vegetation areas followed by application of the region growing algorithm. Experiments were conducted on two sites constructed under different acquisition conditions to confirm the robustness of the proposed method. Various analyses based on visual and quantitative evaluation of the experimental results were performed from geometric and radiometric perspectives. The results evidence the successful integration of the image registration and RRN preprocessing steps by achieving a reasonable and stable performance.

**Keywords:** multitemporal very-high-resolution satellite image; image registration; relative radiometric normalization; integrated preprocessing; pseudo invariant features; conjugate points

## 1. Introduction

Very-high-resolution (VHR) satellite images provide highly reliable information based on detailed descriptions of complex terrain and features. VHR satellite images have attracted considerable research attention for generating high-value data in the field of remote sensing. The practical value of such images is acknowledged globally; hence, various countries are developing and operating satellites equipped with VHR sensors such as WorldView, GeoEye, IKONOS, and KOMPSAT. As the accessibility of VHR satellite images increases, they are being used in a wide range of applications such as mapping, object extraction, disaster monitoring, and change detection [1].

During multitemporal VHR satellite images acquisition, geometric misalignment and radiometric dissimilarity occur owing to the satellite acquisition angle and attitude, absorption and scattering of the atmosphere, seasonal effects, and solar surface sensor interaction [2]. These geometric and radiometric dissimilarities cause fatal problems, particularly when using VHR multitemporal images that describe an object in detail [3,4]. For example, in the change detection field, geometric misalignment and radiometric dissimilarities result in false detection [5]. In the vegetation-monitoring field, radiometric



Citation: Kim, T.; Han, Y. Integrated Preprocessing of Multitemporal Very-High-Resolution Satellite Images via Conjugate Points-Based Pseudo-Invariant Feature Extraction. *Remote Sens.* 2021, *13*, 3990. https:// doi.org/10.3390/rs13193990

Academic Editor: Hemanth Venkateswara

Received: 18 August 2021 Accepted: 30 September 2021 Published: 6 October 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/). dissimilarities generate a distorted vegetation index (VI) [6]. In addition, they inhibit the mosaic image generation [7] and limit the construction of the homogeneous time-series VHR satellite imagery [8]. Therefore, minimization of these errors is required to effectively use VHR multitemporal images [8].

Image registration is a process that minimizes geometric misalignment by unifying coordinates between bitemporal or multitemporal images [9–12]. Image registration generally involves four steps: (1) feature extraction, (2) feature matching, (3) transformation model construction, and (4) image warping [12]. The processes of feature extraction and matching are integrated into the step for extracting conjugate points (CPs), which are key data for performing the image registration [13]. In manual image registration, when CPs are extracted through user intervention, the accuracy depends highly on the user's proficiency [14]. Moreover, if the image to be registered is large and complex, it demands labor-intensive and time-consuming work because of the difficulties in extracting numerous well-distributed CPs [15]. Therefore, automatic image registration without manual extraction of CPs between images is required.

Automatic image registration is classified into area-based and feature-based matching methods depending on the CPs extraction manner [12,16,17]. The area-based matching method sets a template of a certain size between images. This template moves within images, measures the similarity between regions, and extracts the center point of the highest similarity position as the CPs [18,19]. The area-based matching method can be used effectively in areas without spatial characteristics that are identified in the image, such as forests, water, and farmland. However, this method has low accuracy when the geometric misalignment between images is large [20]. A representative area-based matching method includes normalized cross-correlation, mutual information, least-squares matching, and phase correlation.

Unlike the area-based matching method, the feature-based matching method extracts points, lines, and polygon-based features identified as interest points in the image; then, these interest points are matched to CPs via the description process [13,21,22]. The feature-based matching method can extract CPs that are robust to scale, translation, and rotation; therefore, it can effectively perform registration even when the geometric misalignment between images is large. However, CPs extraction in areas without features is difficult using this method. A representative feature-based matching method includes the scale-invariant feature transform (SIFT) from Lowe [23] and speeded-up robust features (SURF) from Bay et al. [24]. Compared with the area-based matching method, it is more suitable for image registration between VHR images [9,25].

Radiometric correction is the process of minimizing the difference in brightness values between images caused by radiometric dissimilarity. This method includes absolute radiometric correction (ARC) and relative radiometric normalization (RRN) [26–29]. ARC converts the brightness value to the scaled surface reflectance by using an atmospheric correction model with inputs from the spectral profile provided in the satellite images' metadata, atmospheric data, and ground observation data [8,27,30]. However, applying the ARC is limited because of the difficulties of collecting the atmospheric and ground observation data [8,31,32]. On the other hand, RRN can minimize the difference in brightness values by estimating the relative change in brightness values between images without ancillary data, and complex atmospheric models [33–35]. The RRN performs normalization by considering linear characteristics between images [26,36–38] or by reflecting their nonlinear characteristics [8,32,39–41].

In general, RRN is classified into global statistical (GS) and radiometric control set sample (RCSS) methods [36,42]. The GS method uses the brightness value of all the pixels in the image [36,42], whereas the RCSS method uses the selected invariant features between images [8,39]. When the images exhibit similar radiometric characteristics, the GS method can rapidly and simply minimize the difference in brightness values between images [43,44]. However, the accuracy of the GS method is low in images containing different radiometric characteristics, such as a large portion of changed areas or seasonal differences

in vegetation areas [39,45]. A representative GS method includes mean–standard deviation (MS) regression, minimum–maximum (MM) regression [46], simple regression (SR) [47], and histogram matching [48]. Unlike the GS method, the RCSS method uses invariant features with similar radiometric characteristics between images [8]. Therefore, the distortion resulting from performing RRN can be minimized, leading to effective normalization between images [26,39,40]. A representative RCSS method includes a dark–bright set based on the Kauth–Thomas transformation [46], no change based on a scattergram between near-infrared images [49], and pseudo-invariant features (PIFs)-based regression. Among the RCSS methods, PIFs regression is commonly used because it is effective to linearly explain the invariant characteristics between multitemporal images [43].

In performing PIFs-based RRN, PIFs are extracted from the invariant area of the same location between multitemporal images. Therefore, image registration for aligning the coordinates between multitemporal images is performed prior to PIFs extraction [2,31,33,50,51]. Next, the PIFs are manually or automatically extracted. These data extraction processes for image registration and RRN are performed separately, thus complicating the preprocessing [31,50]. The complicated preprocessing impairs data consistency and computational efficiency, particularly when processing large amounts of data such as a large volume of multitemporal VHR satellite images. In addition, manpower and time costs are increased due to complex preprocessing. Thus, Li and Davies [50] incorporated preprocessing by comparing the similarity between CPs and their neighboring pixels to extract PIFs. Klaric et al. [51] used the integrated preprocessing approach devised by Li and Davies [50] in the development of geospatial change detection and exploitation, a VHR satellite imagebased automatic change detection system. Kim et al. [31] selected initial PIFs with similar radiometric characteristics among CPs, and then integrated the preprocessing method by extracting numerous PIFs via analysis of the radiometric similarity between the initial PIFs and neighboring pixels. In addition, Armin et al. [36] used CPs extracted based on a block-wise KAZE detector to extract PIFs by using grid interpolation and the K-nearest neighbor (KNN) algorithm. Comprehensively, previous studies were conducted based on the characteristics of CPs which are extracted from the invariant area extracted using the feature-based matching method [52–54].

The preprocessing method can be effectively integrated via the PIFs extraction method based on the CPs extracted through the feature-based matching method. However, such an integrated preprocessing method has certain issues in terms of RRN. The first issue is that the PIFs should be extracted from invariant areas between images. In the abovementioned integrated methods, however, CPs are extracted from vegetation areas sensitive to seasonal and atmospheric components as well as environmental changes [31]. Therefore, PIFs extracted based on CPs contain substantial vegetation data that cause distortion when applying RRN. The second issue is encountered when extracting the PIFs based on CPs in terms of extracted amounts. If there are few changed areas between images, radiometric similarity can be minimized with a small number of PIFs [8]. However, the portion of the changed area between images varies depending on the acquired scene. Thus, numerous PIFs are generally required for performing RRN [27,28,32,55]. For these reasons, in previous studies, pixels with radiometric characteristics similar to those of CPs were extracted as PIFs by analyzing the radiometric characteristics of pixels within the neighborhood region centering on CPs [31,36,50]. Generally, this method can perform RRN with high accuracy by extracting numerous PIFs. However, if the number of CPs are low, the neighborhood region must be largely adjusted to extract numerous PIFs. In this case, the consistency of the radiometric characteristics of CPs will not be maintained, which will result in an error. The final issue encountered in the existing integrated preprocessing methods is that the quality of the PIFs extracted from the CPs is yet to be analyzed. Previous studies have only considered results used for PIF-based RRN. The lack of CPs-based PIFs quality determination results in low reliability of the experiments. Therefore, developing an improved integrated preprocessing method and various experimental executions are essential for solving the abovementioned issues.

In this study, we propose an improved integrated preprocessing method aimed at reducing the problems of existing integrated preprocessing methods. The main purpose of the proposed method is to efficiently extract numerous high-quality PIFs based on CPs for improving the performance of the integrated preprocessing. In the first step, CPs are extracted between images using the SURF algorithm followed by outlier removal. Image registration is then performed by constructing the improved piecewise linear (IPL) model proposed by Han et al. [56] using the CPs. Subsequently, the CPs extracted from vegetation areas are eliminated by using a normalized difference vegetation index (NDVI) mask. In addition, to consider the radiometric characteristics of all bands when extracting PIFs, Z-score images based on the difference images for each band are integrated and then normalized. Based on the pixels of the integrated normalization Z-score (INZ-score) image corresponding to the CPs location, the region growing algorithm is applied to extract the PIFs. Finally, RRN is performed using numerous PIFs extracted by the proposed method. To analyze the performance of the proposed method, the image registration accuracy and PIFs quality are determined. Furthermore, the proposed method is compared with other RRN algorithms by implementing two datasets constructed with VHR images showing diverse properties.

In summary, this study provides three main contributions: (1) the proposed method effectively integrates the preprocessing process by optimizing the characteristics of CPs suitable for the PIFs selection conditions; (2) the proposed mechanism extracts numerous high-quality PIFs based on CPs between VHR multitemporal satellite images, and thus it is possible to perform a more stable RRN of integrated preprocessing; and (3) detailed analysis and discussion of the proposed method focusing on performance in image registration and radiometric correction are provided by implementing two datasets constructed with VHR images. In particular, we confirm the superiority of the proposed method through the PIFs' quality analysis and comparative analysis with other RRN algorithms.

The rest of this manuscript is organized as follows: Section 2 describes the integrated preprocessing methodology proposed herein; Section 3 describes the data and experimental design used in the experiment; Section 4 describes the experimental results; and a detailed analysis of the results is given in Section 5. Finally, Section 6 presents the conclusions.

#### 2. Methodology

Figure 1 shows the workflow of the integrated automatic preprocessing method proposed herein. The CPs were extracted between the reference and sensed images using the SURF algorithm. The outliers included in the extracted CPs were eliminated using affine transformation coefficients and root mean square error (RMSE) values. Based on the CPs that underwent outlier elimination, the IPL was constructed for warping the sensed image to the reference image's coordinates. Subsequently, the CPs extracted from vegetation areas were eliminated using the NDVI mask. The INZ-score image was generated for PIFs extraction performed by considering the radiometric characteristics of all bands based on the CPs. The PIFs were extracted by applying the region growing algorithm centered around the pixels of the INZ-score image corresponding to CPs. Finally, RRN was conducted by constructing a linear regression using the extracted PIFs.

#### 2.1. Image Registration Using CPs Extracted from the SURF Algorithm and Outlier Removal

To apply the feature-based matching method, which is suitable for automatic image registration between VHR satellite images, interest points must be extracted from images. Representative interest point extraction methods include the SIFT algorithm proposed by Lowe [23] and the SURF algorithm proposed by Bay et al. [24]. The SIFT algorithm can perform highly accurate image registration by extracting CPs based on interest points robust in scale, translation, and rotation. However, the processing time increases when the difference of Gaussian (DoG) operation for the scale space and the 128-dimensional vector component in the extraction and description of interest points are used [53,57]. Conversely, the SURF algorithm uses an integral image and approximated Hessian box filters to extract

interest points faster than the SIFT algorithm. In addition, the SURF algorithm utilizes 64-dimensional vector components, i.e., half the number used by the SIFT algorithm, to improve the processing speed [24]. According to the literature [57–59], the matching performance of the SIFT algorithm is similar to that of the SURF algorithm. Therefore, we used the SURF algorithm to extract the CPs due to its good computational efficiency.



Figure 1. Workflow of the proposed preprocessing method.

CPs extracted from the SURF method include outliers unsuitable for estimating transformation coefficients. In this study, we considered an affine transformation model to eliminate the outliers. The affine transformation model is defined by Equation (1):

$$X = a_0 + a_1 x + a_2 x Y = b_0 + b_1 y + b_2 y$$
(1)

where *X*, *Y* represent the coordinates of reference image's CPs; *x*, *y* represent the coordinates of the sensed image's CPs; and  $a_0, b_0 \dots a_2, b_2$  are independent affine transformation coefficients.

First, the RMSE was estimated to determine the difference in the distance between the affine transformation coefficients estimated on the basis of the reference and sensed images' CPs. Then, the CPs with the largest RMSE were identified as outliers and thus eliminated. Finally, the remaining CPs were used again to estimate the affine transformation coefficients to remove outliers by comparing the RMSE. By repeating this process, the CPs corresponding to the RMSE that were lower than the threshold value were selected as final CPs. The threshold value was set to 5 pixels based both on our previous analysis [60] and empirical analysis on the study sites.

Image registration minimizes the geometric misalignment between images using a CPs-based transformation model. The piecewise linear (PL) model proposed by Goshtasby [61] generates Delaunay triangulation based on CPs, estimates affine transformation coefficients for each triangulation, and performs local image registration at each corresponding triangular region between images. In particular, the PL is effective for image registration between VHR satellite images, affected by relief displacement and local distortion [62]. However, severe distortion may occur in areas where triangulation is not constructed [63]. Therefore, we used the IPL transformation model proposed in [56]. The IPL transformation model additionally extracts pseudo-CPs along the boundary of the image for constructing the triangulation in outlying areas of the image so that it improves the registration performance on the areas.

### 2.2. Extraction of Initial PIFs from CPs on Non-Vegetation Areas

PIFs, the core data for performing RRN, should be extracted from the invariant areas with similar radiometric characteristics [38,64]. Typically, most CPs based on the featurebased matching method are extracted from regions having invariant characteristics between images [36,50,65], thus satisfying the criteria of PIFs selection. However, some CPs are extracted from vegetation areas where the radiometric characteristics sensitively change according to the season and environment. According to previous studies [30,36,66], it is noticed that PIFs extracted from vegetation areas have a negative effect on the RRN accuracy. Therefore, we eliminated CPs extracted from vegetation areas by using the NDVI defined as Equation (2):

$$NDVI = \frac{NIR - Red}{NIR + Red}$$
(2)

where *NIR* and *Red* are the brightness values of the NIR and red bands, respectively.

The threshold for each NDVI was then estimated using the Otsu algorithm to generate the binary NDVI images. Through the two binarized NDVI images, the pixels showing the vegetation in both the reference and geo-rectified sensed images were allocated to the vegetation pixels to generate the final NDVI mask. A median filter was then applied to eliminate the noise included in the NDVI mask. Finally, CPs extracted from the NDVI mask were eliminated and the remaining CPs were used as initial PIFs.

#### 2.3. Extraction of PIFs Using Z-Score Image and Region Growing Algorithm

Eliminating the CPs extracted from vegetation areas using the NDVI mask reduces the number of samples. The small number of samples can cause distortion when the normalization coefficients required for performing RRN are estimated [27,28,32,55]. Therefore, additional PIFs should be extracted; here we used the region growing algorithm to this end.

The region growing algorithm selects seed points (SPs) and segments the region by extracting pixels with similar radiometric characteristics between the SPs and their neighborhood pixels [67]. The initial PIFs were selected as SPs for region growing so that the number of PIFs can be increased because a plurality of pixels with radiometric characteristics similar to those can be extracted. However, the VHR satellite images were generally composed of multispectral bands and had an n-dimensional pixel space, where n is the number of bands. Therefore, to extract PIFs that consider all the bands' characteristics, the n-dimensional pixel space should be converted to a one-dimensional one [50].

To convert an n-dimensional pixel space into a one-dimensional space, difference images of each band were generated using Equation (3). Subsequently, the mean and standard deviation of the difference images for each band were estimated and the Z-score images for each band were generated as shown in Equation (4). The Z-score images of each band were integrated and normalized based on Equation (5). The INZ-score image expressed the radiometric characteristics based on the intensity of changes in the corresponding bands between images and expressed near-zero values for the invariant areas between images.

$$D_i = X_i - Y_i \tag{3}$$

$$Z_i = \frac{D_i - \overline{D}_i}{\sigma_{D_i}} \tag{4}$$

$$INZ - score = \sqrt{\sum_{i=1}^{N} (Z_i)^2}$$
(5)

where  $X_i$  and  $Y_i$  represent the reference and geo-rectified sensed images of band *i*, respectively;  $D_i$  is the difference image of the geo-rectified sensed image and reference image for band *i*;  $\overline{D_i}$  is the mean value of the difference image for band *i*;  $\sigma_{D_i}$  is the standard deviation value of the difference image for band *i*;  $Z_i$  is the Z-score of the difference image for band *i*; N is the total number of bands in the reference image and geo-rectified image; and *INZ* is the integrated normalization Z-score image.

The INZ-score value of the SPs, which are initial PIFs, was compared with its neighboring pixels' values in four directions to perform the region growing as shown in Equation (6). A neighboring pixel having the smallest value difference was selected as the pixel with similar radiometric characteristics to SPs, as shown in Equation (7). If the RD(r,s) lies within the threshold value (we set this to 0.2 in consideration of computational efficiency and extraction accuracy), the pixel *s* is selected as the region for extracting PIFs.

$$RD(r, n_i) = |INZ(r) - INZ(n_i)|$$
(6)

$$RD(r,s) = \min_{i=1,\dots,4} \{RD(r,n_i)\}$$
(7)

where  $RD(r, n_i)$  represents the difference of the INZ-score values between the reference pixel *r* and *i*th neighborhood pixel  $n_i$ ; *i* is the number of neighborhood pixels (*i* = 1, ..., 4); and RD(r, s) is the difference of the INZ-score values between the reference pixel *r* and pixel *s*, showing the smallest INZ-score value.

After finding the radiometrically closest pixel to the SP, the reference INZ-score value was repeatedly updated by calculating the average value of the INZ-score. After that, the region growing process was repeatedly performed by Equations (6) and (7). Whenever a new pixel was extracted, the reference INZ-score value was updated as shown in Equation (8):

$$INZ_{k+1}(r) = \frac{kINZ_k(r) + INZ(s)}{k+1}$$
(8)

where  $INZ_{k+1}(r)$  is the new reference INZ-score value,  $INZ_k(r)$  is the *k*th reference INZ-score value, INZ(s) is the INZ-score value of the pixel with the smallest INZ-score difference; and *k* is the number of previously extracted pixels. This updating repeated until no additional new neighborhood pixels were extracted. Numerous pixels can be extracted as PIFs through the process being conducted on every SP.

#### 2.4. Relative Radiometric Normalization Using PIFs Based on CPs

In this study, a linear regression model was constructed to perform RRN using PIFs based on CPs. A linear regression model was defined as Equation (9), where the gain  $(a_i)$  and offset  $(b_i)$  of normalization coefficients were estimated using the PIFs (Equation (10)). Based on the linear regression model constructed, the radiometric dissimilarity between images was minimized by normalizing the radiometric characteristics of the geo-rectified sensed image to those of the reference image.

$$Y'_i = a_i X_i + b_i \tag{9}$$

$$a_{i} = \frac{\sigma_{Y_{i}}^{PIFs}}{\sigma_{X_{i}}^{PIFs}}, \ b_{i} = \overline{Y_{i}^{PIFs}} - a_{i}\overline{X_{i}^{PIFs}}$$
(10)

where  $a_i$  and  $b_i$  are the gain and offset of the normalization coefficients for band *i*;  $X_i$  is the sensed image of band *i*,  $Y'_i$  is the normalized image of band *i*;  $\sigma_{Y_i}^{PIFs}$  and  $\sigma_{X_i}^{PIFs}$  are standard deviation values of PIFs in band *i* for the reference and sensed images, respectively; and  $\overline{Y_i}^{PIFs}$  and  $\overline{X_i}^{PIFs}$  are mean values of PIFs in band *i* for the reference and sensed images, respectively.

#### 3. Dataset Description and Experimental Design

## 3.1. Dataset Description

To analyze the performance of the proposed method, the Gwangju downtown and industrial areas, South Korea, were selected as experimental sites, as shown in Figure 2. The multitemporal VHR satellite images of the Gwangju downtown and industrial areas acquired from KOMPSAT-3A and WorldView-3, respectively, were used as the experimental data, as shown in Figure 3. The specifications of the experimental dataset are listed in Table 1.



**Figure 2.** Experimental sites: site 1 and site 2 are Gwangju downtown and industrial areas, South Korea, respectively.



**Figure 3.** VHR satellite images considered in the experiment: (**a**) reference image of site 1 (2018.09.25), (**b**) sensed image of site 1 (2015.10.19), (**c**) reference image of site 2 (2015.10.19), and (**d**) sensed image of site 2 (2018.05.04).

Site	Site	1	Site 2		
Sensor	KOMPSAT-3A		WorldV	liew-3	
Acquisition date	09/25/2018 (Reference image)	10/19/2015 (Sensed image)	05/26/2017 (Reference image)	05/04/2018 (Sensed image)	
Azimuth angle	260.52°	$152.64^{\circ}$	$180.40^{\circ}$	133.30°	
Incidence angle	11.42°	0.43°	24.10°	31.20°	
Spatial resolution	Panchromatic: 0.55 m Multispectral: 2.2 m Middle Infrared: 5.5 m		Panchromatic: 0.31 m Multispectral: 1.24 m		
Spectral bands	Panchromatic: 450–900 nm Blue: 450–520 nm Green: 520–600 nm Red: 630–690 nm Near Infrared: 760–900 nm Middle Infrared: 3.3–5.2 μm		Coastal: 400–452 nm Blue: 448–510 nm Green: 518–586 nm Yellow: 590–630 nm Red: 632–692 nm Red-edge: 706–746 nm Near Infrared 1: 772–890 nm Near Infrared 2: 866–954 nm		
Radiometric resolution	14 bit		11 b	it	
Location	Gwangju downtown area, South Korea		Gwangju ind South F	ustrial area, Korea	
Processing level	Level	1R	Level 2A		
Image size	$3000 \times 3000$ pixels		4643 × 5030 pixels		

Table 1. Specifications of KOMPSAT-3A and WorldView-3 satellite images.

KOMPSAT-3A is a satellite equipped with a VHR sensor developed and operated by the Korea Aerospace Research Institute (KARI). This satellite provides panchromatic (spatial resolution: 0.55 m), multispectral (2.2 m), and high spatial resolution in middle infrared (5.5 m) images with a radiometric resolution of 14 bits. In this study, we used Level 1R processed multispectral images of the Gwangju downtown area. The Level 1R data is the product processed for coarse radiometric and sensor distortions. The size of the experimental images was 3000 imes 3000 pixels. To perform the experiment, we visually and statistically checked the KOMPSAT-3A images acquired at site 1 and selected the image showing the better contrast as the reference image and the other as the sensed image. Accordingly, the image acquired on 25 September 2018 (Figure 3a) was selected as the reference image, and the image acquired on 19 October 2015 (Figure 3b) was selected as the sensed image. The reference and sensed images of site 1 were acquired during the same season. However, radiometric characteristics of the reference image are clear and bright, whereas those of the sensed image are relatively blurry and dark; thus, strong radiometric dissimilarity exists between these images, rendering site 1 suitable for evaluating the RRN performance of the proposed method.

The WorldView-3 satellite equipped with a VHR sensor, operated by DigitalGlobe, provides panchromatic (spatial resolution: 0.31 m) and multispectral (1.24 m) images with a radiometric resolution of 11 bits. The WorldView-3 provides multispectral images comprising eight bands, but as same with the case of site 1, the blue, green, red, and near-infrared images of site 1 were used in this experiment. The image used in the experiment is a Level 2A product that has been processed with coarse geometric and atmospheric corrections. The size of the experimental images was  $4643 \times 5030$  pixels. The reference and sensed images for the experiment were selected in the same manner as those of site 1. Accordingly, images acquired on 26 May 2017 (Figure 3c) and 4 May 2018 (Figure 3d) were selected as the reference and sensed images, respectively. The site 2 images were acquired during the same season. Unlike the KOMPSAT-3A images of site 1, the Worldview-2 images were

processed with coarse atmospheric correction; hence, the radiometric characteristics were relatively similar. However, the areas between the reference and sensed images changed markedly. Various conditions must be considered to extract PIFs from the remaining invariant areas, excluding widely changed areas. Therefore, site 2 is suitable for evaluating the PIFs extraction performance focusing on these properties.

#### 3.2. Method for the Analysis of Experimental Results

In this study, three types of assessments were performed to analyze the performance of the proposed method. These assessments focused on the following points of aspect:

Assessment 1: image registration accuracy of the proposed method.

Assessment 2: overall characteristics and quality of the PIFs extracted by the proposed method. Assessment 3: performance analysis of the proposed method through comparative analysis with other RRN algorithms.

In Assessment 1, we analyzed the image registration accuracy by generating a chessboard image and estimating the correlation coefficient (CC). The chessboard image, generated by combining two images, is effective for visual analysis because the geometric misalignment between the images is intuitively depicted along the boundaries between the images. The CC is an index representing the similarity between images and is defined by Equation (11) as follows:

$$CC(I_1, I_2) = \frac{\sigma_{I_1 I_2}}{\sqrt{\sigma_{I_1} \sigma_{I_2}}}$$
(11)

where  $CC(I_1, I_2)$  represents the CC between the two images  $I_1$  and  $I_2$ ;  $\sigma_{I_1}$  and  $\sigma_{I_1}$  are the standard deviations of the two images; and  $\sigma_{I_1I_2}$  is the covariance between the two images.

The CC ranges from 0 to 1; values close to 1 indicate that the images are significantly similar, while values close to 0 indicate the opposite. Therefore, if there is a geometric misalignment between the images, the CC index is low owing to different land cover characteristics in the same location. Conversely, if the geometric misalignment between the images is minimized, the CC index is high owing to similar land cover characteristics in the same location.

In Assessment 2, we statistically analyzed the overall characteristics and quality of the PIFs extracted from each site. By calculating the CC and the coefficient of determination ( $R^2$ ), the overall characteristics of PIFs were analyzed, and a *t*-test and F-test were performed for PIFs quality analysis. The CC, which was defined in Equation (11), increases with increasing similarity in the radiometric characteristics of the extracted PIFs. The  $R^2$ value, which ranges from 0 to 1, is a statistical index indicating the appropriateness of the data used in constructing the regression model. If the extracted PIFs fit the linear regression model, a relatively high  $R^2$  value is obtained. The quality of the normalized images can be evaluated via *t*-test and F-test [30,68]. Since PIFs are core data in performing RRN, they directly affect the normalized image quality. Therefore, the quality of the PIFs can be evaluated by analyzing the quality of the normalized images using the PIFs via the *t*-test and F-test. These tests are performed based on the null hypothesis that the mean and standard deviation of the invariant areas between the images are equal [30]. Generally, both the tests accept a null hypothesis if the *p*-value is larger or equal to 0.05 for the 5% significance level. In other words, when reliable results are obtained, t-stat and F-stat are estimated to be close to 0 and 1, respectively [30,68]. In this study, we performed normalization using 70% of the PIFs extracted through the proposed method and performed the *t*-test and F-test by selecting the remaining 30% of PIFs.

In Assessment 3, we used a histogram and various indices to determine whether the proposed method outperforms the other RRN algorithms. The comparison was performed using algorithms such as the MM regression, MS regression, HM, iteratively re-weighted multivariate alteration detection (IR-MAD) [30], and the recently introduced CPs-based RRN method proposed by Armin et al. [36], which extracts PIFs based on CPs using grid interpolation and the KNN algorithm. Although it is intuitive to visually compare and analyze the results, slight changes in brightness values between images are difficult to

see clearly [29,31]. Therefore, qualitative analysis was conducted using a histogram based on the selected test pixels. Moreover, the Euclidean distance between each histogram frequency density was calculated to analyze the histogram distance (HD). A lower HD value represents that the histogram of each image is similar. Subsequently, we used various indices to conduct a quantitative comparison. The normalized absolute error (NAE), structural content (SC), peak signal-to-noise ratio (PSNR), and RMSE were utilized to compare their performance. NAE, SC, PSNR, and RMSE are defined as follows:

$$NAE(R_{T,i}, S_{T,i}) = \frac{\sum_{j=1}^{n} \left| R_{T,i}^{j} - S_{T,i}^{j} \right|}{\sum_{j=1}^{n} \left| R_{T,i}^{j} \right|}$$
(12)

$$SC(R_{T,i}, S_{T,i}) = \frac{\sum_{j=1}^{n} \left| R_{T,i}^{j} - S_{T,i}^{j} \right|}{\sum_{j=1}^{n} \left| R_{T,i}^{j} \right|}$$
(13)

$$PSNR(R_{T,i}, S_{T,i}) = 10 \log_{10} \left( \frac{(2^{n_b} - 1)^2}{\frac{1}{n} \sum_{j=1}^n (R_{T,i}^j - S_{T,i}^j)^2} \right)$$
(14)

$$RMSE(R_{T,i}, S_{T,i}) = \sqrt{\frac{\sum_{j=1}^{n} (R_{T,i}^{j} - S_{T,i}^{j})^{2}}{n}}$$
(15)

where  $R_{T,i}$  and  $S_{T,i}$  represent the test pixels of the reference and sensed images for band i, respectively; n is the total number of test pixels; and  $n_b$  is the number of bits. A lower value of the NAE and RMSE, ca loser value of SC to 1, and a higher value of PSNR indicate that the radiometric characteristics between images are similar and thus the RRN result is superb.

The proposed method and other RRN algorithms were implemented under MATLAB R2020b from a PC with a Ryzen 9 5900X at 4.2 GHz with a 12-core CPU, a GeForce RTX 3060 graphics card, and 64 GB of RAM.

#### 4. Experimental Results

## 4.1. Image Registration Results

To perform integrated preprocessing based on the proposed method, interest points in each dataset were extracted using the SURF algorithm. The first to second distance ratio for extracting the CPs was set to 0.6, which is a generally recommended value [24]. Subsequently, to eliminate outliers, the RMSE of the CPs relating to the affine transformation model was estimated. Next, CPs of the estimated RMSE lower than 5 were considered as outliers and eliminated. Consequently, 3080 and 4304 CPs were extracted in sites 1 and 2, respectively. As shown in Figure 4, CPs were mostly extracted from invariant areas comprising various artificial structures such as roads, flat roofs, and low-rise rooftops (notably, images were contrasted by performing 2% linear stretch to clearly visualize the CPs distribution).

To effectively minimize the geometric misalignment between images, using a transformation model suitable for images is necessary. Here, we performed image registration by constructing an IPL transformation model [56] (Figure 5). In Figure 5, green and purple points represent CPs and pseudo-CPs, respectively. Owing to the pseudo-CPs selected from the outlying area of the image, the Delaunay triangle network was evenly formed to the entire image over a triangular region. Image registration was performed using the Delaunay triangle network constructed on the basis of the CPs and pseudo-CPs extracted from each image.



Figure 4. Extracted CPs: (a) site 1 and (b) site 2.



Figure 5. IPL transformation model constructed based on the CPs: (a) site 1 and (b) site 2.

To visually analyze the registration result, we generated the chessboard image as shown in Figure 6. The reference image was expressed as a true-color image, and the sensed and geo-rectified sensed images were expressed as a false-color image. Objects and lines were inconsistent between chessboard blocks formed between the reference and sensed images (Figure 6a,c). This geometric inconsistency was effectively minimized after performing the image registration, which can be seen from chessboard blocks formed between the reference and geo-rectified sensed images (Figure 6b,d). Some parts of the chessboard image before and after conducting the image registration were magnified, as shown in Figure 7. The geometric inconsistency existing between the images prior to image registration was effectively minimized through the proposed method (denoted using orange circles in Figure 7).

To quantitatively analyze the geometric accuracy, we estimated the CC between the images before and after performing image registration. Table 2 lists the estimated CC for each site. Before performing image registration, the CC values of sites 1 and 2 were 0.260 and 0.680, respectively. After performing image registration, the CC values of sites 1 and 2 were improved to be 0.665 and 0.731, respectively. As a result, the geometric accuracy of sites 1 and 2 was improved by 60.90% and 6.98%, respectively, compared to that of the raw image. The geometric accuracy of site 1 was significantly improved compared to that of site 2 because the WorldView-3 image was processed to coarse geometric correction, whereas the KOMPSAT-3A image did not perform any geometric correction as mentioned in Section 3.1.



**Figure 6.** Chessboard images for the visual inspection of the image registration result: (**a**) before registration of site 1, (**b**) after registration of site 1, (**c**) before registration of site 2, and (**d**) after registration of site 2.

Table 2. Image registration accuracy obtained using CC.

Site	CC				
Site	Raw Image	Geo-Rectified Sensed Image			
Site 1	0.260	0.665			
Site 2	0.680	0.731			

# 4.2. PIFs Extraction Result

To extract PIFs, the NDVI mask was generated, and CPs extracted on the mask were eliminated. The remaining CPs were then used as SPs to perform the region growing-based PIFs extraction by using the INZ-score values. The number of SPs used to extract PIFs was 2006 and 2628 for sites 1 and 2, respectively. Figure 8 shows the distribution of SPs over the INZ-score images. The invariant areas in the INZ-score images were assigned values close to zero; thus, they are presented as dark areas. As one can see from Figure 8, the SPs are located in dark areas of the INZ-score images that show a high probability of invariant areas.

Adjacent pixels with the SPs having similar values to the INZ-score were selected as PIFs by applying the region growing algorithm. Consequently, numerous PIFs were extracted at each site as shown in Figure 9 (PIFs and SPs were denoted using yellow and green marks, respectively). As a result, a total of 30,544 and 211,295 PIFs were extracted from sites 1 and 2, respectively. Table 3 summarizes the extracted numbers of CPs, SPs, and PIFs in each site. In the case of site 1, PIFs were intensively extracted from low-rise rooftops

and roads among various invariant areas (Figure 9a,b). In site 2, PIFs were primarily extracted from roads among various types of invariant areas (Figure 9c,d). Moreover, dramatically changed regions from large-scale agricultural areas to barren areas, located in the lower-left part of the site, and vegetation regions including mountainous areas, located in the upper left part of the site, were excluded from the PIFs extraction.



**Figure 7.** Magnified examples of the chessboard images: (a) before registration of site 1, (b) after registration of site 1, (c) before registration of site 2, and (d) after registration of site 2.



Figure 8. INZ-score images and location of SPs: (a) site 1 and (b) site 2.



**Figure 9.** PIFs extracted by using the proposed method: (**a**) reference image of site 1, (**b**) geo-rectified sensed image of site 1, (**c**) reference image of site 2, and (**d**) geo-rectified sensed image of site 2.

Site	CPs	SPs	PIFs
Site 1	3080	2006	30,544
Site 2	4304	2628	211,295

Table 3. Number of CPs, SPs, and PIFs extracted using the proposed method.

## 4.3. Analysis of PIFs Characteristics and Quality

To analyze the overall characteristics and quality of the PIFs, we randomly sampled 70% of the PIFs for modeling the RRN and the remaining 30% was used for testing the model. Accordingly, 21,381 and 147,907 PIFs were extracted for modeling and 9163 and 63,388 PIFs were used as test pixels in site 1 and site 2, respectively. Figure 10 presents a band-by-band linear regression model of each site, and the CC and  $R^2$  of each linear regression model are estimated in Table 4. Each linear regression model was generated with a positive correlation showing a linear characteristic by estimating the CC and  $R^2$  as a high value of over 0.9 and over 0.8, respectively.

**Table 4.** CC and  $R^2$  estimation results of band-by-band PIFs for each site.

Site 1					Sit	e 2		
	Band 1	Band 2	Band 3	Band 4	Band 1	Band 2	Band 3	Band 4
CC	0.919	0.933	0.927	0.935	0.986	0.968	0.985	0.978
$R^2$	0.844	0.871	0.858	0.873	0.972	0.937	0.969	0.957

We further analyzed the quality of PIFs in detail by conducting the RRN for each site (Figure 11). Images in Figure 11 were displayed as the raw and normalized images (i.e., no

image stretch performed) to intuitively express radiometric variations. Before performing RRN, the radiometric characteristics were expressed differently between reference and geo-rectified sensed images (Figure 11a,b,d,e). As shown in Figure 11c,f, the radiometric characteristics of the geo-rectified sensed images were normalized similar to those of the reference images. In particular, the proposed method effectively minimized the radiometric dissimilarity, even at site 1, where the radiometric dissimilarity was severe compared to that of site 2.



Figure 10. Band-by-band linear regression model based on PIFs: (a) site 1 and (b) site 2.

To evaluate the quality of PIFs from a statistical perspective, we performed via *t*-test and F-test based on the test pixels. Table 5 summarizes the *t*-test and F-test results obtained using the test pixels of site 1. There was a difference between the mean value of the georectified sensed and the reference image, whereas the mean value of the normalized image was estimated similarly to that of the reference image. As a result of the *t*-test, the t-stat of bands 1 and 2 was close to 0, and the *p*-value was estimated to be larger than 0.85. The t-stat of bands 3 and 4 was -0.6696 and -0.5272, respectively, which had relatively large deviations compared to other bands, but the *p*-value was estimated to be over 0.5. The standard deviation value of the geo-rectified sensed image was about 2 times smaller than that of the reference image, whereas the normalized image was similar to the reference image. As a result of the F-test, the F-stat of each band was close to 1, and the *p*-value was also derived to be more than 0.6.

Table 6 summarizes the *t*-test and F-test results obtained using the test pixels of site 2. The mean and standard values of the normalized image were estimated similarly to the reference image. Although different values of t-stat for each band were estimated, the *p*-value was estimated to be over 0.14; thus, it can be considered as statistically similar because the null hypothesis has been accepted. In the case of the F-test, the F-stat of all bands was estimated to be close to 1, and the *p*-value was estimated to be more than 0.79, which is a reliable result.

In summary, performing RRN using the PIFs extracted via the proposed method became significantly similar to the mean and standard deviation between the reference and normalized images in both sites. These results are demonstrated by the *p*-values of the *t*-test and F-test, estimated as higher than 0.05 in both sites. In other words, the null hypothesis that the mean and standard deviation of the invariant areas between the images



were equal was accepted. Therefore, these results demonstrate that the PIFs extracted by the proposed method were superb.

**Figure 11.** Radiometric normalized images obtained using the proposed method: (**a**) reference image of site 1, (**b**) raw geo-rectified sensed image of site 1, (**c**) normalized geo-rectified sensed image of site 1, (**d**) reference image of site 2, (**e**) raw geo-rectified sensed image of site 2, and (**f**) normalized geo-rectified sensed image of site 2.

	Band 1	Band 2	Band 3	Band 4
Geo. mean	2828.16	2964.95	2372.27	3124.70
Ref. mean	2734.95	3035.89	2508.81	3502.35
Nor. mean	2734.25	3036.00	2513.02	3506.85
t-stat	0.1490	-0.0181	-0.6696	-0.5272
<i>p</i> -value	0.8816	0.9855	0.5032	0.5981
Geo. std	464.93	751.69	892.05	1288.52
Ref. std	1098.07	1621.19	1660.16	2102.99
Nor. std	1097.37	1625.77	1667.01	2114.51
F-stat	1.0013	0.9944	0.9918	0.9891
<i>p</i> -value	0.9525	0.7908	0.6992	0.6083

Table 5. Comparison of means and standard deviations with statistical tests in site 1.

# 4.4. Relative Radiometric Normalization Performance

Through the quality assessment of the PIFs with the proposed method, we confirmed that the extracted PIFs have superior quality; however, it was unknown whether the proposed method has superior performance compared with other RRN algorithms. To this end, we performed a comparative analysis of various RRN algorithms such as MM regression, MS regression, HM, IR-MAD, and the method of [36].

To comparatively analyze various RRN algorithms, geo-rectified sensed images of each site were normalized using the proposed method and various RRN algorithms. Subsequently, histograms were constructed based on test pixels to intuitively analyze changes in the radiometric characteristics of the normalized images. The histogram of the reference image was presented in red, that of the geo-rectified sensed image in blue, and that of the normalized image of each RRN algorithm in black. As the distribution of bins in the histogram for each image was different, the sum of the density frequency constituting the histogram was normalized to be less than 1. Subsequently, the HD value between histograms was estimated to analyze their similarity. In addition, to quantitatively perform the comparative analysis on the proposed and other RRN methods, the NASE, SC, PSNR, and RMSE were estimated based on the test pixels. In the case of the NASE, SC, and PSNR, line graphs were generated based on the value of each index to intuitively compare and analyze the performance. Moreover, the computation time of each RRN algorithm was estimated for comparison.

	Band 1	Band 2	Band 3	Band 4
Geo. mean	365.74	455.05	312.13	446.38
Ref. mean	339.04	419.15	292.53	408.56
Nor. mean	339.01	419.00	292.35	408.71
t-stat	0.4655	1.1801	1.4578	0.8003
<i>p</i> -value	0.6415	0.2379	0.1449	0.4235
Geo. std	89.53	122.28	135.48	167.88
Ref. std	100.87	137.73	152.45	177.72
Nor. std	100.96	137.81	152.37	177.62
F-stat	0.9980	0.9988	1.0011	1.0011
<i>p</i> -value	0.7977	0.8719	0.8889	0.8878

Table 6. Comparison of means and standard deviations with statistical tests in site 2.

The histogram and HD value of site 1 are shown in Figure 12 and Table 7, respectively. Although the histogram of the geo-rectified sensed image was generated with a similar structure to the reference image, the frequency density was different due to the radiometric dissimilarity. The HD values of the corresponding bands between the reference and georectified sensed images were estimated as a range from 0.2355 to 0.3896, and the average value was estimated to be 0.3269. In the case of using MM regression, the overall brightness value increased or decreased; hence, the HD values were slightly higher than those of the geo-rectified sensed images, except for band 2. By applying the MS regression, the brightness value was irregularly normalized, resulting in higher HD values than those of the MM regression. In the case of using HM, the histograms of the normalized image were constructed similarly to the reference image in bands 3 and 4, but histograms of bands 1 and 2 were represented differently. However, the HD values decreased compared to the values of the geo-rectified sensed image in all bands. This is because the range of high-density frequency of the histogram in the normalized image is similar to that of the reference image. The histograms when applying IR-MAD represented that the brightness value was biased in one or both directions. As a result, the HD values were estimated higher than those of the geo-rectified sensed image, except for band 3. The histograms of the method of [36] were constructed similarly to HM, and the HD values were also estimated to be low. In the case of using the proposed method, although the histograms were constructed similarly to HM and the method of [36], the HD value was estimated to be on average 0.1694, which is the lowest value compared to the other RRN algorithms.

The graphs of NAE, SC, and PSNR before and after performing the RRN methods in site 1 are shown in Figure 13, and the RMSE and computation time are summarized in Table 8. Although the MM regression and MS regression methods' had short computation times, the values of all the indices achieved compared to those of geo-rectified sensed image were poor. In the case of HM, the NAE and RMSE values were lower than those of the geo-rectified sensed image, the SC value was close to 1, the PSNR value was high, and the computation time was relatively short. The NAE and RMSE values of IR-MAD were lower than those of the geo-rectified sensed image, but the SC values were higher than 1.2, except for band 1. Moreover, the computation time was 45.547 s, the highest time among the algorithms. When applying the method of [36], the computation time and NAE, SC, and RMSE values were similar to HM, but the PSNR value tended to be slightly higher than that of HM. The RMSE of the proposed method was estimated to have the lowest value around 602.89, and the NAE, SC, and PSNR was measured to have similar or improved values compared to those of HM and the method of [36]. However, it had a relatively longer computation time, similar to that of IR-MAD.

Method Band 1 Band 2 Band 3 Band 4 Average Geo-rectified sensed image 0.3896 0.3408 0.3415 0.2355 0.3269 MM regression [46] 0.4288 0.3167 0.3341 0.2283 0.3270 MS regression [46] 0.4836 0.4701 0.4435 0.3245 0.4304 HM [48] 0.2290 0.2149 0.2202 0.0954 0.1899 IR-MAD [30] 0.4252 0.3418 0.3258 0.2507 0.3359 Method of [36] 0.2398 0.2031 0.1970 0.1345 0.1936 Proposed method 0.2270 0.1815 0.1705 0.0987 0.1694

Table 7. Result of estimating HD for each band of site 1.



Band 2











Band 1







Band 4



(a)

Figure 12. Cont.



Figure 12. Histogram based on test pixels of site 1: (a) MM regression, (b) MS regression, (c) HM, (d) IR-MAD, (e) method of [36], and (f) proposed method.



Figure 13. Various evaluation indices for accuracy analysis by algorithms in site 1: (a) NAE, (b) SC, and (c) PSNR.

The histogram and HD values of site 2 are shown in Figure 14 and Table 9, respectively. The Worldview-3 images of site 2 were processed to coarse atmospheric correction as described in Section 3.2. Therefore, the histogram of the geo-rectified sensed image was generated similarly to that of the reference image, and the HD values were estimated to be relatively lower than those of site 1. Histograms of MM regression and MS regression were constructed with a different range from the reference image, and their HD values were estimated to be higher than those of the geo-rectified sensed image. In the case of HM, the histogram of the normalized image was constructed similarly to that of the reference image; thus, the HD value was significantly improved compared to the geo-rectified sensed image. Similarity, the IR-MAD method achieved reliable normalization results by reducing the HD values, except for band 4. Although the histograms of the method of [36] were constructed similarly to the reference image, the HD values were estimated to be higher than those of

HM. The proposed method obtained a considerable improvement in normalization results by achieving an HD value of 0.0218 on average, which was the lowest value compared to the other RRN algorithms.

Mathad		Comp Time				
Wiethod	Band 1	Band 2	Band 3	Band 4	Average	Comp. Time
Geo-rectified sensed image	700.69	957.98	891.41	1094.04	911.03	-
MM regression [46]	760.01	1040.02	895.12	1063.76	939.73	0.313 s
MS regression [46]	905.35	1279.20	1216.94	1582.62	1246.03	0.274 s
HM [48]	468.22	582.70	698.75	871.43	655.28	3.233 s
IR-MAD [30]	506.81	772.55	721.06	1018.93	754.84	45.547 s
Method of [36]	481.08	656.85	654.24	838.13	657.50	3.988 s
Proposed method	438.45	582.30	589.67	801.14	602.89	44.412 s

Table 8. Comparison of RRN methods based on RMSE and computation time (site 1).

Table 9. Result of estimating HD for each band of site 1.

Method	Band 1	Band 2	Band 3	Band 4	Average
Geo-rectified sensed image	0.0788	0.2576	0.0695	0.1336	0.1349
MM regression [46]	0.587	0.4411	0.0695	0.124	0.3054
MS regression [46]	0.2422	0.3914	0.1145	0.5635	0.3279
HM [48]	0.0222	0.0550	0.0173	0.0117	0.0266
IR-MAD [30]	0.0606	0.2756	0.0784	0.2818	0.1741
Method of [36]	0.0113	0.0559	0.0272	0.0372	0.0329
Proposed method	0.0118	0.0138	0.0326	0.0289	0.0218

The graphs of NAE, SC, and PSNR before and after applying the RRN methods in site 2 are shown in Figure 15, and the RMSE and computation time are reported in Table 10. The MM regression and MS regression could not minimize the radiometric dissimilarities as in site 1. HM, IR-MAD, and the method of [36] could minimize radiometric dissimilarity, but the accuracy of HM was lower than that of IR-MAD and the method of [36]. This occurred mainly due to the large-scale changed area between images because HM performs RRN utilizing the all of the pixels including changed areas, unlike IR-MAD and the method of [36], which utilize PIFs only. The accuracy of the proposed method was superior to those of the other RRN algorithms as in site 1. However, the longest computation time was measured as 152.957 s.



(a)

Figure 14. Cont.



Figure 14. Histogram based on test pixels of site 2: (a) MM regression, (b) MS regression, (c) HM, (d) IR-MAD, (e) method of [36], and (f) proposed method.





Figure 15. Various evaluation indices for accuracy analysis by algorithms in site 2: (a) NAE, (b) SC, and (c) PSNR.

Mathad		Comp Time				
Method	Band 1	Band 2	Band 3	Band 4	Average	Comp. Time
Geo-rectified sensed image	33.31	47.64	37.87	60.09	44.73	-
MM regression [46]	71.46	120.32	37.87	59.04	72.17	0.921 s
MS regression [46]	54.17	78.92	58.16	102.68	73.48	0.742 s
HM [48]	24.01	51.32	34.94	48.24	39.63	3.935 s
IR-MAD [30]	26.80	38.12	23.61	55.23	35.94	92.906 s
Method of [36]	18.94	31.18	30.15	52.08	33.09	4.976 s
Proposed method	18.03	29.31	29.46	47.77	31.14	152.957 s

Table 10. Comparison of RRN methods based on RMSE and computation time (site 2).

# 5. Discussion

In this study, we proposed an integrated preprocessing method by extracting numerous PIFs based on CPs. In addition, the accuracy and performance of the proposed method were analyzed through various assessments: (1) image registration accuracy, (2) evaluation of the overall characteristics of PIFs and their quality, and (3) performance analysis through comparative analysis with other RRN algorithms. The detailed discussion of the achieved results will be given in this section.

With respect to the results of image registration, the proposed method effectively reduced the geometric misalignment between images. This was demonstrated through geometric consistency represented by chessboard images generated between reference and geo-rectified images and increased CC values after image registration. Although the applied SURF-based CPs extraction with geometric restriction-based outlier removal is not a novel approach in terms of the image registration, the contribution can be found from the fact that the CPs were used as seed points to extract PIFs.

The overall characteristics of PIFs were analyzed via measurement of CC and  $R^2$  values of each band-by-band linear regression model estimated from the PIFs. The CC and  $R^2$  of PIFs were estimated as high values, between approximately 0.8 and 0.9. Moreover, the quality of PIFs was statistically analyzed by performing a *t*-test and F-test based on test pixels. The *p*-values of the *t*-test and F-test were estimated to be higher than 0.05, and the t-state and F-state were estimated to be close to 0 and 1, respectively. Therefore, the overall characteristics and quality of extracted PIFs were found to be reliable for performing RRN.

Finally, to analyze the RRN performance of the proposed method, a comparative analysis was performed using the proposed method and other RRN algorithms: MM regression, MS regression, HM, IR-MAD, and the method of [36]. As a result, the proposed method obtained a superlative RRN performance over the other methods in both sites. Furthermore, we can confirm several facts through the comparative analysis results in Section 4.3.

First, the MM regression and MS regression are not suitable for minimizing the radiometric dissimilarity between VHR satellite images. These methods forcefully convert the brightness value of entire pixels only using the calculated maximum–minimum and average–standard deviation values without considering the radiometric characteristics between images. Therefore, the brightness value of the normalized image obtained using the MM and MS regressions was distorted so poor results were obtained even before applying the RRN.

Second, in the case of severe radiometric dissimilarity showing between images, such as in site 1, PIFs should be extracted more carefully. Regions such as vegetation that can be affected by even subtle changes of the radiometric properties due to the difference of acquisition conditions or seasons should be removed from the PIFs extraction process. The proposed method excluded the vegetation areas to extract the PIFs, whereas IR-MAD and the method of [36] extracted PIFs without consideration of the vegetation areas. Accordingly, these methods showed lower accuracy than the proposed one.

Third, a nonlinear method such as HM could effectively minimize radiometric dissimilarity. HM utilizes the entire pixel brightness values of the image like MM regression and MS regression but considers them nonlinearly. Therefore, it was relatively less affected by vegetation data, and radiometric dissimilarity was minimized even in site 2, where a large portion of changed areas occurred between images.

The method of [36] and the proposed method, extracting PIFs based on CPs, effectively minimized both the geometric and radiometric dissimilarity at both sites. These results indicate that PIFs extracted based on CPs are suitable to perform RRN. In terms of accuracy, the proposed method outperformed the method of [36] at both sites. This result was related to the number of extracted PIFs and consideration of the vegetation areas for their extraction. The proposed method extracted approximately two and nine times more PIFs than the method of [36] in site 1 and site 2, respectively, although the vegetation areas were removed for extracting them. Consequently, the proposed method could estimate normalization coefficients with characteristics more robust to radiometric dissimilarity than the comparison method. However, the proposed method requires more computation time because it extracts PIFs in a way that repeatedly considers CPs and their neighboring pixels. Particularly, the proposed method in site 2 required about 40 times more computation time than the method of [36], though the accuracy was slightly improved. This is because the simple linear regression model estimated by the PIFs has a limitation of entirely showing an effect in performing normalization while considering complex elements of nonlinearity. These problems remain to be addressed in our future work.

#### 6. Conclusions

In this study, we proposed an integrating preprocessing method by extracting PIFs necessary for performing RRN based on CPs required for image registration. In the first step of the process, the CPs were extracted using the SURF algorithm, and outliers were eliminated. IPL transformation was implemented to perform image registration. Subsequently, the CPs without vegetation areas were selected as SPs for the region growing-based PIFs extraction approach. Finally, RRN was performed by constructing a linear regression model based on the extracted PIFs. To analyze the performance of the proposed method, VHR images acquired from KOMPSAT-3A and WorldView-3 were used for constructing two experiment sites showing different properties.

We showed that the proposed method was effective in reducing geometric misalignment between images of each site. The overall characteristics and quality of PIFs were analyzed by using various statistical indices. As a result, the PIFs of the proposed method were found to be reliable for performing RRN. In addition, we confirmed that the proposed method effectively minimized the radiometric dissimilarity through a comparative analysis with various RRN algorithms by using radiometric normalization assessment measures such as HD, NAE, SC, PSRN, and RMSE. However, the proposed method had a problem with respect to the computational efficiency due to the region growing-based extraction of PIFs. Future research will proceed to improve the efficiency of the proposed method when extracting PIFs. Furthermore, we will conduct research related to nonlinear modeling using the CPs-based PIFs, reflecting the potential of nonlinear normalization confirmed through this study.

**Author Contributions:** Conceptualization, Y.H.; methodology, T.K. and Y.H.; software, T.K.; validation, T.K.; formal analysis, T.K.; investigation, T.K.; data curation, T.K.; writing—original draft preparation, T.K.; writing—review and editing, Y.H.; visualization, T.K.; supervision, Y.H.; funding acquisition, Y.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by 'Satellite Information Application' of the Korea Aerospace Research center (KARI). This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (2021R1A2C2093671).

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

#### References

- 1. Han, Y.; Bovolo, F.; Bruzzone, L. An approach to fine coregistration between very high resolution multispectral images based on registration noise distribution. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 6650–6662. [CrossRef]
- Li, C.; Xiong, H. A geometric and radiometric simultaneous correction model (GRSCM) framework for high-accuracy remotely sensed image preprocessing. *Photogramm. Eng. Remote Sens.* 2017, 83, 621–632. [CrossRef]
- 3. Canty, M.J.; Nielsen, A.A.; Schmidt, M. Automatic radiometric normalization of multitemporal satellite imagery. *Remote Sens. Environ.* **2004**, *91*, 441–451. [CrossRef]
- 4. Johnson, R.D.; Kasischke, E.S. Change vector analysis: A technique for the multispectral monitoring of land cover and condition. *Int. J. Remote Sens.* **1998**, *19*, 411–426. [CrossRef]
- 5. Bruzzone, L.; Bovolo, F. A novel framework for the design of change-detection systems for very-high-resolution remote sensing images. *Proc. IEEE* 2013, 101, 609–630. [CrossRef]
- 6. Chen, X.; Vierling, L.; Deering, D. A simple and effective radiometric correction method to improve landscape change detection across sensors and across time. *Remote Sens. Environ.* **2005**, *98*, 63–79. [CrossRef]
- 7. Afek, Y.; Brand, A. Mosaicking of orthorectified aerial images. *Photogramm. Eng. Remote Sens.* 1998, 64, 115–124.
- Solano-Correa, Y.T.; Bovolo, F.; Bruzzone, L. Generation of homogeneous VHR time series by nonparametric regression of multisensor bitemporal images. *IEEE Trans. Geosci. Remote Sens.* 2019, 57, 7579–7593. [CrossRef]
- Huo, C.; Pan, C.; Huo, L.; Zhou, Z. Multilevel SIFT matching for large-size VHR image registration. *IEEE Geosci. Remote Sens. Lett.* 2011, 9, 171–175. [CrossRef]
- 10. Lee, I.H.; Choi, T.S. Accurate registration using adaptive block processing for multispectral images. *IEEE Trans. Circuits Syst. Video Technol.* **2013**, 23, 1491–1501. [CrossRef]
- 11. Toutin, T. Geometric processing of remote sensing images: Models, algorithms and methods. *Int. J. Remote Sens.* 2004, 25, 1893–1924. [CrossRef]
- 12. Zitová, B.; Flusser, J. Image registration methods: A survey. Image Vis. Comput. 2003, 21, 977–1000. [CrossRef]
- 13. Hong, G.; Zhang, Y. Wavelet-based image registration technique for high-resolution remote sensing images. *Comput. Geosci.* 2008, 34, 1708–1720. [CrossRef]
- 14. Bentoutou, Y.; Taleb, N.; Kpalma, K.; Ronsin, J. An automatic image registration for applications in remote sensing. *IEEE Trans. Geosci. Remote Sens.* **2005**, 43, 2127–2137. [CrossRef]
- 15. Kennedy, R.E.; Cohen, W.B. Automated designation of tie-points for image-to-image coregistration. *Int. J. Remote Sens.* 2003, 24, 3467–3490. [CrossRef]
- 16. 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**, *19*, 15–26. [CrossRef]
- 17. Wang, S.; Quan, D.; Liang, X.; Ning, M.; Guo, Y.; Jiao, L. A deep learning framework for remote sensing image registration. *ISPRS J. Photogramm. Remote Sens.* **2018**, *145*, 148–164. [CrossRef]
- Chen, H.M.; Arora, M.K.; Varshney, P.K. Mutual information-based image registration for remote sensing data. *Int. J. Remote Sens.* 2003, 24, 3701–3706. [CrossRef]
- 19. Okorie, A.; Makrogiannis, S. Region-based image registration for remote sensing imagery. *Comput. Vis. Image Underst.* 2019, 189, 102825. [CrossRef]
- Ye, Y.; Shan, J. A local descriptor based registration method for multispectral remote sensing images with non-linear intensity differences. *ISPRS J. Photogramm. Remote Sens.* 2014, 90, 83–95. [CrossRef]
- 21. Chang, X.; Du, S.; Li, Y.; Fang, S. A coarse-to-fine geometric scale-invariant feature transform for large size high resolution satellite image registration. *Sensors* **2018**, *18*, 1360. [CrossRef]
- 22. Ma, W.; Wen, Z.; Wu, Y.; Jiao, L.; Gong, M.; Zheng, Y.; Liu, L. Remote sensing image registration with modified sift and enhanced feature matching. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 3–7. [CrossRef]
- 23. Lowe, D.G. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. 2004, 60, 91–110. [CrossRef]

- 24. Bay, H.; Ess, A.; Tuytelaars, T.; Van Gool, L. Speeded-up robust features (SURF). *Comput. Vis. Image Underst.* **2008**, *110*, 346–359. [CrossRef]
- Du, S.; Wang, M.; Fang, S. Block-and-octave constraint SIFT with multi-thread processing for VHR satellite image matching. *Remote Sens. Lett.* 2017, *8*, 1180–1189. [CrossRef]
- De Carvalho Júnior, O.A.; Guimarães, R.F.; Silva, N.C.; Gillespie, A.R.; Gomes, R.A.T.; Silva, C.R.; De Carvalho, A.P.F. Radiometric normalization of temporal images combining automatic detection of pseudo-invariant features from the distance and similarity spectral measures, density scatterplot analysis, and robust regression. *Remote Sens.* 2013, 5, 2763–2794. [CrossRef]
- 27. Du, Y.; Teillet, P.M.; Cihlar, J. Radiometric normalization of multitemporal high-resolution satellite images with quality control for land cover change detection. *Remote Sens. Environ.* 2002, *82*, 123–134. [CrossRef]
- Hong, G.; Zhang, Y. A comparative study on radiometric normalization using high resolution satellite images. *Int. J. Remote Sens.* 2008, 29, 425–438. [CrossRef]
- 29. Kim, D.; Pyeon, M.; Eo, Y.; Byun, Y.; Kim, Y. Automatic pseudo-invariant feature extraction for the relative radiometric normalization of hyperion hyperspectral images. *GISci. Remote Sens.* **2012**, *49*, 755–773. [CrossRef]
- 30. Canty, M.J.; Nielsen, A.A. Automatic radiometric normalization of multitemporal satellite imagery with the iteratively re-weighted MAD transformation. *Remote Sens. Environ.* **2008**, *112*, 1025–1036. [CrossRef]
- 31. Kim, T.; Lee, W.H.; Yeom, J.; Han, Y. Integrated automatic pre-processing for change detection based on SURF algorithm and mask filter. *J. Korean Soc. Surv. Geod. Photogramm. Cartogr.* **2019**, *37*, 209–219.
- 32. Seo, D.K.; Kim, Y.H.; Eo, Y.D.; Park, W.Y.; Park, H.C. Generation of radiometric, phenological normalized image based on random forest regression for change detection. *Remote Sens.* 2017, *9*, 1163. [CrossRef]
- 33. Kim, D.S.; Kim, Y.I. Relative radiometric normalization of hyperion hyperspectral images through automatic extraction of pseudo-invariant features for change detection. *J. Korean Soc. Surv. Geod. Photogramm. Cartogr.* 2008, 26, 129–137.
- 34. Biday, S.G.; Bhosle, U. Radiometric correction of multitemporal satellite imagery. J. Comput. Sci. 2010, 6, 1019–1028. [CrossRef]
- Schott, J.R.; Salvaggio, C.; Volchok, W.J. Radiometric scene normalization using pseudoinvariant features. *Remote Sens. Environ.* 1988, 26, 1–16. [CrossRef]
- Moghimi, A.; Sarmadian, A.; Mohammadzadeh, A.; Celik, T.; Amani, M.; Kusetogullari, H. Distortion robust relative radiometric normalization of multitemporal and multisensor remote sensing images using image features. *IEEE Trans. Geosci. Remote Sens.* 2021. early access. [CrossRef]
- 37. Moghimi, A.; Mohammadzadeh, A.; Celik, T.; Amani, M. A novel radiometric control set sample selection strategy for relative radiometric normalization of multitemporal satellite images. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 2503–2519. [CrossRef]
- 38. Xu, H.; Wei, Y.; Li, X.; Zhao, Y.; Cheng, Q. A novel automatic method on pseudo-invariant features extraction for enhancing the relative radiometric normalization of high-resolution images. *Int. J. Remote Sens.* **2021**, *42*, 6155–6186. [CrossRef]
- 39. Sadeghi, V.; Ebadi, H.; Ahmadi, F.F. A new model for automatic normalization of multitemporal satellite images using artificial neural network and mathematical methods. *Appl. Math. Model.* **2013**, *37*, 6437–6445. [CrossRef]
- 40. Denaro, L.G.; Lin, C.H. Hybrid canonical correlation analysis and regression for radiometric normalization of cross-sensor satellite imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2020, 13, 976–986. [CrossRef]
- 41. Yin, Z.; Zou, L.; Sun, J.; Zhang, H.; Zhang, W.; Shen, X. A nonlinear radiometric normalization model for satellite images time series based on artificial neural networks and greedy algorithm. *Remote Sens.* **2021**, *13*, 933. [CrossRef]
- 42. Seo, D.K.; Eo, Y.D. Multilayer perceptron-based phenological and radiometric normalization for high-resolution satellite imagery. *Appl. Sci.* **2019**, *9*, 4543. [CrossRef]
- 43. Zhou, H.; Liu, S.; He, J.; Wen, Q.; Song, L.; Ma, Y. A new model for the automatic relative radiometric normalization of multiple images with pseudo-invariant features. *Int. J. Remote Sens.* **2016**, *37*, 4554–4573. [CrossRef]
- 44. Schultz, M.; Verbesselt, J.; Avitabile, V.; Souza, C.; Herold, M. Error sources in deforestation detection using BFAST monitor on Landsat time series across three tropical sites. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *9*, 3667–3679. [CrossRef]
- 45. Yang, X.; Lo, C.P. Relative radiometric normalization performance for change detection from multi-date satellite images. *Photogramm. Eng. Remote Sens.* **2000**, *66*, 967–980.
- 46. Hall, F.G.; Strebel, D.E.; Nickeson, J.E.; Goetz, S.J. Radiometric rectification: Toward a common radiometric response among multidate, multisensor images. *Remote Sens. Environ.* **1991**, *35*, 11–27. [CrossRef]
- 47. Jensen, J.R. Urban/suburban land use analysis. In *Manual of Remote Sensing*, 2nd ed.; American Society of Photogrammetry: Bethesda, MA, USA, 1983; pp. 1571–1666.
- 48. Richards, J.A.; Richards, J.A. *Remote Sensing Digital Image Analysis*; Springer: Berlin/Heidelberg, Germany, 1999; Volume 3, pp. 10–38.
- Elvidge, C.D.; Yuan, D.; Weerackoon, R.D.; Lunetta, R.S. Relative radiometric normalization of Landsat multispectral scanner (MSS) data using an automatic scattergram controlled regression. *Photogramm. Eng. Remote Sens.* 1995, 61, 1255–1260.
- 50. Li, Y.; Davis, C.H. Pixel-based invariant feature extraction and its application to radiometric co-registration for multi-temporal high-resolution satellite imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2010**, *4*, 348–360. [CrossRef]
- Klaric, M.N.; Claywell, B.C.; Scott, G.J.; Hudson, N.J.; Sjahputera, O.; Li, Y.; Barratt, S.T.; Keller, J.M.; Davis, C.H. GeoCDX: An automated change detection and exploitation system for high-resolution satellite imagery. *IEEE Trans. Geosci. Remote Sens.* 2013, 51, 2067–2086. [CrossRef]

- 52. Mikolajczyk, K.; Schmid, C. A performance evaluation of local descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.* 2005, 27, 1615–1630. [CrossRef] [PubMed]
- 53. Sedaghat, A.; Mohammadi, N. High-resolution image registration based on improved SURF detector and localized GTM. *Int. J. Remote Sens.* **2019**, 40, 2576–2601. [CrossRef]
- 54. Wu, Y.; Ma, W.; Gong, M.; Su, L.; Jiao, L. A novel point-matching algorithm based on fast sample consensus for image registration. *IEEE Geosci. Remote Sens. Lett.* **2014**, *12*, 43–47. [CrossRef]
- 55. Zhang, L.; Wu, C.; Du, B. Automatic radiometric normalization for multitemporal remote sensing imagery with iterative slow feature analysis. *IEEE Trans. Geosci. Remote Sens.* 2014, 52, 6141–6155. [CrossRef]
- 56. Han, Y.; Kim, T.; Yeom, J. Improved piecewise linear transformation for precise warping of very-high-resolution remote sensing images. *Remote Sens.* 2019, *11*, 2235. [CrossRef]
- 57. Yang, F.; Ji, L.E.; Liu, S.; Feng, P. A fast and high accuracy registration method for multi-source images. *Optik* **2015**, *126*, 3061–3065. [CrossRef]
- Kim, T.; Lee, K.; Lee, W.H.; Yeom, J.; Jun, S.; Han, Y. Coarse to fine image registration of unmanned aerial vehicle images over agricultural area using SURF and mutual information methods. *Korean J. Remote Sens.* 2019, 35, 945–957.
- 59. Oh, J.; Han, Y. A double epipolar resampling approach to reliable conjugate point extraction for accurate Kompsat-3/3A stereo data processing. *Remote Sens.* 2020, 12, 2940. [CrossRef]
- 60. Han, Y.; Choi, J.; Byun, Y.; Kim, Y. Parameter optimization for the extraction of matching points between high-resolution multisensor images in urban areas. *IEEE Trans. Geosci. Remote Sens.* **2013**, *52*, 5612–5621. [CrossRef]
- 61. Goshtasby, A. Piecewise linear mapping functions for image registration. Pattern Recognit. 1986, 19, 459–466. [CrossRef]
- 62. Arévalo, V.; González, J. An experimental evaluation of non-rigid registration techniques on QuickBird satellite imagery. *Int. J. Remote Sens.* 2008, *29*, 513–527. [CrossRef]
- 63. Arévalo, V.; González, J. Improving piecewise linear registration of high-resolution satellite images through mesh optimization. *IEEE Trans. Geosci. Remote Sens.* 2008, 46, 3792–3803. [CrossRef]
- Bao, N.; Lechner, A.M.; Fletcher, A.; Mellor, A.; Mulligan, D.; Bai, Z. Comparison of relative radiometric normalization methods using pseudo-invariant features for change detection studies in rural and urban landscapes. J. Appl. Remote Sens. 2012, 6, 063578.
   [CrossRef]
- 65. Liu, T.; Yang, L.; Lunga, D. Change detection using deep learning approach with object-based image analysis. *Remote Sens. Environ.* **2021**, *256*, 112308. [CrossRef]
- Syariz, M.A.; Lin, B.Y.; Denaro, L.G.; Jaelani, L.M.; Van Nguyen, M.; Lin, C.H. Spectral-consistent relative radiometric normalization for multitemporal Landsat 8 imagery. *ISPRS J. Photogramm. Remote Sens.* 2019, 147, 56–64. [CrossRef]
- 67. Adams, R.; Bischof, L. Seeded region growing. IEEE Trans. Pattern Anal. Mach. Intell. 1994, 16, 641–647. [CrossRef]
- 68. Byun, Y.; Han, D. Relative radiometric normalization of bitemporal very high-resolution satellite images for flood change detection. *J. Appl. Remote Sens.* 2018, 12, 026021. [CrossRef]