

Article

kCCA Transformation-Based Radiometric Normalization of Multi-Temporal Satellite Images

Yang Bai ^{1,2}, Ping Tang ² and Changmiao Hu ^{2,*}

¹ University of the Chinese Academy of Sciences, Beijing 100049, China; baiyang@radi.ac.cn

² Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China; tangping@radi.ac.cn

* Correspondence: hucm@radi.ac.cn; Tel.: +86-159-0116-4748

Received: 14 February 2018; Accepted: 06 March 2018; Published: 10 March 2018

Abstract: Radiation normalization is an essential pre-processing step for generating high-quality satellite sequence images. However, most radiometric normalization methods are linear, and they cannot eliminate the regular nonlinear spectral differences. Here we introduce the well-established kernel canonical correlation analysis (kCCA) into radiometric normalization for the first time to overcome this problem, which leads to a new kernel method. It can maximally reduce the image differences among multi-temporal images regardless of the imaging conditions and the reflectivity difference. It also perfectly eliminates the impact of nonlinear changes caused by seasonal variation of natural objects. Comparisons with the multivariate alteration detection (CCA-based) normalization and the histogram matching, on Gaofen-1 (GF-1) data, indicate that the kCCA-based normalization can preserve more similarity and better correlation between an image-pair and effectively avoid the color error propagation. The proposed method not only builds the common scale or reference to make the radiometric consistency among GF-1 image sequences, but also highlights the interesting spectral changes while eliminates less interesting spectral changes. Our method enables the application of GF-1 data for change detection, land-use, land-cover change detection etc.

Keywords: kernel version of canonical correlation analysis; radiometric normalization; Gaofen-1 satellite; nonlinear invariant features

1. Introduction

There is an increasing demand for image sequence analysis, because of the growing use of multi-sensor and multi-temporal remote sensing data to monitor the land-use and land-cover change (LUCC) and climate change, as well as to analyze the global resource environment [1–4]. A satellite image sequence is usually constructed from multi-temporal images generated by one common sensor or multiple sensors. Radiometric normalization is a fundamental method used in the pre-processing of satellite image sequence analysis, especially in the pre-processing of change detection [5]. Radiometric normalization can directly make use of the pixel values of an image to establish the corresponding transformation equation for each multi-spectral band in multi-temporal remote sensing data, without the request of any other parameters such as the atmospheric conditions when the remote sensing data obtained [6]. In such a context, radiometric normalization is called spectral alignment [7]. Radiometric normalization builds not only the common radiometric scale/reference but also the radiometric consistency among an image sequence.

Most radiometric normalization methods make use of the pseudo-invariant feature (PIF) to obtain the common scale among an image sequence (PIF)-based [5,8,9]. It is well-known that these invariant features—the targets whose spectral reflectance does not change over time, are necessary

for the atmospheric normalization of temporal image sequences, and thus they are the basis of image sequence analysis [10]. All the PIF-based methods are linear, and they minimize the image radiometric differences caused by different acquisition conditions, such as atmospheric conditions, earth-sun distance, detector calibration, illumination angles, viewing angles and sensor oscillation [5,6,11–13].

The core problem of PIF-based methods is the selection of appropriate PIF points [13]. From simple to sophisticated, various PIF-based methods were developed for image normalization, such as the Ridge method [2,14–16], Simple image regression [4], Dark set-Bright set (DB) [17] and Automatic Scattergram Control Regression (ASAR) [18,19], etc. Specifically, Canty and Nielsen developed a powerful and widely used method, the iteratively reweighted multivariate alteration detection (iRMAD) transformation, which is invariant to linear and affine scaling. One advantage of this method is that only one parameter needs to be specified, i.e., the chi-squared percentile [20–23].

However, such a linear radiometric normalization cannot eliminate the nonlinear spectral differences, especially the surface reflectance differences [24]. In fact, beside PIFs, there are another two kinds of features/targets in image sequences. One kind of features/targets is the feature whose spectral reflectance changes regularly or expectably over time, such as vegetation growth. The other kind of features/targets is the feature whose spectral reflectance changes irregularly over time, such as city expansion and human activity trails. When these irregularly changed features are analyzed in an image sequence, the regularly changed features are the disturbances that need to be eliminated or attenuated by radiometric normalization. The changes caused by the growth of the vegetation are the most typical non-linear changes in radiometric value of image sequence. Especially, they may induce serious disturbance when the human induced changes are detected. Therefore, different spectral changes need different radiometric normalization methods. If the change detection involves the vegetation variation which needs to be preserved and enhanced, the impact from other features' changes should be eliminated as much as possible.

Eliminating the effects of the regularly nonlinear spectral differences caused by seasonal variation of natural objects requires a nonlinear radiometric normalization method. Histogram matching is a commonly used nonlinear method, but it only uses global statistical information. Thus, it is less flexible and so difficult to adapt to specific applications [4,25].

In this paper, we propose a novel method of nonlinear radiometric normalization to extract nonlinear relationship between multi-temporal image-pairs. The proposed method is based on the kernel version of canonical correlation analysis (kCCA) [1,26–31], that consists of two steps. First, the image-pairs are quantitatively evaluated to extract the unchanged pixels whose Features are Nonlinear Invariant in kernel space (NIFs). Second, the nonlinear transformation of image-pairs is obtained from the nonlinear regression of the NIFs extracted previously. We will analyze the characteristics of NIFs distribution, nonlinear transformations and radiometric normalization results from different temporal image-pairs and compare the results with PIFs-based (CCA-based) method and histogram matching method. The PIFs-based method is derived from CCA of the spectral alignment of linear transformations, and the histogram matching is derived from statistical distribution models.

The paper is organized as follows: In Section 2, we describe the principle of the kernel version of CCA, as well as the selection of NIFs. Section 3 presents the test data and the experimental results, Section 4 is the discussion about the results. Section 5 we give the conclusion of the paper.

2. Relative Radiometric Normalization Based on kCCA Transformation

The proposed method contains two steps: the extraction of NIFs using kCCA and the use of NIFs to derive radiometric normalization. In this section, we review the kCCA transformation and how to extract NIFs based on it.

2.1. kCCA Transformation and NIFs Extraction

kCCA is the kernel version of CCA [7]. CCA is a multivariate feature extraction method that aims at finding the rotation of two sets of variables that maximizes their joint correlation [32]. In

principle, CCA tends to select a group of representative indicators (linear combination of variables) from the two sets of random variables. These indicators express the correlation between the two sets of variables.

kCCA introduces the theory of kernel functions into CCA, which maps data from a low-dimension to a high-dimensional feature space. Each kernel function converts an n-dimensional inner product in low-dimensional space into an m-dimensional inner product in high-dimensional space. Let $\mathbf{x}, \mathbf{z} \in R^n$, the nonlinear function $\phi(\mathbf{x})$ converts R^n to R^m : $\mathbf{x} \rightarrow \phi(\mathbf{x})$ ($n \ll m$), using the kernel function:

$$\mathbf{k}(\mathbf{x}, \mathbf{z}) = \langle \phi(\mathbf{x}), \phi(\mathbf{z}) \rangle, \tag{1}$$

\langle, \rangle is the inner product, and $\mathbf{k}(\mathbf{x}, \mathbf{z})$ is a kernel matrix [33,34]. It is not necessary to know the form and parameters of the nonlinear transformation function $\phi(\mathbf{x})$ in advance. The kCCA method maps the target image and the reference image into high dimensional space in which their linear combination can be obtained. The canonical variable in high dimensional is defined as follows,

$$\mathbf{U} = \mathbf{c}^T \phi_x(\mathbf{x}), \tag{2}$$

$$\mathbf{V} = \mathbf{d}^T \phi_z(\mathbf{z}), \tag{3}$$

where \mathbf{x} represents the target image matrix. Each row of \mathbf{x} is a sample vector, and thus \mathbf{x} has size $n \times N$ where n is the number of bands and N is the number of sample points of the image-pair. \mathbf{x} can be rewritten as $\mathbf{x}_{n \times N} = (\mathbf{x}_1, \mathbf{x}_2 \dots \mathbf{x}_N)$ by column vectors \mathbf{x}_i . \mathbf{z} represents the reference image matrix and similarly \mathbf{z} can be rewritten as $\mathbf{z}_{n \times N} = (\mathbf{z}_1, \mathbf{z}_2 \dots \mathbf{z}_N)$. \mathbf{c} and \mathbf{d} are constant vectors of R^m . \mathbf{U} denotes the linear combination of \mathbf{x} in the high-dimensional feature space; and \mathbf{V} denotes the linear combination of \mathbf{z} in the high-dimensional feature space. The workflow of kCCA is shown in Figure 1. Similar to the solving process of linear correlation analysis, the first step is to solve \mathbf{c} and \mathbf{d} so that the correlation coefficients of the combined variables \mathbf{U} and \mathbf{V} are maximized.

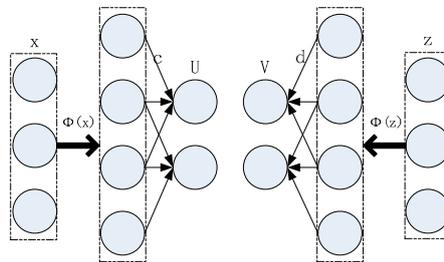


Figure 1. The workflow of kernel CCA.

To solve the vectors \mathbf{c} and \mathbf{d} in a lower dimensional space, \mathbf{c} and \mathbf{d} (of length m) are projected onto the vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ (of length n) with a lower dimension.

$$\mathbf{c} = \sum_i \alpha_i \phi_x(\mathbf{x}_i), \tag{4}$$

$$\mathbf{d} = \sum_i \beta_i \phi_z(\mathbf{z}_i), \tag{5}$$

where $\boldsymbol{\alpha}_{n \times N} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \dots \boldsymbol{\alpha}_N)$, $\boldsymbol{\beta}_{n \times N} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2 \dots \boldsymbol{\beta}_N)$. Then the problem of solving the high-dimensional vectors \mathbf{c} and \mathbf{d} is simplified as a problem of solving the low-dimensional vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. By substituting into Equations (2) and (3), \mathbf{U} and \mathbf{V} can be rewritten as:

$$\mathbf{U} = \sum_i \alpha_i \langle \phi_x(\mathbf{x}_i), \phi_x(\mathbf{x}) \rangle, \tag{6}$$

$$\mathbf{V} = \sum_i \beta_i \langle \phi_z(\mathbf{z}_i), \phi_z(\mathbf{z}) \rangle, \tag{7}$$

According to Equation (1) of the kernel function and the description followed, the kernel matrix can be defined,

$$\mathbf{k}_x(i, j) = k(\mathbf{x}_i, \mathbf{x}_j) = \langle \phi_x(\mathbf{x}_i), \phi_x(\mathbf{x}_j) \rangle, \tag{8}$$

$$\mathbf{k}_z(i, j) = k(\mathbf{z}_i, \mathbf{z}_j) = \langle \phi_x(\mathbf{z}_i), \phi_x(\mathbf{z}_j) \rangle, \tag{9}$$

Using the definition of kernel matrix in Equations (8) and (9), we obtain that the variance and covariance of \mathbf{U} and \mathbf{V} are as follows:

$$\text{Var}(\mathbf{U}) = \mathbf{c}^T \text{Var}(\boldsymbol{\varphi}_x(\mathbf{x})) \mathbf{c} = \boldsymbol{\alpha}^T \mathbf{k}_x \mathbf{k}_x \boldsymbol{\alpha}, \tag{10}$$

$$\text{Var}(\mathbf{V}) = \boldsymbol{\beta}^T \mathbf{k}_z \mathbf{k}_z \boldsymbol{\beta}, \tag{11}$$

$$\text{Cov}(\mathbf{U}, \mathbf{V}) = \boldsymbol{\alpha}^T \mathbf{k}_x \mathbf{k}_z \boldsymbol{\beta}, \tag{12}$$

$\text{Var}(\)$ is the variance and $\text{Cov}(\)$ is the covariance. The key step in the kCCA transformation is to solve the corresponding $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ when the correlation coefficient is maximized. According to the above formula Equations (10)–(12), we obtain,

$$\rho = \max_{\boldsymbol{\alpha}, \boldsymbol{\beta}} \frac{\text{Cov}(\mathbf{U}, \mathbf{V})}{\sqrt{\text{Var}(\mathbf{U})\text{Var}(\mathbf{V})}} = \max_{\boldsymbol{\alpha}, \boldsymbol{\beta}} \frac{\boldsymbol{\alpha}^T \mathbf{k}_x \mathbf{k}_z \boldsymbol{\beta}}{\sqrt{\boldsymbol{\alpha}^T \mathbf{k}_x \mathbf{k}_x \boldsymbol{\alpha} \boldsymbol{\beta}^T \mathbf{k}_z \mathbf{k}_z \boldsymbol{\beta}}}, \tag{13}$$

where ρ is the Pearson correlation coefficient. The problem can be simplified as an optimization problem,

$$\max_{\boldsymbol{\alpha}, \boldsymbol{\beta}} \boldsymbol{\alpha}^T \mathbf{k}_x \mathbf{k}_z \boldsymbol{\beta} \quad \text{s. t. : } \boldsymbol{\alpha}^T \mathbf{k}_x \mathbf{k}_x \boldsymbol{\alpha} = \boldsymbol{\beta}^T \mathbf{k}_z \mathbf{k}_z \boldsymbol{\beta} = 1, \tag{14}$$

This optimization problem has the equality constraints that can be solved by the Lagrange multiplier,

$$L = \boldsymbol{\alpha}^T \mathbf{k}_x \mathbf{k}_z \boldsymbol{\beta} - \frac{\rho_a}{2} (\boldsymbol{\alpha}^T \mathbf{k}_x \mathbf{k}_x \boldsymbol{\alpha} - 1) - \frac{\rho_\beta}{2} (\boldsymbol{\beta}^T \mathbf{k}_z \mathbf{k}_z \boldsymbol{\beta} - 1), \tag{15}$$

By taking derivatives with respect to $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, we obtain,

$$\frac{\partial L_{\boldsymbol{\alpha}, \boldsymbol{\beta}}}{\partial \boldsymbol{\alpha}} = 0 \implies \mathbf{k}_x \mathbf{k}_z \boldsymbol{\beta} = \rho_a (\varepsilon_1 \mathbf{k}_x \mathbf{k}_x + (1 - \varepsilon_1) \mathbf{k}_x) \boldsymbol{\alpha}, \tag{16}$$

$$\frac{\partial L_{\boldsymbol{\alpha}, \boldsymbol{\beta}}}{\partial \boldsymbol{\beta}} = 0 \implies \mathbf{k}_z \mathbf{k}_x \boldsymbol{\alpha} = \rho_\beta (\varepsilon_2 \mathbf{k}_z \mathbf{k}_z + (1 - \varepsilon_2) \mathbf{k}_z) \boldsymbol{\beta}, \tag{17}$$

ε_1 and ε_2 are both regularization coefficients. Together with the constraints, Equations (16) and (17) implies that $\rho_a = \rho_\beta$ and we define $\rho = \rho_a = \rho_\beta$. The above equation can be converted into the eigenvalue problem,

$$\mathbf{k}_{xx} \mathbf{k}_{zz} (\varepsilon_2 \mathbf{k}_{zz} \mathbf{k}_{zz} + (1 - \varepsilon_2) \mathbf{k}_{zz})^{-1} \mathbf{k}_{zz} \mathbf{k}_{xx} \boldsymbol{\alpha} = \rho^2 (\varepsilon_1 \mathbf{k}_{xx} \mathbf{k}_{xx} + (1 - \varepsilon_1) \mathbf{k}_{xx}) \boldsymbol{\alpha}, \tag{18}$$

$$\mathbf{k}_{zz} \mathbf{k}_{xx} (\varepsilon_1 \mathbf{k}_{xx} \mathbf{k}_{xx} + (1 - \varepsilon_1) \mathbf{k}_{xx})^{-1} \mathbf{k}_{xx} \mathbf{k}_{zz} \boldsymbol{\beta} = \rho^2 (\varepsilon_2 \mathbf{k}_{zz} \mathbf{k}_{zz} + (1 - \varepsilon_2) \mathbf{k}_{zz}) \boldsymbol{\beta}, \tag{19}$$

By solving the above equation, we get ρ and the corresponding canonical vectors \mathbf{U} and \mathbf{V} in Equations (2) and (3).

The non-linear NIFs can be chosen according to Nielsen’s [23,35] *MAD* method. The *MAD* method uses the components of \mathbf{U} and \mathbf{V} , i.e., U_i and V_i to define the variation,

$$MAD_i = U_i - V_i, \tag{20}$$

and we get,

$$\text{Cov}(U_i - V_i, U_j - V_j) = 2\delta_{ij}(1 - \rho), \tag{21}$$

$$\delta_{ij} = \begin{cases} 0 & i = j; \\ 1 & i \neq j; \end{cases} \tag{22}$$

The covariance matrix of *MAD* variables is a diagonal matrix:

$$\Sigma_{MAD} = \text{Cov}\{MAD\} = \begin{bmatrix} \gamma_1 & 0 & \cdots & 0 \\ 0 & \gamma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \gamma_N \end{bmatrix}. \tag{23}$$

The diagonal element γ_j is the variance of MAD_j , ordered decreasingly $\gamma_1 \geq \gamma_2 \geq \dots \geq \gamma_N$, which gives $\gamma_i = 2(1 - \rho)$. Using the concept of contribution proposed by Nielsen, the contribution of MAD_i is,

$$\omega_i = \frac{\gamma_i}{\sum_{i=1}^N \gamma_i} \quad i = 1, 2, \dots, N, \quad (24)$$

The formula Equation (24) indicates that the variable carries information of the ratio of the variation. Since the accumulated contribution of the first k variables ($\sum_{i=1}^k \omega_i$) is called the accumulative contribution ratio, the last few variables in $U_i - V_i$ contain the unchanged pixels between the image-pair, and we can select them as NIFs. If we denote,

$$Z = \sum_{i=1}^N \frac{(MAD_i)^2}{\gamma_i}, \quad (25)$$

the points that satisfy the following probability conditions are selected as NIFs:

$$Pr(\text{nochange}) = 1 - P_{\chi^2, N}(Z), \quad (26)$$

where $P_{\chi^2, N}()$ refers to a chi-square test of the N degrees of freedom. $Pr(\text{no change})$ is used to select NIFs in the kernel space. Pixels above the fixed threshold τ , i.e., $Pr(\text{no change}) > \tau$, are NIFs. Normally, the threshold τ needs to be larger than 0.9 to masked out the water, cloud pixels, shadow points in the NIFs. The value of the threshold τ affects the number of the extracted NIFs, as discussed in detail in Section 3.4.2.

2.2. Fitting Non-Linear Transformation for Radiometric Normalization

Based on NIFs, the non-linear transformation can be derived by means of the least squares regression and the key is the selection of transformation. From Section 2.1, we know that

$$\text{NIFs} = \{(x_{i_k}, y_{i_k}), 1 - P_{\chi^2, N}(Z) > \tau\}, \quad (27)$$

If we set the subspace $U_{x_{i_k}} = \{f, f = \sum_k a_k k(x_{i_k}, x), (x_{i_k}, y_{i_k}) \in \text{NIFs}\}$ as the reproducing kernel Hilbert space of $k(x_{i_k}, x)$, and $V_{y_{i_k}} = \{g, g = \sum_k b_k k(y_{i_k}, y), (x_{i_k}, y_{i_k}) \in \text{NIFs}\}$ as the reproducing kernel Hilbert space of $k(y_{i_k}, y)$, we can regard $x_{i_k} \in U_x, y_{i_k} \in V_y$.

The available kernels include:

$$k_{lin}(x, z) = \mathbf{x}^T \mathbf{z}, \quad (28)$$

$$k_{poly}(x, z) = (\gamma \mathbf{x}^T \mathbf{z} + c)^n, \quad (29)$$

$$k_{rbf}(x, z) = \exp(-\gamma \|x - z\|^2), \quad (30)$$

$$k_{sig}(x, z) = \tanh(\gamma \mathbf{x}^T \mathbf{z} + c). \quad (31)$$

In this paper, we choose polynomial kernel k_{poly} with $n = 3, \gamma = 1, c = 2$ as an example, and we use a more general representation,

$$\mathbf{y}(x) = a_0 + a_1 \mathbf{x} + a_2 \mathbf{x}^2 + \dots + a_n \mathbf{x}^n, \quad (32)$$

where \mathbf{x} is the pixel value of NIFs in the test data. We assume that $\hat{\mathbf{y}}(x_{i_k})$ is the fitted value of the above equation. \mathbf{y} is the corresponding pixel value of the reference image. The least square method is used to minimize $|\mathbf{y} - \hat{\mathbf{y}}(x_{i_k})|^2$ to solve the coefficients $\{a_0, a_1, a_2, a_3\}$ of the above polynomial Equation (32). Once the coefficients are resolved, we use Equation (32) do the radiometric normalization.

3. Data and Results

3.1. Test Data

In this work, we use the subsets of 2025×2205 pixels of the Gaofen-1(GF-1) data with a spatial resolution of 16m as the test set. GF-1 was launched by China on 26 April 2013. It carries two cameras with 2 m resolution (panchromatic) and 8 m resolution (multi-spectral), and four multi-spectral cameras with 16 m resolution. The band setting of the GF-1 multi-spectral sensor is similar to that of other high-resolution and wide-spectrum satellite sensors, such as QuickBird, IKONOS, and GeoEye-1, with blue, green, red, and near-infrared bands. We refer the parameters of payload of GF-1 to Table 1.

Table 1. Parameters of payload of GF-1.

Satellite Performance		Technical Capability
Satellite Orbit	type	Solar synchronous circular orbit
	Average orbit height	644.5 km
	descending/ascending nod sun-synchronous	10:30 a.m.
	regressive period	41 days
	Revisit/Coverage characteristic	Revisit: 4 days for 2/8 m camera under side sway Coverage:4 days for 16 m camera; 41 days for 2/8 m camera;
High resolution imaging	Spectrum/ μm	Panchromatic: 0.45–0.90 μm B1: 0.45–0.52 μm ; B2: 0.52–0.59 μm ; B3: 0.63–0.69 μm ; B4: 0.77–0.89 μm ;
	resolution	Panchromatic: better than 2 m Multispectral: better than 8 m
	Swath width (km)	>60
Wide imaging	Spectrum/ μm	B1: 0.45–0.52 μm ; B2: 0.52–0.59 μm ; B3: 0.63–0.69 μm ; B4: 0.77–0.89 μm ;
	resolution	Better than 16 m
	Swath width (km)	>800

To investigate the nonlinear transformation effects derived from NIFs in high dimensional space, 6 GF-1 images obtained from 2013 to 2014 each of which contains all 4 seasons (see Table 2 for details) are tested, as shown in Figure 2. The geographical area of the images is in the northern region of Beijing, China. Most of the land is covered by vegetation, man-made objects, water.

Table 2. Acquisition date of test data.

Image Name	Image Date
Image0705	5 July 2013
Image0717	17 July 2013
Image1002 _{reference}	2 October 2013
Image1214	14 December 2013
Image0330	30 March 2014
Image0403	3 April 2014

The preprocessing of test data includes radiometric calibration, ortho-rectification, and image registration. Here radiometric calibration converts the DN value to the top of atmosphere reflectance (TOA). The GF-1 image acquired in 2 October 2013 is cloud free and has good visibility. There are visible seasonal differences between the image and the others. It is chosen as the reference in geometric space and radiometric space. Each of other test images is at first registered to it

geometrically, and then GF-1 and the test image are used as the image-pair to extract NIFs and do radiometric normalization.

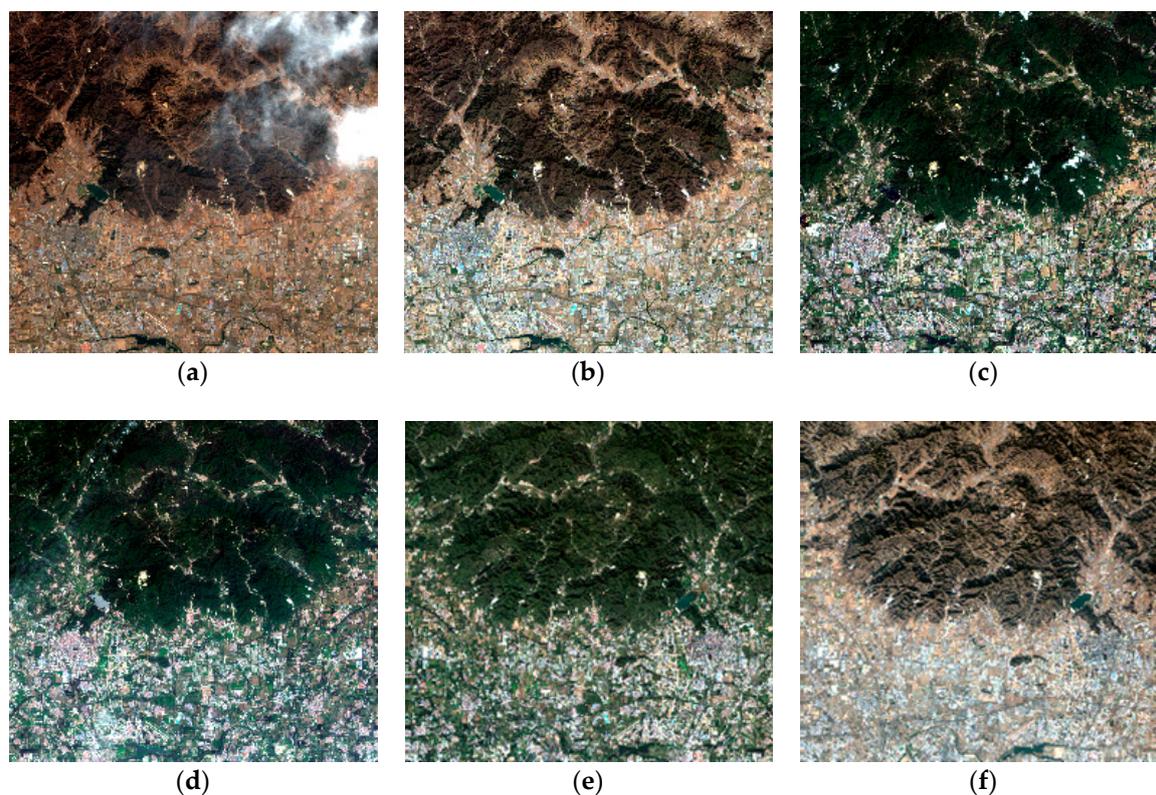


Figure 2. The GF-1 images of the test area (Beijing, China). The GF-1 data, (a–f) are image0330, image 0403, image0705, image0717, image1002 and image1214, acquired in 30 March, 3 April, 5 July, 17 July, 2 October, 14 December 2013.

3.2. NIFs Distribution Map

The parameters of the kCCA-based method used to extract NIFs include the regularization parameter $\varepsilon_1 = 0.0001$ and threshold of $Pr(\text{nochange}) \tau = 99\%$.

The distribution map of the extracted NIFs is indicated in Figure 3, where the NIFs are represented by green points. The total numbers of NIFs in image 0330, image 0403, image 0705, image 0717 and image 1214 are 31027, 14839, 124071, 11466 and 10078, respectively. Most of NIFs locate in the vegetation area. The proposed method can extract much more NIFs than that of CCA-based method, which only extracts 2432, 2382, 1014, 2074 and 1069 PIFs, respectively. Here the points with a normalized vegetation index (NDVI) larger than 0.2 are considered as vegetation targets. Figure 4 shows the ratio of NIFs in vegetation area to the total number of NIFs. The blue one in pie chart represents the percentage of NIFs in vegetation area. The ratio of NIFs in vegetation area are 55.1%, 88.5%, 83.4%, 92.1% and 78.4%, respectively.

According to the distribution map of NIFs, kCCA can select more NIFs than CCA, which indicates that the spectral variation of the corresponding points between the image-pair are mostly nonlinear. Moreover, since most NIFs are distributed in the vegetation area, it indicates that the spectral changes caused by vegetation growth over time are regular and predictable, and thus they can be detected by NIFs.

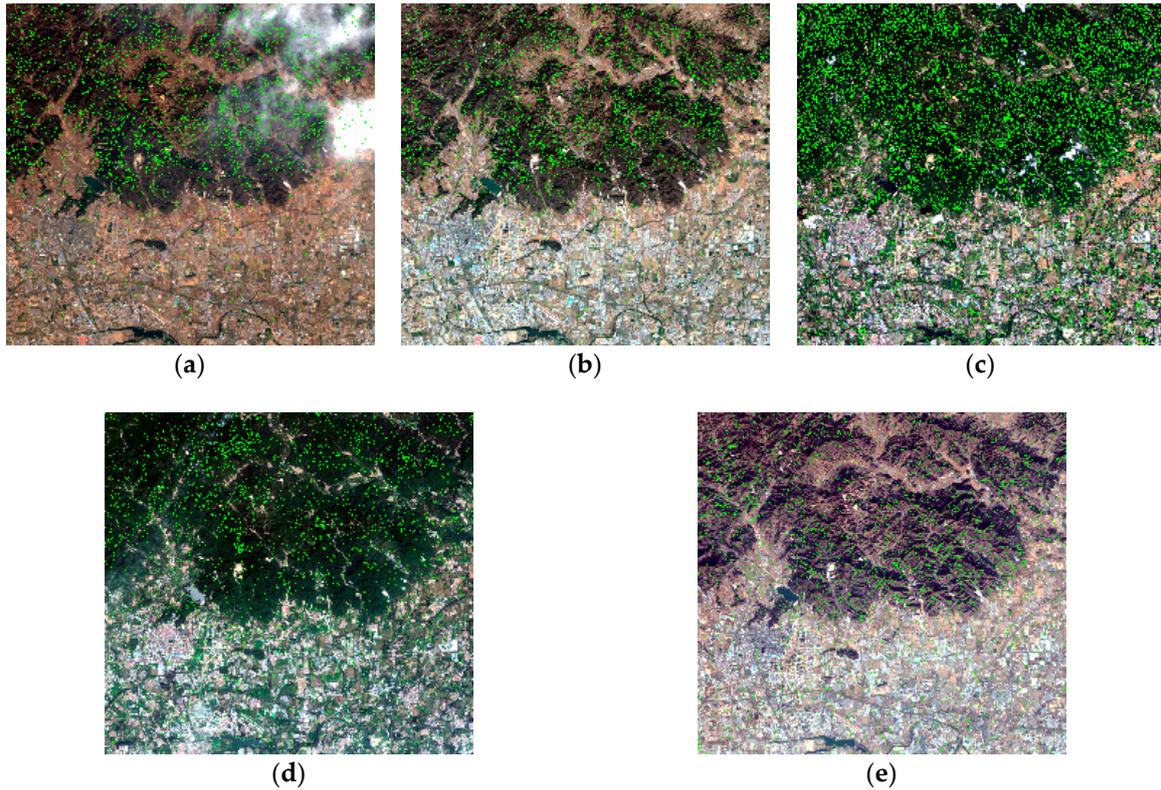


Figure 3. The distribution map of NIFs selected by the kCCA transformation. (a–e) are, in order, image0330, image0403, image0705, image0717 and image1214. The green points are the NIFs of each image.

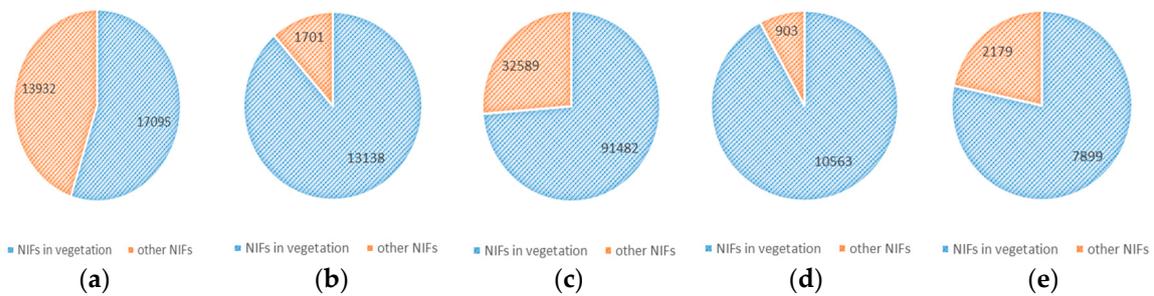


Figure 4. The ratio of NIFs in vegetation area to the total number of NIFs. (a–e) refer to the result of image0330, image0403, image0705, image0717 and image1214, respectively. The blue represents the NIFs in vegetation area and the yellow is the other NIFs.

3.3. Derive Nonlinear Transformations from NIFs

Figure 5 is the scatter plot of the density of the NIFs in blue band of test data. The horizontal axis is the TOA value of the test data that linearly stretched from 0–1 to 0–255. The vertical axis is the corresponding value of the reference image. This density map apparently shows the non-linearity of NIFs between image-pairs.

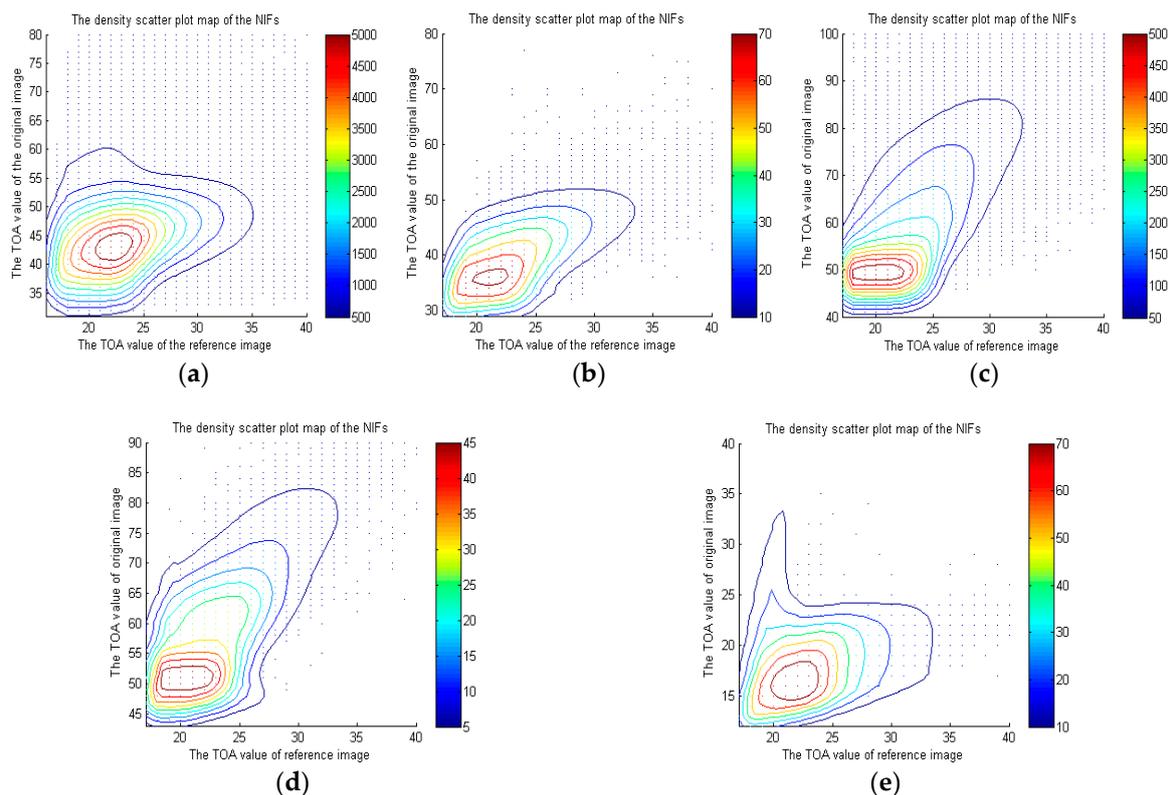


Figure 5. The density scatter plot map of the NIFs. The horizontal axis in the figure is the TOA value of the original image that linearly stretched from 0–1 to 0–255, and the vertical axis is that of the reference image. (a–e) refer to the result of image0330, image0403, image0705, image0717 and image1214, respectively.

Table 3 shows the fitting coefficients and regression error of Equation (27). M_r is the average of the residuals and S_r is the standard deviation of the residuals. As indicated in Table 3, the absolute values of the cubic coefficients of image0330, image0403 and image1214 are significantly larger than those of image0705 and image0717.

Table 3. The coefficients and error ratios of regression equation a_0 , a_1 , a_2 and a_3 represent the coefficients of the constant term, the linear term, the quadratic term, and the cubic term, respectively. M_r is the average of the difference between the function value and the expected value, and S_r is the standard deviation.

		a_3	a_2	a_1	a_0	M_r	S_r
Image 0330	Band1	18.3284	-13.3953	3.1748	-0.1532	0.019194	0.000002
	Band2	21.5494	-14.4433	3.1393	-0.1489	0.019906	0.000003
	Band3	10.9105	-7.7441	1.8255	-0.0855	0.02473	0.000004
	Band4	1.4785	-1.7238	0.7561	0.072	0.035371	0.000007
Image 0403	Band1	-7.367	5.2422	-0.5349	0.0748	0.011784	0.000012
	Band2	-3.0287	2.4336	-0.0403	0.0372	0.01312	0.000013
	Band3	-1.8955	2.0456	-0.099	0.0288	0.018398	0.000015
	Band4	7.3705	-6.866	2.1832	-0.046	0.030512	0.000037
Image 0705	Band1	-0.7959	0.4918	0.194	0.0297	0.010231	0.000024
	Band2	-0.4582	0.0524	0.3273	0.0071	0.011725	0.00002
	Band3	0.173	-0.4517	0.4073	0.0041	0.015785	0.000023
	Band4	0.0878	-0.2848	0.3699	0.0379	0.0341	0.000083
Image 0717	Band1	-2.2136	2.4316	-0.445	0.0895	0.008238	0.000082
	Band2	-0.7765	0.7466	0.1471	0.0187	0.009817	0.00005

	Band3	-0.0018	0.0113	0.3381	0.0061	0.013293	0.000033
	Band4	0.6642	-1.2344	0.9188	-0.0681	0.031373	0.000042
Image 1214	Band1	-87.0617	28.3288	-1.1949	0.0675	0.010096	0.000023
	Band2	-44.4402	12.7071	0.2623	0.02687	0.011614	0.000023
	Band3	-41.5801	14.5948	-0.1363	0.0267	0.017183	0.000032
	Band4	13.8154	-9.1873	2.0127	0.07651	0.026603	0.000044

Figure 6 shows the transformation function between test image-pair. The red, green, purple, black, and cyan-blue curve denotes the transformation function between the image-pair (image0330, image1002), (image0403, image1002), (image0705, image1002), (image0717, image1002), and (image1214, image1002), respectively. (a–d) are the results fitted by NIFs, (e–h) are results fitted by PIFs. From (a–h), we find that the nonlinear relationship fitted by NIFs between image-pairs increases with the time span. Similarly, the absolute value of the cubic coefficient in Table3 is relatively larger and the slope of the linear relationship fitted by PIFs also increase with the time span.

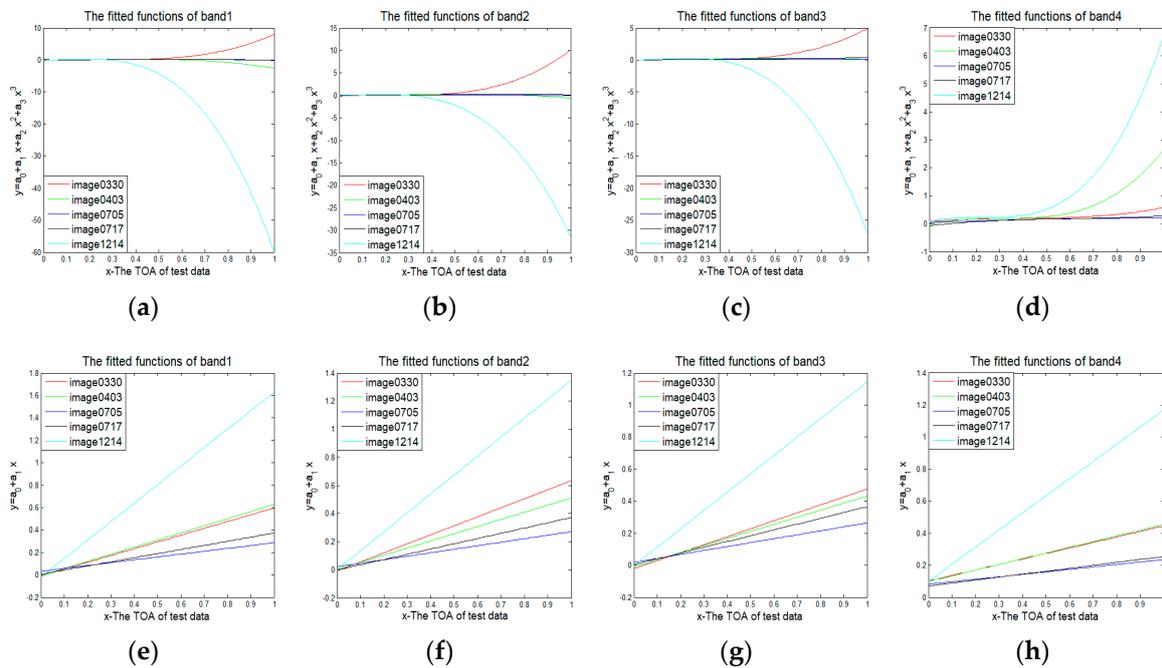
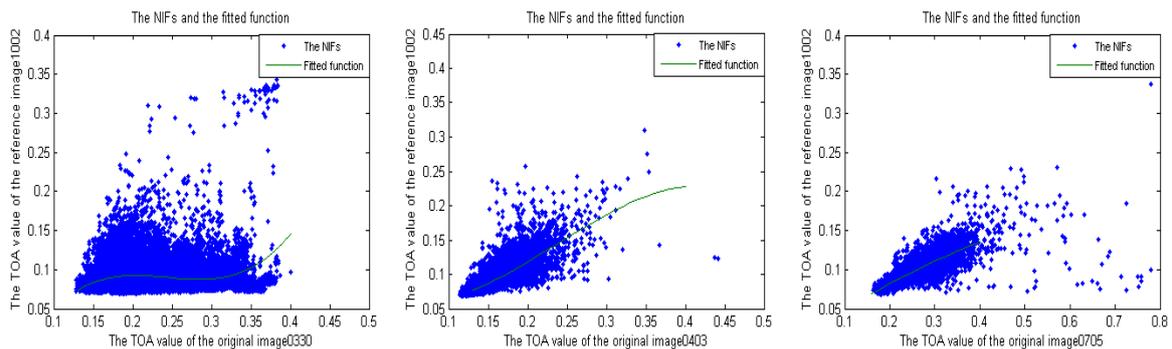


Figure 6. The fitted functions of the four band for the 5 test GF-1 image-pairs. (a–d) refer to the results of band1, band2, band3 and band4 based on the proposed method. (e–h) are the linear regression equations of the PIFs-based method. — is the fitted function of image0330 and the reference, — is that of image0403 and the reference, — is that of image0705 and the reference, — is that of image0717 and the reference, — is that of image1214 and the reference.

Figure 7a–e shows a superimposed display of the NIFs and the fitted curve of NIFs and Figure 7f–j shows that of the PIFs and the curve of the fitted functions based on the PIFs described in Figure 6.



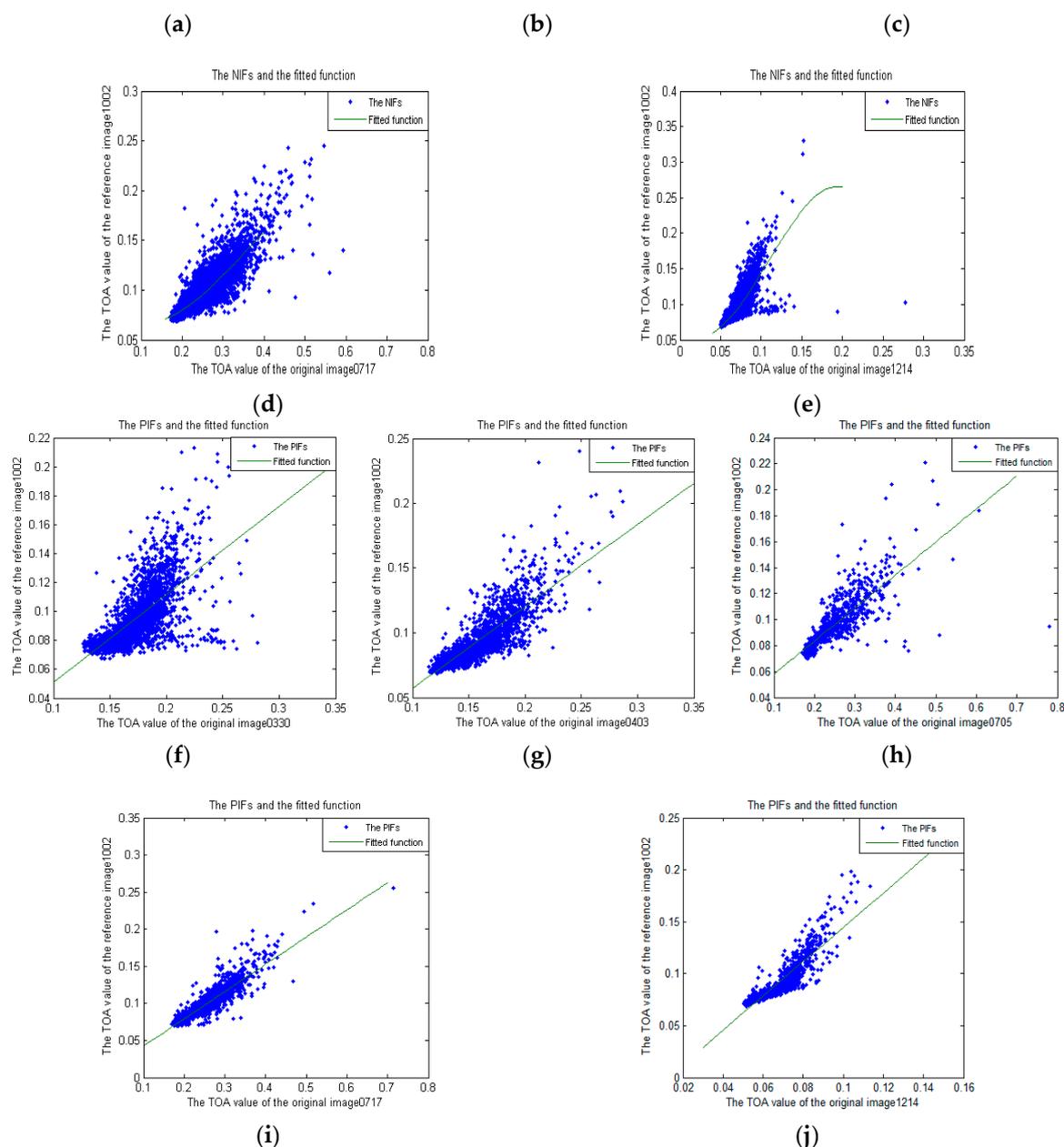
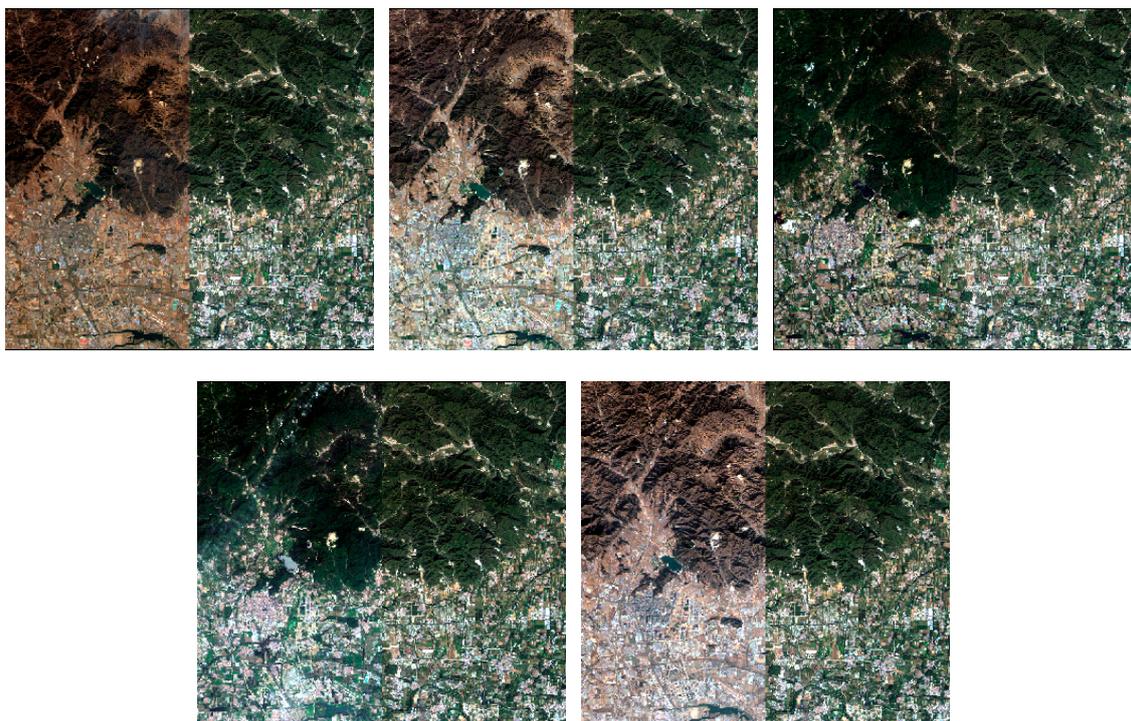


Figure 7. The superimposed display of the NIFs and PIFs distribution map and the curve of the fitted functions based on the NIFs and PIFs described above. The horizontal axis in the figure is the TOA value of the original image, and the vertical axis is that of the reference image. (a–e) refer to the NIFs results of image0330, image0403, image0705, image0717 and image1214, respectively. (f–j) are that of PIFs results.

3.4. Radiometric Normalization Results

According to the transformation given in Table 3, the image can be radiometrically normalized. Figure 8 shows the radiation normalization results of the CCA, kCCA and the histogram matching. Each image in Figure 8 is a mosaic image of the reference image (right) and the normalization result (left). Figure 8a shows the original mosaic images and the reference, in which the spectrums of the image-pairs are visually different and in particular, the spectral difference of vegetation increases with the time span between the image-pair. The normalization results of the PIFs-based method (Figure 8b) show that the spectral differences still exist, especially in the vegetation area. The normalization results of the kCCA-based method (Figure 8c) show the spectral differences are significantly reduced, even though the vegetation in spring of image0330 and in winter of image1214 shows similar spectral

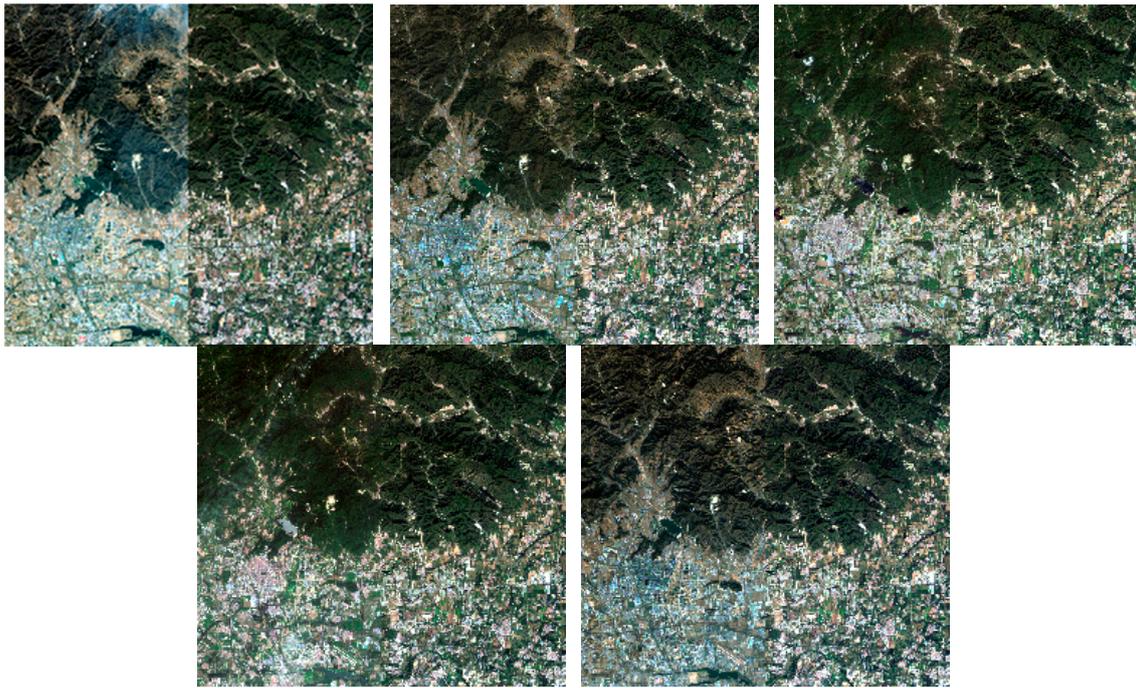
characteristics with the vegetation in early autumn of image1002. The results of the histogram matching (Figure 8d) show that spectral differences are well reduced.



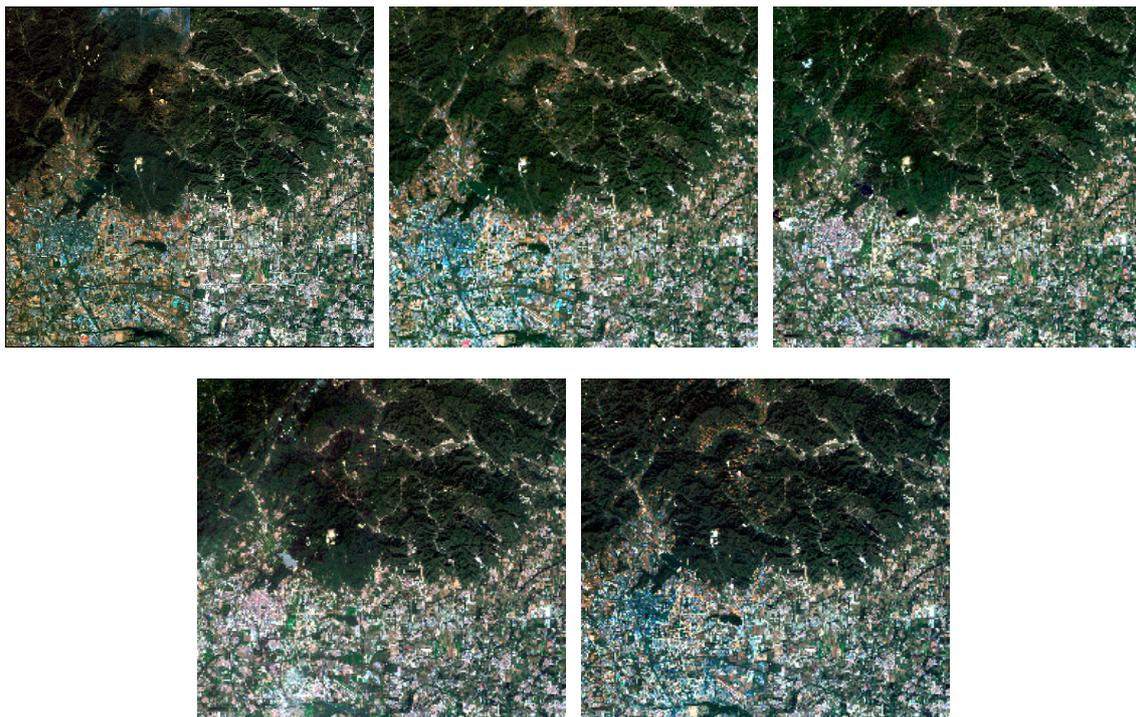
(a)



(b)



(c)



(d)

Figure 8. The comparison of the results of radiometric normalization based on the kCCA transformation, the CCA method and the histogram matching. Each image is a mosaic image. The left one is the corrected image, and the right one is the reference image. (a) is the test data and the reference image; (b) is the results of the CCA algorithm and the reference images; (c) is the results based on NIFs extracted by the kCCA transformation; (d) is the results of the histogram matching.

3.4.1. Clouds Pixels in the Image

Towards a quantitative comparison of the proposed radiometric normalization with the CCA-based method and with the histogram matching method, we first discuss the effects of cloud pixels and threshold τ for NIFs extraction and radiometric normalization transformation. The effect caused

by cloud pixels is an unavoidable problem that must be considered in remote sensing image radiometric normalization. The usual approach is to mask out the cloud pixels from an image-pair. We use the automatic threshold method to do cloud detection [36]. In fact, if cloud pixels are falsely selected as NIFs, bias will be induced into transformation. The image0330 with clouds is randomly selected as a test image. Following the NIFs extraction steps in Section 3.2, if the cloud pixels have not been masked out, some cloud pixels are selected as NIFs, as shown in the distribution map of NIFs (Figure 3a).

Figure 9a,b show the radiometric normalization results of image0330, showing NIFs with and without cloud pixels. Visually, the brightness of the result including cloud pixels is relatively lower. This can be explained by the normalization transformation derived from NIFs. Figure 9c–f show the histogram of image0330, where we can see that the pixel values are mainly concentrated in [0.1, 0.4]. Figure 9g–j are the fitting curves of the four bands of the image-pair (image0330, image1002). The red and green curves denote the fitted curve from NIFs with and without the cloud pixels, respectively. According to the histogram, we notice that the TOA values of NIFs are mostly distributed between 0.1 and 0.4. In this interval, the fitting values of the NIFs without cloud pixels in Figure 9 are obviously larger than that of the NIFs with cloud pixels, while out of this interval, the fitting values of the NIFs without cloud pixels is smaller. The resulting normalized image of NIFs with cloud fitting is darker than that without cloud fitting.

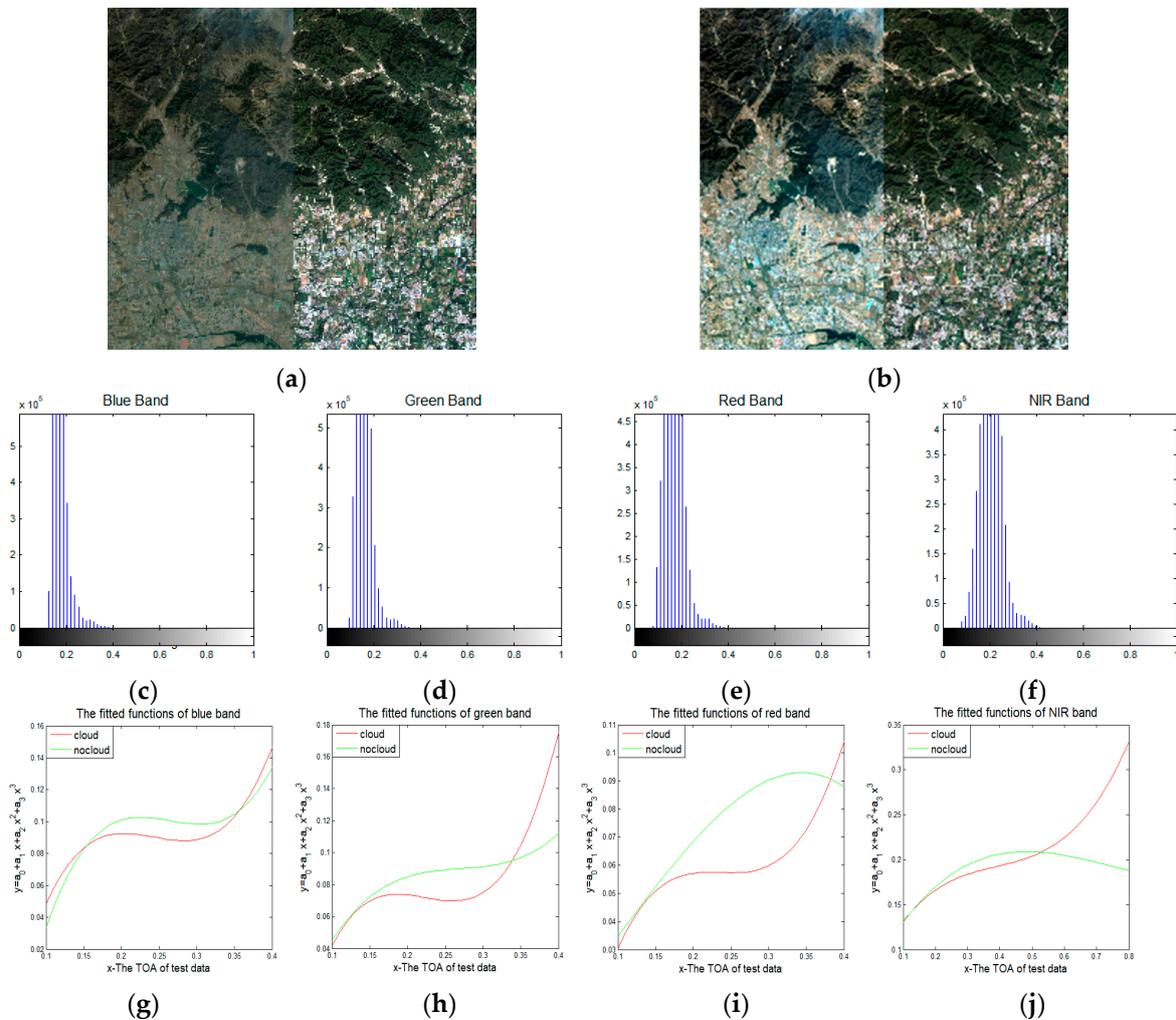


Figure 9. The results of the radiometric normalization of image0330 based on two NIFs set and the fitted curve of the two normalized results. (a) is the result based on all NIFs (including the cloud-NIFs), (b) is the result based on NIFs without cloud pixels, (c–f) are histogram of image0330. (g–j) are the fitted curve of four bands of image-pair (image0330 and image1002). — denotes the fitted curve using all NIFs including cloud pixels. — is the fitted curve using NIFs without the cloud pixels.

3.4.2. The Threshold Parameters τ for NIFs Extraction

We use the formula $Pr(nochange) = 1 - P_{\chi^2, N}(Z) > \tau$ to extract the NIFs. The threshold τ directly determines the number of NIFs selected. The number of NIFs increases with the value of τ . In Section 3, $\tau = 0.99$ is used. Here we set the threshold to 0.95, 0.90 and 0.85 respectively to investigate the changes of NIFs distribution map and the normalization function.

Table 4 is the number of NIFs with different τ of the image pair (image0717, image1002). It shows that the total number of NIFs and the number of NIFs in the vegetation area increase as the τ decreases, and the ratio change of NIFs in vegetation area changes is $<0.5\%$, which is relatively stable. Figure 10a–c show the distribution maps of NIFs with different threshold τ . Figure 10d–g show the histogram of image0717. We can see that the pixel values of the red, green, and blue bands are mainly concentrated in $[0.1, 0.4]$, and the pixel values of near infrared band are distributed between 0.1 and 0.8. Figure 10h–k are the fitting curves of the four bands using the above NIFs. It shows that the fitting curves for different thresholds τ are very similar in the corresponding interval range of the TOA values.

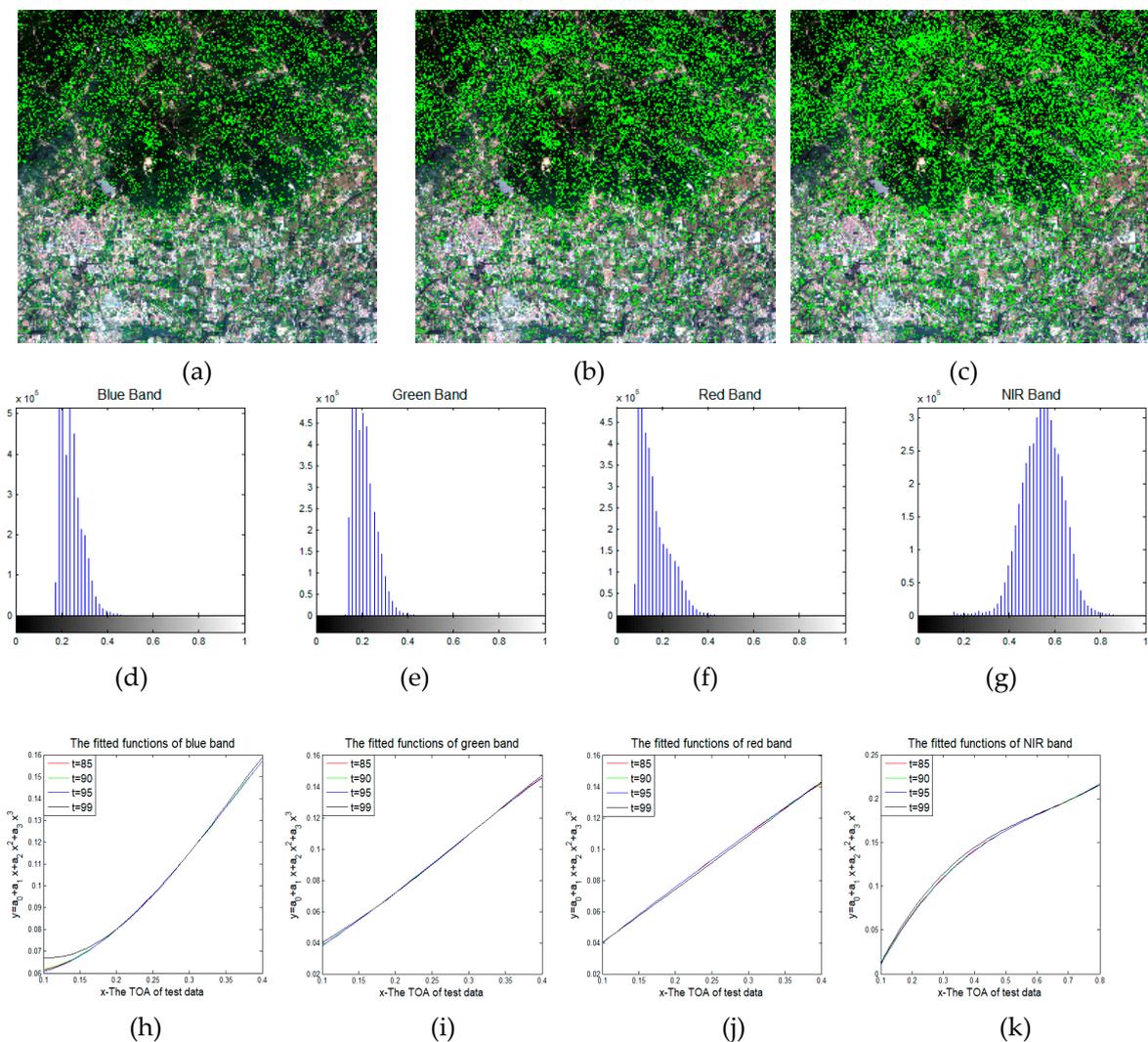


Figure 10. The distribution map of NIFs based on different threshold τ . The data is image0717. (a–c) are the distribution maps when $\tau = 0.95$, $\tau = 0.90$ and $\tau = 0.85$. (d–g) are histogram of four bands of image0717. (h–k) are fitting curves of four bands based on the above NIFs.

Table 4. The number of NIFs with different τ .

	$\tau = 0.99$	$\tau = 0.95$	$\tau = 0.90$	$\tau = 0.85$
Total number of NIFs	11466	27596	41421	53453
NIFs in vegetation area	10563	25319	38021	49097
Ratio of NIFs in vegetation area (%)	92.12	91.75	91.79	91.85

3.4.3. Quantitative Comparison of the Radiometric Normalization Results

We use the root mean square error (RMSE), Pearson correlation coefficient and histogram similarity to evaluate the similarity of the radiometric normalization results of proposed method with the CCA-based method and the histogram matching method.

RMSE is used to evaluate the similarity between the normalized image and the reference image. RMSE is defined as follows,

$$\text{RMSE}_k = \sqrt{\frac{1}{|scene|} \sum_{scene} (S_k - R_k)^2}, \quad (33)$$

where X_i is the pixel value of the normalized data, Y_i is the pixel value of the reference image, and $|scene|$ is the total number of pixels. The smaller the value of RMSE is, the higher the similarity of the two images is.

Pearson correlation coefficient is used to measure the linear correlation between the normalized image and the reference image. The Pearson correlation coefficient is defined as follows,

$$\rho = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2 \sum(Y_i - \bar{Y})^2}}, \quad (34)$$

where X_i denotes pixel value of the normalized data, \bar{X} denotes the average of the normalized data, Y_i denotes the corresponding pixel value of the reference data, and \bar{Y} denotes the average of the reference data.

Histogram similarity measured by the histogram correlation which is used to calculate the linear correlation between the two histograms of the normalized image and the reference image. The histogram correlation is defined as follows,

$$\rho(H1, H2) = \frac{\sum(H1_i - \bar{H1})(H2_i - \bar{H2})}{\sqrt{\sum(H1_i - \bar{H1})^2 \sum(H2_i - \bar{H2})^2}}, \quad (35)$$

where $H1$ is the histogram of the normalized data, $\bar{H1}$ is the mean of $H1$, $H2$ is the histogram of the reference data, $\bar{H2}$ is the mean of $H2$. The similarity positively correlates to the histogram correlation.

Table 5. RMSE is the RMSE of the test data and reference image, ρ is the Pearson correlation coefficients of the test data and reference image, ρ_h is histogram correlation of the test data and reference image.

	RMSE	ρ	ρ_h
<u>image pair (imag0330, image1002)</u>			
Band 1	0.09	0.32	0.04
Band 2	0.09	0.40	0.03
Band3	0.12	0.45	-0.05
Band 4	0.06	0.36	0.74
<u>image pair (imag0403, image1002)</u>			
Band 1	0.07	0.80	0.02
Band 2	0.07	0.77	0.05
Band3	0.10	0.71	-0.02
Band 4	0.06	0.51	0.69
<u>image pair (imag0705, image1002)</u>			
Band 1	0.15	0.80	0.07
Band 2	0.15	0.77	0.09
Band3	0.13	0.78	0.09
Band 4	0.42	0.40	0.24
<u>image pair (imag0717, image1002)</u>			
Band 1	0.15	0.89	0.07
Band 2	0.14	0.87	0.08
Band3	0.11	0.86	0.07

Band 4	0.38	0.46	0.23
image pair (imag1214, image1002)			
Band 1	0.03	0.81	0.43
Band 2	0.02	0.81	0.78
Band3	0.02	0.73	0.75
Band 4	0.11	0.67	-0.04

Tables 5 and 6 are the RMSE, Pearson correlation coefficient and histogram similarity between each multi-temporal image and the reference image before and after radiometric normalization. Here $RMSE^{CCA}$ is RMSE of the radiometric normalization results from the CCA-based method, $RMSE^{kCCA}$ is RMSE of the results from the kCCA-based transformation, and $RMSE^H$ is RMSE of the results from the histogram matching method. ρ^{CCA} , ρ^{kCCA} and ρ^H are Pearson correlation coefficients of the radiometric normalization results from the CCA-based, kCCA-based and the histogram matching methods, respectively. ρ_h^{CCA} , ρ_h^{kCCA} and ρ_h^H are histogram correlation of the radiometric normalization results from the CCA-based, kCCA-based and the histogram matching methods, respectively.

Table 6. $RMSE^{CCA}$ is the RMSE of the radiometric normalization results based on CCA algorithm, $RMSE^{kCCA}$ is the RMSE of the results based on NIFs extracted by the kCCA transformation, $RMSE^H$ is the RMSE of the results based on the histogram matching method. ρ^{CCA} , ρ^{kCCA} and ρ^H are the Pearson correlation coefficients of the radiometric normalization results based on the CCA algorithm, kCCA transformation and the histogram matching method. ρ_h^{CCA} , ρ_h^{kCCA} and ρ_h^H are histogram correlation of the radiometric normalization results based on the CCA algorithm, kCCA transformation and the histogram matching method.

	$RMSE^{CCA}$	$RMSE^{kCCA}$	$RMSE^H$	ρ^{CCA}	ρ^{kCCA}	ρ^H	ρ_h^{CCA}	ρ_h^{kCCA}	ρ_h^H
image pair (imag0330, image1002)									
Band 1	0.02	0.02	0.02	0.32	0.42	0.32	0.58	0.69	0.98
Band 2	0.02	0.02	0.02	0.40	0.47	0.39	0.77	0.78	0.97
Band3	0.03	0.02	0.03	0.45	0.47	0.41	0.73	0.75	0.98
Band 4	0.04	0.03	0.04	0.36	0.37	0.37	0.81	0.82	0.94
image pair (imag0403 image1002)									
Band 1	0.01	0.01	0.01	0.81	0.81	0.80	0.95	0.95	0.98
Band 2	0.01	0.01	0.01	0.78	0.78	0.77	0.93	0.93	0.99
Band3	0.02	0.02	0.02	0.71	0.72	0.71	0.86	0.93	0.99
Band 4	0.06	0.03	0.04	0.51	0.55	0.55	0.74	0.89	0.95
image pair (imag0705, image1002):									
Band 1	0.02	0.01	0.01	0.80	0.83	0.81	0.92	0.93	1.00
Band 2	0.02	0.01	0.01	0.77	0.80	0.78	0.98	0.94	0.91
Band3	0.03	0.01	0.02	0.78	0.79	0.77	0.97	0.91	0.88
Band 4	0.06	0.03	0.04	0.41	0.42	0.38	0.82	0.81	0.99
image pair (imag0717, image1002):									
Band 1	0.01	0.01	0.01	0.89	0.89	0.89	0.94	0.94	0.99
Band 2	0.01	0.01	0.01	0.87	0.87	0.87	0.94	0.94	0.98
Band3	0.02	0.01	0.01	0.86	0.86	0.86	0.93	0.93	0.98
Band 4	0.06	0.03	0.04	0.47	0.47	0.46	0.88	0.88	0.99
image pair (imag1214, image1002):									
Band 1	0.01	0.01	0.01	0.83	0.85	0.85	0.81	0.85	0.85
Band 2	0.01	0.01	0.01	0.81	0.82	0.82	0.86	0.88	0.88
Band3	0.02	0.02	0.02	0.74	0.76	0.74	0.74	0.79	0.86
Band 4	0.03	0.03	0.03	0.67	0.68	0.67	0.77	0.90	0.92

4. Discussion

The paper focused on the radiometric normalization of satellite image sequence analysis, especially, the elimination of the spectral difference among the images caused by different image acquisition conditions and the effects of the regularly nonlinear spectral differences caused by seasonal variation of natural objects. We proposed a novel method, i.e., the kCCA, to select nonlinear radiation control points-NIFs and to derive the nonlinear transformation of image-pairs using the least squares regression method.

The total number of the selected NIFs were larger than that of the linear radiometric normalization method (CCA). This indicates that the spectral variation of the corresponding points between each image-pair is mostly nonlinear. Moreover, since most NIFs were distributed in the vegetation area, it indicates that the spectral changes caused by vegetation growth over time are regular and predictable, and thus they can be detected by NIFs. Figure 9 showed us that the effect of cloud pixels was a problem that must be considered in the extraction of NIFs and that the removal of cloud pixels can improve the accuracy of radiometric normalization. The chi-squared percentile τ directly determined the number of NIFs. Table 4 showed that the total number of NIFs and the number of NIFs in the vegetation area increased as the τ decreased. However, the ratio of NIFs in vegetation area was relatively stable and the fitting curves for different thresholds τ were similar. Therefore, we should appropriately choose the value of τ so that we can get enough invariant points to fit the nonlinear relationship. If the value of τ is too small, the numbers of NIFs will be quite large, which leads to higher computation afford.

By using of the least squares regression, we have derived relationships between image-pairs from NIFs. The scatter plot map of the NIFs density (Figure 5) showed us that the relationships were nonlinear. Table 3 and Figure 6 showed that the nonlinear relationship positively correlated with the time span between image-pairs.

Table 5 and Table 6 showed that the kCCA-based nonlinear radiometric normalized images had the smallest RMSE and the largest correlation coefficient with the reference image. As the time span increased, the RMSE increased and the correlation coefficient decreased, and so was that of the normalized images based on the CCA method. Radiometric normalized images of the histogram matching method had the largest RMSE and the smallest correlation coefficient with the reference image. It means the kCCA-based normalization can preserve more similarity and better correlation between image-pairs than the CCA algorithm and histogram matching. Histogram similarity is an index that measures the color difference between an image-pair. The histograms of the radiometric normalized images using the histogram matching had the highest histogram similarity with the reference image, those of kCCA transformation had the second highest similarity, while the CCA-based linear radiometric normalized results had the lowest similarity. However, the histogram matching uses global statistical information for normalization which eliminates all the spectral difference between an image-pair, including the difference must be preserved. Compared with the CCA-based linear normalization, the proposed method can better avoid the color error propagation.

In general, the kCCA-based normalization can either preserve more similarity and correlation between each image-pair or effectively avoid the color error propagation. The proposed method not only builds the common scale or reference to make the radiometric consistency among GF-1 image sequences, but also highlights the interesting spectral changes while eliminates less interesting spectral changes. Our method enables the application of GF-1 data for change detection, land-use, land-cover change detection etc.

5. Conclusions

In this paper, we presented a novel method to perform nonlinear relative spectral alignment and radiometric normalization for high-resolution wide-spectrum multi-temporal satellite images. The method used the kernel canonical correlation analysis transformation (kCCA) to extract NIFs, a set of samples belonging to unchanged area in the kernel space, and to align the relative spectrum of two images using the NIFs-fitted nonlinear relationship. The proposed approach does not need full supervision. The presented approach may be used as a preprocessing technique for change detection

of multi-temporal images, especially when the vegetation spectral changes of multi-temporal images is a disturbance of change detection. Furthermore, we qualitatively and quantitatively compared the radiometric normalization results of the proposed method with CCA-based PIFs method and the histogram matching method derived from statistical distribution. The comparison indicated that the radiometric normalization result of the proposed method had the best linear correlation and least bias with the reference, whereas the radiometric normalization of the histogram matching had the best histogram similarity with the reference image. Finally, in this paper we dealt only with pixel-wise approaches for both the relative spectral alignment and radiometric normalization using the polynomial kernel canonical correlation analysis transformation. In future, we will address the use of different kernel function space, mixed kernel function space and the spatial context to improve both the alignment and the radiometric normalization. This would open many applications of very high-resolution imagery, because high-resolution multi-temporal image-pairs have more complex non-linear relationship and the semantic classes cannot be modeled pixel-wise.

Acknowledgments: This work has been supported by the National Natural Science Fund (41601384) and the National High Technology Research and Development Program (863) of China under Grant 2016YFB0501501.

Author Contributions: Yang Bai, Ping Tang and Changmiao Hu conceived and designed the experiments; Yang Bai performed the experiments and analyzed the data; Changmiao Hu contributed analysis tools; Yang Bai wrote the paper.

Conflicts of Interest: The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

MDPI: Multidisciplinary Digital Publishing Institute

DOAJ: Directory of open access journals

TLA: Three letter acronym

LD: linear dichroism

References

1. Bai, Y.; Tang, P.; Hu, C.M. Kernel Mad Algorithm for Relative Radiometric Normalization. *Int. Arch. Photogramm.* **2016**, *3*, 49–53, doi:10.5194/isprsannals-III-1-49-2016.
2. Song, C.; Woodcock, C.E.; Seto, K.C.; Lenney, M.P.; Macomber, S.A. Classification and change detection using Landsat TM data: When and how to correct atmospheric effects? *Remote Sens. Environ.* **2001**, *75*, 230–244, doi:10.1016/S0034-4257(00)00169-3.
3. Yang, Q.-H.; Qi, J.-W.; Sun, Y. The application of high-resolution satellite remotely sensed data to landuse dynamic monitoring. *Remote Sens. Land Resour.* **2002**, *4*, 20–27.
4. Yang, X.J.; Lo, C.P. Relative radiometric normalization performance for change detection from multi-date satellite images. *Photogramm. Eng. Remote Sens.* **2000**, *66*, 967–980.
5. Barazzetti, L.; Gianinetto, M.; Scaioni, M. Radiometric Normalization with Multi-image Pseudo-invariant Features. In Proceedings of the Fourth International Conference on Remote Sensing and Geoinformation of the Environment (Rscy2016), Paphos, Cyprus, 4–8 April 2016.
6. Hu, C.M.; Tang, P. Automatic algorithm for relative radiometric normalization of data obtained from Landsat TM and HJ-1A/B charge-coupled device sensors. *J. Appl. Remote Sens.* **2012**, *6*, doi:10.1117/1.JRS.6.063509.
7. Volpi, M.; Camps-Valls, G.; Tuia, D. Spectral alignment of multi-temporal cross-sensor images with automated kernel canonical correlation analysis. *J. Photogramm. Remote Sens.* **2015**, *107*, 50–63, doi:10.1016/j.isprsjprs.2015.02.005.
8. Berk, A.; Anderson, G.P. Impact of MODTRAN@5.1 on atmospheric compensation. In Proceedings of the 2008 IEEE International Geoscience and Remote Sensing Symposium, Boston, MA, USA, 6–11 July 2008; pp. 127–129.
9. El Hajj, M.; Begue, A.; Lafrance, B.; Hagolle, O.; Dedieu, G.; Rumeau, M. Relative radiometric normalization and atmospheric correction of a SPOT 5 time series. *Sensors* **2008**, *8*, 2774–2791.

10. Philpot, W.; Ansty, T. Analytical Description of Pseudoinvariant Features. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 2016–2021, doi:10.1109/Tgrs.2012.2227328.
11. Canty, M.J.; Nielsen, A.A.; Schmidt, M. Automatic radiometric normalization of multitemporal satellite imagery. *Remote Sens. Environ.* **2004**, *91*, 441–451, doi:10.1016/j.rse.2003.10.024.
12. De Carvalh, O.A.; Guimaraes, 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, doi:10.3390/rs5062763.
13. Schott, J.R.; Salvaggio, C.; Volchok, W.J. Radiometric Scene Normalization Using Pseudo-invariant Features. *Remote Sens. Environ.* **1988**, *6*, 1–16.
14. Andrefouet, S.; Muller-Karger, F.E.; Hochberg, E.J.; Hu, C.M.; Carder, K.L. Change detection in shallow coral reef environments using Landsat 7 ETM+ data. *Remote Sens. Environ.* **2001**, *78*, 150–162.
15. Chen, X.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, doi:10.1016/j.rse.2005.05.021.
16. Kennedy, R.E.; Cohen, W.B. Automated designation of tie-points for image-to-image coregistration. *Int. J. Remote Sens.* **2003**, *24*, 3467–3490, doi:10.1080/0143116021000024249.
17. 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.
18. 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.
19. Yu, X.; Chen, Y. Relative radiometric normalization of remotely sensed images based on improved automatic scattergram-controlled regression. *Opt. Tech.* **2007**, *22*, 185–188.
20. Canty, M.J.; Nielsen, A.A. Unsupervised classification of changes in multispectral satellite imagery. *Proc. SPIE* **2004**, *5573*, 356–363, doi:10.1117/12.565090.
21. 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, doi:10.1016/j.rse.2007.07.013.
22. Nielsen, A.A. Multiset canonical correlations analysis and multispectral, truly multitemporal remote sensing data. *IEEE Trans. Image Process.* **2002**, *11*, 293–305.
23. Nielsen, A.A.; Conradsen, K.; Simpson, J.J. Multivariate alteration detection (MAD) and MAF postprocessing in multispectral, bitemporal image data: New approaches to change detection studies. *Remote Sens. Environ.* **1998**, *64*, 1–19, doi:10.1016/S0034-4257(97)00162-4.
24. 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, doi:10.1016/j.apm.2013.01.006.
25. Helmer, E.H.; Ruefenacht, B. A comparison of radiometric normalization methods when filling cloud gaps in Landsat imagery. *Can. J. Remote Sens.* **2007**, *33*, pp. 457–458.
26. Canty, M.J.; Nielsen, A.A. Linear and kernel methods for multivariate change detection. *Comput. Geosci.* **2012**, *38*, 107–114, doi:10.1016/j.cageo.2011.05.012.
27. Huang, S.Y.; Lee, M.H.; Hsiao, C.K. Nonlinear measures of association with kernel canonical correlation analysis and applications. *J. Stat. Plan. Inference* **2009**, *139*, 2162–2174, doi:10.1016/j.jspi.2008.10.011.
28. Nielsen, A.A. Kernel Methods in Orthogonalization of Multi- and Hypervariate Data. In Proceedings of the 2009 16th IEEE International Conference on Image Processing (ICIP), Cairo, Egypt, 7–10 November 2009; pp. 3729–3732.
29. Nielsen, A.A.; Vestergaard, J.S. A Kernel Version of Multivariate Alteration Detection. In Proceedings of the 2013 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Melbourne, VIC, Australia, 21–26 July 2013; pp. 3451–3454.
30. Yoshida, K.; Yoshimoto, J.; Doya, K. Sparse kernel canonical correlation analysis for discovery of nonlinear interactions in high-dimensional data. *BMC Bioinform.* **2017**, *18*, doi:10.1186/s12859-017-1543-x.
31. Shao, J.; Wang, L.Q.; Zhao, Z.C.; Su, F.; Cai, A.N. Deep canonical correlation analysis with progressive and hypergraph learning for cross-modal retrieval. *Neurocomputing* **2016**, *214*, 618–628, doi:10.1016/j.neucom.2016.06.047.

32. Hotelling, H. Relations between two sets of variates. *Biometrika* **1936**, *28*, 321–377, doi:10.1093/biomet/28.3-4.321.
33. Camps-Valls, G.; Bruzzone, L. *Kernel Methods for Remote Sensing Data Analysis*; Wiley: Hoboken, NJ, USA, 2009.
34. Shawe-Taylor, J.; Cristianini, N. *Kernel Methods for Pattern Analysis*; Cambridge University Press: Cambridge, UK, 2004.
35. Nielsen, A.A.; Conradsen, K.; Andersen, O. A change oriented extension of EOF analysis applied to the 1996–1997 AVHRR sea surface temperature data. *Phys. Chem. Earth* **2002**, *27*, 1379–1386.
36. Hu, C.M.; Bai, Y.; Tang, P. Automatic cloud detection for GF-4 series images. *J. Remote Sens.* **2018**, *22*, 132–142.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).