

## Article

# Assessment of Pansharpening Methods Applied to WorldView-2 Imagery Fusion

Hui Li \*, Linhai Jing \* and Yunwei Tang

Key Laboratory of Digital Earth Sciences, Institute of Remote Sensing and Digital Earth,  
Chinese Academy of Sciences, Beijing 100094, China; tangyw@radi.ac.cn

\* Correspondence: lihui@radi.ac.cn (H.L.); jinglh@radi.ac.cn (L.J.); Tel.: +86-10-8217-8106 (L.J.)

Academic Editor: Vittorio M. N. Passaro

Received: 11 October 2016; Accepted: 26 December 2016; Published: 5 January 2017

**Abstract:** Since WorldView-2 (WV-2) images are widely used in various fields, there is a high demand for the use of high-quality pansharpened WV-2 images for different application purposes. With respect to the novelty of the WV-2 multispectral (MS) and panchromatic (PAN) bands, the performances of eight state-of-art pan-sharpening methods for WV-2 imagery including six datasets from three WV-2 scenes were assessed in this study using both quality indices and information indices, along with visual inspection. The normalized difference vegetation index, normalized difference water index, and morphological building index, which are widely used in applications related to land cover classification, the extraction of vegetation areas, buildings, and water bodies, were employed in this work to evaluate the performance of different pansharpening methods in terms of information presentation ability. The experimental results show that the Haze- and Ratio-based, adaptive Gram-Schmidt, Generalized Laplacian pyramids (GLP) methods using enhanced spectral distortion minimal model and enhanced context-based decision model methods are good choices for producing fused WV-2 images used for image interpretation and the extraction of urban buildings. The two GLP-based methods are better choices than the other methods, if the fused images will be used for applications related to vegetation and water-bodies.

**Keywords:** pansharpening; WorldView-2; quality indices; information indices

## 1. Introduction

The WorldView-2 (WV-2) satellite, launched in October 2009, offers eight multispectral (MS) bands of 1.84-m spatial resolution and a panchromatic (PAN) band of 0.46 m spatial resolution [1]. The MS bands cover the spectrum from 400 nm to 1050 nm, and include four conventional visible and near-infrared MS bands: blue (B, 450–510 nm), green (G, 510–580 nm), red (R, 630–690 nm), and near-IR1 (NIR1, 770–895 nm); and four new bands: coastal (C, 400–450 nm), yellow (Y, 585–625 nm), red edge (RE, 705–745 nm), and near-IR2 (NIR2, 860–1040 nm). The PAN band has a spectral response range of 450–800 nm, which covers shorter NIR spectral range than some common PAN bands of 450–900 nm. The WV-2 images have been widely used in various fields, e.g., geological structure interpretation [1], Antarctic land cover mapping [2], bamboo patch mapping [3], high density biomass estimation for wetland vegetation [4], mapping natural vegetation on a coastal site [5], predicting forest structural parameters [6], and especially for the detection of urban objects. Since numerous applications need high-spatial-resolution (HSR) MS images, it is highly desirable to fuse the eight MS bands and the PAN band to produce HSR MS imagery for better monitoring the Earth's surface.

Numerous pansharpening methods have been proposed in the last decades to produce spatially enhanced MS images by fusing the MS and PAN images. These methods are divided into two categories: the component substitution (CS) family and multi-resolution analysis (MRA) family. The CS approaches focus on the substitution of a component that is obtained by a spectral transformation of the

MS bands with the PAN image. The representative CS methods are the intensity-hue-saturation [7,8], principal component analysis [9], and Gram-Schmidt spectral sharpening (GS) [10,11] methods. The CS methods are easy to implement, and the generated fused MS images yield high spatial quality. However, the CS methods suffer from spectral distortions since the local dissimilarities between the PAN and MS channels, which are caused by different spectral response ranges, are not considered by them. The MRA-based techniques rely on the injection of the spatial details that are obtained through a multi-resolution decomposition of the PAN image into the up-sampled MS bands. Multi-resolution decomposition methods, such as “à trous” wavelet transform [12,13], undecimated or decimated Wavelet transform [14–16], Laplacian pyramids [17], Contourlet [18–20], and Curvelet [21], are often employed to extract spatial details of the PAN image. Although the MRA-based methods better preserve spectral information of the original MS images than the CS methods, they may cause spatial distortions, such as ringing or aliasing effects, originating shifts or blurred contours and textures [22]. Numerous hybrid schemes combining CS and MRA-based methods are developed to maximize spatial improvement and minimize spectral distortions [23–26]. In addition, several new pansharpening methods were proposed for the fusion of WV-2 imagery, i.e., the Hyperspherical Color Sharpening (HCS) [27] method and the improved Non-subsampled Contourlet Transform (NSCT) method [28]. These methods were proved to be better than early CS methods, such as GS, PCA.

Several studies have performed comparisons and analyses of some widely used state-of-the-art pansharpening methods, using test images covering different regions from several sensors. Previous studies showed that a pansharpening method may give different performances for test images from different sensors [29,30]. A noticeable point for the WV-2 is that the spectral ranges of the PAN band overlap limited part of the spectral ranges of the C, NIR1, and NIR2 bands. This will result in relative low correlation coefficients between these bands and the PAN bands, which may lead to spectral distortions of the fused version of these bands [31]. Regarding the wide use of the fused WV-2 images, it is urgent to evaluate the performances of different state-of-the-art pansharpening methods applied to WV-2 imagery. Some of the previous comparisons also used test images recorded by WV-2 [30,32–35] and other sensors. In these works, the early pansharpening methods, such as GS, PANSHARP, Ehlers, modified intensity-hue-saturation (M-IHS), high pass filter (HPF), principal component analysis (PCA), and wavelet-PCA (W-PCA) methods were assessed regarding quality indexes and visual inspection, usually using one or two test images covering urban areas. However, a fusion product providing the best performance in terms of quality indexes and visual inspection may be the best choice for applications such as image interpretation, but it may be not the best choice for applications related to classification and objects identification, i.e., the extraction of buildings, vegetation, and water-bodies [29,36,37]. Consequently, it is important to evaluate the widely used state-of-the-art pansharpening methods from the point of applications, such as land cover classification and object extraction. The purpose of this study was to assess the performances of the existing state-of-the-art pansharpening methods applied to WV-2 imagery, using information indices related to land cover classification and information extraction, as well as quality indexes and visual inspection. Several test images, presenting typical image scenes covering urban, suburban, and rural regions, are employed in the experiments. In addition, the newly proposed HCS, and NSCT methods, which are rarely included in previous comparisons, will be included in this work.

In this study, eight state-of-the-art algorithms, most of which have been demonstrated to outperform some other methods were assessed using both quality indices and information indices, along with visual inspection. The selected algorithms include four methods belonging to the CS family and four methods belonging to the MRA family. The four CS methods including Gram-Schmidt (GS) [10], adaptive GS (GSA) [38], Haze- and Ratio-based (HR) [39], and HCS [27] were compared. The four MRA methods include undecimated “à trous” wavelet transform (ATWT) using additive injection model [40,41], Generalized Laplacian pyramids (GLP) using spectral distortion minimal model (SDM) and context-based decision model (CBD) [42,43], and the improved NSCT method introduced in [28]. Traditional image quality indices couple with visual inspection were adopted

to assess the quality of the fused images. Four comprehensive indices, including Dimensionless Global Relative Error of Synthesis (ERGAS) [44], Spectral Angle Mapper (SAM) [45],  $Q2^n$  [46,47], and spatial correlation coefficient (SCC) [48] were employed to measure the spectral distortion between the fused and the original MS bands. Regarding the application purpose of the high-resolution fused images, which includes land cover classification of urban or suburban areas, bamboo and forest mapping, and so on, some widely used indexes, derived from the fusion products, were assessed to evaluate the information presentation ability of the fusion products. The employed indexes include morphological building index (MBI) [49], normalized difference vegetation index (NDVI), and normalized difference water index (NDWI). The information presentation of a fusion product was assessed using the correlation coefficient (CC) between an index derived from the fusion product and the same index derived from the corresponding original MS image. A higher CC value implies a better information preservation ability of the fusion product.

This paper is organized as follows: the eight selected pansharpening methods are introduced in Section 2, as well as the quality indexes; the experimental results with visual and quantitative comparisons with other outstanding fusion methods are presented in Section 3. Discussions are presented in Section 4, whereas the conclusions are summarized in Section 5.

## 2. Methodology

### 2.1. Algorithms

The algorithms used for the comparisons are introduced in the following subsections.  $MS$  and  $P$  represent the original low-resolution MS image and high-resolution PAN image, respectively.  $\widetilde{MS}$  and  $\widehat{MS}$  represent the up-sampled MS and the fused MS images, respectively. A general formulation of CS fusion is given by:

$$\widehat{MS}_i = \widetilde{MS}_i + g_i(P - I_L), i = 1, \dots, N \quad (1)$$

in which the subscript  $i$  indicates the  $i$ th spectral band;  $g_i$  is the injection gains of the  $i$ th band, while the intensity image  $I_L$  is defined by Equation (2):

$$I_L = \sum_{i=1}^N w_i \widetilde{MS}_i \quad (2)$$

where  $w_i$  is the weight of the  $i$ th MS band, and  $N$  is the number of MS bands.

Similarity, a general formulation for MRA-based methods can be given by Equation (3):

$$\widehat{MS}_i = \widetilde{MS}_i + g_i(P - P_L), i = 1, \dots, N \quad (3)$$

where  $P_L$  is the low-frequency component of the PAN band.  $P_L$  can be obtained by different approaches, such as low-pass filter, Laplacian pyramid and wavelet decomposition.

#### 2.1.1. GS and GSA

GS is a representative method of the CS family, the fusion process of which is described by Equation (1), with the injection gains  $g_i$  given by Equation (4):

$$g_i = \frac{\text{cov}(MS_i, I_L)}{\text{var}(I_L)} \quad (4)$$

where  $\text{cov}(X, Y)$  is the covariance between the two images  $X$  and  $Y$ , and  $\text{var}(X)$  is the variance of  $X$ .

Several versions of GS can be achieved by changing the method for generating  $I_L$ . One way to obtain  $I_L$  is simply averaging the MS components (i.e., using Equation (2) with setting  $w_i = 1$  for all  $i = 1, \dots, N$ ). This modality is referred as GS or GS mode 1. Another way is using the low-pass version of  $P$  as  $I_L$ . This modality is referred as GS mode 2.

An enhanced version, called adaptive GS (GSA), is proposed by assigning  $I_L$  as a weighted average of the original MS bands, as Equation (2). The weights  $w_i$  in Equation (2) are calculated with the minimum mean square error (MSE) solution of Equation (5):

$$P_0 = \sum_{i=1}^N w_i MS_i + \varepsilon \quad (5)$$

where  $P_0$  is the degraded version of  $P$ , with the same pixel sizes of the original MS bands.  $P_0$  is generated by low-pass filtering of  $P$ , followed by decimation. Both the GS and GSA methods are included in the experiments in this study.

### 2.1.2. HR

The HR method is based on the assumption that the ratio of a HSR MS band to a low-spatial-resolution (LSR) MS band is equal to the ratio of a HSR PAN image to an assumed LSR PAN image a fusion method that considers haze [50–52]. The fused  $i$ th MS band  $\widehat{MS}_i$  can be calculated by using Equation (6) according to HR fusion method:

$$\widehat{MS}_i = \left( \widetilde{MS}_i - H_i \right) \frac{P - H_p}{P_L - H_p} + H_i, i = 1, \dots, N \quad (6)$$

where  $P_L$  is a low-pass version of  $P$ ;  $H_i$  and  $H_p$  denote the haze values in the  $i$ th MS band and the PAN band, respectively. The values of  $H_i$  and  $H_p$  can be determined using the minimum grey level values in  $MS_i$  and  $P$  according to an image-based dark-object subtraction method [50–52].

### 2.1.3. HCS

HCS is a pansharpening method designed for WV-2 imagery based on transforming the MS bands into hyperspherical color space. The HCS method offers two modes, including the native mode and the smart mode. The process of the native mode is described as follows:

- (a) The squares of the multispectral intensity ( $I^2$ ) and the PAN ( $P^2$ ) are calculated using Equations (7) and (8), respectively:

$$I^2 = \sum_{i=1}^N \widetilde{MS}_i^2 \quad (7)$$

$$P^2 = (P)^2 \quad (8)$$

- (b) Calculate the mean ( $u_P$ ) and standard deviation ( $\sigma_P$ ) of  $P^2$ , as well as the mean ( $u_I$ ) and standard deviation ( $\sigma_P$ ) of  $I^2$ .
- (c) The  $P^2$  is adjusted to the mean and standard deviation of  $I^2$ , using Equation (9):

$$P^2 = \frac{\sigma_I}{\sigma_P} \left( P^2 - u_P + \sigma_P \right) + u_I - \sigma_I \quad (9)$$

- (d) The square root of the adjusted  $P^2$  is assigned to  $I_{adj}$  (i.e.,  $I_{adj} = \sqrt{P^2}$ ),  $I_{adj}$  is used in the reverse transform from HCS color space back to the original color space, using Equation (10):

$$\widetilde{MS}_i = \begin{cases} I_{adj} \cos \varphi_i, & i = 1 \\ I_{adj} \sin \varphi_1 \sin \varphi_2 \dots \sin \varphi_{i-1} \cos \varphi_i, & i = 2, \dots, N-1 \\ I_{adj} \sin \varphi_1 \sin \varphi_2 \dots \sin \varphi_{i-1}, & i = N \end{cases} \quad (10)$$

in which  $\varphi_i$  is defined using Equation (11):



$$\varphi_i = \begin{cases} \tan^{-1} \left( \sqrt{\sum_{k=2}^N \widetilde{MS}_k^2 / \widetilde{MS}_i} \right), & i = 1 \\ \tan^{-1} \left( \sqrt{\sum_{k=i+1}^N \widetilde{MS}_k^2 / \widetilde{MS}_i} \right), & i = 2, \dots, N-2 \\ \tan^{-1} \left( \widetilde{MS}_{i+1} / \widetilde{MS}_i \right), & i = N-1 \end{cases} \quad (11)$$

For the smart case, similarity to  $P^2$ , a  $PS^2$  is calculated by using the low-pass version of the PAN image ( $P_L$ ), i.e.,  $PS^2 = (P_L)^2$ . The  $P_L$  is generated by an average filtering with size of  $7 \times 7$ . Both the means and standard deviations of  $PS^2$  and  $P^2$  are adjusted to those of  $I^2$ . Then, the adjusted intensity  $I_{adj}$  is assigned by using Equation (12):

$$I_{adj} = \sqrt{(P^2 / PS^2) I^2} \quad (12)$$

Finally, the fused image can be generated by using Equations (10) and (11). The HCS method using the smart mode is considered in the experiment in this study.

#### 2.1.4. ATWT

The ATWT method is a MRA fusion approach that extracting spatial details using “à trous” wavelet transform. The fusion scheme utilizing the additive injection model can be formulated as:

$$\widetilde{MS}_i = \widetilde{MS}_i + P - P_L, i = 1, \dots, N \quad (13)$$

where  $P_L$  is the low frequency component of  $P$  and is generated by the “à trous” wavelet.

The “à trous” wavelet is a kind of non-orthogonal wavelet that is different from orthogonal and biorthogonal. It is a redundant transform, since decimation is not implemented during the process of wavelet transform while the orthogonal or biorthogonal wavelet can be carried out using either decimation or undecimation mode.

The whole process of the ATWT fusion can be divided into two steps [40]:

- (1) Use the à trous wavelet transform to decompose the PAN image to  $n$  wavelet planes. Usually,  $n = 2$  or 3.
- (2) Add the wavelet planes (i.e., spatial details) of the decomposed PAN images to each of the spectral bands of the MS image to produce fused MS bands.

#### 2.1.5. GLP

The fusion process of GLP can also be formulated as Equation (3). For the GLP method, the low-frequency component of the PAN band  $P_L$  is generated by up-sampling the down-sampled version of the original PAN band  $P$ . The down-sampling of  $P$  is implemented using a low-pass reduction filter that matches the MTF of the band, whereas the up-sampling of  $P$  is carried out by using an ideal expansion low-pass filter.

For the GLP fusion, different detail injection models are designed to obtain the injection coefficients  $g_i$ . The most used models are the spectral distortion minimizing (SDM) model and the context-based decision (CBD) model.

For the SDM model, the injection coefficients  $g_i$  can be obtained using Equation (14):

$$g_i(m, n) = \beta(m, n) \frac{\widetilde{MS}_i(m, n)}{P_L(m, n)}, i = 1, \dots, N \quad (14)$$

where  $\beta(m, n)$  is equal to 1 for all pixels in the original SDM model [42], whereas  $\beta(m, n)$  is defined as the ratio between average local standard deviations of resampled MS bands and local standard deviation of  $P_L$ , for the enhanced SDM (ESDM) model proposed by Aiazzi [43]:

$$\beta(m, n) = \sqrt{\frac{\frac{1}{N} \sum_{k=1}^N \text{var}[\widetilde{MS}_k](m, n)}{\text{var}[P_L](m, n)}} \quad (15)$$

In the CBD model, the space-varying coefficients  $g_i$  is defined as Equation (16):

$$g_i(m, n) = \begin{cases} \min\left\{\frac{\sigma_{\widetilde{MS}_i}(m, n)}{\sigma_{P_L}(m, n)}, c\right\} & \text{if } \rho_i(m, n) \geq \theta_i, \\ 0 & \text{if } \rho_i(m, n) < \theta_i, \end{cases} \quad (16)$$

in which  $\rho_i(m, n)$  is the local correlation coefficient between  $\widetilde{MS}_i$  and  $P_L$  calculated on a square sliding window of size  $L \times L$  centered on pixel  $(m, n)$ .

The CBD model is uniquely defined by the set of thresholds  $0 < \theta_i \leq 1$ , for  $i = 1, \dots, N$ , generally different for each band, and by the window size  $L$  depending on the spatial resolutions and scale ratio of the images be merged, as well as the landscape characteristics (typically,  $7 \leq L \leq 11$  to avoid loss of local sensitivity with  $L > 11$  and statistical instability with  $L < 7$ ). The thresholds may be related to the spectral content of the Pan image, e.g.,  $\theta_i = 1 - \rho_i$ , where  $\rho_i$  is the global correlation coefficient between the  $k$ th band and the Pan image spatially degraded to the same resolution. A clipping constant  $c$  was introduced to avoid numerical instabilities (empirically,  $2 \leq c \leq 3$ ).

For the Enhanced CBD (ECBD) model proposed by Aiazzi [43], the coefficients  $g_i$  is calculated as Equation (17):

$$g_i(m, n) = \min\left\{\frac{\sigma_{\widetilde{MS}_i}(m, n)}{\sigma_{P_L}(m, n)} \cdot \frac{\rho_i(m, n)}{\bar{\rho}_i}, c\right\} \quad (17)$$

where  $\bar{\rho}_i$  is the global correlation coefficient between  $\widetilde{MS}_i$  and  $P_L$ . Both the GLP methods using the ESDM and the ECBD models are considered in the experiment.

## 2.1.6. NSCT

The contourlet transform (CT) is proved to be a better approach than the wavelet for pan-sharpening. CT is implemented by a multiscale decomposition using the Laplacian pyramid followed by a local directional transform using the directional filter bank (DFB).

The NSCT is a shift-invariant version of CT and has excellent multilevel and multi-direction properties. NSCT is built upon the non-subsampled pyramid filter banks (NSPFBs) and the non-subsampled directional filter banks (NSDFBs). The NSPFB employed by NSCT is a 2-D two-channel non-subsampled filter bank, whereas the NSDFB employed by NSCT is a shift-invariant version of the critically sampled DFB in CT. The details of the NSCT can be attended in [28,53]. An improved version of the standard NSCT-based method is introduced in [28]. The process of the improved NSCT method with mode 2 (NSCT\_M2) in [28] is described as follows.

- Each original MS band  $MS_i$ , is decomposed using 1-level NSCT to get one coarse level,  $C_0^{MS_i}$ , and one fine level,  $C_1^{MS_i}$ ;
- The PAN band is decomposed using 3-level NSCT into one coarse level,  $C_0^{PAN}$ , and three fine levels, which are denoted as  $C_1^{PAN}$ ,  $C_2^{PAN}$ , and  $C_3^{PAN}$ , respectively.
- The coefficients of each MS band,  $C_0^{MS}$  and  $C_1^{MS}$ , are up-sampled to the scale of the PAN band using the bi-linear interpolation algorithm.

- (d) The coarse level of the fused  $i$ th MS band,  $C_0^{F_i}$ , is the up-sampled coarse level of the  $i$ th MS band  $C_0^{MS_i}$ , whereas the fine levels 2 ( $C_2^{F_i}$ ) and 3 ( $C_3^{F_i}$ ) of the fused  $i$ th MS band are the fine levels 2 ( $C_2^{PAN}$ ) and 3 ( $C_3^{PAN}$ ) of the PAN band.
- (e) The fused fine level 1,  $C_1^{F_i}$ , is obtain by fusing the coefficients of the same level obtained from both the  $i$ th MS band and the PAN band. For each pixel  $(x, y)$ , the coefficients of the fused fine level 1,  $C_1^{F_i}(x, y)$ , is determined according to Equation (18):

$$C_1^{F_i}(x, y) = \begin{cases} C_1^{MS_i}(x, y), & LE_{MS_i}(x, y) > LE_{PAN}(x, y) \\ C_1^{PAN}(x, y), & otherwise \end{cases} \quad (18)$$

where  $LE_{MS_i}(x, y)$  and  $LE_{PAN}(x, y)$  are the local energy of pixel  $(x, y)$  for the  $i$ th MS band and the PAN band, respectively, calculated within a  $(2M + 1) \times (2P + 1)$  window using the formula shown in Equation (19):

$$LE(x, y) = \sum_{i=-M}^M \sum_{j=-P}^P (C_1(x + i, y + j))^2 \quad (19)$$

The inverse NSCT is applied to the fused coefficients to provide the fused  $i$ th MS band.

This improved version was demonstrated to provide pansharpened images with a good spectral quality.

## 2.2. Quality Indexes

Quality assessment of fusion products can be performed using two approaches. The first approach considers fusing images at a spatial resolution lower than the original resolution and uses the original MS image as a reference to assess the quality of the fused images. Several indexes have been proposed for evaluating the spatial and spectral distortions of the fused image with respect to an available reference image. The widely used spectral quality indexes include Root Mean Square Error (RMSE), Relative Average Spectral Error (RASE) [54], ERGAS, SAM, Universal Image Quality Index (UIQI) [55], Q4 [46], Q2n, and Peak Signal-to-Noise Ratio (PSNR) [56], whereas the widely used spatial indexes include SCC and Structural SIMilarity (SSIM) [57]. The second approach uses quality indexes that do not require a reference image but operate on relationships among the original images and the fusion products. This approach has the advantage of validating the products at the original scale, thus avoiding any hypothesis on the behavior at different scales. However, appropriate indexes requiring no reference should be exploited to assess the quality of the fusion product. The Quality with no reference index (QNR) was one of the mostly used indexes [58]. It is composed by the product of two separate values that quantify spectral and spatial distortions, respectively. However, QNR is proved to be lower reliability than the indexes belonging to the first approach [30], since it can be affected by slightly mismatches among the original image bands. The acquisition modality of WV-2 can led to a small temporal misalignments among the MS bands, since the eight MS bands of are arranged in two arrays of four bands each. Consequently, the QNR index is not used in this study. Four quality indexes belong to the first approach, including ERGAS, SAM, Q2n, and SCC, are employed to assess the quality of fused images at data level. These indexes are chosen due to they are widely used in literatures related to fusion of remote sensing imagery.

### 2.2.1. ERGAS

The global index ERGAS is an improved version of the index named Relative Average Spectral Error (RASE), which is defined based on Root Mean Square Error (RMSE). The formula of ERGAS [44] is defined as follows:

$$\text{ERGAS} = \frac{100}{R} \sqrt{\frac{1}{N} \sum_{k=1}^N \left( \frac{\text{RMSE}(\widehat{MS}_k, MS_k)}{u(MS_k)} \right)^2} \quad (20)$$

where  $u(MS_k)$  is the mean of the  $k$ th band of the reference image;  $R$  is the spatial resolution ratio between the MS and PAN bands; RMSE is defined as Equation (21):

$$\text{RMSE}(\widehat{MS}_k, MS_k) = \sqrt{E \left[ (\widehat{MS}_k - MS_k)^2 \right]} \quad (21)$$

The optimal value for ERGAS is 0, since it is defined as a weighted sum of RMSE values.

### 2.2.2. SAM

SAM expresses the spectral similarity between a fused image and a reference image with the average spectral angle of all pixels involved [45]. Let two spectral vectors  $V = \{V_1, V_2, \dots, V_N\}$  and  $\widehat{V} = \{\widehat{V}_1, \widehat{V}_2, \dots, \widehat{V}_N\}$  present the reference spectral pixel and the fused spectral pixel, respectively, their spectral angle SAM is defined as in Equation (22):

$$\text{SAM} = \arccos \left( \frac{\langle V, \widehat{V} \rangle}{|V| \cdot |\widehat{V}|} \right) \quad (22)$$

where  $\langle X, Y \rangle$  stands for the inner-product of the two vectors  $X$  and  $Y$ , and  $|X|$  stands for the modulus of a vector  $X$ .

The smaller the spectral angle, the higher the similarity between the two vectors. Since the angle is independent of the magnitudes of the two vectors, the index SAM is not affected by solar illumination factors.

### 2.2.3. Q2<sup>n</sup>

Q2<sup>n</sup> is a generalization of Q index for monoband images [47] and an extension of Q4 [46]. Q2<sup>n</sup> is derived from the theory of hypercomplex numbers, particularly of 2<sup>n</sup>-ones [59,60]. For a 2<sup>n</sup>-on hypercomplex random variable  $z$  (in boldface) is written in the following form:

$$z = z_0 + i_1 \cdot z_1 + \dots + i_{2^n-1} \cdot z_{2^n-1} \quad (23)$$

where  $z_0, z_1, \dots, z_{2^n-1}$  are real numbers, and  $i_0, i_1, \dots, i_{2^n-1}$  are hypercomplex unit vectors, the conjugate  $z^*$  is given by:

$$z^* = z_0 - i_1 \cdot z_1 - \dots - i_{2^n-1} \cdot z_{2^n-1} \quad (24)$$

and the modulus  $|z|$  is defined by Equation (25):

$$|z| = \sqrt{z_0^2 + z_1^2 + \dots + z_{2^n-1}^2} \quad (25)$$

Give two 2<sup>n</sup>-on hypercomplex random variables  $z$  and  $v$ , the hypercomplex covariance between  $z$  and  $v$  is defined as Equation (26):

$$\text{cov}(z, v) \triangleq E[(z - \bar{z})(v - \bar{v})] = E[zv^*] - \bar{z} \cdot \bar{v}^* = \sigma_{z,v} \quad (26)$$

The hypercomplex CC between the two 2<sup>n</sup>-on random variables are defined as the normalized covariance:

$$q(z, v) \triangleq \frac{\sigma_{z,v}}{\sigma_z \sigma_v} \quad (27)$$

in which  $\sigma_z$  and  $\sigma_v$  are the square roots of the variances of  $z$  and  $v$ , and are obtained by Equations (28) and (29), respectively:

$$\sigma_z = E\left[|z|^2\right] - |\bar{z}|^2 \quad (28)$$

$$\sigma_v = E\left[|v|^2\right] - |\bar{v}|^2 \quad (29)$$

The index  $Q2^n$  can be computed from Equation (30):

$$Q2^n_{M \times M} = \frac{\sigma_{z,v}}{\sigma_z \sigma_v} \cdot \frac{2\bar{z}\bar{v}}{(\bar{z})^2 + (\bar{v})^2} \cdot \frac{2\sigma_z \sigma_v}{\sigma_z^2 + \sigma_v^2} \quad (30)$$

where  $M$  is the size of the local window used to calculate  $Q2^n$ .

Finally,  $Q2^n$  is obtained by averaging the magnitudes of all  $Q2^n_{M \times M}$  over the whole image, according to Equation (31):

$$Q2^n = E[|Q2^n_{M \times M}|] \quad (31)$$

According to [46], the value of  $M$  is suggested to be 32.

#### 2.2.4. SCC

To assess the spatial quality of the fusion products, the spatial details presented in the fused images will be compared with those presented in the reference image by calculating the correlation coefficient between the spatial details extracted from the two images. Similar to the procedure proposed by Otazu et al. [48], the spatial information presented in the two images to be compared is extracted by using a Laplacian filter, then, the correlations between the two filtered images are calculated band by band. However, an overall correlation coefficient of the two edge images with eight bands is calculated in this study. A high SCC value indicates that many of the spatial details of the reference image are presented in the fused image.

### 2.3. Information Indexes

In order to assess the ability of information extraction of the fusion products employed in specific remote sensing applications, we assessed the quality of fused images by the use of a series of information indices, which are proved to be useful in land cover classification and information extraction in previous studies. Three indices are employed in this study: MBI, NDVI, and NDWI. The accuracy of an information index derived from a fusion product is assessed using the CC between the information index and the same information index derived from the corresponding reference MS image. A higher CC value implies a better information preservation ability of the fusion product in terms of the information index. Henceforth, the CC values calculated for MBI, NDVI, and NDWI are denoted as  $C_{MBI}$ ,  $C_{NDVI}$ , and  $C_{NDWI}$ , respectively. The employed three information indices are introduced in this subsection.

#### 2.3.1. NDVI

Based on the principle that vegetation has a strong reflectance in the near-infrared (NIR) channel but a strong absorption in the red (Red) channel, the NDVI is defined as Equation (32):

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (32)$$

#### 2.3.2. NDWI

The NDWI is defined using the spectral value of the Green band (G) and the NIR band as Equation (33):

$$NDWI = \frac{G - NIR}{G + NIR} \quad (33)$$

### 2.3.3. MBI

The MBI is proposed by Huang et al. [49], aiming to represent spectral and spatial characteristics of buildings using a series of morphological operators. The MBI is calculated according to the following steps:

- (a) A brightness image  $b$  is generated by setting the value of each pixel  $p$  to be the maximum digital number of the visible bands. Only the visible channels are considered due to they have the most significant contributions to the spectral property of buildings.
- (b) The directional white top-hat (WTH) reconstruction is employed to highlight bright structures that have a size equal to or smaller than the size of the structure element (SE), and meanwhile suppresses other dark structures in the image. WTH with linear SE is defined as Equation (34):

$$WTH(s, \theta) = b - \gamma_{SE}(s, \theta) \quad (34)$$

in which  $\gamma_{SE}(s, \theta)$  is an opening by reconstruction operator using a linear SE with a size of  $s$  and a direction of  $\theta$ .

- (c) The difference morphological profiles (DMP) of white top-hat transforms are employed to model building structures in a multi-scale manner:

$$\begin{cases} DMP_{WTH} = \{DMP_{WTH}(s, \theta) : s_{min} \leq s \leq s_{max}, \theta \in D\} \\ DMP_{WTH}(s, \theta) = |WTH(s + \Delta s, \theta) - WTH(s, \theta)| \end{cases} \quad (35)$$

- (d) Finally, MBI is defined based on DMP using Equation (36):

$$MBI = \frac{\sum_{s, dir} DMP_{WTH}(s, \theta)}{N_D \times N_s} \quad (36)$$

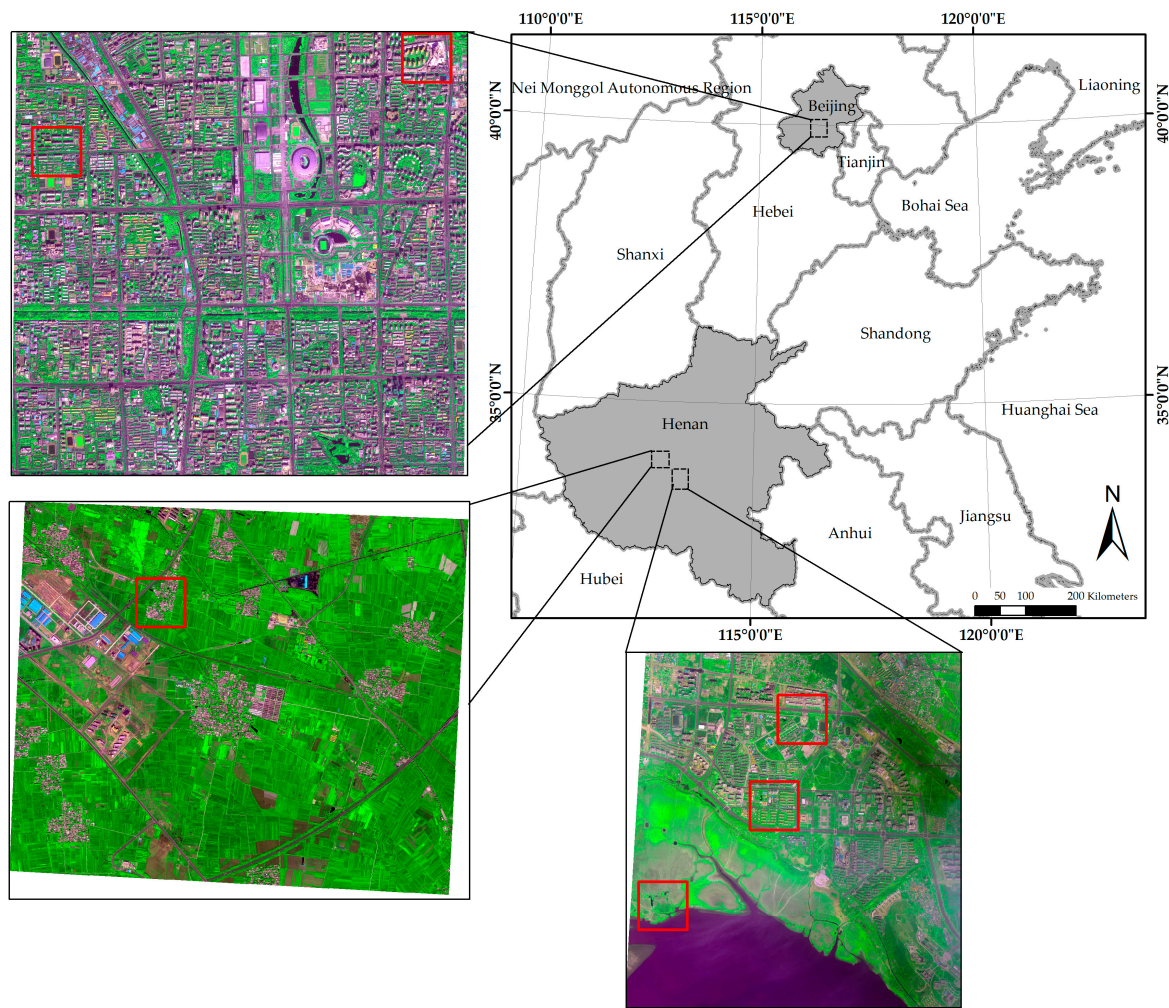
where  $N_D$  and  $N_s$  are the directionality and the scale of the DMPs, respectively. The definition of the MBI is based on the fact that building structures have high local contrast and, hence, have larger feature values in most of the directions of the profiles. Accordingly, buildings will have large MBI values.

## 3. Experimental Results

### 3.1. Datasets

Six datasets clipped from three WV-2 scenes were considered in this study. One scene covers Beijing City, China, whereas the other two scenes cover Pingdingshan City, Hebei Province, China. The Beijing scene was acquired on 21 September 2013, with an off-nadir angle of  $13.7^\circ$ . One of the Hebei scenes was acquired on 29 April 2014, with an average off-nadir angle of  $12.5^\circ$ , whereas the other was obtained on 21 August 2014, with an average off-nadir angle of  $8.6^\circ$ . The locations of the three scenes are shown in Figure 1. The red rectangles in the figure indicate the locations of the six datasets. Two datasets covering urban areas were selected from the Beijing scene, whereas the other four datasets were obtained from the other two WV-2 scenes. Two of them cover suburb areas, whereas the other two cover rural areas. For all the datasets, the radiometric resolution is 16 bits, whereas the spatial resolution ratio  $R$  is 4. Each dataset has a size of  $256 \times 256$  pixels at MS scale. The two urban images are henceforth referred as I1 and I2, respectively, whereas the two suburban images are referred as I3 and I4, respectively. The two rural images are referred as I5 and I6, respectively. The typical image objects shown in the six images are listed in Table 1; the six images are shown in Figure 2.





**Figure 1.** The locations of the three WV-2 scenes employed in this work. The locations of the six test images used in the experiments are marked by the rectangles in red.

To evaluate the quality of the fusion products with respect to the original MS images, the degraded datasets were produced by reducing the original MS and PAN images to a spatial resolution of 8 m and 2 m, respectively. All the fusion experiments were performed on the degraded datasets. The fused 2-m images were compared with the 2-m true MS images to assess their quality. The degraded images were generated using averaging in this study [61]. The spatial resolution ratio  $R$  in this study is 4 for all the tested images.

According to the image objects included in the six test images, different information indices were chosen for each of the images (Table 1). The fusion products of the two urban images were assessed with respect to  $C_{MBI}$  and  $C_{NDVI}$ , due to no water bodies can be observed from these two images. The fusion products of the two suburban images were assessed using all the three indices, whereas the fusion products of the two rural images were assessed using only  $C_{NDVI}$  and  $C_{NDWI}$ .



**Figure 2.** The selected six dataset used in the experiments. (a) I1; (b) I2; (c) I3; (d) I4; (e) I5; (f) I6.

**Table 1.** The description of the selected six datasets.

Image	Location	Type	Objects	Information Indices
I1	Beijing	Urban	High buildings, squares, roads, vegetation, shadows	MBI, NDVI
I2	Beijing	Urban	Moderate buildings, squares, roads, vegetation, shadows	MBI, NDVI
I3	Pingdingshan	Suburban	Low buildings, squares, roads, vegetation, shadows, water bodies	MBI, NDVI, NDWI
I4	Pingdingshan	Suburban	Low buildings, squares, roads, vegetation, shadows, water bodies	MBI, NDVI, NDWI
I5	Pingdingshan	Rural	Building, roads, farms, water bodies	NDVI, NDWI
I6	Pingdingshan	Rural	Vegetation, water bodies, bare soils	NDVI, NDWI

### 3.2. Fusing Using the Selected Algorithms

All the fusion algorithms were implemented in MATLAB version 2014b. For all the selected eight fusion methods, the up-sampled MS images were produced using the bi-cubical interpolation approach. For the HR method, the low-spatial-resolution version of PAN image ( $P_L$ ) was obtained by averaging the PAN pixels in an  $R \times R$  window, followed by the bi-cubically up-sampled to the resolution of the original PAN image. The GLP methods using the ESDM model and the ECBD model were used in the fusion experiment; the two methods are referred as GLP\_ESDM and GLP\_ECB, respectively, in this study. For the HR method, the haze values for the MS and PAN bands were determined using the values of the pixel that offering the lowest value in the PAN image.

### 3.3. Quality Indexes

#### 3.3.1. Assessment for the Two Urban Images

The image quality indices of the fusion products of the two urban images are shown in Table 2. The fusion products of these two images are partly shown in Figures 3 and 4, respectively, in order to

facilitate the observation of the details of the fused images. The images in each figure in this work are stretched by using an identical histogram obtained from the corresponding reference MS images.

**Table 2.** Quality indices of the fused images generated by the selected algorithms for the two urban images.

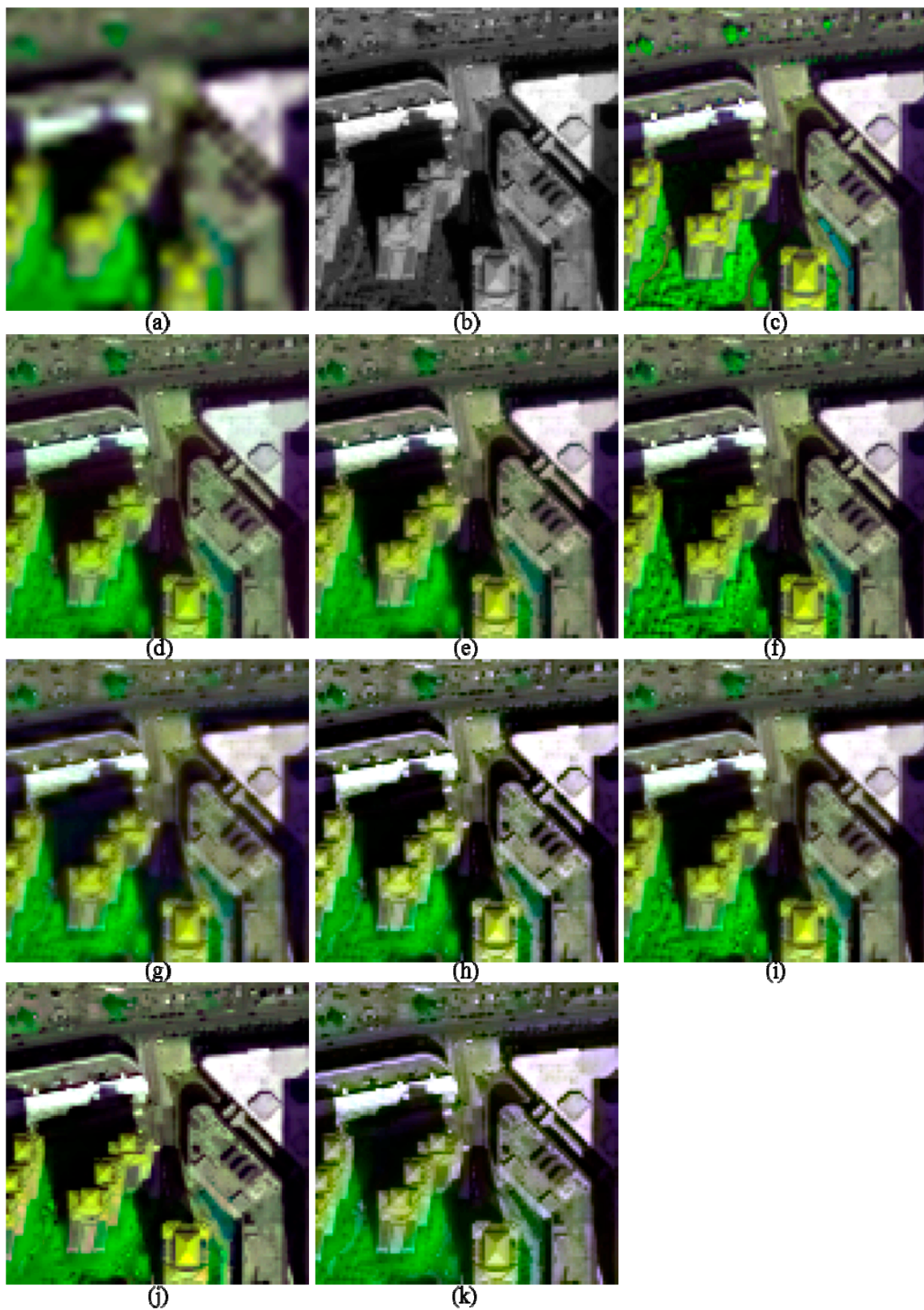
Image	Method	ERGAS	SAM	Q8	SCC
I1	GS	1.64	2.19	0.957	0.911
	GSA	1.28	1.92	0.974	0.908
	HR	1.26	1.76	0.977	0.911
	HCS	1.95	2.58	0.897	0.881
	ATWT	2.09	2.72	0.879	0.863
	GLP_ESDM	1.56	2.05	0.956	0.898
	GLP_ECBD	1.93	2.36	0.956	0.888
	NSCT_M2	2.22	3.66	0.849	0.861
	EXP	1.64	2.19	0.857	0.582
I2	GS	2.70	3.79	0.909	0.844
	GSA	2.22	3.45	0.946	0.839
	HR	2.20	3.14	0.951	0.849
	HCS	2.87	3.99	0.850	0.809
	ATWT	2.76	4.02	0.861	0.799
	GLP_ESDM	2.51	3.47	0.916	0.824
	GLP_ECBD	3.14	4.26	0.905	0.785
	NSCT_M2	3.04	5.03	0.831	0.795
	EXP	3.84	3.99	0.797	0.499

The HR method offers the highest Q8 and SCC values and the lowest ERGAS and SAM values for the two urban images, indicating that the HR method gives the best performances in both spectral and spatial quality indices. The NSCT\_M2 methods yields the lowest Q8 and SCC values, the highest ERGAS and SAM values, indicating the poorest performance in spectral preservation. The GSA method provides Q8 values slight lower than the highest values provided by the HR method, followed by the GS, GLP\_ESDM, GLP\_ECBD, ATWT, and HCS methods. The GLP\_ESDM performs the best among the MRA methods in terms of both spectral and spatial quality indices. The HCS method yields the poorest performance among the CS methods, in terms of all the four quality indices for the two urban images.

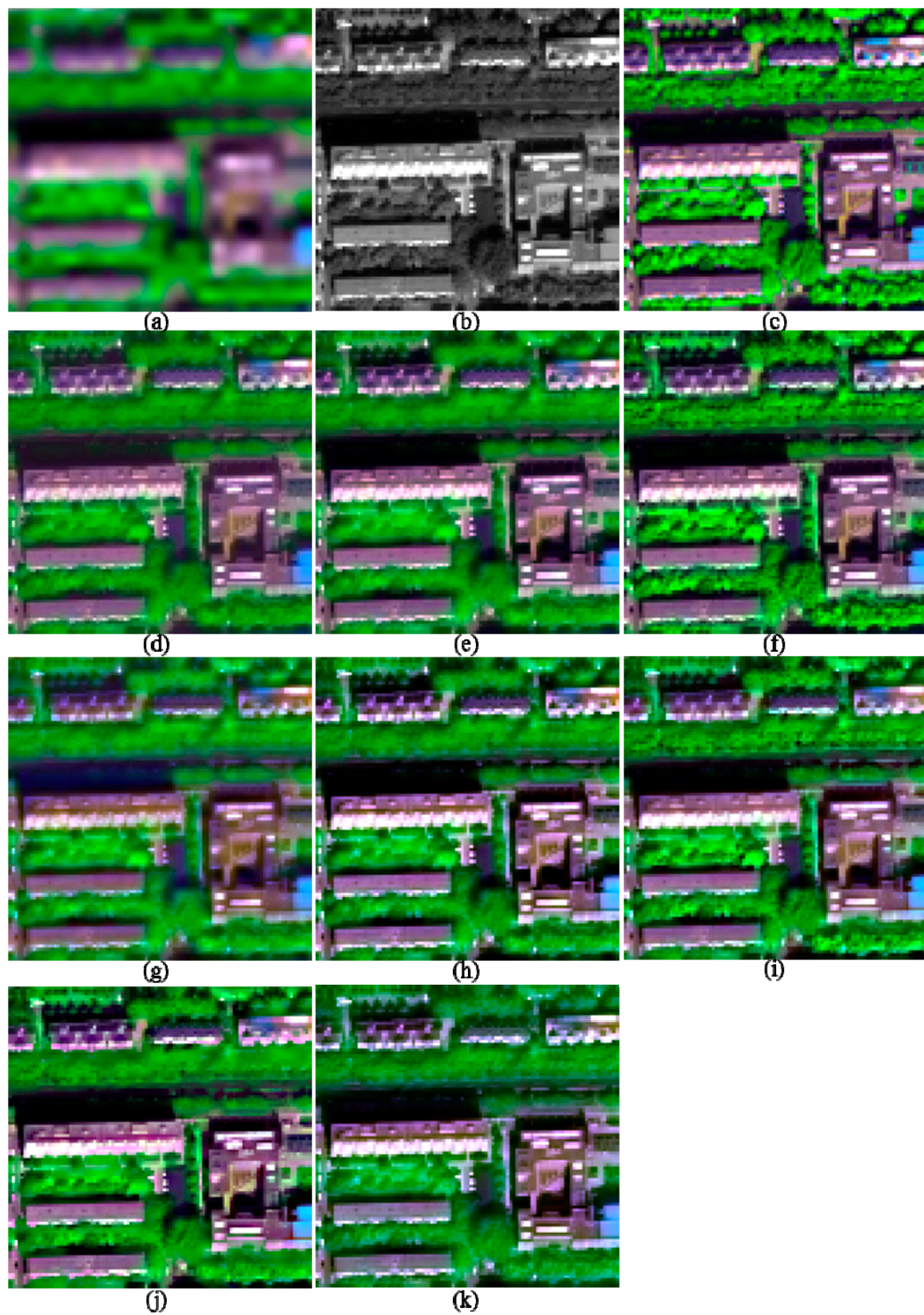
Visual comparisons for the fusion products of the two urban images (Figures 3 and 4) show that the fusion products generated by the GS, HCS, and NSCT\_M2 methods show significant spectral distortions in shadow covered regions; this is especially obvious for the fusion products of I2. In addition, the fused image generated by the GLP\_ECBD method for I2 also show significant spectral distortions in shadow covered regions, which is consistent with the poor performance of the GLP\_ECBD method in terms of quality indexes. Although the fused images generated by the ATWT and GLP\_ECBD methods show more sharpened boundaries between different objects than the other fusion products, they seem to be over sharpened and show spectral distortions. This is very obvious for the fusion products of I2. The two HCS-fused images are more blurred than other fusion products. It also can be observed that the fusion products generated by GSA, HR, and GLP\_ESDM methods yield lower spectral distortions and more sharpened boundaries between different objects than other fusion products. In addition, the HR-fused images provide more texture details in vegetation covered areas.

According to the quality indexes and visual inspection, the HR, GSA, and GLP\_ESDM methods give better performances than the other methods, whereas the NSCT\_M2 and HCS methods offer the poorest performances, for the two urban images.





**Figure 3.** The degraded and 2-m fused images of I1, shown in band 5-7-2 composition. (a) 8-m MS; (b) 2-m PAN; (c) True 2-m MS; (d) GS; (e) GSA; (f) HR; (g) HCS; (h) ATWT; (i) GLP\_ESDM; (j) GLP\_ECBD; and (k) NSCT\_M2.



**Figure 4.** The degraded and 2-m fused images of I2, shown in band 5-7-2 composition. (a) 8-m MS; (b) 2-m PAN; (c) True 2-m MS; (d) GS; (e) GSA; (f) HR; (g) HCS; (h) ATWT; (i) GLP\_ESDM; (j) GLP\_ECBD; and (k) NSCT\_M2.



### 3.3.2. Assessment for the Two Suburban Images

The image quality indices of the fusion products of the two suburban images are shown in Table 3. Similarly, parts of the fusion products of the two images are shown in Figures 5 and 6, respectively.

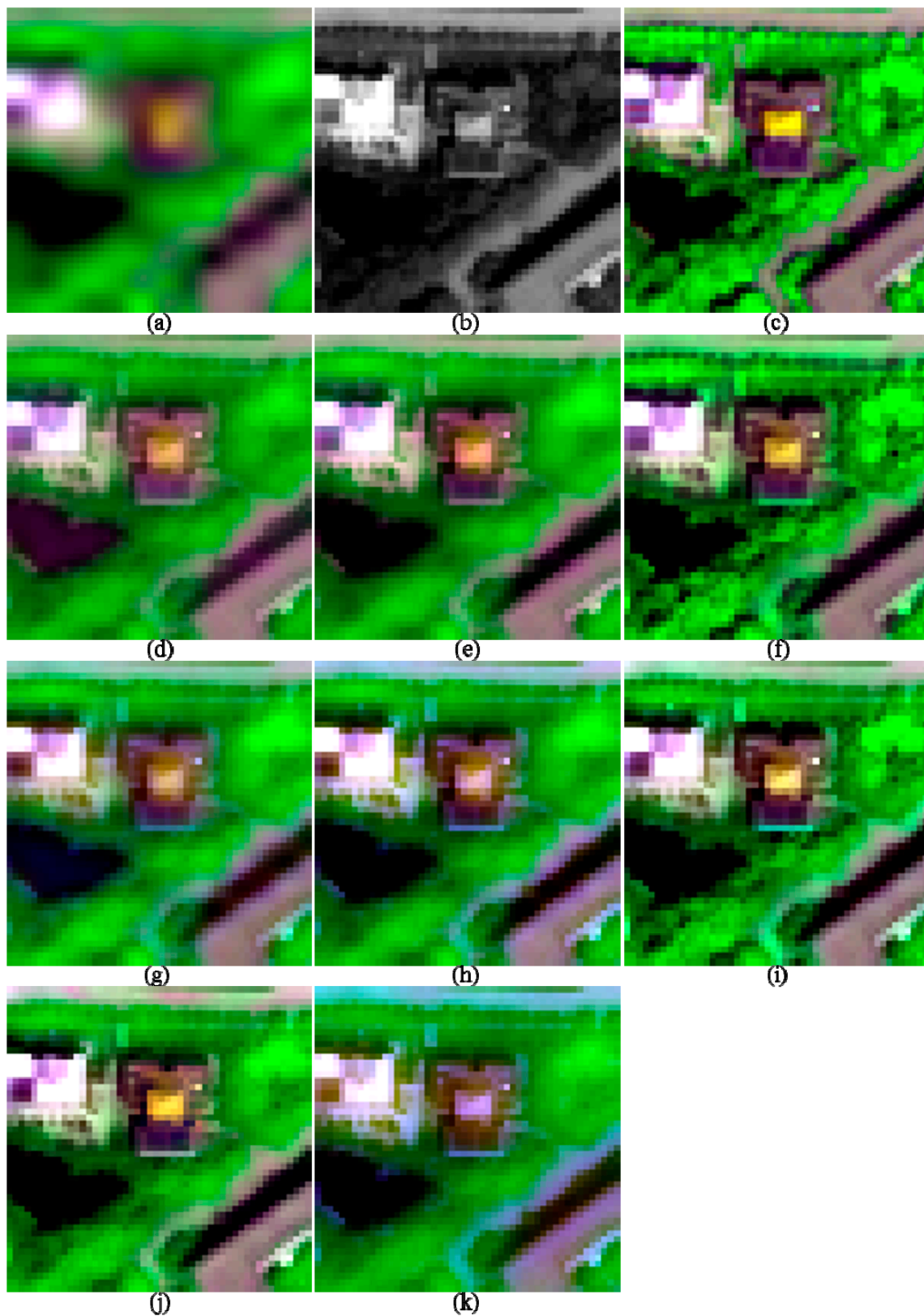
**Table 3.** Quality indices of the fused images generated by the selected algorithms for the two suburban images.

Image	Method	ERGAS	SAM	Q8	SCC
I3	GS	1.31	1.90	0.908	0.885
	GSA	1.03	1.76	0.942	0.883
	HR	1.38	2.23	0.927	0.849
	HCS	1.28	1.98	0.881	0.871
	ATWT	1.21	1.93	0.907	0.872
	GLP_ESDM	1.39	1.91	0.902	0.835
	GLP_ECBD	1.33	1.91	0.921	0.856
	NSCT_M2	1.39	2.41	0.887	0.869
	EXP	1.62	1.98	0.837	0.796
I4	GS	1.74	2.85	0.871	0.841
	GSA	1.43	2.74	0.919	0.831
	HR	1.84	3.25	0.893	0.779
	HCS	1.74	3.01	0.850	0.817
	ATWT	1.66	2.98	0.868	0.821
	GLP_ESDM	1.82	2.82	0.876	0.796
	GLP_ECBD	1.79	2.88	0.892	0.805
	NSCT_M2	1.88	3.64	0.843	0.818
	EXP	2.21	3.01	0.759	0.669

The GSA method offers best performances in terms of ERGAS, SAM, and Q8, whereas the GS method provides the highest SCC values, for the two suburban images. The NSCT\_M2 method gives the poorest performances in terms of ERGAS, SAM, and Q8. Among the MRA methods, the GLP\_ECBD method offers the highest Q8 value, whereas the ATWT method provides the highest SCC values. Although the HR method provides Q8 values slightly lower than the highest value provided by the GSA method, the former offers the lowest SCC value among the CS methods. The HR method also provides the highest SAM values among the CS methods. Although the HR method gives the best performances for the two urban images, it offers relative poor performances for the two suburban images.

Visual comparisons of the fusion products of the two suburban images show that obvious spectral distortions can be found from the water-body and shadow covered regions in the fused images generated by GS, HCS, ATWT, and NSCT\_M2 methods. In addition, the fusion products of the HCS and NSCT\_M2 methods are significant more blurred than other products, due to few spatial details are injected into the fused images. The fusion products generated by HR, GLP\_ESDM, and GLP\_ECBD methods are more sharpened and yield lower spectral distortions than other fusion products. Although the GSA method gives the best performance in terms of Q8, the corresponding fusion products seem to be more blurred than the fusion products generated by the HR, GLP\_ESDM, GLP\_ECBD methods. Although the GS method offers the highest SCC values, the two GS-fused images show significant spectral distortions. The two HR-fused images also show very blurred boundaries between vegetation and non-vegetation objects, which may contribute to the relative low SCC values and high SAM values provided by this method. According to the quality indexes and visual inspection, the GLP\_ECBD method gives better performances than other methods for the two suburban images.





**Figure 5.** The original and 2-m fused images of I3, shown in band 5-7-2 composition. (a) 8-m MS; (b) 2-m PAN; (c) True 2-m MS; (d) GS; (e) GSA; (f) HR; (g) HCS; (h) ATWT; (i) GLP\_ESDM; (j) GLP\_ECBD; and (k) NSCT\_M2.



**Figure 6.** The original and 2-m fused images of I4, shown in band 5-7-2 composition. (a) 8-m MS; (b) 2-m PAN; (c) True 2-m MS; (d) GS; (e) GSA; (f) HR; (g) HCS; (h) ATWT; (i) GLP\_ESDM; (j) GLP\_ECBD; and (k) NSCT\_M2.

### 3.3.3. Assessment for the Two Rural Images

The image quality indices of the fusion products of the two rural images are shown in Table 4, whereas the sub-images of fusion products of the two images are shown in Figures 7 and 8, respectively.

**Table 4.** Quality indices of the fused images generated by the selected algorithms for the two rural images.

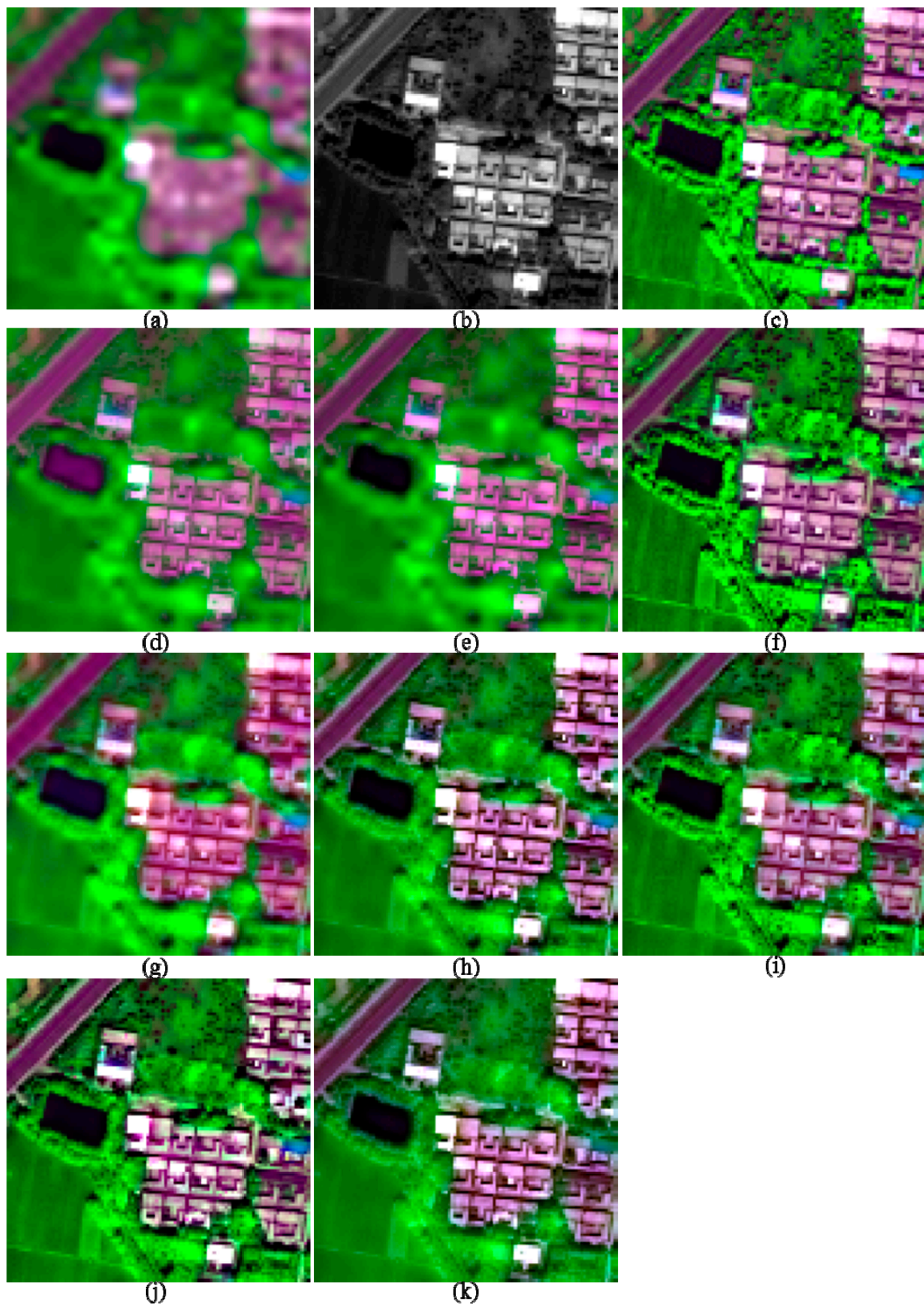
Image	Method	ERGAS	SAM	Q8	SCC
I5	GS	1.66	2.41	0.822	0.838
	GSA	1.56	2.66	0.887	0.799
	HR	1.32	2.06	0.914	0.882
	HCS	1.61	2.23	0.837	0.861
	ATWT	1.59	2.37	0.853	0.857
	GLP_ESDM	1.49	1.99	0.875	0.874
	GLP_ECBD	1.90	2.45	0.874	0.845
	NSCT-M2	1.71	2.83	0.829	0.855
	EXP	2.30	2.23	0.748	0.665
I6	GS	1.88	2.95	0.762	0.756
	GSA	1.16	1.75	0.857	0.785
	HR	1.08	1.60	0.873	0.812
	HCS	1.05	1.51	0.735	0.848
	ATWT	0.97	1.49	0.858	0.868
	GLP_ESDM	1.12	1.52	0.857	0.809
	GLP_ECBD	0.99	1.44	0.868	0.854
	NSCT-M2	1.10	1.77	0.820	0.865
	EXP	1.09	1.51	0.818	0.849

The eight fusion methods give different performances for I5 and I6, which may due to the fact that the land cover types of the two scenes are obviously different. For I5, the HR and GLP\_ESDM methods give the highest Q8 and SCC values; the GS and NSCT\_M2 methods provide the lowest Q8 values; the GSA method yields the lowest SCC values. However, for I6, the HR, GLP\_ECBD, and ATWT methods offer the highest Q8 values, whereas the ATWT, NSCT\_M2, and GLP\_ECBD methods provide the highest SCC values. The HCS and GS methods provide the lowest Q8 and SCC values, respectively, for I6. The GLP\_ESDM gives a slightly better performance than the GLP\_ECBD method for I5. In contrast, the former offers a poorer performance than the latter for I6.

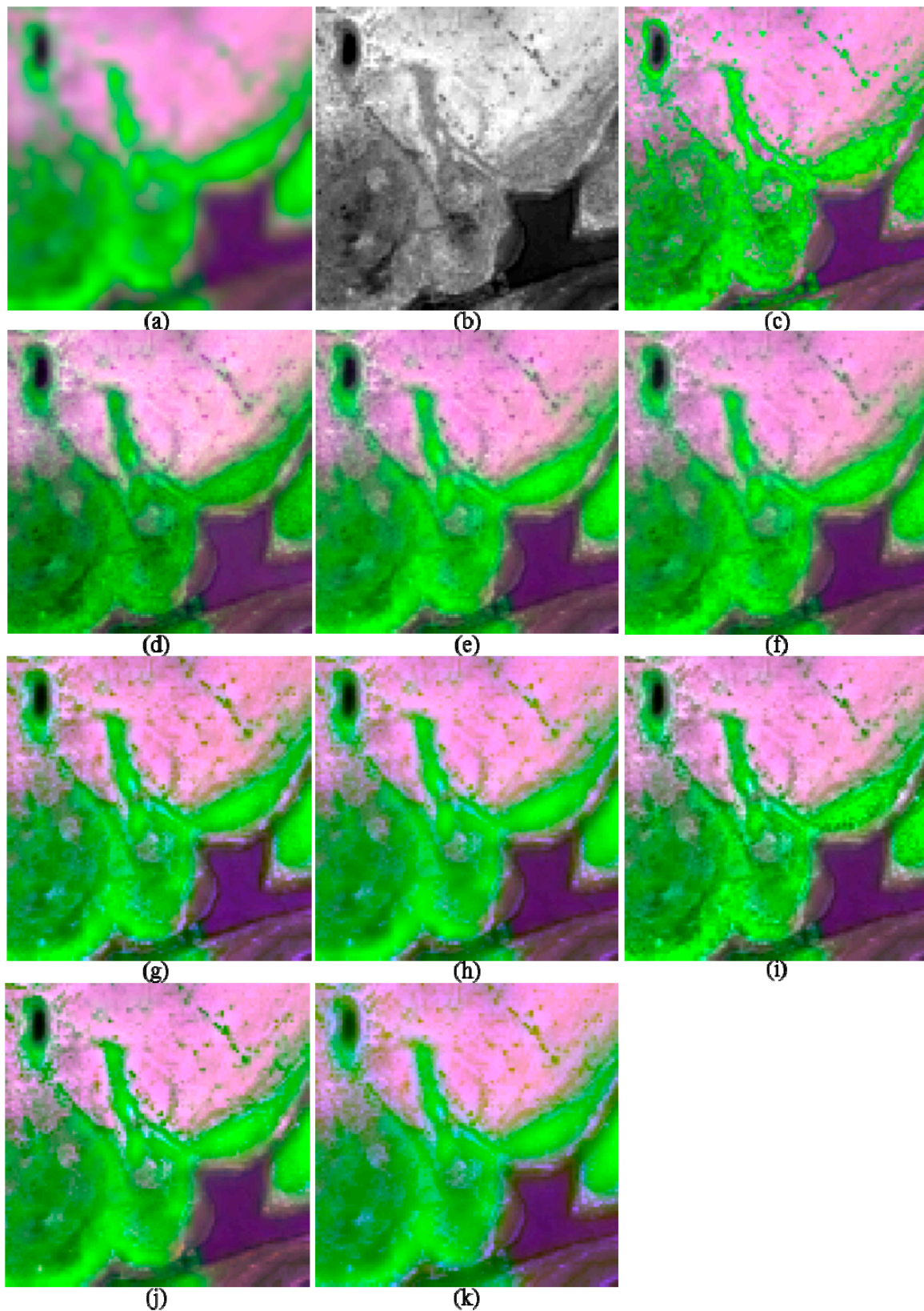
Obvious spectral distortions can be observed from the fused images generated by the GS, GSA, and HCS methods for I5, especially for the water-body covered regions and rooftop regions. The fusion products generated by the GS, GSA, and HCS methods for I5 are also more blurred than other fusion products. Conversely, the fused image generated by GLP\_ECBD for I5 seems to be too sharpened, due to over injection of the spatial details. Consequently, the HR and GLP\_ESDM methods give the best performances for I5, with respect to both the quality indexes and visual inspection.

The eight fusion products for I6 show no significant differences. Among the eight fusion products for I6, the product generated by the NSCT\_M2 method is the most blurred; whereas pixels corresponding to vegetation in the product generated by the GLP\_ESDM method seem to be too light colored. The HR-fused image shows slightly more details than other products, especially in vegetation covered regions. With respect to the quality indexes and visual inspection, the HR, GLP\_ECBD, and ATWT methods yield the best performances for I6.





**Figure 7.** The original and 2-m fused images of I5, shown in band 5-7-2 composition. (a) 8-m MS; (b) 2-m PAN; (c) True 2-m MS; (d) GS; (e) GSA; (f) HR; (g) HCS; (h) ATWT; (i) GLP\_ESDM; (j) GLP\_ECBD; and (k) NSCT\_M2.



**Figure 8.** The original and 2-m fused images of I6, shown in band 5-7-2 composition. (a) 8-m MS; (b) 2-m PAN; (c) True 2-m MS; (d) GS; (e) GSA; (f) HR; (g) HCS; (h) ATWT; (i) GLP\_ESDM; (j) GLP\_ECBD; and (k) NSCT\_M2.



### 3.4. Information Preservation

For each of the fusion products, a CC between an information index derived from the fused image and the same index derived from the corresponding reference MS image was calculated. The CC values for MBI ( $C_{\text{MBI}}$ ), NDVI ( $C_{\text{NDVI}}$ ) and NDWI ( $C_{\text{NDWI}}$ ) for fusion products of the six test datasets are listed in Table 5.

**Table 5.** The CC values for the information indices for the fused images generated by the selected eight algorithms.

Method	$C_{\text{MBI}}$				$C_{\text{NDVI}}$					$C_{\text{NDWI}}$				
	I1	I2	I3	I4	I1	I2	I3	I4	I5	I6	I3	I4	I5	I6
GS	0.973	0.972	0.969	0.934	0.922	0.877	0.929	0.891	0.912	0.946	0.915	0.858	0.876	0.950
GSA	0.975	0.979	0.981	0.959	0.915	0.855	0.924	0.879	0.884	0.966	0.908	0.853	0.830	0.977
HR	0.976	0.978	0.977	0.950	0.923	0.885	0.899	0.816	0.927	0.969	0.876	0.780	0.895	0.981
HCS	0.962	0.949	0.923	0.848	0.926	0.883	0.917	0.880	0.940	0.974	0.911	0.859	0.917	0.986
ATWT	0.965	0.956	0.934	0.882	0.911	0.849	0.928	0.890	0.928	0.975	0.912	0.851	0.915	0.986
GLP_ESDM	0.965	0.960	0.945	0.899	0.928	0.887	0.915	0.871	0.942	0.972	0.901	0.841	0.922	0.984
GLP_ECBD	0.964	0.957	0.961	0.926	0.892	0.811	0.925	0.876	0.906	0.977	0.909	0.851	0.860	0.988
NSCT_M2	0.965	0.948	0.905	0.842	0.860	0.778	0.898	0.862	0.896	0.968	0.873	0.800	0.871	0.982
EXP	0.825	0.798	0.829	0.783	0.926	0.883	0.917	0.880	0.940	0.974	0.911	0.859	0.917	0.986

#### 3.4.1. $C_{\text{MBI}}$

The fused products generated from the two urban (I1 and I2) and two suburban (I3 and I4) images were assessed in terms of  $C_{\text{MBI}}$ . All the fusion products of the two urban images offer  $C_{\text{MBI}}$  values that are significantly higher than those of the up-sampled MS images (EXPs), indicating that all the fusion products show obvious improvements in terms of  $C_{\text{MBI}}$ . This is due to the fact that spatial details extracted from the PAN bands were injected to produce these images. Generally, the performances of the eight methods in terms of  $C_{\text{MBI}}$  are consistent with those in terms of Q8 and SCC. The HR, GSA, and GS methods offer the highest  $C_{\text{MBI}}$  values, whereas the NSCT\_M2 and HCS methods offer the lowest  $C_{\text{MBI}}$  values, for both the urban and suburban test images. The CS methods give slightly better performances than the MRA methods in terms of  $C_{\text{MBI}}$ , for the four test images. This is consistent with the performances of the eight methods in terms of quality indexes, for the four test images. This is also consistent with the conclusion of previous studies, which imply that the fusion products generated by the CS methods offer good visual and geometrical impression [31].

Among the MRA methods, the GLP\_ESDM method offers the best performance for the two urban images, whereas the GLP\_ECBD method performs the best for the two suburban images, in terms of  $C_{\text{MBI}}$ . This is also consistent with the performances of the two methods in terms of quality indexes. According to the assessment results with respect to the quality indexes and visual inspection, the HR, GSA, and GLP\_ESDM methods outperform other methods for the two urban images, whereas the GLP\_ECBD method gives the best performances for the two suburban images. Actually, although the four methods give different performances vary different image scenes, they outperform the other methods for all the four images, in terms of  $C_{\text{MBI}}$ , as well as the quality indexes and visual inspection. Consequently, the HR, GSA, GLP\_ESDM, and GLP\_ECBD methods may be the best choice for producing fused urban/suburban WV-2 images used for image interpretation and buildings extraction.

#### 3.4.2. $C_{\text{NDVI}}$ and $C_{\text{NDWI}}$

The fusion products of all the six images were assessed using  $C_{\text{NDVI}}$ , whereas only the fusion products for the two suburban images and the two rural images were assessed using  $C_{\text{NDWI}}$ . We discuss the two indexes together because both of the two indexes measure the differences between inter-band relationships of the fused image and those of the reference MS image. In addition, both of them are related to the NIR bands.

It can be observed that some of the fusion products offer  $C_{\text{NDVI}}$  or  $C_{\text{NDWI}}$  values that are lower than those of provided by the corresponding up-sampled MS images (EXPs), and the highest  $C_{\text{NDVI}}$



values are just slightly higher than those of the EXPs. This indicates that the fusion products show limited improvements in terms of  $C_{NDVI}$  and  $C_{NDWI}$ , which may be caused by the spectral distortions of the fused NIR bands.

The HCS method offers relative high  $C_{NDVI}$  and  $C_{NDWI}$  values for all the test images. This may be due to the fact that the inter-band relationships of up-sampled LSR MS bands are preserved in the corresponding HCS-fused images. Since the HCS-fused images show significant spectral distortions in terms of both quality indexes and visual inspection, it is regarded as offering poor performance for specific applications. Hence, we do not discuss about the performance of this method henceforth in this section.

For the two urban images, the GLP\_ESDM, HR, and GS methods offer the highest  $C_{NDVI}$  and  $C_{NDWI}$  values. With respect to the fact the GS-fused images show obvious spectral distortions and only the GLP\_ESDM method offers higher  $C_{NDVI}$  and  $C_{NDWI}$  values than those of the corresponding EXPs, the GLP\_ESDM method is a better choice for producing fusion products that will be used in applications related to urban vegetation and water-bodies.

For I3 and I4, the GS, ATWT, and GLP\_ECBD methods provide the highest  $C_{NDVI}$  and  $C_{NDWI}$  values, whereas the HR and NSCT\_M2 methods offer the lowest  $C_{NDVI}$  and  $C_{NDWI}$  values. With respect to the fact that the GS and ATWT methods give poor performances in terms of quality indexes and visual inspection, the GLP\_ECBD method gives the best performances for the two suburban images in terms of NDVI and NDWI information preservation, as well as quality indexes and visual inspection.

For I5, the GLP\_ESDM, ATWT, and HR methods offer higher  $C_{NDVI}$  and  $C_{NDWI}$  values than other methods. With respect to the fact that the ATWT method gives poor performances in terms of quality indexes and visual inspection and only the GLP\_ESDM method offers higher  $C_{NDVI}$  and  $C_{NDWI}$  values than those of the corresponding EXPs, the GLP\_ESDM method is the better choice for rural WV-2 images with similar image objects with I5, in terms of NDVI and NDWI information preservation, as well as quality indexes and visual inspection.

For I6, the GLP\_ECBD, ATWT, and GLP\_ESDM methods provide higher  $C_{NDVI}$  and  $C_{NDWI}$  values than the other methods. With respect to the fact that the HR, GLP\_ECBD, and ATWT methods outperform the other methods in terms of quality indexes and visual inspection only the GLP\_ECBD method offers higher  $C_{NDVI}$  and  $C_{NDWI}$  values than those of the corresponding EXPs, the GLP\_ECBD method is the best choice for rural WV-2 images with similar image objects with I6, in terms of NDVI and NDWI information preservation, as well as quality indexes and visual inspection.

#### 4. Discussion

Generally, the comparisons of different pansharpening methods are performed by assessing fusion products using spectral and spatial quality indexes, as well as visual inspection. However, a good performance in terms of quality indexes and visual inspection does not always result in a good choice for different application purposes. The NDVI, NDWI, and MBI index, which are widely used in applications related land cover classification, the extraction of vegetation area, buildings, and water bodies, were employed in this study to evaluate the performances of the selected pansharpening methods in terms of the information presentation ability. In this study, the performances of eight selected state-of-art pan-sharpening methods were assessed using information indices (NDVI, NDWI and MBI), along with current image quality indices (ERGAS, SAM,  $Q2^n$  and SCC) and visual inspection, with six datasets from two WV-2 scenes.

##### 4.1. General Performances of the Selected Pansharpening Methods

Generally, the HR, GSA, GLP\_ESDM, and GLP\_ECBD methods give better performances than the other methods, whereas the NSCT and HCS methods offer the poorest performances, for most of the test images, in terms of quality indexes and visual inspection. The four methods also give slightly different performances for images including different image objects. For example, the HR, GSA, GLP\_ESDM methods give the best performances for the two urban images, whereas the GLP\_ECBD provides

the best performances for the two rural images. However, the fusion products of the four methods offer good visual quality for most images. Consequently, the HR, GSA, GLP\_ESDM, and GLP\_ECBD methods are good choices if the fused WV-2 images will be used for image interpretation.

The results of the assessments using the three information indices show that the rank of the selected eight fusion methods in terms of  $C_{MBI}$  is a little similar with those in terms of Q8 and SCC. This may indicate that the assessment using only the quality indexes and visual inspection is sufficient for selecting a best fusion method for producing fused urban WV-2 images used for image interpretation and applications related to urban buildings. The order of eight methods for in terms of  $C_{NDVI}$  is similar with that in terms of  $C_{NDWI}$ . This is due to the fact that both  $C_{NDVI}$  and  $C_{NDWI}$  measure the differences between the inter-band relationships of a fused image and those of the corresponding reference MS image. In contrast, the orders of the eight methods in terms of  $C_{NDVI}$  and  $C_{NDWI}$  are significant different from those in terms of Q8 and SCC. This indicates that a fusion method offering the best performance for a certain image in terms of quality indexes and visual inspection does not always provide the highest  $C_{NDVI}$  and  $C_{NDWI}$  values. Generally, the GLP\_ESDM method outperforms the other methods for I1, I2 and I5, whereas the GLP\_ECBD method provides the best performances for I3, I4 and I6, in terms of  $C_{NDVI}$  and  $C_{NDWI}$ , as well as quality indexes and visual inspection. This indicates that the GLP\_ESDM is the best choice for images with similar objects with I1, I2 and I5, whereas the GLP\_ECBD is the best choice for images with similar objects with I3, I4 and I6, for producing fusion products used for applications related to vegetation or water-bodies. In addition, the fusion products show limited improvements in terms of  $C_{NDVI}$  and  $C_{NDWI}$ . This indicates that it is hard for the fusion products to preserve the NDVI and NDWI information obtained from the corresponding up-sampled MS images. Consequently, it is necessary to evaluate fusion products using information indices (i.e., NDVI and NDWI) if fused WV-2 images will be used for applications related to vegetation and water-bodies.

#### 4.2. Effects for Different Spectral Ranges between the PAN and MS Bands

A noticeable point for the WV-2 is that the spectral range of the PAN band covers limited portion of the spectral ranges of the C, NIR1 and NIR2 bands. This results in relative low correlation coefficients between these bands and the PAN band. It is interesting to see the performances of the selected pansharpening methods on the two NIR bands and the C band of WV-2. In order to assess the spectral distortion of each fused band, the CC value between each fused band and the corresponding reference band was calculated for each fusion product. The CC values for the fusion products of I1, I3, I5 and I6 are shown in Table 6. The CC values of I2 and I4 are not presented because they are similar with those of I1 and I3, respectively.

It can be seen from Table 6 that the CC values of the two NIR bands are significantly lower than those of the other bands for all the fusion products, indicating that the fused NIR bands show more spectral distortions than the other bands. This is caused by the relative low correlation coefficients between the two NIR bands and the PAN band. This is also revealed by previous studies, the higher the correlation between the PAN band and each MS band, the better the success of fusion [31]. Generally, the four CS methods offer higher CC values for the two NIR bands than the four MRA methods for most of the test images. This is consistent with the result of visual inspection of these fusion products. However, the two GLP-based methods, which provide good performances in terms of NDVI and NDWI information preservation, offer relative low CC values for the fused NIR1 band, due to the low CC between the PAN and the NIR1 band. This proves again that it is necessary to evaluate fusion products using information indices (i.e., NDVI and NDWI) if fused WV-2 images will be used for applications related to vegetation and water-bodies.

**Table 6.** The CC values for the fused bands for I1, I3, I5 and I6.

Image	Method	CC								
		C	B	G	Y	R	RE	NIR1	NIR2	Avg
I1	GS	0.985	0.985	0.988	0.992	0.990	0.990	0.978	0.978	0.986
	GSA	0.987	0.987	0.991	0.995	0.993	0.992	0.978	0.977	0.988
	HR	0.990	0.990	0.993	0.996	0.994	0.993	0.977	0.977	0.989
	HCS	0.819	0.945	0.982	0.989	0.987	0.984	0.970	0.969	0.956
	ATWT	0.849	0.931	0.985	0.991	0.986	0.987	0.974	0.971	0.959
	GLP_ESDM	0.959	0.979	0.986	0.990	0.990	0.988	0.975	0.973	0.980
	GLP_ECBD	0.982	0.982	0.986	0.989	0.986	0.985	0.967	0.959	0.980
	NSCT_M2	0.805	0.906	0.985	0.993	0.987	0.988	0.971	0.965	0.950
	EXP	0.952	0.948	0.944	0.942	0.939	0.928	0.917	0.916	0.936
I3	GS	0.958	0.957	0.962	0.967	0.962	0.952	0.899	0.903	0.945
	GSA	0.977	0.977	0.986	0.990	0.984	0.977	0.906	0.906	0.963
	HR	0.976	0.975	0.983	0.986	0.980	0.969	0.830	0.823	0.940
	HCS	0.836	0.946	0.977	0.979	0.969	0.968	0.892	0.893	0.933
	ATWT	0.918	0.932	0.980	0.982	0.975	0.971	0.900	0.898	0.945
	GLP_ESDM	0.919	0.964	0.978	0.985	0.980	0.966	0.842	0.844	0.935
	GLP_ECBD	0.962	0.963	0.975	0.982	0.975	0.965	0.876	0.861	0.945
	NSCT_M2	0.886	0.906	0.980	0.978	0.969	0.968	0.872	0.863	0.928
	EXP	0.922	0.921	0.921	0.923	0.920	0.910	0.886	0.885	0.911
I5	GS	0.958	0.951	0.945	0.947	0.953	0.908	0.868	0.875	0.926
	GSA	0.969	0.964	0.963	0.962	0.967	0.981	0.949	0.957	0.964
	HR	0.968	0.966	0.974	0.968	0.967	0.984	0.954	0.965	0.968
	HCS	0.739	0.892	0.954	0.966	0.967	0.984	0.962	0.973	0.930
	ATWT	0.901	0.924	0.976	0.967	0.965	0.987	0.966	0.975	0.958
	GLP_ESDM	0.913	0.962	0.975	0.970	0.970	0.984	0.949	0.963	0.961
	GLP_ECBD	0.942	0.935	0.959	0.967	0.968	0.985	0.964	0.974	0.962
	NSCT_M2	0.858	0.888	0.970	0.960	0.957	0.981	0.957	0.968	0.943
	EXP	0.940	0.932	0.923	0.925	0.931	0.976	0.966	0.974	0.946
I6	GS	0.957	0.959	0.967	0.967	0.966	0.948	0.802	0.800	0.921
	GSA	0.970	0.974	0.985	0.982	0.980	0.963	0.750	0.747	0.919
	HR	0.979	0.980	0.987	0.987	0.985	0.969	0.858	0.854	0.950
	HCS	0.879	0.966	0.978	0.971	0.964	0.936	0.848	0.846	0.924
	ATWT	0.896	0.924	0.980	0.979	0.971	0.950	0.858	0.862	0.928
	GLP_ESDM	0.905	0.962	0.976	0.978	0.979	0.955	0.869	0.848	0.934
	GLP_ECBD	0.956	0.960	0.972	0.971	0.967	0.945	0.790	0.829	0.924
	NSCT_M2	0.873	0.905	0.982	0.978	0.968	0.953	0.801	0.829	0.911
	EXP	0.904	0.899	0.893	0.899	0.900	0.788	0.747	0.748	0.848

#### 4.3. How to Extend the Selected Pansharpening Methods to Other HSR Satellite Images

As introduced in the previous sections, the HR, GSA, GLP\_ESDM, and GLP\_ECBD methods are good choices for producing fused WV-2 images used for image interpretation and applications related to urban buildings. The two GLP-based methods outperform other methods for generating fused WV-2 images used for applications related to vegetation and water-bodies. It is interesting for the readers that whether these methods give similar performances to the sensors having a similar PAN spectral range with WV-2, such as GeoEye-1, and WorldView-3/4.

Actually, the selected pansharpening methods can be categorized into two groups, according to the approaches employed to generate the synthetic PAN band  $P_L$ , which mainly contains the low-frequency component of the original PAN band. For the first group,  $P_L$  is generated by applying filters to the original PAN band, or by up-sampling the degraded version of the original PAN band. In contrast, for the second group, the intensity image  $I_L$ , which can be seen as another approach for generating the synthetic PAN band  $P_L$ , is generated using the weighted combination of the LSR MS bands. The methods belong to the first group include HR, ATWT, NSCT and the two GLP-based

methods, whereas the methods belong to the second group include GS, GSA and HCS methods. For the first group, the low-frequency component of  $P_L$  has relative low correlations with the C, NIR1 and NIR2 bands, but has relative high correlations with the other spectral bands. This result in the fact that the details of the PAN band have relative high correlations with the B, G, Y, R and RE bands, but relative low correlations with the C, NIR1 and NIR2 bands. This may result in the fact that a large amount of spatial details are injected into the B, G, Y, R and RE bands, but only a small amount of the spatial details are injected into the C, NIR1 and NIR2 bands, especial for the case the injection gains are determined considering the relationship between each MS band and the PAN band. For the second group, the low-frequency component of the intensity image is related or partly related to the C, NIR1 and NIR2 bands. This may result in the fact that the low-frequency component of the C, NIR1 and NIR2 bands may be injected into the B, G, Y, R and RE bands, and hence may lead to spectral distortions of these bands. An exception occurs for GSA, since the intensity image  $I_L$  employed the GSA method have low CC with the C, NIR1 and NIR2 bands, due to the weights  $w_i$  obtained using Equation (5) are very low for these bands.

According to the introduction about the algorithms of the selected methods, different injection gains  $g_i$  are employed by these methods. The GS and GSA methods use a band-dependent model considering the relationship between each MS band and the PAN band. The GSA method outperform the GS method due to the intensity image  $I_L$  employed the former have low CC values with the C, NIR1 and NIR2 bands. It can be seen from Table 6 that the CC values for the B, G, Y, R, and RE bands of the GSA-fused images are significantly higher than those of the GS-fused image. The HR method uses the SDM model, which is also band-dependent. The ATWT method employs a simple additive injection model with weights for each band equal to 1, whereas the two GLP-based methods use the ESDM and ECBD models, respectively. Among these models, only the ESDM and ECBD models consider the local dissimilarity between the MS and PAN bands. According to the experimental results, the two GLP-based methods give good performances in terms of NDVI and NDWI information preservation. This may due to the fact that only the ESDM and ECBD models consider the local dissimilarity between the MS and PAN bands. It is also demonstrated by previous studies that local dissimilarity between the MS and PAN bands should be considered by pansharpening methods to reduce spectral distortions.

As a result of the above analyses about the algorithms of the selected pansharpening methods, we can obtain the following conclusions. Firstly, for the spectral bands with relative high correlations with the PAN band, the synthesized PAN band should be obtained using the original PAN band and the injection gains should considering the relationship between each MS band and the PAN band. Secondly, for the spectral bands with relative low correlations with the PAN band, further experiments should be designed to evaluate which approach is better for generating the synthesized PAN band. However, there is no doubting that local dissimilarity between the MS and PAN bands should be considered for the fusion of these bands, i.e., the NIR band, especially for the case that the fused images will be used in applications related to vegetation and water-bodies.

According to the analysis, we can conclude that the GSA, HR, GLP\_ESDM, and GLP\_ECBD methods can also provide good performances for similar sensors, such as GeoEye-1, WorldView-3, WorldView-4, for the cases that the fusion products will be used in image interpretation or urban buildings. Actually, it is proved by previous studies that the performances of these newly proposed methods are sensor independent [30]. However, for the case that the fusion products will be used in applications related to vegetation or water-bodies, the GLP-ESDM and GLP\_ECBD methods or other fusion methods consider local dissimilarity between the MS and PAN bands are better choices.

## 5. Conclusions

The performances of eight state-of-art pan-sharpening methods for WV-2 imagery were assessed using information indices (NDVI, NDWI, and MBI), along with current image quality indices (ERGAS, SAM,  $Q2^n$ , and SCC) and visual inspection, with six WV-2 datasets. The main findings and conclusions derived from our analyses are as follows:

- (1) Generally, the HR, GSA, GLP\_ESDM and GLP\_ECBD methods give better performances than the other methods, whereas the NSCT and HCS methods offer the poorest performances, for most of the test images, in terms of quality indexes and visual inspection. Some of the fusion products generated by the GS and ATWT methods show significant spectral distortions. In addition, the performances of the eight methods in terms of  $C_{MBI}$  are consistent with those in terms of Q8 and SCC. Consequently, the HR, GSA, GLP\_ESDM, and GLP\_ECBD methods are good choices if the fused WV-2 images will be used for image interpretation and applications related to urban buildings. The four methods can also provide good performances for other WV-2 image scenes, for producing fused images used for image interpretation.
- (2) The order of the pansharpening methods in terms of  $C_{NDVI}$  is consistent with that in terms of  $C_{NDWI}$ . This is because both of the two indexes measure the differences between inter-band relationships of the fused image and those of the reference MS image, and both of them are related to the quality of the fused NIR1 bands. The GLP\_ESDM method offers higher  $C_{NDVI}$  and  $C_{NDWI}$  values for I1, I2 and I5, whereas the GLP\_ECBD method provides higher  $C_{NDVI}$  and  $C_{NDWI}$  values for I3, I4 and I6, as well as good performances in terms of quality indexes and visual inspection. Consequently, the GLP\_ESDM and GLP\_ECBD methods are better than other methods, if the fused WV-2 images will be used for applications related to vegetation and water-bodies. However, for this case, it is better to select a best method by comparing the indexes  $C_{NDVI}$  and  $C_{NDWI}$ , as well as quality indexes and visual inspection, since the GLP\_ESDM and GLP\_ECBD methods may give different performances for images with different land cover objects.
- (3) According to the experimental results of this work and the analyses the algorithms of the selected pansharpening methods, we can offer two suggestions for the fusion of images obtained by sensors similar with WV-2, such as Geoeye-1 and Worldview-3/4. Firstly, for the spectral bands with relative high correlations with the PAN band, the synthesized PAN band should be obtained using the original PAN band and the injection gains should considering the relationship between each MS band and the PAN band. The HR, GSA, GLP\_ESDM, and GLP\_ECBD method also can offer good performances for scenes obtained by GeoEye-1 and Worldview-3/4, for producing fused images used for interpretation and applications related to urban buildings. Secondly, for the spectral bands with relative low correlations with the PAN band, local dissimilarity between the MS and PAN bands should be considered for the fusion of these bands, i.e., the NIR band, especially for the case that the fused images will be used in applications related to vegetation and water-bodies.

**Acknowledgments:** The authors wish to acknowledge the anonymous reviewers for providing helpful suggestions that greatly improved the manuscript. This research was supported by the One Hundred Person Project of Chinese Academy of Sciences (grant No. Y34005101A), Youth Foundation of Director of Institution of Remote Sensing and Digital Earth, Chinese Academy of Sciences (grant No. Y6SJ1100CX), the Key Research Program of Chinese Academy of Sciences (Grant No. ZDRW-ZS-2016-6-1-3), the National Science and Technology Support Program of China (grant No. 2015BAB05B05-02), and the “Light of West China” Program of Chinese Academy of Sciences (grant No. 2015-XBQN-A-07).

**Author Contributions:** Hui Li drafted the manuscript and was responsible for the research design, experiments and the analyses. Linhai Jing provided technical guide and reviewed the manuscript. Yunwei Tang processed the WV-2 data.

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

## References

1. Zhang, R.; Zeng, M.; Chen, J. Study on geological structural interpretation based on Worldview-2 remote sensing image and its implementation. *Procedia Environ. Sci.* **2011**, *10*, 653–659.
2. Jawak, S.D.; Luis, A.J. A spectral index ratio-based Antarctic land-cover mapping using hyperspatial 8-band Worldview-2 imagery. *Polar Sci.* **2013**, *7*, 18–38. [[CrossRef](#)]

3. Ghosh, A.; Joshi, P.K. A comparison of selected classification algorithms for mapping bamboo patches in lower gangetic plains using very high resolution Worldview-2 imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *26*, 298–311. [[CrossRef](#)]
4. Mutanga, O.; Adam, E.; Cho, M.A. High density biomass estimation for wetland vegetation using Worldview-2 imagery and random forest regression algorithm. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *18*, 399–406. [[CrossRef](#)]
5. Rapinel, S.; Clement, B.; Magnanon, S.; Sellin, V.; Hubert-Moy, L. Identification and mapping of natural vegetation on a coastal site using a Worldview-2 satellite image. *J. Environ. Manag.* **2014**, *144*, 236–246. [[CrossRef](#)] [[PubMed](#)]
6. Ozdemir, I.; Karnieli, A. Predicting forest structural parameters using the image texture derived from Worldview-2 multispectral imagery in a dryland forest, Israel. *Int. J. Appl. Earth Obs. Geoinf.* **2011**, *13*, 701–710. [[CrossRef](#)]
7. Carper, W.J.; Lillesand, T.M.; Kiefer, R.W. The use of intensity-hue-saturation transformations for merging SPOT panchromatic and multispectral image data. *Photogramm. Eng. Remote Sens.* **1990**, *56*, 459–467.
8. Zhou, X.; Liu, J.; Liu, S.; Cao, L.; Zhou, Q.; Huang, H. A GIHS-based spectral preservation fusion method for remote sensing images using edge restored spectral modulation. *ISPRS J. Photogramm. Remote Sens.* **2014**, *88*, 16–27. [[CrossRef](#)]
9. Welch, R.; Ehlers, M. Merging multiresolution SPOT HRV and Landsat TM data. *Photogramm. Eng. Remote Sens.* **1987**, *53*, 301–303.
10. Laben, C.A.; Brower, B.V. Process for Enhancing the Spatial Resolution of Multispectral Imagery Using Pan-Sharpening. U.S. Patent 6,011,875, 4 January 2000.
11. Aiazzi, B.; Baronti, S.; Selva, M.; Alparone, L. MS + Pan image fusion by an enhanced Gram-Schmidt spectral sharpening. In *New Developments and Challenges in Remote Sensing*; Bochenek, Z., Ed.; Millpress: Rotterdam, The Netherlands, 2007; pp. 113–120.
12. Shensa, M.J. The discrete wavelet transform—Wedding the à trous and Mallat algorithms. *IEEE Trans. Sign. Process.* **1992**, *40*, 2464–2482. [[CrossRef](#)]
13. Teggi, S.; Cecchi, R.; Serafini, F. TM and IRS-1C-PAN data fusion using multiresolution decomposition methods based on the ‘à trous’ algorithm. *Int. J. Remote Sens.* **2003**, *24*, 1287–1301. [[CrossRef](#)]
14. Mallat, S.G. Multifrequency channel decompositions of images and wavelet models. *IEEE Trans. Acoust. Speech Sign. Process.* **1989**, *37*, 2091–2110. [[CrossRef](#)]
15. Yocky, D.A. Multiresolution wavelet decomposition image merger of Landsat Thematic Mapper and SPOT panchromatic data. *Photogramm. Eng. Remote Sens.* **1996**, *62*, 1067–1074.
16. Amolins, K.; Zhang, Y.; Dare, P. Wavelet based image fusion techniques—An introduction, review and comparison. *ISPRS J. Photogramm. Remote Sens.* **2007**, *62*, 249–263. [[CrossRef](#)]
17. Aiazzi, B.; Alparone, L.; Baronti, S.; Lotti, F. Lossless image compression by quantization feedback in a content-driven enhanced Laplacian pyramid. *IEEE Trans. Image Process.* **1997**, *6*, 831–843. [[CrossRef](#)] [[PubMed](#)]
18. Yang, X.H.; Jiao, L.C. Fusion algorithm for remote sensing images based on nonsubsampling Contourlet transform. *Acta Autom. Sin.* **2008**, *34*, 274–281. [[CrossRef](#)]
19. Saeedi, J.; Faez, K. A new pan-sharpening method using multiobjective particle swarm optimization and the shiftable contourlet transform. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 365–381. [[CrossRef](#)]
20. Dong, W.; Li, X.E.; Lin, X.; Li, Z. A bidimensional empirical mode decomposition method for fusion of multispectral and panchromatic remote sensing images. *Remote Sens.* **2014**, *6*, 8446–8467. [[CrossRef](#)]
21. Dong, L.; Yang, Q.; Wu, H.; Xiao, H.; Xu, M. High quality multi-spectral and panchromatic image fusion technologies based on Curvelet transform. *Neurocomputing* **2015**, *159*, 268–274. [[CrossRef](#)]
22. Alparone, L.; Wald, L.; Chanussot, J.; Thomas, C.; Gamba, P.; Bruce, L.M. Comparison of pan-sharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 3012–3021. [[CrossRef](#)]
23. Gonzalez-Audicana, M.; Saleta, J.L.; Catalan, R.G.; Garcia, R. Fusion of multispectral and panchromatic images using improved IHS and PCA mergers based on wavelet decomposition. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 1291–1299. [[CrossRef](#)]
24. Guo, Q.; Liu, S. Performance analysis of multi-spectral and panchromatic image fusion techniques based on two wavelet discrete approaches. *Optik Int. J. Light Electron Opt.* **2011**, *122*, 811–819. [[CrossRef](#)]



25. Chen, F.; Qin, F.; Peng, G.; Chen, S. Fusion of remote sensing images using improved ICA mergers based on wavelet decomposition. *Procedia Eng.* **2012**, *29*, 2938–2943. [[CrossRef](#)]
26. Liu, Y.; Liu, S.P.; Wang, Z.F. A general framework for image fusion based on multi-scale transform and sparse representation. *Inf. Fusion* **2015**, *24*, 147–164. [[CrossRef](#)]
27. Padwick, C.; Deskevich, M.; Pacifici, F.; Smallwood, S. Worldview-2 pan-sharpening. In Proceedings of the 2010 Conference of American Society for Photogrammetry and Remote Sensing, San Diego, CA, USA, 21 May 2010.
28. El-Mezouar, M.C.; Kpalma, K.; Taleb, N.; Ronsin, J. A pan-sharpening based on the non-subsampled Contourlet transform: Application to Worldview-2 imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 1806–1815. [[CrossRef](#)]
29. Du, Q.; Younan, N.H.; King, R.; Shah, V.P. On the performance evaluation of pan-sharpening techniques. *IEEE Geosci. Remote Sens. Lett.* **2007**, *4*, 518–522. [[CrossRef](#)]
30. Vivone, G.; Alparone, L.; Chanussot, J.; Dalla Mura, M.; Garzelli, A.; Licciardi, G.A.; Restaino, R.; Wald, L. A critical comparison among pansharpening algorithms. *IEEE Trans. Geosci. Remote Sens.* **2015**, *33*, 2565–2586. [[CrossRef](#)]
31. Thomas, C.; Ranchin, T.; Wald, L.; Chanussot, J. Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 1301–1312. [[CrossRef](#)]
32. Zhang, Y.; Mishra, R.K. A review and comparison of commercially available pan-sharpening. In Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS), Munich, Germany, 22–27 July 2012.
33. Yuhendra; Alimuddin, I.; Sumantyo, J.T.S.; Kuze, H. Assessment of pan-sharpening methods applied to image fusion of remotely sensed multi-band data. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *18*, 165–175. [[CrossRef](#)]
34. Nikolakopoulos, K. Quality assessment of ten fusion techniques applied on Worldview-2. *Eur. J. Remote Sens.* **2015**, *48*, 141–167. [[CrossRef](#)]
35. Maglione, P.; Parente, C.; Vallario, A. Pan-sharpening Worldview-2: IHS, Brovey and Zhang methods in comparison. *Int. J. Eng. Technol.* **2016**, *8*, 673–679.
36. Huang, X.; Wen, D.; Xie, J.; Zhang, L. Quality assessment of panchromatic and multispectral image fusion for the ZY-3 satellite: From an information extraction perspective. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 753–757. [[CrossRef](#)]
37. Ghassemian, H. A review of remote sensing image fusion methods. *Inf. Fusion* **2016**, *32*, 75–89. [[CrossRef](#)]
38. Aiazzi, B.; Baronti, S.; Selva, M. Improving component substitution pansharpening through multivariate regression of MS + Pan data. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 3230–3239. [[CrossRef](#)]
39. Jing, L.; Cheng, Q. Two improvement schemes of PAN modulation fusion methods for spectral distortion minimization. *Int. J. Remote Sens.* **2009**, *30*, 2119–2131. [[CrossRef](#)]
40. Nunez, J.; Otazu, X.; Fors, O.; Prades, A.; Pala, V.; Arbiol, R. Multiresolution-based image fusion with additive Wavelet decomposition. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 1204–1211. [[CrossRef](#)]
41. Vivone, G.; Restaino, R.; Dalla Mura, M.; Licciardi, G.; Chanussot, J. Contrast and error-based fusion schemes for multispectral image pansharpening. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 930–934. [[CrossRef](#)]
42. Garzelli, A.; Nencini, F. Interband structure modeling for Pan-sharpening of very high-resolution multispectral images. *Inf. Fusion* **2005**, *6*, 213–224. [[CrossRef](#)]
43. Aiazzi, B.; Alparone, L.; Baronti, S.; Garzelli, A. MTF-tailored multiscale fusion of high-resolution MS and PAN imagery. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 591–596. [[CrossRef](#)]
44. Ranchin, T.; Aiazzi, B.; Alparone, L.; Baronti, S.; Wald, L. Image fusion—The ARSIS concept and some successful implementation schemes. *ISPRS J. Photogramm. Remote Sens.* **2003**, *58*, 4–18. [[CrossRef](#)]
45. Yuhas, R.; Goetz, A.; Boardman, J. Discrimination among Semi-Arid Landscape Endmembers Using the Spectral Angle Mapper (sam) Algorithm. In *Summaries of the Third Annual JPL Airborne Geoscience Workshop*; Jet Propulsion Laboratory: Pasadena, CA, USA, 1992; pp. 147–149.
46. Alparone, L.; Baronti, S.; Garzelli, A.; Nencini, F. A global quality measurement of pan-sharpened multispectral imagery. *IEEE Geosci. Remote Sens. Lett.* **2004**, *1*, 313–317. [[CrossRef](#)]
47. Garzelli, A.; Nencini, F. Hypercomplex quality assessment of multi/hyperspectral images. *IEEE Geosci. Remote Sens. Lett.* **2009**, *6*, 662–665. [[CrossRef](#)]

48. Otazu, X.; Gonzalez-Audicana, M.; Fors, O.; Nunez, J. Introduction of sensor spectral response into image fusion methods: Application to wavelet-based methods. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 2376–2385. [[CrossRef](#)]
49. Huang, X.; Zhang, L. Morphological building/shadow index for building extraction from high-resolution imagery over urban areas. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5*, 161–172. [[CrossRef](#)]
50. Chavez, P.S., Jr. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sens. Environ.* **1988**, *24*, 459–479. [[CrossRef](#)]
51. Chavez, P.S. Image-based atmospheric corrections—Revisited and improved. *Photogramm. Eng. Remote Sens.* **1996**, *62*, 1025–1036.
52. Moran, M.S.; Jackson, R.D.; Slater, P.N.; Teillet, P.M. Evaluation of simplified procedures for retrieval of land surface reflectance factors from satellite sensor output. *Remote Sens. Environ.* **1992**, *41*, 169–184. [[CrossRef](#)]
53. Zhang, Q.; Guo, B.L. Multifocus image fusion using the nonsubsampling contourlet transform. *Sign. Process.* **2009**, *89*, 1334–1346. [[CrossRef](#)]
54. Wald, L. Quality of high resolution synthesised images: Is there a simple criterion? In Proceedings of the International Conference on Fusion Earth Data, Sophia Antipolis, France, 26–28 January 2000.
55. Wang, Z.; Bovik, A.C. A universal image quality index. *IEEE Signal Process. Lett.* **2002**, *9*, 81–84. [[CrossRef](#)]
56. Liu, J.; Wang, X.; Chen, M.; Liu, S.G.; Shao, Z.F.; Zhou, X.R.; Liu, P. Illumination and contrast balancing for remote sensing images. *Remote Sens.* **2014**, *6*, 1102–1123. [[CrossRef](#)]
57. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Geosci. Remote Sens.* **2004**, *13*, 600–612. [[CrossRef](#)]
58. Alparone, L.; Alazzi, B.; Baronti, S.; Garzelli, A.; Nencini, F.; Selva, M. Multispectral and panchromatic data fusion assessment without reference. *Photogramm. Eng. Remote Sens.* **2008**, *74*, 193–200. [[CrossRef](#)]
59. Smith, W.D. Quaternions, Octonions, and Now, 16-ons and 2<sup>n</sup>-ons, New Kinds of Numbers, 2004. Available online: <http://www.scorevoting.net/WarrenSmithPages/homepage/nce2.pdf> (accessed on 30 September 2014).
60. Ebbinghaus, H.B. *Numbers*; Springer: New York, NY, USA, 1991.
61. Jing, L.; Cheng, Q.; Guo, H.; Lin, Q. Image misalignment caused by decimation in image fusion evaluation. *Int. J. Remote Sens.* **2012**, *33*, 4967–4981. [[CrossRef](#)]



© 2017 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/>).