

Article

IHS-GTF: A Fusion Method for Optical and Synthetic Aperture Radar Data

Zhenfeng Shao ¹, Wenfu Wu ^{2,*} and Songjing Guo ³

¹ State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China; shaozhenfeng@whu.edu.cn

² School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China

³ School of Geography and Information Engineering, China University of Geosciences (Wuhan), Wuhan 430074, China; guosongjing@cug.edu.cn

* Correspondence: wuwfw09140818@whu.edu.cn

Received: 28 July 2020; Accepted: 26 August 2020; Published: 28 August 2020



Abstract: Optical and Synthetic Aperture Radar (SAR) fusion is addressed in this paper. Intensity–Hue–Saturation (IHS) is an easily implemented fusion method and can separate Red–Green–Blue (RGB) images into three independent components; however, using this method directly for optical and SAR images fusion will cause spectral distortion. The Gradient Transfer Fusion (GTF) algorithm is proposed firstly for infrared and gray visible images fusion, which formulates image fusion as an optimization problem and keeps the radiation information and spatial details simultaneously. However, the algorithm assumes that the spatial details only come from one of the source images, which is inconsistent with the actual situation of optical and SAR images fusion. In this paper, a fusion algorithm named IHS-GTF for optical and SAR images is proposed, which combines the advantages of IHS and GTF and considers the spatial details from the both images based on pixel saliency. The proposed method was assessed by visual analysis and ten indices and was further tested by extracting impervious surface (IS) from the fused image with random forest classifier. The results show the good preservation of spatial details and spectral information by our proposed method, and the overall accuracy of IS extraction is 2% higher than that of using optical image alone. The results demonstrate the ability of the proposed method for fusing optical and SAR data effectively to generate useful data.

Keywords: optical and SAR; image fusion; pixel saliency; impervious surface

1. Introduction

With the rapid development of Earth observation technology, various remote sensing sensors have begun to play a role, bringing a wealth of available data for research [1]. However, many sensors have some limitations due to technical bottlenecks and defects in principle. For example, hyperspectral remote sensing has a high spectral resolution, while its spatial resolution is low. For another example, SAR image is difficult to be interpreted and its application is limited because of the inherent speckle [2]. In addition, with the increasing complexity of observation tasks and the high heterogeneity of observation scenes [3], information from a single data source cannot meet the requirements and data from different images need to be collected and combined into a single image in order to extract additional information [4]. At this point, image fusion can come into play. Image fusion can be divided into three levels, namely pixel level, feature level, and decision level, among which pixel level fusion is considered in this paper. Formally, image fusion at the pixel level is combining two or more images covering the same scene into a high-quality image through a certain algorithm [5].

Multi-spectral (MS) and SAR are two typical and widely used remote sensing sensors. MS is a passive sensor which has some advantages of multiple spectral bands and simple interpretation. However, its imaging is easily affected by cloud, fog, and illumination [6,7]. SAR, as an active detector, can penetrate cloud cover, haze, dust, and other climatic conditions because of its long wavelength, which is not easily affected by meteorological conditions and sunshine level. Due to this property, SAR can observe the Earth during all weather, day and night. Furthermore, SAR is sensitive to moisture and geometric characteristics, which can provide very useful information different from MS images. However, SAR image is difficult to interpret and apply because of side-looking imaging and the speckle. Information in MS and SAR images can complement each other. Fusing them can produce a higher quality image which has been used successfully in many fields, such as land use mapping and monitoring and identification of human settlements [8–11]. Therefore, it is very meaningful to develop a fusion algorithm for optical and SAR images.

During the pixel-level fusion of MS and SAR images with high spatial resolution, generally, the ideal fusion product will retain the spectral information from MS image and reserve the spatial information from SAR image. Many methods have been proposed for fusing MS and SAR images. These methods can be grouped into four categories: component substitution methods, multi-scale decomposition methods, hybrid methods, and model-based methods [4]. The idea of component substitution methods is to firstly carry out space transformation on the MS images with low spatial resolution to realize the space-spectrum separation; then replace the spatial components of the multi-spectral images with high-resolution SAR images; and finally carry out the corresponding inverse space transformation to obtain the fused images. IHS transform [12], Principal Component Analysis (PCA) [13], Gram–Schmidt (GS) [14], Brovery Transform (BT) [15], etc. are all representatives of component substitution methods. In multi-scale decomposition methods, the source images are decomposed into different levels and then fuse them at each level. Wavelet [16], Contourlet [17] and Shearlet transform [18] are the most commonly used multi-scale decomposition methods. Hybrid methods combine component substitution methods and multi-scale decomposition methods, such as IHS combined with à Trous wavelet (AWT) [19] and PCA combined with AWT [20], which make full use of the advantages of the two methods. Model-based methods have two types: variational models [21] and sparse representation-based models [22]. Reviews of optical and SAR images fusion can be found in [4,23].

In pixel-level fusion, the design of fusion rules is very important but difficult. Ma et al. proposed a method named GTF to fuse infrared and gray visible images, which formulates the image fusion as an optimization problem, avoiding design fusion rules and keeping radiation and spatial information simultaneously [24]. However, this method assumes that the spatial details only come from one of the source images, which is inconsistent with the actual situation of optical image and SAR image. Optical and SAR sensors are two different imaging mechanisms, and they contain different details which should be considered in the fusion process. IHS fusion method is a kind of classic, easy to implement method and can separate RGB images into three independent components, but it is inappropriate to use this method directly for optical and SAR images fusion, as it will cause serious spectral distortion. Therefore, in this paper, we propose a fusion algorithm, named IHS-GTF, to bridge the gap between IHS and GTF. Furthermore, the proposed method was tested by an application example of urban IS extraction from the fused image. The reason for choosing IS extraction as an example is that bright impervious surface and bare soil, as well as dark impervious surface and shadow, in optical images usually cause spectral confusion. Previous studies have indicated that fusing optical and SAR images can help solve this problem, which can provide us with a reference [25,26]. In addition, our group also performed some work in this field [18,27].

The rest of the papers is organized as follows. The study area and dataset used in this study are depicted in detail in Section 2. In Section 3, we provides an overview of IHS, GTF, and IHS-GTF algorithms. Section 4 shows the results of fusion and impervious surface extraction. The conclusions of this study are given in Section 5.

2. Study Area and Dataset

2.1. Study Area

Wuhan city, located at $113^{\circ}41'E-115^{\circ}05'E$, $29^{\circ}58'N-31^{\circ}22'N$, is one of the regions with the most frequent and strongest rainstorms in China. The annual mean precipitation ranges from 1150 to 1450 mm. The Yangtze River and its largest tributary, the Han River, meet in Wuhan, making Wuhan convenient in transportation and developed in shipping. Since the Ming and Qing dynasties, Wuhan has been an important economic city in China. By the founding of new China in 1949, Wuhan entered a period of high-speed development and gradually developed into the center city of city clusters in the middle reaches of the Yangtze River, which was accompanied by a sharp expansion of the impervious surface. The increase of IS area brings a series of problems, such as urban heat island effect [28–31], urban flood disaster [32–34], and the decrease in cultivated land [35]. Continuous monitoring IS using satellite remote sensing can be conducive to develop urban sustainably. To sum up, Wuhan is an relatively ideal study area and two sites covering it were carefully selected to test our proposed method in this study. Figure 1 shows the geographic location of the study area.

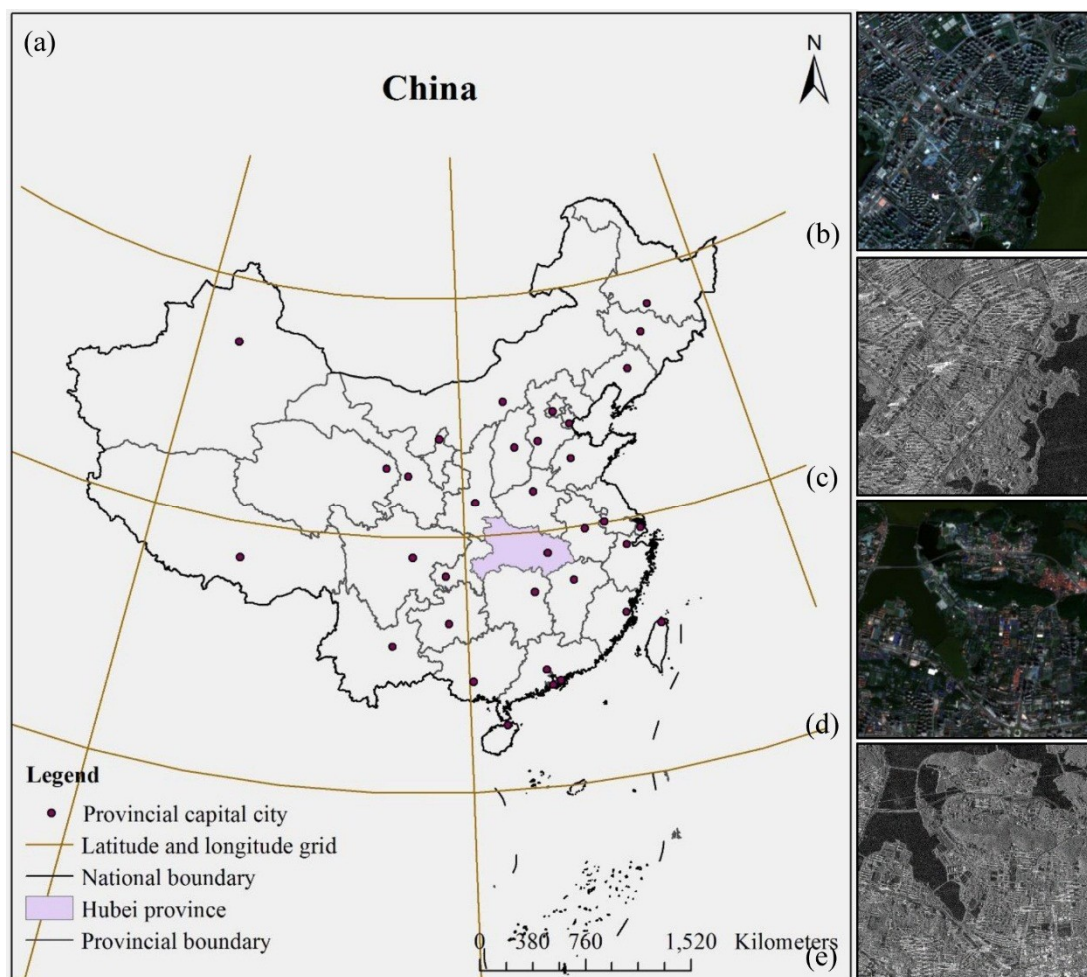


Figure 1. The geographical location of the study area: (a) map of China; (b,c) Sentinel-2A image and TerraSAR image of Site 1, respectively; and (d,e) the Sentinel-2A image and TerraSAR image of Site 2, respectively.

2.2. Dataset and Preprocessing

TerraSAR and Sentinel-2A (S2) were employed as SAR and optical data, respectively. TerraSAR is Germany's first satellite, launched on 15 June 2007, which is currently in orbit and operates in the X

band (9.6 GHz). The imaging can be done in three modes, bunching, striping, and sweeping, and can be polarized in many ways. In this study, the TerraSAR image used is the single look complex data of StripMap mode with HH (horizontal transmit and horizontal receive) polarization in the orbital descending direction. The image was acquired on 8 August 2017. The spatial resolution of the TerraSAR image is 3 m. Speckle filter, radiometric calibration, terrain correction, and other preprocessing steps were done with the software of Sentinel Application Platform (SNAP) 6.0. Note that the refined Lee filter with a window size of 5×5 was selected to filter the speckle.

The Sentinel-2A satellite, the second of the global environment and security monitoring satellites, was launched on 23 June 2015. Since 3 December 2016, Sentinel-2A has been providing data services to users worldwide through the European Space Agency (<https://scihub.copernicus.eu/>). Sentinel-2A carries a multispectral imager covering 13 spectral bands, which includes three bands with 60-m resolution, six bands with 20-m resolution, and four bands with 10-m resolution. In this study, bands 2–4 (10-m resolution) of Sentinel-2A MSI Level-1C dataset with nearly cloud-free conditions acquired on 15 March 2017 were used. There is a time difference of approximately five months between the acquisition dates of the Sentinel-2A and TerraSAR data. However, the study sites located at the central urban part of Wuhan are developed and well-planned urban areas. Thus, the land cover at the study sites did not change significantly between March 2017 and August 2017, which was also the case, as seen from high-resolution Google Earth images. Note that, since Sentinel-2A MSI Level-1C data were corrected for topography, but not for atmosphere, we used SNAP's plugin *sen2cor* to conduct atmospheric correction on Level-1C data and obtained Level-2A data.

Image fusion requires both images registration. In our study, we used the automatic registration module in ENVI 5.3 software with the aid of manual inspection to perform image registration. The Root Mean Square Error (RMSE) for registration was less than one pixel. After the above series of preprocessing, both images were registered to the same geo-reference system of the Universal Transverse Mercator (UTM) projection (Zone 49N) and Datum World Geodetic System 84 (WGS84).

3. Methodology

3.1. Framework of the Proposed Method

Figure 2 shows the framework of the proposed method. The IHS-GTF algorithm has five main steps:

- (1) Perform IHS transformation on optical image and obtain I component.
- (2) Obtain the detail layers of SAR image (SAR_d) and I component (I_d) through two-scale image decomposition (TSID). Note that histogram matching between SAR image and I component should be performed before decomposition. Besides, Gaussian filter is used to reduce the speckle in the detail layer of the SAR image.
- (3) Combine the information of I component detail layer with that of SAR image detail layer through the pixel saliency to obtain F_{SAR_I} .
- (4) I component and F_{SAR_I} are input into the GTF algorithm to obtain the fused I component.
- (5) Carry out inverse IHS transformation (I-IHS) to obtain the fused image.

IS extraction, the last work in this study, was used to further test our proposed fusion method. The following subsections describe the technical details of each part in the framework.

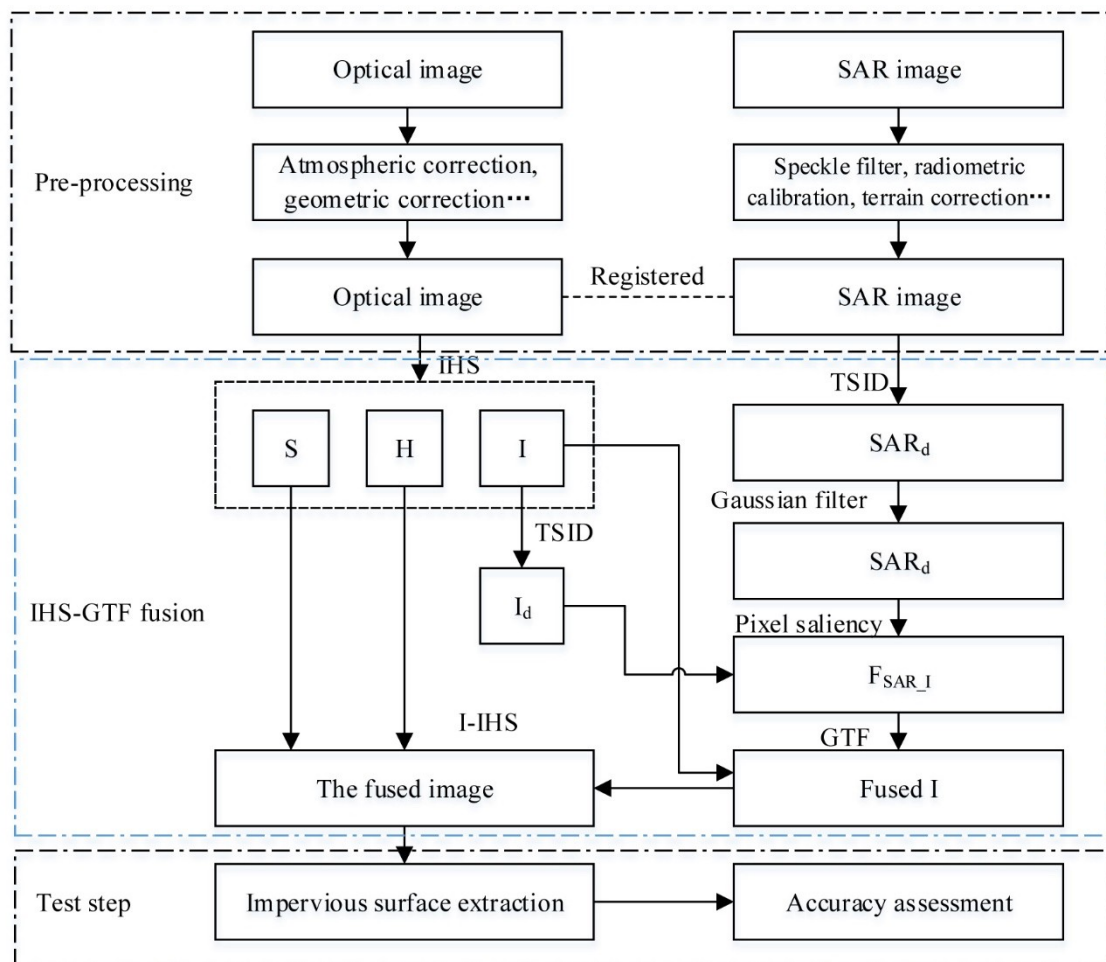


Figure 2. The framework of the proposed method.

3.2. Overview of IHS Fusion Method

IHS fusion method is a kind of classic, easy to implement method, which is based on IHS transform. The IHS transform can separate an RGB image into three independent components: intensity (I), hue (H), and saturation (S) [36]. Among them, most spatial information is isolated in I component, while the H and S components mainly contain spectral information. Then, we can process independently the three components. When using IHS to fuse optical and SAR images, usually the I component is replaced with the high resolution SAR image and then the inverse IHS transform is carried out to obtain the fused image [5]. The IHS fusion method can be expressed as follows:

$$\begin{bmatrix} I \\ V_0 \\ V_1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ -\frac{\sqrt{2}}{6} & \frac{\sqrt{2}}{6} & \frac{2\sqrt{2}}{6} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

Substitution operation and inverse IHS transform lead to:

$$\begin{bmatrix} R_{new} \\ G_{new} \\ B_{new} \end{bmatrix} = \begin{bmatrix} 1 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 1 & \sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} SAR \\ V_0 \\ V_1 \end{bmatrix} = \begin{bmatrix} R + (SAR - I) \\ G + (SAR - I) \\ B + (SAR - I) \end{bmatrix} \quad (2)$$

According to Equations (1) and (2), IHS fusion method makes full use of spatial information of the SAR image and improves the spatial resolution. After the substitution operation and IHS transform,

the intensity of the new RGB channel differs from that of the original RGB channel by $(SAR - I)$. Due to the difference in imaging mechanism between optical and SAR images, the value of $(SAR - I)$ may be very large, which will lead to serious spectral distortion. Therefore, it is not appropriate to use IHS method directly for optical and SAR images fusion.

3.3. Overview of GTF Fusion Method

GTF method was proposed by Ma et al. to fuse infrared and gray visible images [24]. A fused image preserving radiation information and spatial details simultaneously is the most ideal, which is similar to optical and SAR images fusion. The radiation information is from infrared image and the spatial details are from visible image, which has a relatively high resolution. Therefore, on the one hand, the fused image should have a similar pixel intensity distribution with the given infrared image. That is, the difference of pixel intensity between the fused image and the given infrared image should be as small as possible. This can be mathematically modeled as:

$$\varepsilon_1(\mathbf{x}) = \frac{1}{p} \|\mathbf{x} - \mathbf{u}\|_p \quad (3)$$

where $\varepsilon_1(\mathbf{x})$ is the empirical error of the intensity are \mathbf{x} and \mathbf{u} are the column-vector form of the fused image and the infrared image, respectively. $\|\bullet\|_p$ is the p norm.

On the other hand, they assumed that the fused image should preserve the spatial details from the visible image. In the field of image, the gradient is usually used to measure the spatial detail. Hence, the fused image should have similar gradients with the given visible image. This step can be mathematically modeled as:

$$\varepsilon_2(\mathbf{x}) = \frac{1}{q} \|\nabla \mathbf{x} - \nabla \mathbf{v}\|_q \quad (4)$$

where $\varepsilon_2(\mathbf{x})$ is the empirical error of the gradient and $\nabla \mathbf{x}$ and $\nabla \mathbf{v}$ are the gradients of the fused image and the infrared image, respectively. Note that the visible image is the gray scale image. $\|\bullet\|_q$ represents the q norm.

Thus, the fusion problem can be transformed into an optimization problem, i.e., minimizing the following objective function:

$$\min \varepsilon(\mathbf{x}) = \varepsilon_1(\mathbf{x}) + \lambda \varepsilon_2(\mathbf{x}) = \frac{1}{p} \|\mathbf{x} - \mathbf{u}\|_p + \lambda \frac{1}{q} \|\nabla \mathbf{x} - \nabla \mathbf{v}\|_q \quad (5)$$

where λ is a parameter controlling the amount of gradient information of the visible image injected into the infrared image. Considering $\mathbf{x} - \mathbf{u}$ should be Laplacian or impulsive, set $p = 1$. Because 0 norm is NP-hard, set $q = 1$. Since $\mathbf{y} = \mathbf{x} - \mathbf{v}$, Equation (5) can be rewritten as:

$$\begin{aligned} \mathbf{y}^* &= \operatorname{argmin} \left\{ \sum_{i=1}^{mn} |\mathbf{y}_i - (\mathbf{u}_i - \mathbf{v}_i)| + \lambda J(\mathbf{y}) \right\} \\ J(\mathbf{y}) &= \sum_{i=1}^{mn} |\nabla_i \mathbf{y}| = \sum_{i=1}^{mn} \sqrt{(\nabla_i^h \mathbf{y})^2 + (\nabla_i^v \mathbf{y})^2} \end{aligned} \quad (6)$$

where $\nabla_i = (\nabla_i^h, \nabla_i^v)$ represents the horizontal and vertical gradients at the pixel i . mn is the size of image.

Equation (6) describes a standard problem of minimizing total variation with a norm of 1. In Equation (6), only \mathbf{y} is unknown, which can be calculated by directly using the algorithm proposed in [37]. If \mathbf{y} is calculated, then the fused image \mathbf{x} can be expressed as $\mathbf{x} = \mathbf{y} + \mathbf{u}$.

3.4. IHS-GTF Fusion Algorithm

According to the above description of IHS and GTF, in summary, the IHS fusion method is an easy to implement fusion algorithm and can separate RGB image into three independent components,

but it cannot be directly used for optical and SAR images fusion, otherwise it will cause spectral distortion problem. The GTF method transforms the image fusion into the optimization problem, seeking a balance between the maintenance of spatial details and spectral information. The spectral distortion in IHS fusion method depends on the difference between the I component of the fused image and that of the optical source image, which can be seen as an optimization problem similar to GTF. However, GTF assumes that the spatial details come from only one of the given source images, which is inconsistent with the actual situation of optical image and SAR image. Optical and SAR sensors have two different imaging mechanisms and contain different details, which should be considered in the fusion process. Therefore, we propose a method named IHS-GTF bridging the gap between IHS and GTF to fuse optical and SAR data. The idea of this algorithm is that the spatial details of optical image are combined firstly with those of the SAR image and then they are input into GTF.

An image can be decomposed into the detail layer and the base layer through the two-scale decomposition, in which the detail layer contains a large amount of information in the source image. The detail layer of an image can be obtained as follows:

$$\begin{aligned} B &= S * Z \\ D &= I - B \end{aligned} \quad (7)$$

where S denotes the source image, Z represents the average filter, B is the base layer of the source image, and D is the detail layer of the source image. The size of the average filter is conventionally selected as 31×31 [38]. It is worth noting that, due to much noise in the detail layer of SAR image, we carry out Gaussian filtering with a window size of 3×3 on it.

After IHS transform, the detail layers of the I component and the SAR image are obtained through the two-scale decomposition. We combine the details of the I component with that of the SAR image through the pixel saliency. The specific process is as follows:

$$S^k = \max(S_I^k, S_{SAR}^k) \quad (8)$$

$$P_I^k = \begin{cases} 1 & \text{if } S^k = S_I^k \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

$$P_{SAR}^k = \begin{cases} 1 & \text{if } S^k = S_{SAR}^k \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

$$F_{SAR_I}^k = S_{SAR}^k \bullet P_{SAR}^k + S_I^k \bullet P_I^k \quad (11)$$

In Equations (8)–(11), S_I^k and S_{SAR}^k represent the pixel intensity of the detail layers of I component and SAR image at pixel k . S^k is the significant value of the both detail layers at pixel k . P_I^k and P_{SAR}^k denote the weight of the detail layers of I component and SAR image at pixel k . $F_{SAR_I}^k$ is the result of combining SAR image details with I component details.

At this point, we input F_{SAR_I} and I component into the GTF and get the fused I component, which simultaneously preserves the radiation information from the I component and the details from F_{SAR_I} . This process can be modeled as:

$$\min \varepsilon(x) = \varepsilon_1(x) + \lambda \varepsilon_2(x) = \frac{1}{p} \|x - I\|_p + \lambda \frac{1}{q} \|\nabla x - \nabla F_{SAR_I}\|_q \quad (12)$$

The solution of Equation (12) can be found in Section 3.3. After the fused I component is obtained, the inverse IHS transform can be performed to obtain the fused image.

3.5. Impervious Surface Extraction and Accuracy Assessment

In this study, impervious surface is extracted from the fused image to test the proposed fusion method. Many methods for IS extraction using satellite remote sensing have been proposed,

among which random forest (RF) algorithm as an ensemble learning algorithm has been successfully applied [27,32,39]. The overall framework of the algorithm was first proposed by Breiman in 2001 [40]. Multiple decision trees generated by bootstrap resampling technique make up a forest. The final decision is made through a voting mechanism. Compared with the state-of-the-art methods, such as deep learning, RF has some significant advantages. Firstly, it needs short training time and is easy to implement. Secondly, it is a white box and not a black box such as deep learning methods and can yield explicable results. Finally, it can also obtain considerable performance with small samples. Given the advantages of RF and its considerable performance in the field of remote sensing, we also selected RF to extract the IS in this study. In urban area, four texture features derived from gray level co-occurrence matrix, namely the homogeneity, dissimilarity, entropy, and angular second moment, can identify different urban land cover types effectively [39]. Therefore, the three bands of the fused image and their corresponding four texture features, a total of 15 features, were used as the input of the RF model. In RF, the number of decision trees and the number of variables for splitting each node are the two most important parameters which can affect the performance of RF. We set them to 20 and 4 empirically. Although the IS is our final output, in this study, we still used the common classification scheme in the optical images to extract the IS. Five land cover types, water bodies (WB), vegetation (VG), bare soil (BS), bright impervious surface (BIS), and dark impervious surface (DIS), were identified firstly. Then, bright impervious surface and dark impervious surface were combined into IS, while the rest was combined into pervious surface (NIS).

Four accuracy indices based on the confusion matrix were used to assess the accuracy of IS extraction: overall accuracy (OA), Kappa coefficient, user's accuracy (UA), and producer's accuracy (PA). With the aid of high resolution Google Earth images, samples uniformly distributed throughout the study sites were randomly selected through visual interpretation in this study. It is note that two-thirds of these samples were used as training samples, and the rest were used as testing samples. Table 1 shows the numbers of samples for Sites 1 and 2 in detail.

Table 1. The numbers of samples for Sites 1 and 2.

	Training Samples					Testing Samples				
	BIS	DIS	BS	VG	WB	BIS	DIS	BS	VG	WB
Site 1	176	115	133	116	228	88	57	66	57	114
Site 2	134	105	131	182	486	66	52	65	90	242

4. Results and Discussion

We used our proposed method to fuse Sentinel-2A and TerraSAR images for two sites, and compared it with IHS, GTF, and discrete wavelet transform (DWT) methods. Then, we also tested our proposed method through extracting IS.

4.1. Visual Analysis of Fusion Results

SAR images can well maintain the contour of ground objects, and optical images can well maintain the spectral information of the ground objects, thus the ideal results of optical and SAR images fusion should look like the optical image after image enhancement. The qualitative comparison of our proposed method and the other three methods including IHS, GTF, and DWT, is shown in Figures 3 and 4. According to the previous subsection, λ is the only parameter in the GTF algorithm. According to the experimental results and the work in [24], we set it to 4. During DWT fusion, the scale of decomposition was set to 4. The fusion rule is to average the wavelet coefficients of the approximate sub-bands to produce the approximate sub-band of the fused image and select the maximum of detail sub-bands as the detail sub-bands of fused image. From these results, some interesting findings can be found. Firstly, all fusion results look brighter than the corresponding Sentinel-2A image. Secondly, the fusion results of IHS contain a lot of structural information of SAR image, which looks like a color

SAR image, but the spectral information is quite different from the Sentinel-2A image. Therefore, the result of IHS is not ideal for SAR and optical images fusion. Thirdly, the results of the GTF are better than the fusion results of IHS, but they are not very satisfactory. There are obvious SAR image features in the fused images, and the spectral information is preserved to a certain extent, but serious spectral distortion still exists. In addition, the fusion image also retains a large amount of SAR image speckle noise. Finally, the results of DWT and IHS-GTF appear to be the closest to the Sentinel-2A image, but there are significant differences between them. The spectral of vegetation in DWT result is markedly different from that in the Sentinel-2A image. The result of IHS-GTF is the most ideal fusion result. Its spectral information is very close to the source optical image, which can be found from water, vegetation, and other land covers. In addition, it also retains significant features of SAR images, such as obvious edge information. It is worth noting that some buildings in the optical image have low spectral reflectance, but they are bright in the fused images, which is a typical feature of buildings in SAR images. This phenomenon may be very beneficial to the extraction of IS. Besides, the result of IHS-GTF has less speckle noise.

For the convenience of comparison, we selected and zoomed in two sub-regions from Sites 1 and 2, as shown in the rectangular regions in Figures 3 and 4. The overpass in Site 1 is obvious in the optical image, but only the main features of the cross can be observed in the SAR image, while other fine features are not shown. Among the fusion results, only the that of IHS-GTF retains the features of the overpass completely, while the others only retained the cross-over features. This indicates that, in the fusion process, the spatial characteristics of optical and SAR images should be considered simultaneously, which is consistent with the starting point of our proposed method. From the rectangular area in Figure 4, we can find that the fusion result of IHS-GTF can well retain the edge information in SAR image and enhance the pixel values, which is similar to an image after an enhancement. The biggest difference between our method and the GTF method is that the spatial details of both optical and SAR images and the effect of speckle in SAR image are all taken into account. The results prove the effectiveness of our proposed method. In summary, the fusion effect through visual analysis is in the sequence: HIS-GTF > DWT > GTF > IHS.

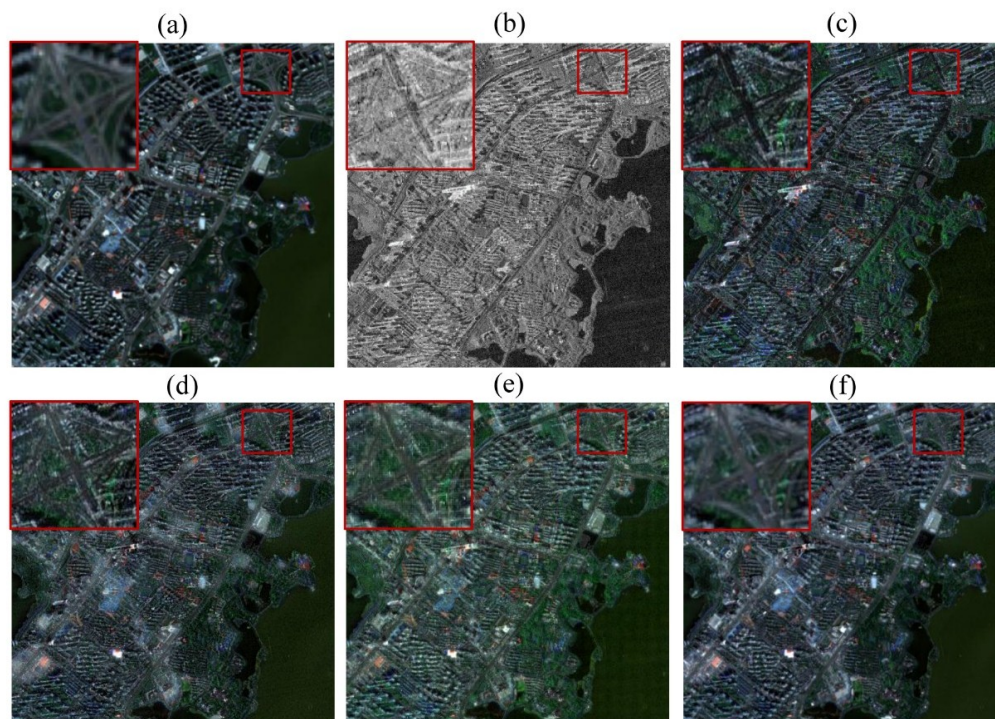


Figure 3. The source images and the fused images with different methods for Site 1: (a–f) Sentinel-2A image, TerraSAR image, and the fusion results of IHS, GTF, DWT, and IHS-GTF, respectively.

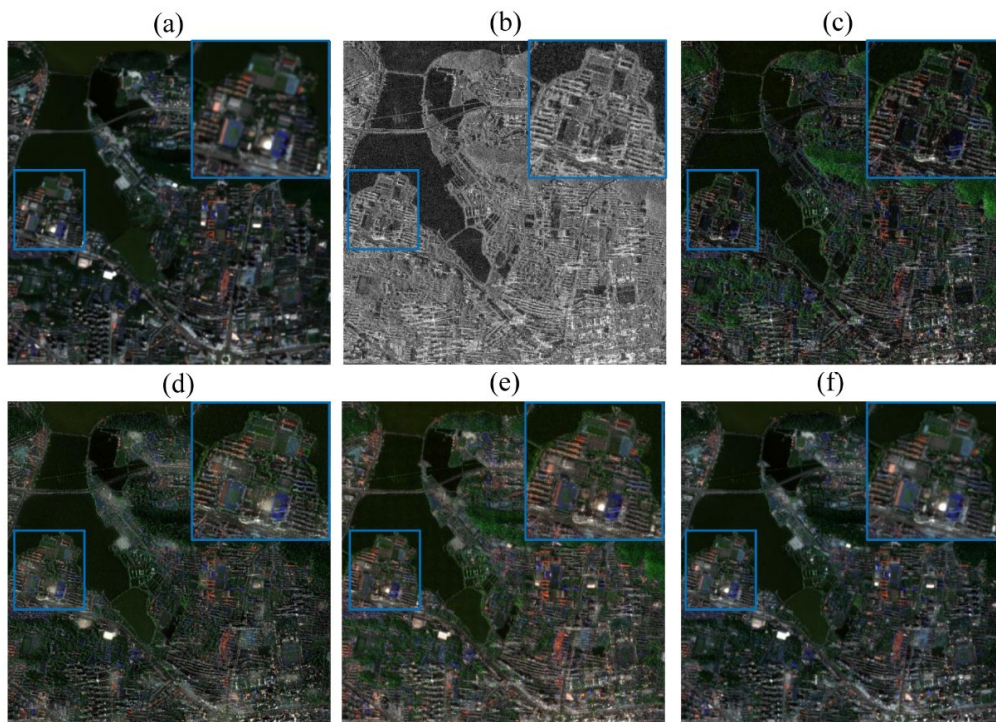


Figure 4. The source images and the fused images with different methods for Site 2: (a–f) Sentinel-2A image, TerraSAR image, and the fusion results of IHS, GTF, DWT, and IHS-GTF, respectively.

4.2. Quantitative Comparison of Fusion Results

The previous subsection presents the qualitative evaluation of the fusion effect of each method, and this section evaluates the fusion effect from a quantitative perspective. According to the principle of IHS transform, the difference between the I component of the fused image and the I component of the source optical image can be used as an indicator for fusion effect. If the difference is low, the scatters formed by the I component value of the source optical image and the corresponding I component of the fusion image should be distributed near 1:1 reference line. Therefore, we randomly selected 200 pixels from the images and compared the corresponding I component values. The results are shown in Figures 5 and 6. Most of the scatters are above the 1:1 line, indicating that the fusion methods overestimate the I component, which is consistent with the fusion images looking brighter compared with the Sentinel-2A image, as described above. In Figure 5 (Site 1), the I component of IHS fusion image (IHS-I) deviates greatly from the I component of Sentinel-2A (S2-I), and the coefficient of determination (R^2) is only 0.1092. In the IHS transform fusion, the I component of the optical image is directly replaced with the SAR image, and the result shows that the SAR image is quite different from the I component of the optical image, thus direct replacement is not appropriate. Compared with IHS, the differences between the I component of GTF (GTF-I) and DWT (DWT-I) image and S2-I are reduced, but the R^2 is still low, 0.4166 and 0.5514, respectively. The scatters formed by the I component of IHS-GTF (IHS-GTF-I) and S2-I are tightly distributed around the 1:1 reference line, with a R^2 as high as 0.8017. This demonstrates that the IHS-GTF method has the best fusion performance. In Figure 6 (Site 2), the results are the same as those in Figure 5, except that the values are different. According to Figures 5 and 6, the sequence of the fusion effect is: IHS-GTF > DWT > GTF > HIS. This is consistent with the visual result shown in Section 4.1.

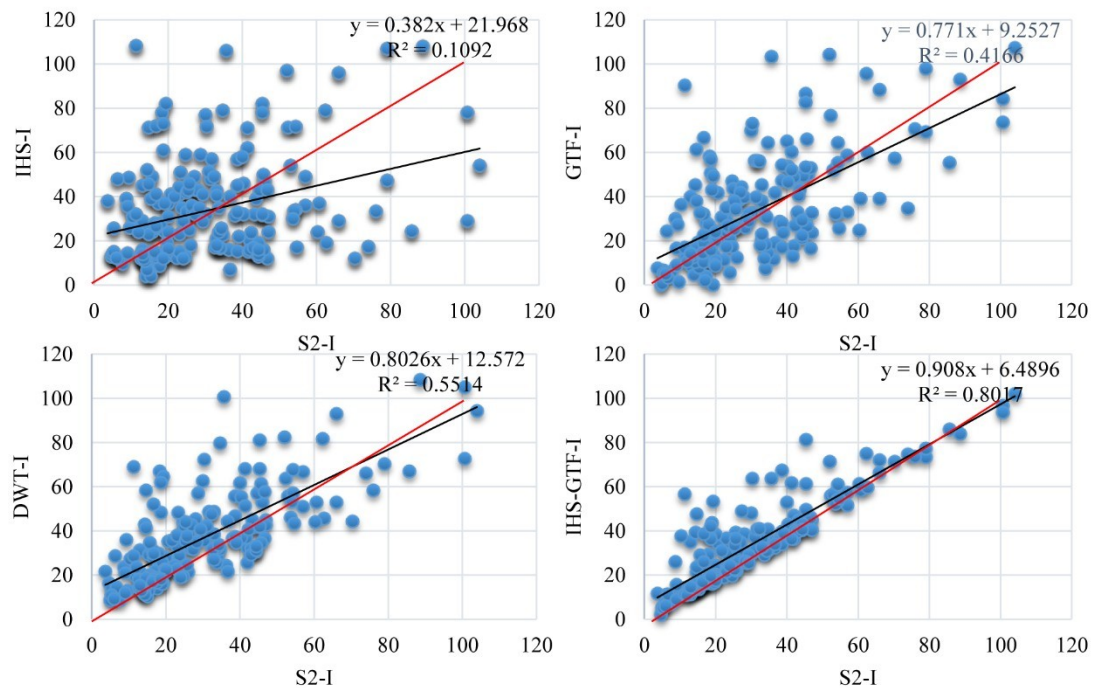


Figure 5. The difference between the I component of the fused images and the I component of Sentinel-2A for Site 1. The black and red lines represent the fitted and 1:1 lines, respectively.

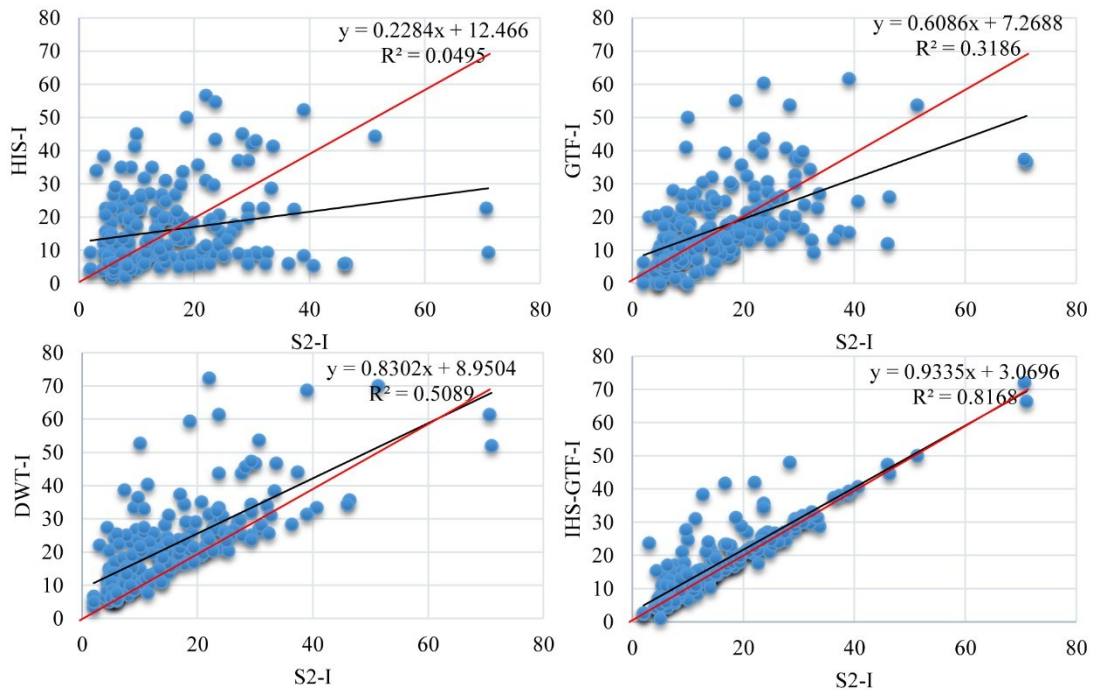


Figure 6. The difference between the I component of the fused images and the I component of Sentinel-2A for Site 2. The black and red lines represent the fitted and 1:1 lines, respectively.

For quantitative comparison, ten indices are commonly used to evaluate the fusion effect: standard deviation (STD), average gradient (GRAD), the peak signal to noise ratio (PSNR), structural similarity (SSM), the root mean squared error (RMSE), mutual information (MI), the Shannon entropy (EN), the spectral angle mapper (SAM), the relative global synthesis error (ERGAS), and correlation coefficient (CC).

(1) Standard Deviation (STD): It is a measure of image contrast. The larger is the value, the higher is the image contrast and the richer is the information. It can be calculated as:

$$STD = \sqrt{\frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N (F(m, n) - \bar{F})^2} \quad (13)$$

where $F(m, n)$ represents the value of the fused image F at pixel (m, n) and \bar{F} stands for the mean value of the fused image.

(2) Average Gradient (GRAD): It reflects sharpness of images, and a larger average gradient indicates a clearer image. It can be written as:

$$GRAD = \frac{1}{M \times N} \sum_{m=1}^{M-1} \sum_{n=1}^{N-1} \left[\frac{\left(\frac{\partial F(m, n)}{\partial m} \right)^2 + \left(\frac{\partial F(m, n)}{\partial n} \right)^2}{2} \right]^{\frac{1}{2}} \quad (14)$$

where $\frac{\partial F(m, n)}{\partial m}$ and $\frac{\partial F(m, n)}{\partial n}$ represent gradients in the m and n directions, respectively.

(3) Peak Signal to Noise Ratio (PSNR): It measures the noise in the fused image. The higher is the value, the lower is the noise. It can be calculated as:

$$PSNR = 10 \lg \left| \frac{F_{\max}^2}{\frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N [F(m, n) - MS(m, n)]^2} \right| \quad (15)$$

where F_{\max} denotes the max value of the fused image F and $MS(m, n)$ denotes the pixel value of the multi-spectral image MS .

(4) Structural Similarity (SSIM): It measures the structural difference between two images, with a value ranging from 0 to 1. The larger is the value, the smaller is the difference. It can be calculated according to the following formula:

$$SSIM(F, MS) = \frac{(2\mu_F \mu_{MS} + c_1)(2\delta_{FMS} + c_2)}{(\mu_F^2 + \mu_{MS}^2 + c_1)(\delta_F^2 + \delta_{MS}^2 + c_2)} \quad (16)$$

where μ_F and δ_F denote the mean value and the standard deviation of the fused image F , respectively; μ_{MS} and δ_{MS} denote the mean value and the standard deviation of the multi-spectral image MS , respectively; δ_{FMS} stands for the covariance of the fused image F and the multi-spectral image MS ; and c_1 and c_2 are constants.

(5) Root mean squared error (RMSE): It is an indicator of the degree of difference between the fused result F and the multi-spectral image MS . A smaller RMSE indicates a better fusion result. It can be defined as:

$$RMSE(F, MS) = \frac{1}{M \times N} \sqrt{\sum_{m=1}^M \sum_{n=1}^N [F(m, n) - MS(m, n)]^2} \quad (17)$$

(6) Mutual Information (MI): It measures the distance between joint probability distribution of the fused result F and the multi-spectral image MS . The larger is the MI, the richer is the information the fusion image obtains from the source image, and the better is the fusion effect. It can be written as:

$$MI_{MSF}(i, f) = \sum_{i, f} P_{MSF}(i, f) \log \frac{P_{MSF}(i, f)}{P_{MS}(i)P_F(f)} \quad (18)$$

where $P_{MS}(i)$ and $P_F(f)$ denote the possibility distribution of the image MS and F, respectively. $P_{MSF}(i, f)$ is the joint probability distribution of MS and F.

(7) Entropy (EN): It reflects the average information abundance in the fusion image, and a larger entropy indicates larger information richness. It can be written as:

$$EN = - \sum_{g=0}^{L-1} p(g) \times \log_2 p(g) \quad (19)$$

where $p(g)$ is the distribution of gray value of the image F. $L - 1$ is the grayscale.

(8) Spectral Angle Mapper (SAM): It measures the similarity between the spectra by calculating the included angle between two vectors. The smaller is the included angle, the more similar are the two spectra. It can be written as:

$$SAM(v, \hat{v}) = \arccos\left(\frac{\langle v, \hat{v} \rangle}{\|v\|_2 \cdot \|\hat{v}\|_2}\right) \quad (20)$$

where v is the spectral pixel vector of the original image, and \hat{v} is the spectral pixel vector of the fused image.

(9) Relative Global Synthesis Error (ERGAS): It is a measure of the global quality of the fused image F. A smaller ERGAS indicates a better fusion result. It can be written as:

$$ERGAS = 100 \frac{h}{l} \sqrt{\frac{1}{k} \sum_{i=1}^k \left[\frac{RMSE(i)}{Mean(i)} \right]^2} \quad (21)$$

where h is the spatial resolution of SAR image, l is the spatial resolution of the MS image, k is the number of bands of the fused image, $Mean(i)$ is the mean value of the i th band of the MS image, and $RMSE(i)$ is the root mean squared error of the i th band of the MS image.

(10) Correlation Coefficient (CC): It reflects the degree of correlation between two images, with a larger correlation coefficient indicating a better fusion effect. It can be defined as:

$$CC(F, MS) = \frac{\sum_{m=1}^M \sum_{n=1}^N [F(m, n) - \bar{F}] [MS(m, n) - \overline{MS}]}{\sqrt{\sum_{m=1}^M \sum_{n=1}^N [F(m, n) - \bar{F}]^2 \times \sum_{m=1}^M \sum_{n=1}^N [MS(m, n) - \overline{MS}]^2}} \quad (22)$$

The quantitative comparisons of the two sites are given in Tables 2 and 3. At the two sites, the IHS-GTF method has several indices that are significantly superior to the three other approaches. These indices include PSNR, SSIM, RMSE, MI, SAM, ERGAS, and CC. For the other metrics, such as STD, GRAD, and EN, the best values are not for the IHS-GTF method. This indicates that no fusion method can have all the optimal metrics, and the process of fusion is to balance various metrics to achieve the optimal fusion effect comprehensively. It is interesting that, through the analysis of the indicators of the IHS-GTF method, it is found that the capability of the IHS-GTF method in spectral information retention is better than that in spatial information retention. Although the IHS-GTF method performs slightly worse in the retention of spatial details, it is still acceptable. Overall, the IHS-GTF method is the best. In addition, the result of DWT is also acceptable. The results of IHS and GTF have good spatial information retention capability, but they have serious spectral distortion. The quantitative analysis and qualitative analysis of the fusion results are completely consistent.

Table 2. Evaluation of fusion results for Site 1. The best results for each quality measure are in bold.

	STD	GRAD	PSNR	SSIM	RMSE	MI	EN	SAM	ERGAS	CC
IHS	21.763	7.235	19.337	0.275	0.031	0.163	6.105	0.730	0.038	0.184
GTF	22.093	6.997	22.105	0.522	0.022	0.343	6.255	0.520	0.028	0.575
DWT	19.455	4.983	24.333	0.744	0.017	0.606	6.108	0.385	0.022	0.740
IHS-GTF	21.639	4.045	27.351	0.836	0.012	1.279	6.210	0.2618	0.015	0.886

Table 3. Evaluation of fusion results for Site 2. The best results for each quality measure are in bold.

	STD	GRAD	PSNR	SSIM	RMSE	MI	EN	SAM	ERGAS	CC
IHS	12.214	4.034	24.751	0.515	0.016	0.160	5.173	0.754	0.042	0.236
GTF	13.305	4.066	27.074	0.693	0.013	0.400	5.419	0.540	0.032	0.604
DWT	13.343	3.242	27.510	0.787	0.012	0.638	5.470	0.402	0.031	0.749
IHS-GTF	12.358	2.050	33.021	0.906	0.006	1.299	5.283	0.258	0.016	0.902

4.3. Impervious Surface Extraction and Comparison

The IS extraction in this study was used as an example of applications to test the proposed fusion method. For comparison, we extracted IS with the fusion results of IHS, GTF, DWT, IHS-GTF, and Sentinel-2A. In this study, we still used the common classification scheme in the optical images to extract the IS. Five land cover types, WB, VG, BS, BIS, and DIS, were identified firstly. The accuracy of classification of the two sites are shown in Figures 7 and 8. From the results, some interesting findings can be observed. Firstly, IHS-GTF has the highest OA and Kappa coefficient. Compared with the classification result of Sentinel-2A, the improvement is about 2%. This indicates that fusing optical and SAR images at the pixel level can improve urban land cover classification, which is consistent with previous studies [18,41,42]. Secondly, the OA and Kappa of IHS, GTF, and DWT fusion results are significantly lower than those of the Sentinel-2A image, which indicates that the design of fusion algorithm is very important for pixel-level fusion of optical and SAR images, and an improper algorithm may not be able to improve the accuracy of land cover classification. Thirdly, through the analysis and comparison of UA and PA of land cover types, we found that the improvement of classification results of IHS-GTF fusion images is mainly concentrated in BIS and BS; there is no improvement in DIS, or it is even lower than for Sentinel-2A; and there is no significant improvement in VG and WB. The reason may be that SAR images are sensitive to soil moisture and roughness, which is conducive to distinguishing BIS and BS with similar spectra in optical images. However, the fused image retain the information in SAR images, making some original DIS in the source optical image become BIS, so their classification results of DIS are poor. Finally, IHS, GTF, and DWT are poor at distinguishing BIS and DIS, but are better at distinguishing BS. This is because the fused images of IHS, GTF, and DWT retain more spatial details of SAR image, but the definition of BIS and DIS in optical images is not suitable for SAR image. BIS and BS with the spectral similarity are easily confused in optical images, but they can be distinguished in SAR image because SAR is sensitive to soil moisture and roughness. The classification results are consistent with the fusion results.

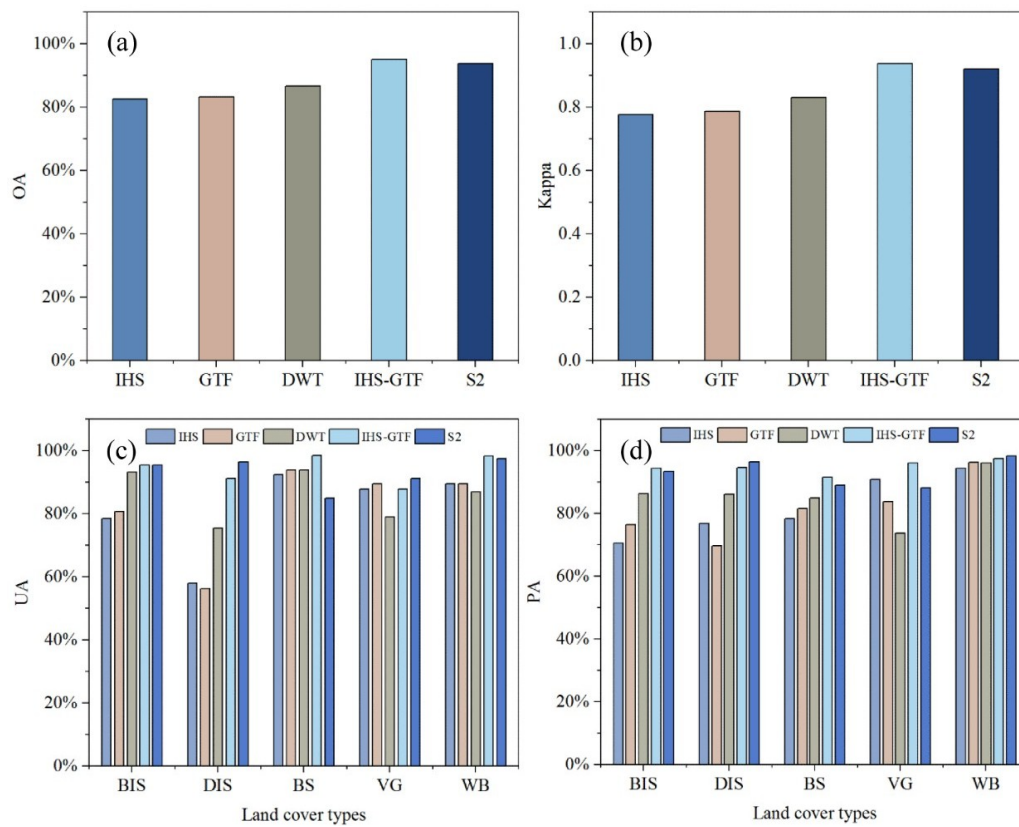


Figure 7. Comparison of classification results with different images in Site 1: (a) OA; (b) Kappa coefficients; (c) UA; and (d) PA.

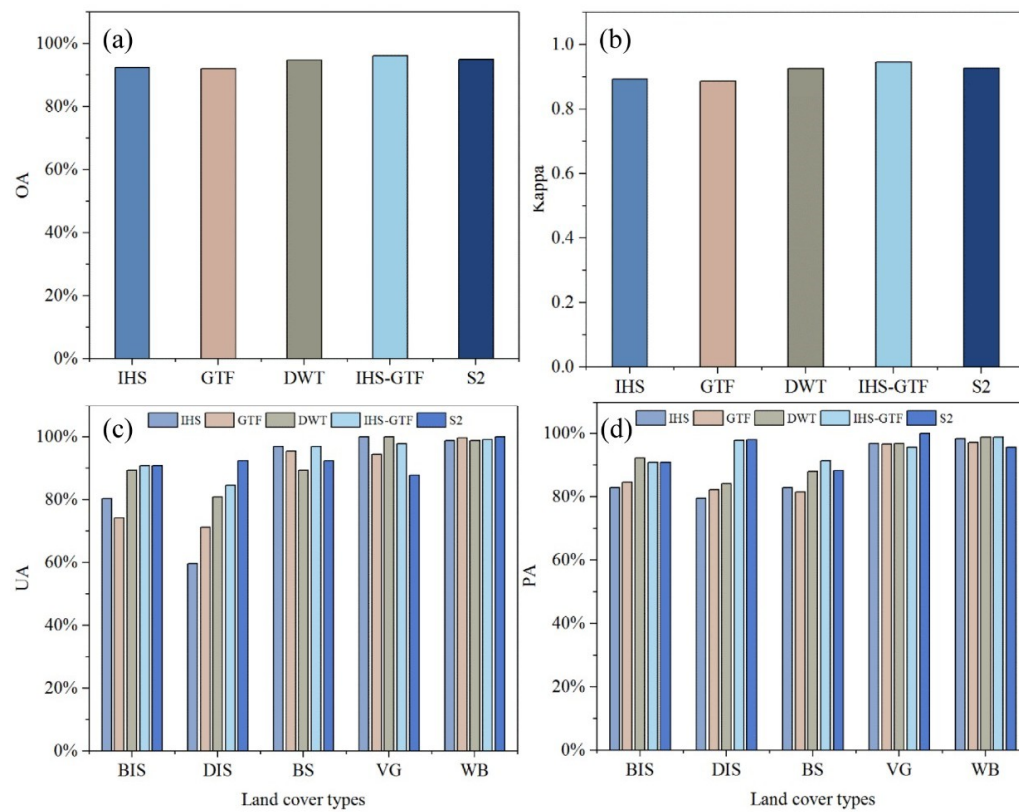


Figure 8. Comparison of classification results with different images in Site 2: (a) OA; (b) Kappa coefficients; (c) UA; and (d) PA.

Figures 9 and 10 show the impervious surface extraction results using different images. They are similar visually and can reflect the distribution of IS. However, through careful observation, we can find that the extraction results of IHS, GTF, and DWT are relatively fragmented (Figures 9a–c and 10a–c). This is mainly because the fused images preserve spatial details as well as speckle noise in SAR images. This again proves that our proposed fusion method is superior to the three other fusion methods. Table 4 shows the corresponding confusion matrices for the IS extraction using different images. We can find that the accuracy of IS extraction from IHS–GTF fusion result is improved, while those of the other fusion methods are all lower than that of Sentinel-2A, which is completely consistent with the previous classification results.

To sum up, we can conclude that the appropriate pixel-level fusion algorithm of optical and SAR images can improve the accuracy of IS extraction, especially in the recognition of BIS and BS, which is consistent with the previous studies [18,42]. These results demonstrate the effectiveness of the propose method, which can be used to generate useful data.

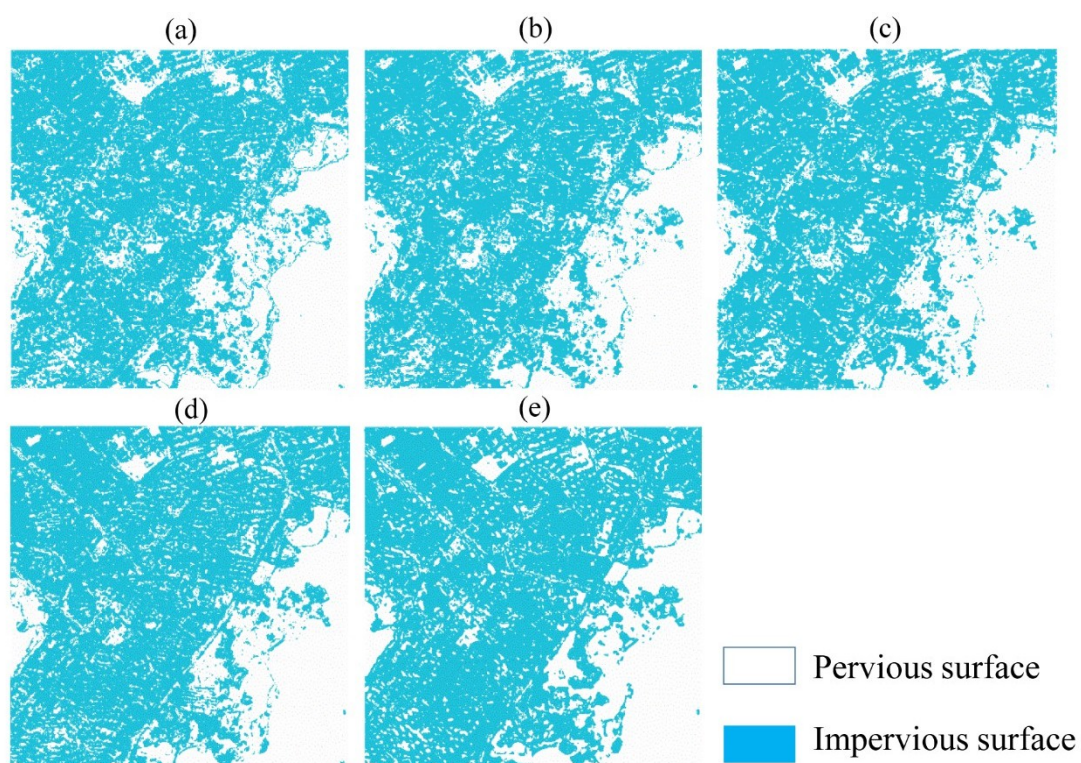


Figure 9. The IS extraction results using different images in Site 1: (a–e) IS results of IHS, GTF, DWT, IHS–GTF, and the Sentinel-2A image, respectively.

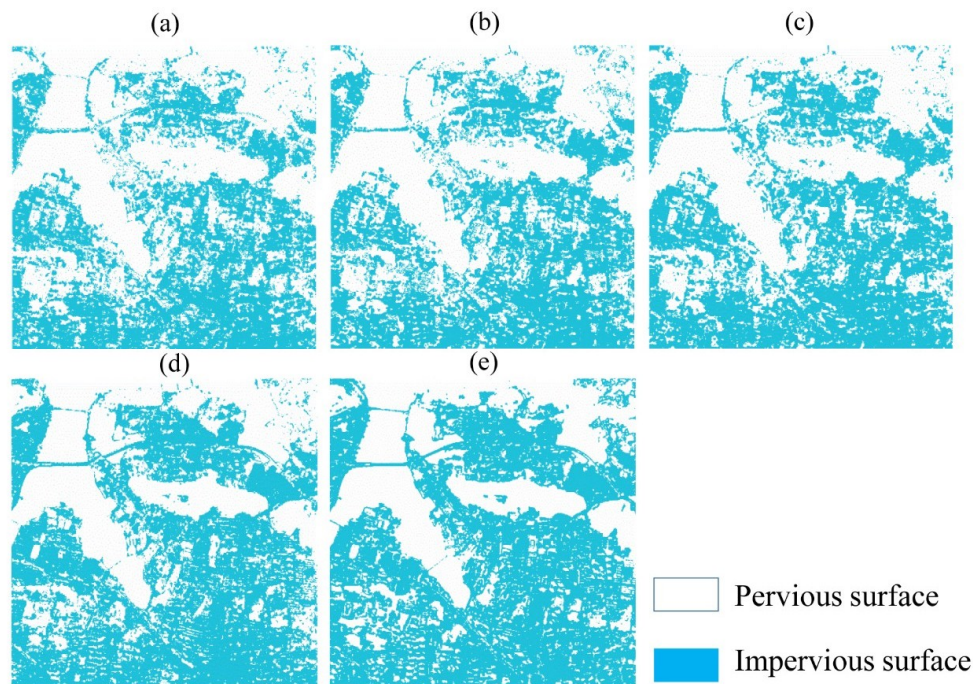


Figure 10. The IS extraction results using different images in Site 2: (a–e) IS results of IHS, GTF, DWT, IHS-GTF, and the Sentinel-2A image, respectively.

Table 4. Confusion matrices for the IS extraction using different images.

		Site 1			Site 2			
IHS		IS	NIS	UA	IS	NIS	UA	
	IS	129	12	91.49%	IS	101	2	98.06%
	NIS	16	225	93.36%	NIS	17	395	95.87%
	PA	88.97%	94.94%		PA	85.59%	99.50%	
	OA	92.67%	Kappa	0.8435	OA	96.31%	Kappa	0.8907
GTF		IS	NIS	UA	IS	NIS	UA	
	IS	132	7	94.96%	IS	99	4	96.12%
	NIS	13	230	94.65%	NIS	19	393	95.39%
	PA	91.03%	97.05%		PA	83.90%	98.99%	
	OA	94.76%	Kappa	0.8879	OA	95.53%	Kappa	0.8677
DWT		IS	NIS	UA	IS	NIS	UA	
	IS	136	9	93.79%	IS	107	7	93.86%
	NIS	9	228	96.20%	NIS	11	390	97.26%
	PA	93.79%	96.20%		PA	90.68%	98.24%	
	OA	95.29%	Kappa	0.9000	OA	96.50%	Kappa	0.8999
IHS-GTF		IS	NIS	UA	IS	NIS	UA	
	IS	139	5	96.53%	IS	110	1	99.10%
	NIS	6	232	97.48%	NIS	8	396	98.02%
	PA	95.86%	97.89%		PA	93.22%	99.75%	
	OA	97.12%	Kappa	0.9388	OA	98.25%	Kappa	0.9495
S2		IS	NIS	UA	IS	NIS	UA	
	IS	139	8	94.56%	IS	110	5	95.65%
	NIS	6	229	97.45%	NIS	8	392	98.00%
	PA	95.86%	96.62%		PA	93.22%	98.74%	
	OA	96.34%	Kappa	0.9224	OA	97.48%	Kappa	0.9279

5. Conclusions

Optical and SAR images fusion can make full use of the complementary information of both images to generate a higher quality image, which has been widely used in various fields, especially in land cover classification. It is very meaningful to develop image fusion algorithms for them. In this paper, a fusion algorithm named IHS-GTF is proposed. The impervious surface extraction was selected as an application example to further test the proposed method. From the experiment results, the proposed method shows good preservation of spatial details and spectral information and improves the overall accuracy of IS extraction by 2% compared with that of using optical image alone. The results demonstrate the ability of the proposed method for fusing optical and SAR data effectively to generate useful data which can be used in some subsequent applications. In future work, we can develop new image fusion algorithms at the pixel level, such as using deep learning and sparse representation, to further improve the fusion effect.

Author Contributions: Data curation, W.W. and S.G.; funding acquisition, Z.S.; methodology, Z.S. and W.W.; supervision, Z.S.; validation, W.W.; writing—original draft, W.W.; and writing—review and editing, Z.S. and S.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National key R & D plan on strategic international scientific and technological innovation cooperation special project under Grant 2016YFE0202300, the National Natural Science Foundation of China under Grants 61671332, 41771452, and 41771454, the Natural Science Fund of Hubei Province in China under Grant 2018CFA007, and the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing Special Research Funding.

Acknowledgments: We thank the editor and anonymous reviewers for their constructive comments and suggestions that improved the manuscript.

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

References

1. Shao, Z.; Zhou, W.; Deng, X.; Zhang, M.; Cheng, Q. Multilabel Remote Sensing Image Retrieval Based on Fully Convolutional Network. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 318–328. [[CrossRef](#)]
2. Dekker, R.J. Texture Analysis and Classification of ERS SAR Images for Map Updating of Urban Areas in The Netherlands. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1950–1958. [[CrossRef](#)]
3. Shao, Z.; Tang, P.; Wang, Z.; Saleem, N.; Yam, S.; Sommai, C. BRRNet: A Fully Convolutional Neural Network for Automatic Building Extraction From High-Resolution Remote Sensing Images. *Remote Sens.* **2020**, *12*, 1050. [[CrossRef](#)]
4. Samadhan, C.K.; Priti, P.R. Pixel Level Fusion Techniques for SAR and Optical Images: A Review. *Inf. Fusion* **2020**, *59*, 13–19.
5. Pohl, C.; Van Genderen, J. Multisensor Image Fusion in Remote Sensing: Concepts, Methods and Applications. *Int. J. Remote Sens.* **1998**, *19*, 823–854. [[CrossRef](#)]
6. Shao, Z.; Pan, Y.; Diao, C.; Cai, J. Cloud Detection in Remote Sensing Images Based on Multiscale Features-Convolutional Neural Network. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 4062–4076. [[CrossRef](#)]
7. Zhenfeng, S.; Juan, D.; Lei, W.; Yewen, F.; Neema, S.; Qimin, C. Fuzzy AutoEncode Based Cloud Detection for Remote Sensing Imagery. *Remote Sens.* **2017**, *9*, 311.
8. Calabresi, G. The Use of ERS SAR for Flood Monitoring: AN Overall Assessment. In Proceedings of the Second ERS Application Workshop, London, UK, 6–8 December 1995; pp. 237–241.
9. Floyd, M.H.; Xia, Z.G. SAR applications in human settlement detection, population estimation and urban land use pattern analysis: A status report. *IEEE Trans. Geosci. Remote Sens.* **1997**, *35*, 79–85.
10. Joshi, N.; Baumann, M.; Ehammer, A.; Reiche, J. A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring. *Remote Sens.* **2016**, *8*, 70. [[CrossRef](#)]
11. Gamba, P.; Dell’Acqua, F. Fusion of radar and optical data for identification of human settlements. In *Remote Sensing Impervious Surface*; Taylor & Francis: Boca Raton, FL, USA, 2008; pp. 143–159.
12. Guo, Q.; Li, A.; Zhang, H.Q.; Feng, Z.K. Remote Sensing Image Fusion Based on Discrete Fractional Random Transform for Modified IHS. *Pet. Chem.* **2013**, *XL-7/W1*, 19–22. [[CrossRef](#)]

13. Pal, S.K.; Majumdar, T.J.; Bhattacharya, A.K. ERS-2 SAR and IRS-1C LISS III data fusion: A PCA approach to improve remote sensing based geological interpretation. *Isprs J. Photogramm. Remote Sens.* **2007**, *61*, 281–297. [[CrossRef](#)]
14. Abd, M.S.; Norwati, M.; Nasir, S.M.; Azura, H.N.; Abdul, H.M.R. Comparison of Classification Techniques on Fused Optical and SAR Images for Shoreline Extraction: A Case Study at Northeast Coast of Peninsular Malaysia. *J. Comput. Sci.* **2016**, *12*, 399–411.
15. Dupas, C. SAR and LANDSAT TM image fusion for land cover classification in the Brazilian Atlantic Forest Domain. *Int. Arch. Photogram. Remote Sens.* **2000**, *33*, 96–103.
16. Chandrakanth, R.; Saibaba, J.; Varadan, G.; Raj, P.A. Feasibility of high resolution SAR and multispectral data fusion. In Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS), IEEE, Vancouver, BC, Canada, 24–29 July 2011.
17. Do, M.N.; Member; Ieee; Fellow; Ieee. The Contourlet Transform: An Efficient Directional Multiresolution Image Representation. *IEEE Trans. Image Process. A Publ. IEEE Signal Process. Soc.* **2006**, *14*, 2091–2106. [[CrossRef](#)] [[PubMed](#)]
18. Wu, W.; Guo, S.; Cheng, Q. Fusing optical and synthetic aperture radar images based on shearlet transform to improve urban impervious surface extraction. *J. Appl. Remote Sens.* **2020**, *14*, 1. [[CrossRef](#)]
19. Chen, S.; Zhang, R.; Su, H.; Tian, J.; Xia, J. SAR and Multispectral Image Fusion Using Generalized IHS Transform Based on à Trouis Wavelet and EMD Decompositions. *IEEE Sens. J.* **2010**, *10*, 737–745. [[CrossRef](#)]
20. Byun, Y.; Choi, J.; Han, Y. An Area-Based Image Fusion Scheme for the Integration of SAR and Optical Satellite Imagery. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2013**, *6*, 2212–2220. [[CrossRef](#)]
21. Zhang, W.; Yu, L. SAR and Landsat ETM+ image fusion using variational model. In Proceedings of the 2010–2010 International Conference on Computer and Communication Technologies in Agriculture Engineering, Chengdu, China, 12–13 June 2010. [[CrossRef](#)]
22. Yin, Z. Fusion algorithm of optical images and SAR with SVT and sparse representation. *Int. J. Smart Sens. Intell. Syst.* **2015**, *8*, 1123–1141. [[CrossRef](#)]
23. Wen, X.; Li, C. Feature Level Image Fusion for SAR and Optical Images. In Proceedings of the IET International Conference on Information Science & Control Engineering, Shenzhen, China, 18 January 2012.
24. Ma, J.; Chen, C.; Li, C.; Huang, J. Infrared and visible image fusion via gradient transfer and total variation minimization. *Inf. Fusion* **2016**, *31*, 100–109. [[CrossRef](#)]
25. Xu, R.; Zhang, H.; Lin, H. Urban Impervious Surfaces Estimation From Optical and SAR Imagery: A Comprehensive Comparison. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 4010–4021. [[CrossRef](#)]
26. Lin, Y.; Zhang, H.; Lin, H.; Gamba, P.E.; Liu, X. Incorporating synthetic aperture radar and optical images to investigate the annual dynamics of anthropogenic impervious surface at large scale. *Remote Sens. Environ.* **2020**, *242*, 111757. [[CrossRef](#)]
27. Shao, Z.; Fu, H.; Fu, P.; Yin, L. Mapping urban impervious surface by fusing optical and SAR data at the decision level. *Remote Sens.* **2016**, *8*, 945. [[CrossRef](#)]
28. Xu, H. Analysis of impervious surface and its impact on Urban heat environment using the normalized difference impervious surface index (NDISI). *Photogramm. Eng. Remote Sens.* **2010**, *76*, 557–565. [[CrossRef](#)]
29. Yang, J.; Sun, J.; Ge, Q.; Li, X. Assessing the impacts of urbanization-associated green space on urban land surface temperature: A case study of Dalian, China. *Urban For. Urban Green.* **2017**, *22*, 1–10. [[CrossRef](#)]
30. Weng, Q.; Lu, D.; Schubring, J. Estimation of land surface temperature–vegetation abundance relationship for urban heat island studies. *Remote Sens. Environ.* **2004**, *89*, 467–483. [[CrossRef](#)]
31. Wang, J.; Zhan, Q.M.; Guo, H.G.; Jin, Z.C. Characterizing the spatial dynamics of land surface temperature-impervious surface fraction relationship. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *45*, 55–65. [[CrossRef](#)]
32. Shao, Z.; Fu, H.; Li, D.; Altan, O.; Cheng, T. Remote sensing monitoring of multi-scale watersheds impermeability for urban hydrological evaluation. *Remote Sens. Environ.* **2019**, *232*, 111338. [[CrossRef](#)]
33. Yu, H.F.; Zhao, Y.L.; Fu, Y.C. Optimization of Impervious Surface Space Layout for Prevention of Urban Rainstorm Waterlogging: A Case Study of Guangzhou, China. *Int. J. Environ. Res. Public Health.* **2019**, *16*, 3613. [[CrossRef](#)]
34. Yu, H.; Zhao, Y.; Fu, Y.; Li, L. Spatiotemporal variance assessment of urban rainstorm waterlogging affected by impervious surface expansion: A case study of Guangzhou, China. *Sustainability* **2018**, *10*, 3761. [[CrossRef](#)]

35. Shao, Z.; Li, C.; Li, D.; Altan, O.; Zhang, L.; Ding, L. An Accurate Matching Method for Projecting Vector Data into Surveillance Video to Monitor and Protect Cultivated Land. *Isprs Int. J. Geo-Inf.* **2020**, *9*, 448. [[CrossRef](#)]
36. Carper, W.; Lillesand, T.; Kiefer, P. The use of intensity-hue-saturation transformations for merging SPOT panchromatic and multispectral image data. *Photogramm. Eng. Remote Sens.* **1990**, *56*, 459–467.
37. Chan, T.F.; Esedoglu, S. Aspects of total variation regularized L1 function approximation. *Siam J. Appl. Math.* **2005**, *65*, 1817–1837. [[CrossRef](#)]
38. Li, S.; Kang, X.; Hu, J. Image Fusion With Guided Filtering. *IEEE Trans. Image Process.* **2013**, *22*, 2864–2875. [[PubMed](#)]
39. Zhang, Y.; Zhang, H.; Lin, H. Improving the impervious surface estimation with combined use of optical and SAR remote sensing images. *Remote Sens. Environ.* **2014**, *141*, 155–167. [[CrossRef](#)]
40. Breiman, L.; Breiman, L.; Cutler, R.A. Random Forests Machine Learning. *J. Clin. Microbiol.* **2001**, *2*, 199–228.
41. Zhang, H.; Xu, R. Exploring the optimal integration levels between SAR and optical data for better urban land cover mapping in the Pearl River Delta. *Int. J. Appl. Earth Observ. Geoinf.* **2018**, *64*, 87–95. [[CrossRef](#)]
42. Zhang, H.; Li, J.; Wang, T.; Lin, H.; Zheng, Z.; Li, Y.; Lu, Y. A manifold learning approach to urban land cover classification with optical and radar data. *Landsc. Urban Plan.* **2018**, *172*, 11–24. [[CrossRef](#)]



© 2020 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/>).