



Article

# Improved Pansharpening with Un-Mixing of Mixed MS Sub-Pixels near Boundaries between Vegetation and Non-Vegetation Objects

Hui Li <sup>1</sup>, Linhai Jing <sup>1,\*</sup>, Liming Wang <sup>1</sup> and Qiuming Cheng <sup>2,3</sup>

<sup>1</sup> Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, No. 9 Dengzhuang South Road, Beijing 100094, China; lihui@radi.ac.cn (H.L.); wanglm@radi.ac.cn (L.W.)

<sup>2</sup> Department of Earth and Space Science and Engineering, York University, Toronto, ON M3J 1P3, Canada; qiuming@yorku.ca

<sup>3</sup> State Key Lab of Geological Processes and Mineral Resources, China University of Geosciences, Beijing 100083, China

\* Correspondence: jinglinhai@hotmail.com; Tel.: +86-10-8217-8106

Academic Editors: Parth Sarathi Roy and Prasad S. Thenkabail

Received: 13 October 2015; Accepted: 16 January 2016; Published: 22 January 2016

**Abstract:** Pansharpening is an important technique that produces high spatial resolution multispectral (MS) images by fusing low spatial resolution MS images and high spatial resolution panchromatic (PAN) images of the same area. Although numerous successful image fusion algorithms have been proposed in the last few decades to reduce the spectral distortions in fused images, few of these take into account the spectral distortions caused by mixed MS sub-pixels (MSPs). Typically, the fused versions of MSPs remain mixed, although some of the MSPs correspond to pure PAN pixels. Due to the significant spectral differences between vegetation and non-vegetation (VNV) objects, the fused versions of MSPs near VNV boundaries cause blurred VNV boundaries and significant spectral distortions in the fused images. In order to reduce the spectral distortions, an improved version of the haze- and ratio-based fusion method is proposed to realize the spectral un-mixing of MSPs near VNV boundaries. In this method, the MSPs near VNV boundaries are identified first. The identified MSPs are then defined as either pure vegetation or non-vegetation pixels according to the categories of the corresponding PAN pixels. Experiments on WorldView-2 and IKONOS images of urban areas using the proposed method yielded fused images with significantly clearer VNV boundaries and smaller spectral distortions than several other currently-used image fusion methods.

**Keywords:** pansharpening; high resolution optical images; mixed MS sub-pixels; un-mixing

## 1. Introduction

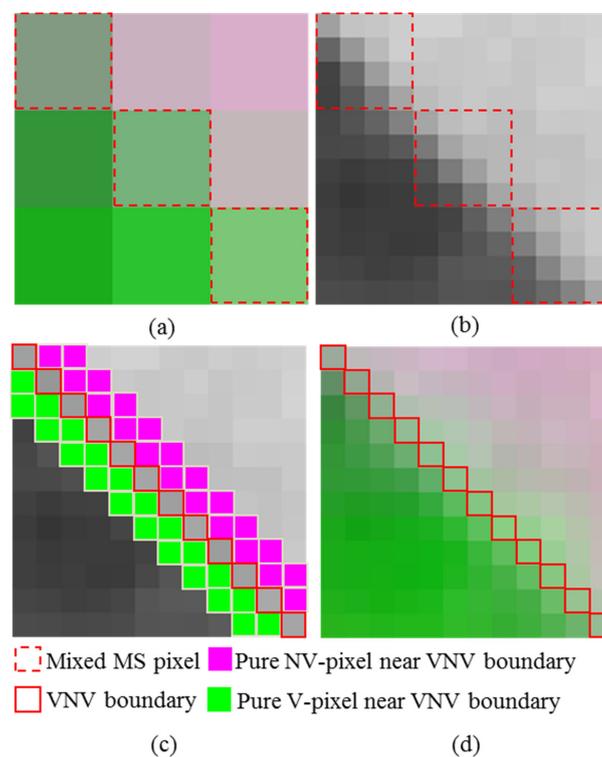
Remote sensing images are recorded in different spectral regions and with different spatial, temporal and spectral resolutions. Current satellite sensors have two physical limitations: the incoming radiation at the sensor and the data volume collected by the sensor [1–4]. In order to collect more energy and simultaneously maintain the signal-to-noise ratio, a multispectral (MS) sensor provides multiple bands with narrow spectral bandwidths and a low spatial resolution (LSR). In contrast, a panchromatic (PAN) sensor covers a wider spectral bandwidth with a higher spatial resolution than the MS sensor. Due to the limitations on data storage and transmission, many Earth observation satellites, such as SPOT, IKONOS, QuickBird and WorldView-2/-3, are equipped with both MS and PAN sensors. The fusion of a high spatial resolution PAN image and an LSR MS image simultaneously acquired over the same area, referred to as pansharpening, can produce high quality high spatial

resolution MS images. These synthesized images are commonly used commercially, e.g., by Google Earth and Bing Maps. The fused products are useful for improving image interpretation, as well as in land cover classification and change detection.

Numerous pansharpening algorithms have been proposed in recent decades. The most effective techniques are, generally speaking, the component substitution (CS) methods, the methods based on arithmetic PAN-modulation and the methods based on multi-resolution analysis (MRA) [5–8]. The CS approaches focus on the substitution of a component that is obtained by a spectral transformation of the MS bands along with the PAN image. The representative CS methods include the intensity-hue-saturation [9,10], principal component analysis [11] and Gram–Schmidt spectral sharpening (GS) [12,13] methods. The CS methods are easy to implement, and the fusion images yield a high spatial quality. However, the CS methods suffer from spectral distortions, since they do not take into account the local dissimilarities between the PAN and MS channels, which are due to the different spectral response ranges of the PAN and MS bands. The PAN modulation methods are based on the assumption that the ratio of a high spatial resolution MS band to an LSR MS band is equal to the ratio of a high spatial resolution PAN image to an assumed LSR PAN image. The assumed LSR PAN image can be obtained either from a spatially-degraded version of the high spatial resolution PAN image or from a linear combination of the original LSR MS bands. The representative PAN modulation methods include the Pradines [14], synthetic variable ratio [15,16], smoothing filter-based intensity modulation (SFIM) [17,18], PANSHARP (PS) [19,20] and haze- and ratio-based (HR) [21] methods. The PAN modulation methods reduce the spectral distortions, since these methods clamp the spectral distortions of the fused images [22]. 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 image. Multi-resolution decomposition methods, such as the “à trous” wavelet transform [23,24], the undecimated or decimated wavelet transform [25–27], Laplacian pyramids [28] and the contourlet [29,30] and curvelet [31] methods, are often employed to extract the spatial details of a PAN image. Although the MRA-based methods better preserve the spectral information of the original MS images than the CS and PAN modulation methods, they may cause spatial distortions, such as ringing or aliasing effects, originating shifts or blurred contours and textures [32]. Numerous hybrid schemes that combine MRA and other methods have been developed to maximize spatial improvement and minimize spectral distortions [33–36]. Although existing image fusion methods work well in some respects, margins still have to be improved in order to minimize spectral distortions while preserving the spatial details of the PAN image [32,37].

A major objective of pansharpening is to synthesize images that have both high spatial and spectral resolutions and are as identical as possible to real high spatial resolution MS images that could be produced by MS sensors at the PAN scale. Pansharpening algorithms can be generalized as the injection of spatial details derived from the PAN image into the up-sampled MS image to produce high spatial resolution MS images. Most of the current methods focus on optimizing the spatial details derived from the PAN image or on optimizing the weights by which the spatial details are multiplied during the injection, in order to reduce spectral distortions of fused images. Recent studies also demonstrate that an up-sampled MS image generated by current up-sampling methods is not spectrally consistent with the real high spatial resolution MS image at PAN scale for preserving spectral information [38]. However, commonly-used image fusion methods rarely consider the spectral distortions introduced by the up-sampled MS image. Due to the differences in spatial resolution between the MS and PAN images, the up-sampled MS image contains a large number of mixed sub-pixels that correspond to pure pixels in the PAN image. However, the fused versions of these mixed sub-pixels remain mixed in the fusion products generated by current fusion methods. Additionally, these fused pixels are significantly spectrally different from the corresponding real MS pixels of the same spatial resolution of the PAN image. Such mixed fused versions of these mixed sub-pixels bring in blurred boundaries between different objects and significant spectral distortions in the fusion products. As shown in Figure 1, each LSR MS pixel (Figure 1a) covers several PAN pixels (Figure 1b) and also covers several sub-pixels

in its up-sampled version at the PAN scale. Some of the sub-pixels of a mixed MS pixel cover mixed PAN pixels (*i.e.*, edge pixels across the boundaries between different objects). The PAN pixels that lie across the boundaries between vegetation and non-vegetation (VNV boundaries) shown in Figure 1c are an example of this. The other sub-pixels of the same mixed MS pixel cover pure PAN pixels near these edge pixels; examples are the pure vegetation (V) and non-vegetation (NV) pixels near VNV boundaries. Typically, the fused versions of all of the sub-pixels of a mixed MS pixel remain mixed in the fused products generated by current fusion methods, although some of them correspond to pure PAN pixels. For example, in the GS-fused image shown in Figure 1d, the fused versions of the mixed sub-pixels that correspond to pure PAN (V and NV) pixels near the VNV boundaries remain mixed. From here on in this paper, the sub-pixels of mixed MS pixels that correspond to pure PAN pixels are referred to as MSPs, whereas the edge pixels that lie across boundaries between different objects are referred to as boundary pixels. It is desirable that these MSPs are set to either pure vegetation pixels or pure non-vegetation pixels in the fused images in order to reduce spectral distortions.



**Figure 1.** Mixed multispectral (MS) sub-pixels and the corresponding PAN and fused images. (a) The mixed MS pixels at MS resolution; (b) the PAN pixels corresponding to the MS pixels in (a); (c) PAN (edge) pixels across vegetation and non-vegetation (VNV) boundaries and pure PAN (V and NV) pixels near the VNV boundaries; (d) Gram-Schmidt (GS)-fused image overlain with VNV boundaries from the PAN image.

Due to the significant spectral differences between vegetation and non-vegetation objects, the mixed fused versions of the MSPs near the edge pixels lying across VNV boundaries contribute much to the spectral distortions of the fused images. In this paper, an improved image fusion method is proposed to realize the spectral un-mixing of MSPs near VNV boundaries during the fusion process. In this study, the word “un-mixing” means to set the fused version of an MSP into a pure pixel that has the same land cover category (*i.e.*, vegetation or non-vegetation category) as the corresponding PAN pixel. This is significantly different from both the spectral endmember un-mixing technique [39] and the un-mixing-based fused method introduced by Zhukov *et al.* [40], which is also referred to as the multisensor multiresolution technique (MMT). The former uses pure spectra of different land

cover classes (*i.e.*, reference spectra of endmembers) to derive the proportion of each endmember in mixed pixels. The latter is based on the classification of high spatial resolution data, followed by the un-mixing of low spatial resolution MS pixels to retrieve signals of the LSR sensor for the classes recognized in the high spatial resolution data and the reconstruction of high spatial resolution data with the same spectral-resolution as the LSR data. The spectral endmember un-mixing technique is widely used to analyze the mixed pixels in multispectral and hyperspectral images, whereas the MMT is used for the fusion of an LSR MS image with a high spatial resolution multispectral image rather than a monochromatic image (*i.e.*, a PAN band). In this study, the proposed method is used to fuse a high spatial resolution PAN image with an LSR MS image obtained by the same satellite.

In the proposed method in this paper, MSPs near VNV boundaries are identified, and their land cover categories (*i.e.*, vegetation or non-vegetation) are determined. Then, each identified MSP is fused to a vegetation pixel or a non-vegetation pixel using the HR fusion method, depending on the corresponding land cover category. Hence, the identified MSPs near VNV boundaries are spectrally un-mixed to pure vegetation and non-vegetation pixels in the resultant fused image. This improved HR method, which includes the un-mixing of the MSPs, is given the name “UHR”.

This paper is organized as follows. The proposed method is introduced in Section 2, and the experimental results, including visual and quantitative comparisons with other fusion methods, are presented in Section 3. The discussion is presented in Sections 4 and 5 concludes the paper.

## 2. Methodologies

Given the registered MS and PAN imagery, MSPs near VNV boundaries are first identified using the coarse VNV boundaries generated from an NDVI (Normalized Difference of Vegetation Index) image derived from the up-sampled MS bands and the fine boundaries (between different objects) generated from the PAN band. A land cover category map for the identified MSPs, which is a prerequisite for the un-mixing of the MSPs, is generated by dividing the identified MSPs into the vegetation and non-vegetation categories based on the categories in the MS and PAN images. Finally, each of the identified MSPs is fused to be either a pure vegetation pixel or a pure non-vegetation pixel, according to the corresponding land cover category. The remaining MS sub-pixels (*i.e.*, the sub-pixels that are not identified as MSPs near VNV boundaries) are fused using the HR method. The details of this process are described in the following sections. The notation used in the next sections is detailed as follows. The high spatial resolution PAN image is denoted as  $P$ , the ratio of the spatial resolutions of the MS and PAN images  $R$ , the number of spectral bands in the MS image  $N$ , the up-sampled MS image  $\widetilde{MS}$ , the fused MS image  $\widehat{MS}$  and the NDVI image derived from the up-sampled MS image  $I_{NDVI}$ .

### 2.1. Identification of the MSPs near the VNV Boundaries

After the up-sampled MS image is produced by up-sampling the MS image to match the pixel size of the PAN image, MSPs near VNV boundaries are identified according to the procedure shown in Figure 2.

Step 1: VNV boundaries in the PAN image are identified using the following steps.

- (1) An edge map,  $E_{NDVI}$ , which includes coarse VNV boundaries from  $I_{NDVI}$ , is produced by applying an Laplacian of Gaussian (LOG) edge detector [41] to a binary image generated from  $I_{NDVI}$ . This binary image has been previously produced by applying a threshold ( $T_V$ ) obtained using the automatic threshold selection method proposed by Otsu [42]. In the  $E_{NDVI}$ , edge and non-edge pixels have values of 1 (true) and 0 (false), respectively. The edge pixels are then considered to be coarse VNV boundaries generated from  $I_{NDVI}$ .
- (2) Because  $\widetilde{MS}$  lacks spatial detail, the VNV boundaries in the PAN image may be displaced relative to the coarse VNV boundaries generated from  $I_{NDVI}$ . To limit the extensions in which VNV boundaries in the PAN image corresponding to the coarse VNV boundaries in  $I_{NDVI}$  are searched for, a mask map,  $E_{NDVI}^W$ , is then generated by applying to the  $E_{NDVI}$  a morphologically-dilated

operator using a disk-shaped structure element (SE) with a diameter of  $L_V$ . Next, a subset of the PAN image,  $PS$ , is generated by clipping the PAN image using the  $E_{NDVI}^W$  as a mask image. Using  $PS$  to find the fine VNV boundaries in the PAN image helps to reduce the amount of calculation. Since the displacements between the VNV boundaries in the PAN image and the corresponding coarse VNV boundaries generated from INDVI are mainly determined by the spatial resolution ratio  $R$ , the value of  $L_V$  can also be set according to  $R$ . Since a mixed MS pixel may overlap all of the sub-pixels within a window of size  $R \times R$ , the displacement between an edge pixel lying across the coarse VNV boundaries and the corresponding edge pixel lying across the fine VNV boundaries in the PAN image is no larger than  $2R - 1$ . Consequently, the value of  $L_V$  should be no higher than  $2R - 1$  in the case of fine alignment between MS and PAN bands.

- (3) An edge map,  $E_P$ , is calculated by applying an LOG edge detector with a standard deviation of  $\delta$  to  $P_S$ . Similar to  $E_{NDVI}$ , the edge and non-edge pixels in  $E_P$  have values of 1 (true) and 0 (false), respectively. The value of  $\delta$  also determines the window size  $S_G$  of the LOG detector, *i.e.*,  $S_G = (\delta \times 3) \times 2 + 1$ . The LOG detector uses a Gaussian filter specified by  $\delta$  and  $S_G$  to smooth the input image before the edge detection, in order to reduce the effects of noises. A large  $S_G$  results in a small number of edge pixels being detected by the LOG algorithm and may lead to incorrect positioning of edge pixels. Hence, the value of  $\delta$  is set to 0.3, which results in an  $S_G$  of 3, in the UHR method.
- (4) To remove edge pixels that do not lie across VNV boundaries in  $E_P$ , an edge map  $E_P^V$  is generated by setting the value of each pixel  $t$ .  $E_P^V(t)$  is determined with respect to the values of both  $E_P(t)$  and  $I_{NDVI}(t)$ , along with the NDVI value of a neighboring pixel  $t'$ , which has the largest difference from pixel  $t$  in the PAN image. This is done according to Equation (1):

$$E_P^V(t) \begin{cases} 0, & \text{if } E_P(t) = 0 \\ 0, & \text{if } \max(I_{NDVI}(t), I_{NDVI}(t')) < T_V \text{ and } E_P(t) = 1, \\ 0, & \text{if } \min(I_{NDVI}(t), I_{NDVI}(t')) > T_V \text{ and } E_P(t) = 1 \\ 1, & \text{otherwise} \end{cases} \quad (1)$$

where  $I_{NDVI}(t)$  and  $I_{NDVI}(t')$  are the NDVI values of pixels  $t$  and  $t'$ , respectively, and  $T_V$  is a threshold obtained using the OTSU method, as introduced in Subsection 2.1. The purposes of the second and third lines of Formula (1) are to eliminate the boundary pixels that do not lie across VNV boundaries in the PAN image. The two lines are based on a fact that once either pixel  $t$  or  $t'$  lies across VNV boundaries in the PAN image, one of them will have a relative high DNVI value, and the other should have a relative low NDVI value. Once both pixels have high or low DNVI values, they lie in a pure vegetation area or a non-vegetation area and should be eliminated from  $E_P$  to yield an edge map  $E_P^V$ . In  $E_P^V$ , only the edge pixels at the VNV boundaries in the PAN image have values of 1 (true).

Step 2: After the identification of the VNV boundaries in the PAN image, the pixels on both sides of the identified VNV boundaries are identified as MSPs near VNV boundaries, as follows.

- (1) The pixels on both sides of the identified VNV boundaries are identified as pure PAN pixels near VNV boundaries by means of a morphological dilation operation. A morphological dilation operation using a disk-shaped SE with a diameter of  $L_P$  is applied to  $E_P^V$  to generate a map,  $E_P^W$ , in which both the VNV boundaries and the pure PAN pixels near the VNV boundaries have values of 1 (true). The value of  $L_P$  determines the maximum distance (in pixels) between the pure PAN pixels near VNV boundaries and the nearest VNV boundaries. For a spatial resolution of  $R$ , a pure PAN pixel corresponding to the same mixed MS pixel as a VNV BP pixel,  $p$ , will lie within a window with a width of  $2R - 1$ , centered at  $p$ . Consequently, the value of  $L_P$  can be set to  $2R - 1$ . In practice, the value of  $L_P$  may be also affected by the interpolation method employed to yield the up-sampled MS image.

- (2) Mark each sub-pixel  $t$  in the up-sampled MS image as an MSP near VNV boundaries only if  $E_P^W(t) = 1$ . A pixel classification map  $M_{PI}$  is produced by assigning the MSPs near VNV boundaries to category  $C_M$  and all others to category  $C_O$ , according to Equation (2):

$$M_{PI}(t) = \begin{cases} C_M, & E_P^W(t) = 1 \\ C_O, & E_P^W(t) = 0 \end{cases} \quad (2)$$

The MSPs belonging to the category  $C_M$  in  $M_{PI}$  will be un-mixed during the fusion process. Two of the three involved parameters, the diameters of the two disk-shaped SEs ( $L_V$  and  $L_P$ ), are required in this step.

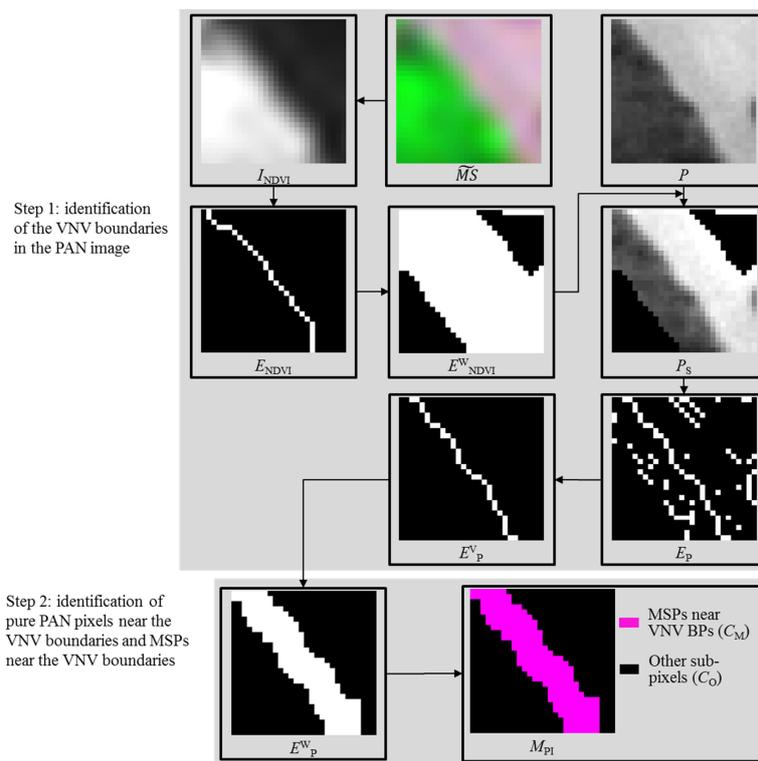


Figure 2. Scheme for the identification of MS sub-pixels (MSPs) near VNV boundaries.

### 2.2. Determining the Categories of the Identified MSPs

In order to set the fused version of the MSPs near VNV boundaries (*i.e.*, the sub-pixels with category  $C_M$ ) as pure vegetation or non-vegetation pixels, the land cover categories (*i.e.*, vegetation or non-vegetation) of the MSPs should be determined first and act as a prerequisite for the un-mixing during the fusion process.

The NDVI can be employed to divide a pixel into the vegetation and non-vegetation categories since a vegetation pixel usually has a higher NDVI than a non-vegetation one. However, MSPs near the VNV boundaries may also have high NDVI values due to the high NDVI values of the corresponding mixed MS pixels, which consist of combinations of vegetation and non-vegetation objects. Consequently, it is difficult to divide the identified MSPs into the vegetation and non-vegetation categories by applying a single threshold to the whole NDVI image. Hence, two maps,  $T_V$  and  $T_{NV}$ , which record the local average NDVI values of vegetation edge pixels and non-vegetation edge pixels, respectively, are produced, and the category of each identified MSP is then determined with respect to the corresponding values recorded in the two maps.

Moreover, ideally, vegetation pixels will lie on one side of the VNV boundaries and the non-vegetation pixels on the other side. Hence, a vegetation pixel map,  $M_V$ , and a non-vegetation pixel map,  $M_{NV}$ , are generated by applying morphological dilation operations to the vegetation edge pixels map,  $E_V$ , and the non-vegetation edge pixels map,  $E_{NV}$ . Finally,  $M_V$ ,  $M_{NV}$ ,  $T_V$  and  $T_{NV}$ , along with the NDVI, are used to determine the land cover category (*i.e.*, vegetation or non-vegetation) of each identified MSP. The flow diagram for the classification of the identified MSPs is shown in Figure 3, and the details of each step of this procedure are listed below.

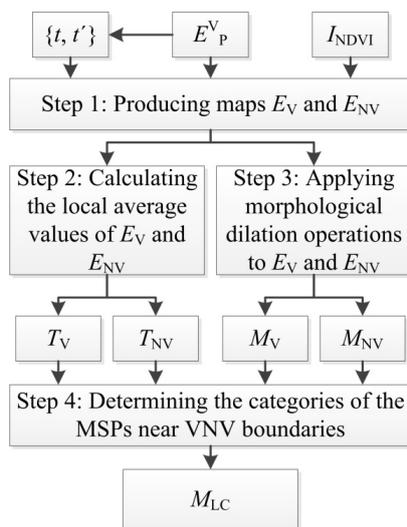


Figure 3. The flow diagram of the classification of the MSPs near VNV boundaries.

Step 1: A vegetation edge pixels map  $E_V$  and a non-vegetation edge pixels map  $E_{NV}$  are generated with respect to  $E_P^V$  in this step. All of the pixels that are true in  $E_V$  are the vegetation pixels that lie across VNV boundaries, whereas all of the pixels that are true in  $E_{NV}$  are the non-vegetation pixels that lie across VNV-boundaries. The details of this step are introduced as follows:

- (1) For each pixel,  $t$ , that is true in  $E_P^V$  (*i.e.*,  $E_P^V(t) = 1$ ), a neighboring pixel,  $t'$ , with the largest spectral difference from the pixel  $t$ , is searched for in the eight-pixel neighborhood of  $t$ . This results in a list of pixel pairs,  $\{t, t'\}$ ;
- (2) Based on the pixel pairs  $\{t, t'\}$  and the NDVI, the vegetation edge pixels map,  $E_V$ , and the non-vegetation edge pixels map,  $E_{NV}$ , are produced according to Equation (3):

$$\begin{aligned} E_V(t^V) &= 1, E_V(t^{NV}) = 0 \\ E_{NV}(t^{NV}) &= 1, E_{NV}(t^V) = 0 \end{aligned} \tag{3}$$

where  $t^V$  and  $t^{NV}$  are determined using Equation (4).

$$\begin{aligned} t^V &= t, t^{NV} = t', \text{ if } I_{NDVI}(t) > I_{NDVI}(t') \\ t^V &= t', t^{NV} = t, \text{ if } I_{NDVI}(t) < I_{NDVI}(t') \end{aligned} \tag{4}$$

The maps  $E_V$  and  $E_{NV}$  are employed to generate the maps  $T_V$  and  $T_{NV}$  in the next step.

Step 2: The maps  $T_V$  and  $T_{NV}$ , which record the local average NDVI values of vegetation edge pixels and non-vegetation edge pixels, respectively, are produced in this step. The maps are generated by assigning the value of each pixel,  $p$ , in the two maps, according to Equations (5) and (6), respectively:

$$T_V(p) = \frac{\sum_{j \in N_{S_p}(p)} (I_{NDVI}(j) \times E_V(j))}{\sum_{j \in N_{S_p}(p)} E_V(j)} \tag{5}$$

$$T_{NV}(p) = \frac{\sum_{j \in N_{S_p}(p)} (I_{NDVI}(j) \times E_{NV}(j))}{\sum_{j \in N_{S_p}(p)} E_{NV}(j)} \quad (6)$$

where  $N_{S_p}(p)$  is a neighboring window with a size of  $S_p \times S_p$  pixels of pixel  $p$ . Since a high value of  $S_p$  will increase the amount of calculation,  $S_p$  can be set equal to  $2R - 1$ .

Step 3: The maps  $M_V$  and  $M_{NV}$  are generated by iteratively applying morphological dilation operations to  $E_V$  and  $E_{NV}$  as follows.

- (1) The vegetation map,  $M_V$ , is produced by applying a morphological dilation operator using a disk-shaped SE with a diameter of 3 pixels to  $E_V$ . The pixels that are true in  $M_V$  and also true in  $E_{NV}$  are then set to be false in  $M_V$ .
- (2) The resultant vegetation map,  $M_V$ , in which the pixels that are true represent vegetation, is obtained by iterating Step 1  $R - 1$  times.
- (3) The non-vegetation map,  $M_{NV}$ , is generated by applying the morphological dilation operator described in Step 1 to  $E_{NV}$ . The pixels that are true in  $M_{NV}$  and also in  $E_V$  are then set to be false in  $M_{NV}$ .
- (4) The resultant non-vegetation map,  $M_{NV}$ , in which the pixels that are true represent non-vegetation, is obtained by iterating Step 3  $R - 1$  times.

Step 4: The category (*i.e.*, vegetation or non-vegetation) of each of the identified MSPs (denoted as pixel  $m$ ) is determined, according to Equation (7):

$$M_{LC}(m) = \begin{cases} C_V, & \text{if } M_V(m) = 1 \text{ and } (M_{NV}(m) = 0 \text{ or } I_{NDVI}(m) > T_V(m)) \\ C_{NV}, & \text{if } M_{NV}(m) = 1 \text{ and } (M_V(m) = 0 \text{ or } I_{NDVI}(m) < T_{NV}(m)) \end{cases} \quad (7)$$

An MSP  $m$  that is true in  $M_V(m)$  is determined to be a vegetation pixel if  $M_{NV}(m) = 0$  or  $I_{NDVI}(m) > T_V(m)$ , whereas an MSP,  $m$ , that is true in  $M_{NV}(m)$  is determined to be a non-vegetation pixel if  $M_V(m) = 0$  or  $I_{NDVI}(m) < T_{NV}(m)$ . Eventually, all of the identified MSPs will be divided into the vegetation and non-vegetation categories, and a classification map,  $M_{LC}$ , is produced to assist the fusion of the identified MSPs.

The size of the neighboring window,  $S_p$ , is the only parameter that needs to be decided upon in the classification.

### 2.3. Fusion of the MS Sub-Pixels

The HR fusion method is employed to yield the fused versions of the MS sub-pixels in the proposed method, for its better performances than other methods [21], such as the SVR, GS and PS methods. In order to reduce spectral distortions, the fused version of each identified MSP (*i.e.*, the sub-pixels with category  $C_M$  in  $M_{PI}$ ) is calculated with reference to the spectra of a neighboring pure (vegetation or non-vegetation) pixel. The fused versions of the sub-pixels with category  $C_O$  in  $M_{PI}$  are calculated with reference to their own spectra.

The HR method is a PAN modulation fusion method that takes haze into account [43–45]. In this method, the fused  $i$ -th spectral band of pixel  $t$ ,  $\widehat{MS}_i(t)$ , is calculated by using Equation (8):

$$\widehat{MS}_i(t) = \left( \widetilde{MS}_i(t) - H_i \right) \frac{P(t) - H_p}{P_L(t) - H_p} + H_i \quad (8)$$

where  $P_L$  is a low-pass filtered version of  $P$ ;  $\widetilde{MS}_i$  and  $\widehat{MS}_i$  are the  $i$ -th bands of the up-sampled and fused MS images, respectively; and  $H_i$  and  $H_p$  denote the haze values in the  $i$ -th MS band and PAN band, respectively. The  $H_i$  and  $H_p$  can be determined using the minimum grey level values in the  $i$ -th MS band and  $P_L$  according to an image-based dark-object subtraction method [43–45]. The HR method uses the spectral distortion minimizing (SDM) model [46] as the injection model and clumps the spectral angles between the fused and up-sampled MS pixels. It can be proven that the spectra

vector of pixel  $t$  in the fused image ( $\hat{V}_t = \{\widehat{MS}_1(t), \widehat{MS}_2(t), \dots, \widehat{MS}_N(t)\}$ ) is parallel to that of  $t$  in the up-sampled image ( $\tilde{V}_t = \{\widetilde{MS}_1(t), \widetilde{MS}_2(t), \dots, \widetilde{MS}_N(t)\}$ ) [46].

As the flow diagram shown in Figure 4, each MS sub-pixel is fused according to the appropriate category (*i.e.*, either the identified MSPs or the remaining sub-pixels) in  $M_{PI}$ .

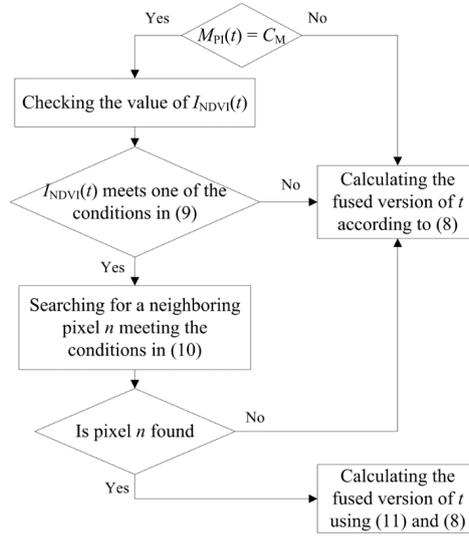


Figure 4. The flow diagram for the fusion of the MS sub-pixels.

- (1) For a sub-pixel,  $t$ , if  $M_{PI}(t) = C_O$ , its fused version is calculated according to Equation (8). Otherwise, its land cover category,  $M_{LC}(t)$ , is examined with respect to the value of  $I_{NDVI}(t)$ , according to the two conditions in Equation (9).

$$\begin{cases} I_{NDVI}(t) \geq \frac{\sum_{j \in N_{S_N}(t)} (I_{NDVI}(j) \times E_V(j))}{\sum_{j \in N_{S_N}(t)} E_V(j)}, & \text{if } M_{LC}(t) = C_V \\ I_{NDVI}(t) \leq \frac{\sum_{j \in N_{S_N}(t)} (I_{NDVI}(j) \times E_{NV}(j))}{\sum_{j \in N_{S_N}(t)} E_{NV}(j)}, & \text{if } M_{LC}(t) = C_{NV} \end{cases}, \quad (9)$$

where  $N_{S_N}(t)$  is a window with a size of  $S_N \times S_N$  pixels centered at pixel  $t$ , and  $E_V$  and  $E_{NV}$  are the vegetation edge pixels map and the non-vegetation edge pixels map, respectively, as introduced in Section 2.2. The first condition is based on the fact that if  $M_{LC}(t) = C_V$ ,  $I_{NDVI}(t)$  should be higher than the average NDVI value of the surrounding vegetation edge pixels in  $E_V$ . The second condition is based on the fact that if  $M_{LC}(t) = C_{NV}$ ,  $I_{NDVI}(t)$  should be lower than the average NDVI value of the surrounding non-vegetation edge pixels in  $E_{NV}$ .

- (2) If  $I_{NDVI}(t)$  meets no conditions in Equation (9), pixel  $t$  may be misclassified in  $M_{LC}$ , and its fused version is calculated using Equation (8), in order to avoid spectral distortions caused by the incorrect substitutions employed. Otherwise, a neighboring pixel  $n$  is searched for among the pixels with the same land cover class as pixel  $t$  (*i.e.*,  $M_{LC}(t) = M_{LC}(n)$ ) in a neighboring window  $N_{S_N}(t)$  according to the conditions in Equation (10):

$$\begin{cases} I_{NDVI}(n) = \max_{j \in N_{S_N}(t) \& M_{LC}(j) = C_V} I_{NDVI}(j) \text{ and } I_{NDVI}(n) > I_{NDVI}(t), & \text{if } M_{LC}(t) = C_V \\ I_{NDVI}(n) = \min_{j \in N_{S_N}(t) \& M_{LC}(j) = C_{NV}} I_{NDVI}(j) \text{ and } I_{NDVI}(n) < I_{NDVI}(t), & \text{if } M_{LC}(t) = C_{NV} \end{cases}. \quad (10)$$

The purpose of this step is to find a substitution pixel  $n$  for  $t$ , depending on  $M_{LC}(t)$ . If  $t$  is a vegetation pixel in  $M_{LC}$  (*i.e.*,  $M_{LC}(t) = C_V$ ), the sub-pixel  $n$  should have

the maximum NDVI value in window  $N_{S_N}(t)$  and belong to the same category as  $t$  in  $M_{LC}$  (i.e.,  $M_{LC}(n) = M_{LC}(t) = C_V$ ). Otherwise, if  $t$  is a non-vegetation pixel in  $M_{LC}$  (i.e.,  $M_{LC}(t) = C_{NV}$ ), the sub-pixel  $n$  should have the minimum NDVI value in the window  $N_{S_N}(t)$  and belongs to the same category as  $t$  (i.e.,  $M_{LC}(n) = M_{LC}(t) = C_{NV}$ ). The spectra of pixel  $n$  will be used to set the fused version of pixel  $t$  as a pure vegetation or non-vegetation pixel. In order to avoid spectral distortions introduced by incorrect substitutions used in the fusion, the substitute pixel  $n$  is restricted to have an NDVI value higher than that of  $t$  if  $M_{LC}(t) = C_V$ , whereas it is restricted to have an NDVI value lower than that of  $t$  if  $M_{LC}(t) = C_{NV}$ .

- (3) If no such pixel  $n$  meets the conditions in Equation (10), the fused version of  $t$  is calculated using Equation (8). Otherwise,  $\widetilde{MS}_i(n)$  and  $P_L(n)$ , the spectral values of the pixel  $n$  in  $\widetilde{MS}_i$  and  $P_L$ , are assigned to  $\widetilde{MS}_i(t)$  and  $P_L(t)$ , respectively, as Equation (10), and then, the fused version of  $t$  is calculated using Equations (8) and (11).

$$\begin{cases} \widetilde{MS}_i(t) = \widetilde{MS}_i(n) \\ P_L(t) = P_L(n) \end{cases} \quad (11)$$

In this case, the pixel vector of pixel  $t$  in the fused image is parallel to that of the up-sampled pixel  $n$ , according to the SDM injection model employed in the HR method. It is assumed that pixel  $n$  is a pure vegetation or non-vegetation pixel, and therefore, MSP  $t$  is fused to be a pure vegetation or non-vegetation pixel in the resultant fused image, realizing spectral un-mixing in the fusion process and reducing the spectral distortion in the fused product.

In this step, parameter  $S_N$  stands for the size of the neighborhood window within which to search for a substitution pixel used to un-mix an MSP in the fusion process. The parameter is mainly determined by the spatial resolution ratio  $R$ , along with the interpolation method employed to generate the up-sampled MS. Specifically, a pure PAN pixel corresponding to the same mixed MS pixel as  $t$  lies within a window with a width of  $2R-1$  centered at  $t$ . Consequently, the value of  $S_N$  can be set to  $2R-1$ , or slightly lower than  $2R-1$ , in order to avoid over-un-mixing.

### 3. Experiments

#### 3.1. Test Data, Fusion Methods for Comparison and Evaluation Criteria

Two datasets acquired by the WorldView-2 (WV-2) and the IKONOS sensors were used to assess the performance of the proposed method. The WV-2 sensor provides data in eight MS bands and a PAN band with 2-m and 0.5-m spatial resolutions, respectively. The MS bands 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 WV-2 PAN band, for which the spectral response range is 450–800 nm, covers a narrower NIR spectral range than the IKONOS PAN band. The WV-2 dataset was acquired in September 2014 over the city of Beijing, China, and had an average off-nadir angle of  $13.7^\circ$ . The IKONOS sensor provides data in four MS bands with 4-m spatial resolution. These comprise the blue (B, 0.45–0.53  $\mu\text{m}$ ), green (G, 0.52–0.61  $\mu\text{m}$ ), red (R, 0.64–0.72  $\mu\text{m}$ ) and near-infrared (NIR, 0.77–0.88  $\mu\text{m}$ ) bands, and there is also a corresponding 1-m PAN band (0.45–0.90  $\mu\text{m}$ ). The IKONOS dataset was acquired over Beijing in May 2000. These two sensors were chosen for this work, since the two PAN bands have different spectral response ranges. Both of the original MS images from the two sensors had a size of  $512 \times 512$  pixels, and both of the original PAN images had a size of  $2048 \times 2048$  pixels. Typical land cover types in the two scenes included water bodies, grasses, trees, buildings, roads and shadows.

To verify the performance of the proposed fusion scheme, the fused images obtained using the proposed method were compared to methods from the literature that were both well known and

known to perform well. These included the HR [21], PS [19,20], GS [12] with Mode 1 (GS1), GS with Mode 2 (GS2) and GS adaptive (GSA) [13] methods.

Two quantitative assessment procedures were adopted to evaluate the performance of the proposed fusion method, along with a visual comparison. The first procedure compared the fused imagery generated from the degraded MS and PAN images of the two sensors with the corresponding original MS images according to Wald's protocol [47]. Three comprehensive indices, the Erreur Relative Globale Adimensionnell de Synthèse (ERGAS) [48], the Spectral Angle Mapper (SAM) [49,50] and  $Q2^n$  [51,52], which is a generalization of the Universal Image Quality Index (UIQI) for monoband images and derived from the theory of hypercomplex numbers, particularly of  $2^n$ -ones, were employed to do this. These indices estimate the global spectral quality of the fused images. The lower the ERGAS and the SAM values, the better the quality of the fused image. The higher the  $Q2^n$  index, the better the quality of the fused product. Given the number of bands in the two datasets, the Q4 and Q8 indices were used for the IKONOS and WV-2 datasets, respectively. The second procedure compared the fused versions of the original MS and PAN images of the two sensors using the Quality index with no reference images (QNR) [53]. The QNR is formed from the product of two variables,  $D_\lambda$  and  $D_S$ , which quantify the spectral and spatial distortions, respectively.  $D_\lambda$  is derived from the difference between the inter-band UIQI [54] values calculated from the fused MS bands and those calculated from the LSR MS bands.  $D_S$  is generated from the difference between the UIQI values calculated using the PAN band and each of the fused MS bands and those calculated using the degraded PAN band and each of the LSR MS bands. The higher the QNR index, the better the quality of the fused product. The maximum theoretical value of the QNR is 1, which occurs when both  $D_\lambda$  and  $D_S$  are equal to 0.

### 3.2. Generation of the Fusion Products for Comparisons

The degraded versions of the WV-2 and IKONOS images were obtained by applying a low-pass filter that matched the modulation transfer function (MTF) of the corresponding sensor and a decimation characterized by a sampling factor equal to the resolution ratio,  $R$ , to each band. The fusion experiments were then carried out on both the degraded and original versions of the two datasets.

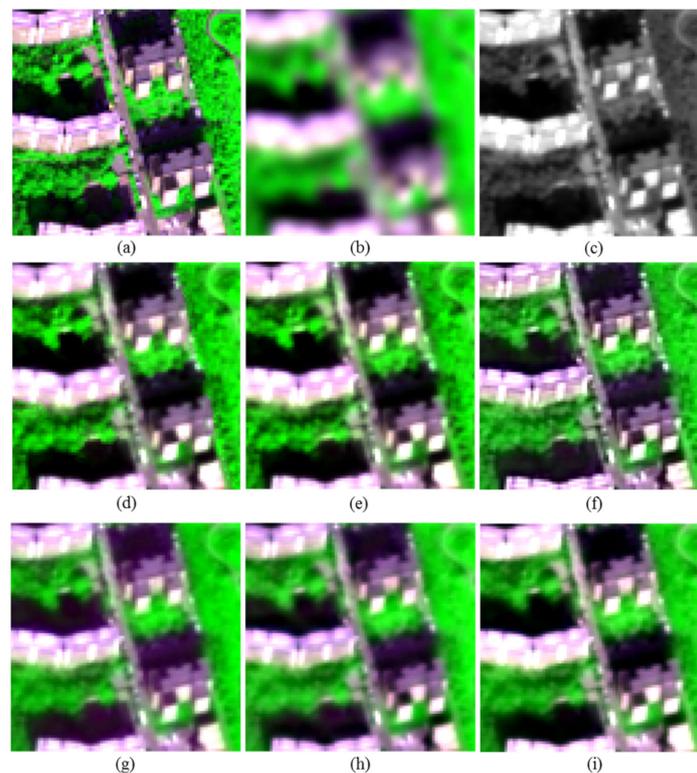
The PS-fused images were generated in PCI Geomatics software, whereas the other fusion products were produced by using the corresponding algorithms developed in MATLAB R2011b (Version 7.13.0.564). For all of the fusion methods, except PS, the up-sampled MS images were generated by using a polynomial kernel with 23 coefficients [55]. For GS2 (GS with Mode 2), GSA, HR and the proposed UHR methods, the low-pass version PAN images ( $P_L$ ) were produced by employing low-pass filters that matched the MTF of the corresponding sensors [46,56]. The HR-fused images were produced by using the same haze values as for the UHR method. For the UHR method, the NDVI images of the original and degraded WV-2 datasets were obtained by using the bands R and NIR1, whereas those of the original and degraded IKONOS datasets were generated by using the bands R and NIR.

The values of the parameters required by the proposed method were determined according to the explanation given in Section 2. Parameter  $\delta$  was set to 0.3; the values of  $S_P$ ,  $L_V$  and  $L_P$  were set to  $7(2R - 1)$ ,  $5(2R - 3)$  and  $7(2R - 1)$ , respectively, for a value of 4 for  $R$ . In addition, the value of  $S_N$  is also mainly determined by the spatial resolution ratio,  $R$ . The value of  $S_N$  can be set to  $2R - 1$ , or slightly lower than  $2R - 1$  in order to avoid over-un-mixing. The  $S_N$  values for the original and degraded datasets were thus set to 7 and 5, respectively, for both the WV-2 and IKONOS datasets.

Finally, the UHR-fused images of the original and degraded WV-2 datasets were produced with 652,048 and 61,901 un-mixed MSPs, respectively. The UHR-fused images of the original and degraded IKONOS datasets were generated with 864,400 and 60,953 un-mixed MSPs, respectively.

### 3.3. Wald's Protocol

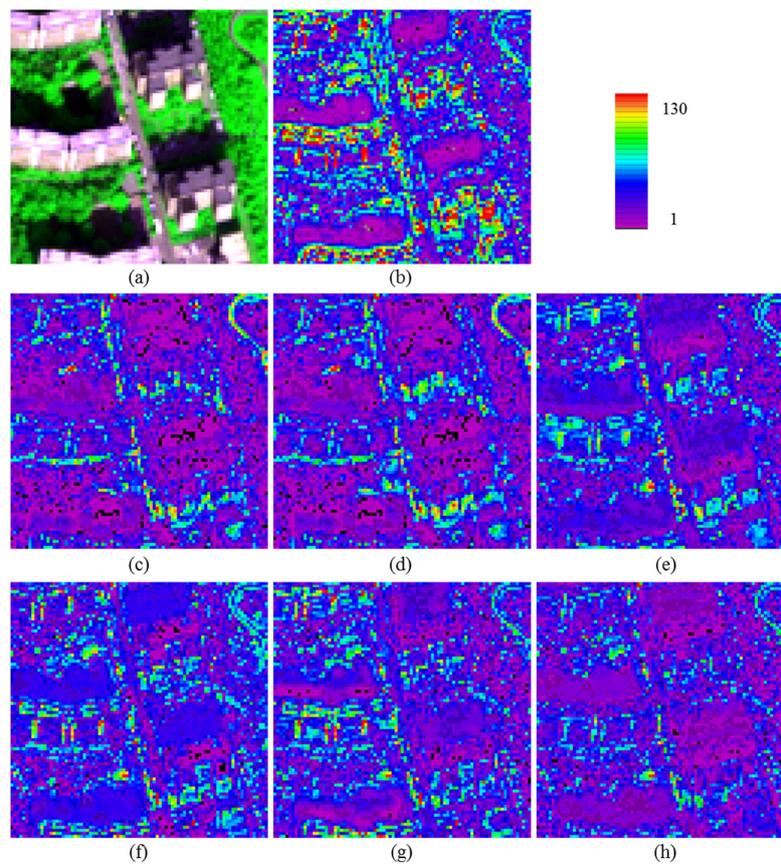
The quality indices for the fused images produced from the WV-2 and IKONOS datasets at the degraded scale are shown in Table 1. "EXP" in the table refers to interpolated MS images at the PAN scale formed without the injection of details derived from the corresponding PAN bands. In order to observe the fusion products in detail, the subsets of the MS image and the corresponding fused images produced from the degraded WV-2 dataset are shown in Figure 5. In addition, the error maps of the fusion products shown in Figure 5 are shown in Figure 6 to better compare the quality of the fused images. These error maps, which record the root mean square error (RMSE) values of the fused pixels in each of the fusion products, are applied to an identical histogram stretching scheme generated from the error map of the up-sampled MS image. The subsets of the MS image and the corresponding fused images produced from the degraded IKONOS dataset are shown in Figure 7. In this study, the original and fused MS images of the WV-2 dataset are shown as a combination of bands R, NIR1 and B, whereas the original and fused MS images of the IKONOS dataset are shown as a combination of bands R, NIR and B. An identical histogram stretching scheme generated from the up-sampled MS image (in the case of the original scale) or the reference MS image (in the case of the degraded scale) was applied to the MS images shown in these two figures.



**Figure 5.** The  $84 \times 84$  pixel subsets of the MS images and the corresponding fused images produced from the degraded WorldView (WV)-2 dataset: (a) true 2-m MS image; (b) up-sampled version of the 8-m MS image; (c) 2-m PAN image; and the results of (d) UHR fusion, (e) HR fusion, (f) PS fusion, (g) GS1 fusion, (h) GS2 fusion and (i) GSA fusion.

**Table 1.** Values of the comprehensive quality indices of the fused images produced from the degraded WV-2 and IKONOS datasets. Q, Quality; UHR, un-mixed haze- and ratio-based (HR); PS, PANSHARP.

	WV-2			IKONOS		
	ERGAS	SAM (°)	Q8	ERGAS	SAM (°)	Q4
UHR	<b>2.459</b>	<b>3.052</b>	<b>0.914</b>	<b>1.577</b>	<b>1.882</b>	<b>0.871</b>
HR	2.496	3.197	0.912	1.581	1.931	0.869
PS	2.651	3.752	0.885	1.962	2.403	0.779
GS1	2.997	3.749	0.846	2.112	2.358	0.751
GS2	2.834	3.613	0.865	2.058	2.315	0.760
GSA	2.473	3.440	0.911	1.652	2.067	0.859
EXP	3.826	3.932	0.722	2.673	2.576	0.584

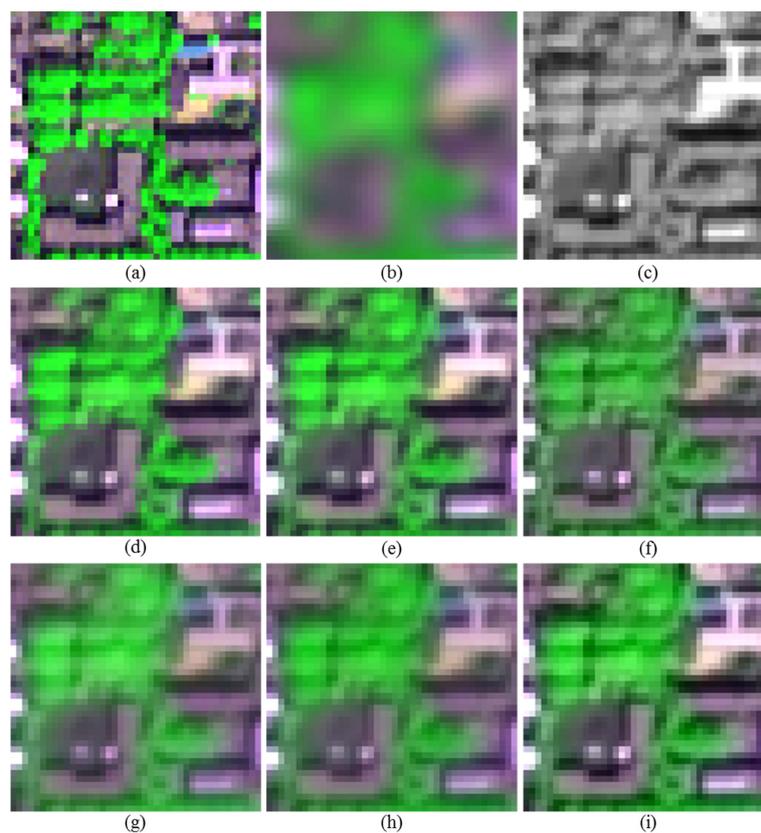


**Figure 6.** The  $84 \times 84$  pixel subsets of the original images and the error maps of the fused images produced from the degraded WV-2 dataset: (a) true 2-m MS image; and the error maps for the results of (b) the up-sampled version of the 8-m MS image; (c) UHR fusion; (d) HR fusion; (e) PS fusion, (f) GS1 fusion; (g) GS2 fusion and (h) GSA fusion.

Comparison between the fusion products (of the two sensors) generated by the six approaches shows that the proposed UHR method provides the best result in terms of ERGAS, SAM and Q8, followed by the HR and GSA methods. The GSA and PS perform better than the GS2 and GS1 methods due to the use of an MSE minimization solution to estimate the weights for the generation of an LSR PAN image from the MS image. However, the GS1 and GS2 methods yield better values of SAM than the PS methods due to the fact that the weights obtained by the MSE minimization approach are not the best for SAM [13]. The GS2 performs better than the GS1 method. This is because the former method uses an LSR PAN image obtained by applying an MTF-matched filter to the original PAN image followed by decimation, which restricts the spectral distortions introduced by the LSR PAN

image. Compared to the HR method, the superiority of the UHR is demonstrated in that it shows that the un-mixing of the MSPs during the fusion is effective at reducing the spectral distortions.

In terms of visual quality, the images produced by the UHR and HR methods contain more spatial details than the other fusion products, especially for vegetation-covered areas. This is also demonstrated by the error maps shown in Figure 6, in which more spatial details can be observed from the error maps of the PS-, GS1- and GS2-fused images. The error maps of the UHR- and HR-fused images offer smaller RMSEs than other error maps, for the vegetation- and shadow-covered areas. In addition, the UHR method produces sharper boundaries between vegetation and non-vegetation objects, such as buildings, squares, roads and shadows, than the HR method does. It also can be seen from Figure 6 that the pixels near VNV boundaries offer lower RMSEs in the UHR-fused image than those in the HR-fused image. For the fusion products derived from the degraded WV-2 dataset, the GS1, GS2 and GSA results show similar tonalities, whereas the PS method produces a well-textured result for vegetation-covered areas. In contrast, for the fusion products derived from the degraded IKONOS dataset, the GSA method produces better results than the fusion products generated by the PS, GS2 and GS1 methods, especially for areas containing buildings. Although the PS-fused images for the two sensors yield well-textured results, they do show spectral distortions, particularly for the PS-fused image produced from the degraded IKONOS dataset. In addition, the fused images generated by the GS1, GS2 and GSA methods seem blurred, as few spatial details extracted from the PAN image were injected into the fused images. These fusion products have very blurred VNV boundaries in particular.



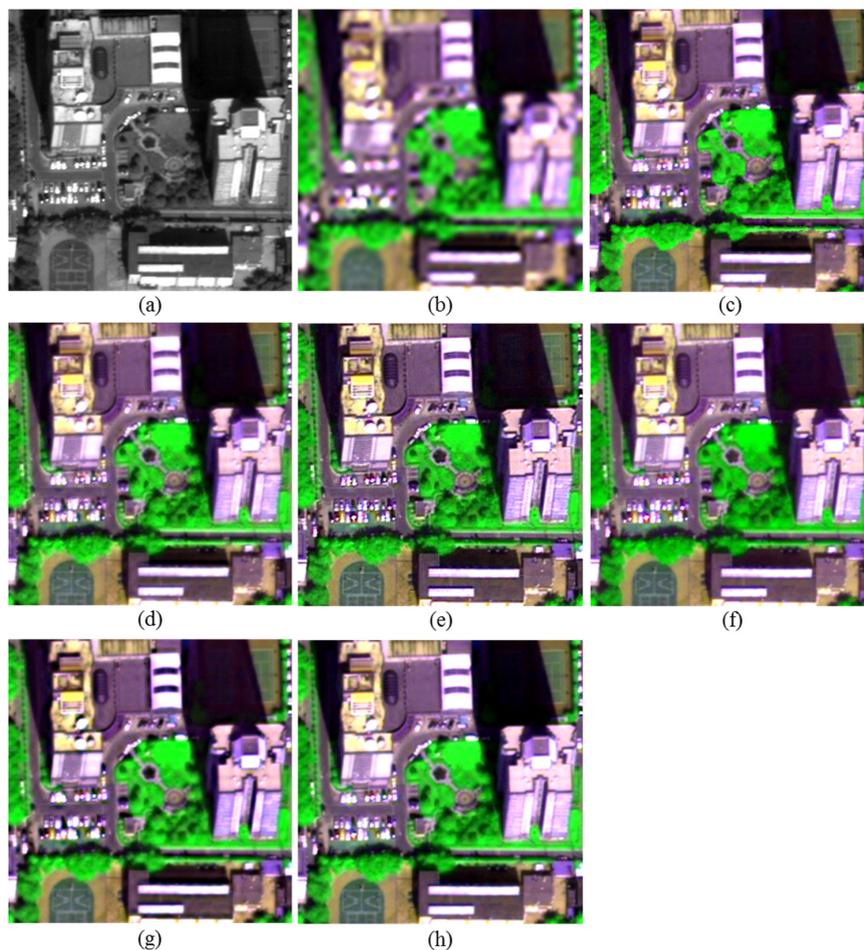
**Figure 7.** The  $64 \times 64$  pixel subsets of the MS images and the corresponding fused images produced from the degraded IKONOS dataset: (a) true 4-m MS image; (b) up-sampled version of the 16-m MS image; (c) 4-m PAN image; and the results of (d) UHR fusion; (e) HR fusion; (f) PS fusion; (g) GS1 fusion; (h) GS2 fusion and (i) GSA fusion.

### 3.4. Full Resolution

The quality indices of the fusion products derived from the two types of imagery at the original scale are shown in Table 2. Figures 8 and 9 show the subsets of the MS and PAN images and the corresponding fused images for the original WV-2 and IKONOS datasets, respectively.

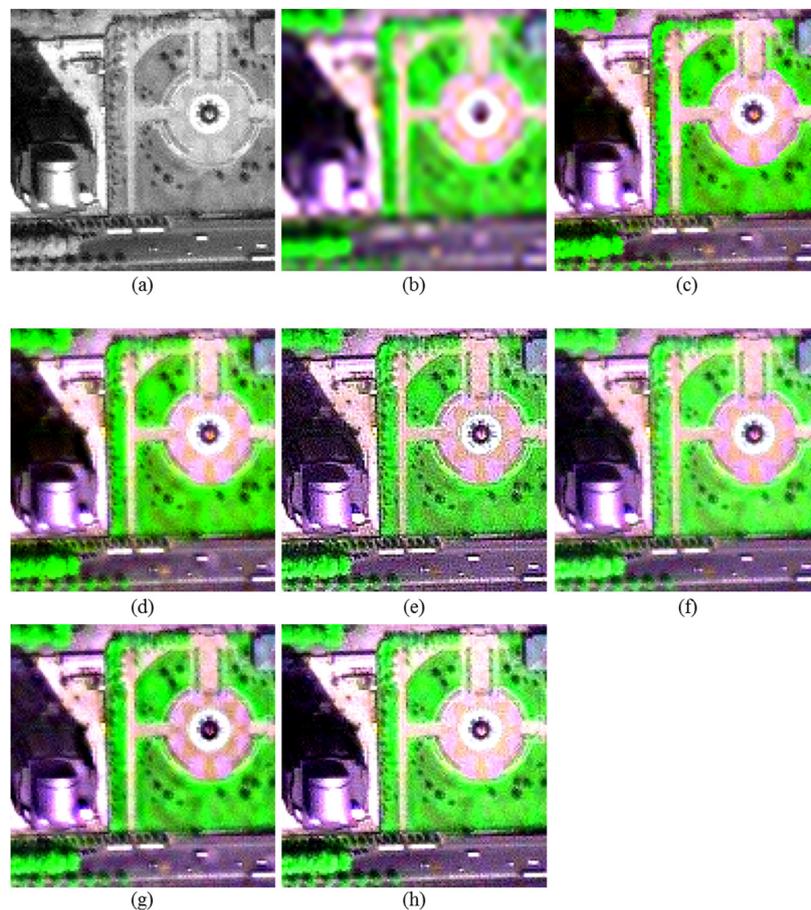
**Table 2.** Values of the comprehensive quality indices of the fused images produced from the original WV-2 and IKONOS datasets. QNR, Quality index with no reference images.

	WV-2			IKONOS		
	$D_\lambda$	$D_S$	QNR	$D_\lambda$	$D_S$	QNR
UHR	<b>0.015</b>	<b>0.050</b>	<b>0.936</b>	<b>0.022</b>	<b>0.075</b>	<b>0.906</b>
HR	0.031	0.062	0.909	0.030	0.087	0.885
PS	0.064	0.075	0.866	0.108	0.121	0.784
GS1	0.019	0.114	0.870	0.051	0.134	0.822
GS2	0.048	0.065	0.890	0.055	0.089	0.860
GSA	0.065	0.090	0.852	0.099	0.157	0.760
EXP	0	0.067	0.933	0	0.095	0.905



**Figure 8.** The  $256 \times 256$  subsets of the MS and PAN images and the corresponding fused images produced from the original WV-2 dataset: (a) 0.5-m PAN image; (b) up-sampled version of the 2-m MS image; and the results of (c) UHR fusion, (d) HR fusion, (e) PS fusion, (f) GS1 fusion, (g) GS2 fusion and (h) GSA fusion.

For the experiments at the original scale, the UHR method gives the highest QNR value followed by the HR and GS2 methods, for both the WV-2 and IKONOS datasets. The UHR performs better than the HR method in terms of QNR,  $D_\lambda$ , and  $D_S$ , indicating that the UHR is effective at reducing both spectral and spatial distortions. It can also be seen from the UHR-fused images shown in Figures 8 and 9 that this method produces sharper VNV boundaries than the other fusion methods.



**Figure 9.** The  $128 \times 128$  subsets of the MS and PAN images and the corresponding fused images produced from the original IKONOS dataset: (a) 1-m PAN image; (b) up-sampled version of the 4-m MS image; and the results of (c) UHR fusion; (d) HR fusion; (e) PS fusion; (f) GS1 fusion; (g) GS2 fusion and (h) GSA fusion.

The GS2 method outperforms the GS1 method in terms of the QNR, because the former has a significantly lower  $D_S$  value than the latter. This may be because GS2 employs a low-pass version of the PAN image obtained by applying an MTF-matched filter to the original PAN image rather than by using a weighted average of the MS bands. The latter approach is also employed by the GSA and PS methods to produce the LSR PAN images. The LSR PAN images produced by the latter approach may introduce spectral distortions due to the difference between the spectral response ranges of the MS bands and that of the PAN band. This is especially obvious in Figure 9, where the PS-, GS1- and GSA-fused images show obvious spectral distortions.

The GS1 method has a higher QNR than the PS and GSA methods because it produces lower  $D_\lambda$  values than the other two methods, especially for the IKONOS dataset. Although PS and GSA perform better than GS2 and GS1 at the degraded scale in terms of ERGAS, SAM and  $Q2^n$ , the former methods perform more poorly than the latter in terms of  $D_\lambda$ . This is because  $D_\lambda$  measures the difference between the inter-band relationships of the fused bands and those of the LSR MS bands, and this difference is not considered by the quality indices.

The GSA yields the lowest QNR for the WV-2 dataset, as it produces the highest  $D_\lambda$  value, whereas it provides the lowest QNR for the IKONOS dataset, as it produces the highest  $D_S$ . This difference between the  $D_\lambda$  and  $D_S$  values of the GSA-fused images of the two sensors may be caused by the different spectral response ranges of the PAN images acquired by the two sensors.

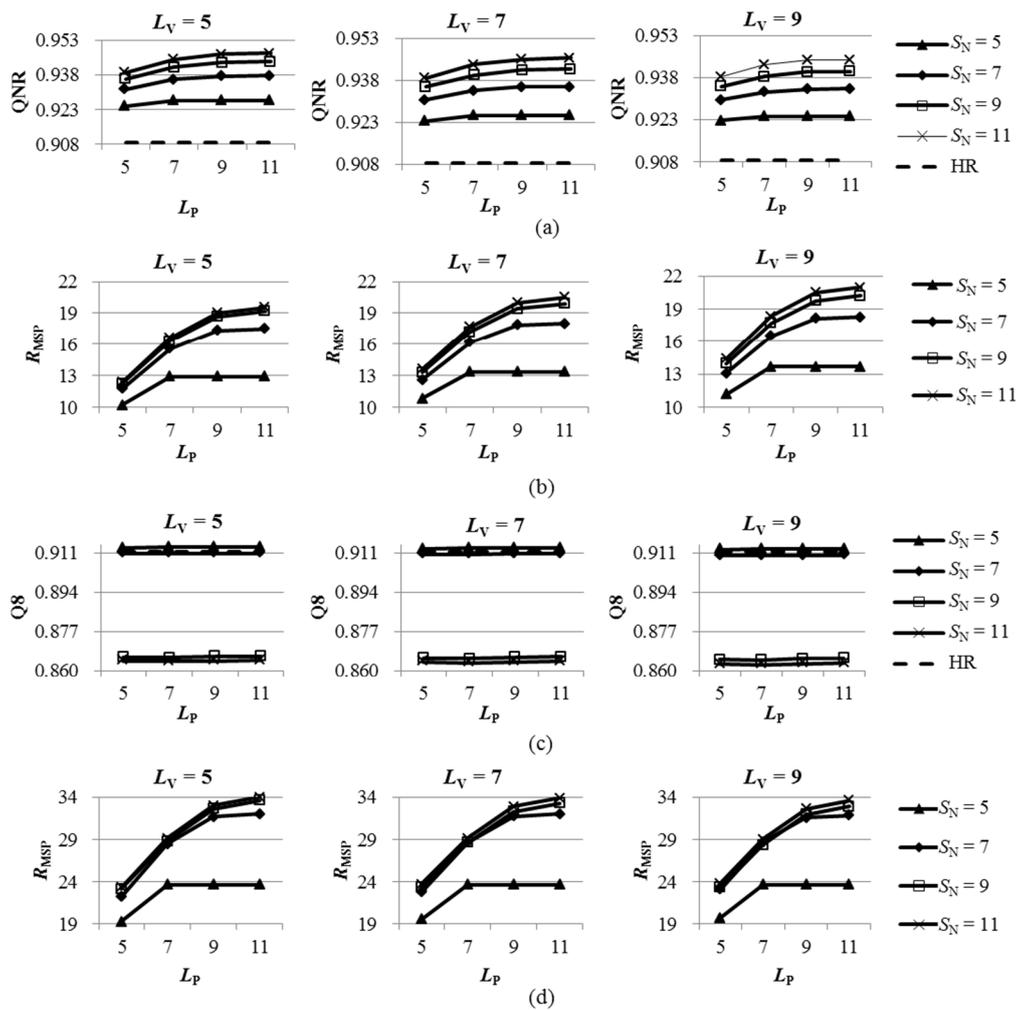
### 3.5. Determination of the Involved Parameters

The determination of the values of the parameters used in the fusion process, on which the quality of the fusion products generated by the proposed method mainly depends, is discussed in this section in order to facilitate the application of the proposed method. Each of the parameters  $L_V$ ,  $L_P$  and  $S_N$  was set to a range of different values, and the quality of the resultant fusion products obtained using both the WV-2 and IKONOS datasets at the original and degraded scales was assessed using both the QNR and  $Q2^n$  indices.

In Section 2, it was already explained that the value of  $L_V$ , which determines the accuracy of the identification of VNV boundaries in the PAN image, can be set no higher than  $2R - 1$ . Given a value of  $R$  of 4 for all of the datasets,  $L_V$  was, therefore, tested with values of 5 ( $2R - 3$ ), 7 ( $2R - 1$ ) and 9 ( $2R + 1$ ) in the experiment.

The value of  $L_P$ , which determines the maximum distance (in pixels) between the identified MSPs and the corresponding VNV boundaries, also depends mainly on the value of  $R$  and, based on this consideration, should be set equal to  $2R - 1$ . However, the value of  $L_P$  is also affected by the interpolation method used to produce the up-sampled MS image. A 23-tag pyramid filter was used to generate the up-sampled MS images in this study. The spectrum of a vegetation pixel that lies across the coarse VNV boundaries derived from the LSR MS image can affect the spectral values of pixels within the neighboring  $23 \times 23$ -pixel window. Since the coefficients with significantly higher values are concentrated at the 11 pixels in the center of the filter,  $L_P$  was set to the odd numbers in the range [5,11], and the quality of the resultant fused images was assessed. The value of  $S_N$  determines the size of the window within which the search for the purest vegetation or non-vegetation pixel to be used to un-mix an MSP during the fusion process is carried out. Its value is also affected by the value of  $R$  and the interpolation method used for producing the up-sampled MS, as well as by the spatial resolution. Consequently, the value of  $S_N$  was also set to odd numbers in the range [5,11].

The fusion products derived from the two image types at both scales were generated by using different values for  $L_V$ ,  $L_P$  and  $S_N$ , along with a  $\delta$  of 3 and an  $S_P$  of 7. The values of QNR and  $Q2^n$  were used to assess the quality of the fused images.  $R_{MSP}$ , the ratio (as a percentage) of the number of un-mixed MSPs in each fusion product to the image size (in pixels) was also recorded as a reference of the determination of these parameters. The values of QNR and  $Q8$  along with the  $R_{MSP}$  values of the fusion products for the two WV-2 datasets are shown in the line graphs in Figure 10; those for the two IKONOS datasets are shown in Figure 11.



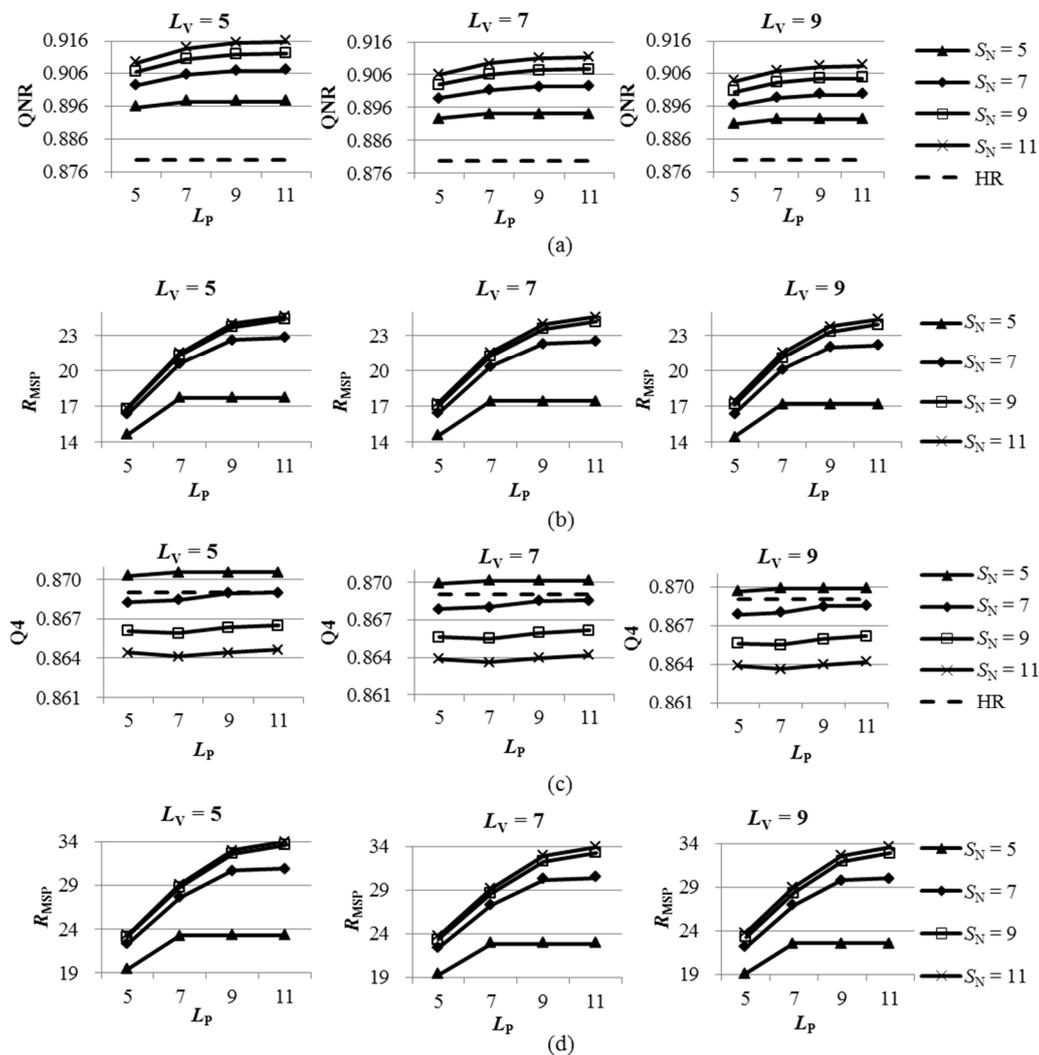
**Figure 10.** (a) QNR and (b) ratio of un-mixed MSPs ( $R_{MSP}$ ) for the UHR-fused images produced from the original WV-2 dataset; (c) Q8 and (d)  $R_{MSP}$  for the UHR-fused images produced from the degraded WV-2 dataset.

It can be observed from Figures 10 and 11 that the quality of the fusion products, in terms of both  $Q2^n$  and QNR, for both the WV-2 and IKONOS datasets at the two scales, is most sensitive to the value of  $S_N$ , followed by the values of  $L_V$  and  $L_P$ . For the two datasets at the degraded scale, only the fusion products generated using a value of 5 for  $S_N$  yield better performances than the HR-fused image in terms of  $Q2^n$ . In contrast, for the two datasets at the original scale, all of the fusion products yield better performances than the HR method in terms of QNR.

$R_{MSP}$  is also more sensitive to the value of  $S_N$  than it is to  $L_V$  and  $L_P$ . Generally, a higher value of  $S_N$  results in a higher  $R_{MSP}$ . However, a higher  $S_N$  does not always result in a better quality fusion product, because a higher  $R_{MSP}$  may introduce spectral distortions caused by over-un-mixing. This is especially significant for the case of the fusion products at the degraded scale, where a value of 11 for  $S_N$  results in the highest  $R_{MSP}$ , but the lowest  $Q2^n$  value.

The best performances for the two degraded datasets are achieved by a value of 5 for  $L_V$  and values of 7, 9 and 11 for  $L_P$ , along with a value of 5 for  $S_N$ . The best performances for the two original datasets are achieved by the fusion products generated using a value of 5 for  $L_V$  and values of 7, 9 and 11 for  $L_P$ , along with a value of 11 for  $S_N$ . Although a value of 11 gives the poorest performance at the degraded scale, it performs the best at the original scale. This may be caused by the fact that the QNR is less sensitive to spectral distortions introduced by over-un-mixing of MSPs than the  $Q2^n$ . Consequently, it is suggested that the values of  $L_V$  and  $S_N$  be set to 5 (*i.e.*,  $2R - 3$ ), whereas the value of

$L_P$  can be set to odd values in the range  $(2R - 1, 11)$  in cases where a 23-tag filter is employed to yield the up-sampled MS image.



**Figure 11.** (a) QNR and (b) ratio of un-mixed MSPs ( $R_{MSP}$ ) for the UHR-fused images produced from the original IKONOS dataset; (c) Q4 and (d)  $R_{MSP}$  for the UHR-fused images produced from the degraded IKONOS dataset.

#### 4. Discussion

Pansharpening can be generalized as the injection of spatial details derived from the PAN image into an up-sampled MS image to produce a high spatial resolution MS image. Numerous methods were proposed to optimize the spatial details derived from the PAN image or to optimize the weights by which the spatial details multiply during the injection, in order to reduce spectral distortions of the fused images. However, commonly-used image fusion methods rarely consider the spectral distortions introduced by the up-sampled MS image. Consequently, the generated fusion products suffer from spectral distortions and blurred boundaries between different objects. Hence, the application of these fused images in mapping, classification and object extraction is limited. It is desirable that the MSPs that correspond to the pure PAN pixels are fused to pure pixels with the same land cover classes (*i.e.*, the same as those of the corresponding pure PAN pixels) to reduce the spectral distortions in the fused images. A pansharpening method including the un-mixing of MSPs is proposed in this study.

However, the un-mixing of MSPs is not an easy task, because it is difficult to identify the MSPs and to decide which land cover categories the corresponding pure PAN pixels belong to using only the LSR MS and PAN images. Moreover, incorrect un-mixing of the MSPs, caused by incorrect decisions about the land cover categories and the use of incorrect substitutions during the un-mixing, may introduce spectral distortions. Since the spectral differences between vegetation and non-vegetation objects are more significant than those between other objects, the MSPs near the VNV boundaries contribute more spectral distortions than those near the other boundaries. Hence, only the MSPs near the VNV boundaries are considered to be un-mixed to pure vegetation or pure non-vegetation in the proposed UHR fusion method.

In the tests on the IKONOS and WV-2 datasets, two quality assessments were performed on the fusion products obtained at both the original and degraded scales in order to evaluate the proposed method. The quality assessments of the datasets acquired by the two sensors demonstrated that the UHR-fused images had the lowest spectral and spatial distortions. Visual comparison demonstrated that the UHR-fused images also had clearer VNV boundaries than the fusion products generated by the PS, GS and HR methods. The UHR method performed better on the original datasets in terms of the QNR than on the degraded datasets in terms of  $Q2^n$ , perhaps because of the relatively low spatial resolution of the degraded imagery, as well as the different quality indexes used at the two scales. The relatively low spatial resolutions of the degraded datasets increased the ratio of the number of MSPs to the number of pixels in the whole image. As a result, it became more difficult to search for correct substitutes for the fusion of the identified MSPs. As a consequence, some MSPs that correspond to pure non-vegetation were defined as being pure vegetation in the fused images, and *vice versa*. The incorrect fusion of these MSPs led to limited improvement in the quality indexes (*i.e.*,  $Q2^n$ ) of the UHR-fused images for the degraded IKONOS and WV-2 datasets. However, the UHR-fused images for the two degraded datasets still had significantly sharpened VNV boundaries, as could be observed by visual inspection.

It was necessary to set the values of several parameters to generate fused images using the proposed method. These parameters include the standard deviation ( $\delta$ ) of an LOG edge detector, the size of a moving window within which to calculate the average NDVI values ( $S_P$ ), the diameters of two disk-shaped SEs ( $L_V$  and  $L_P$ ) and the size of the neighborhood window within which to search for a substitution pixel ( $S_N$ ). The values of  $\delta$  and  $S_P$  were set to 0.3 and  $2R - 1$ , respectively, where  $R$  was the spatial resolution ratio of the MS and PAN images. The parameters  $L_V$ ,  $L_P$ ,  $S_N$  were set to different values, and the quality of the resulting fusion products at the original and degraded scales were accessed using the QNR and  $Q2^n$  indices, respectively. The experimental results showed that the best performances were achieved by the fusion products generated using a value of 5 for both  $L_V$  and  $S_N$ , along with values of 7, 9 and 11 for  $L_P$ . Consequently, it is recommended that  $L_V$  and  $S_N$  are set to a value of  $2R - 3$ , whereas the value of  $L_P$  can be set to any odd number in the range  $(2R - 1, 11)$ . Haze was taken into account in the proposed fusion method. In the fusion experiment, haze values were determined for each of the bands using the minimum grey level values in the bands, based on an image-based dark-object subtraction method [42–44].

The quality of the fused images generated by the proposed fusion method is affected by the accuracies of the three processing steps involved, including identifying MSPs near VNV boundaries, determining land cover categories of the identified MSPs and finding the correct substitution pixels for the un-mixing of these MSPs near VNV boundaries. The second step, determining the categories (*i.e.*, vegetation or non-vegetation) of the MSPs near VNV boundaries (*i.e.*, the sub-pixels with category  $C_M$  in  $M_{PI}$ ) is the most difficult, since these MSPs often have high NDVI values. In this study, the VNV boundaries, along with the NDVI image, were used to determine the categories of the identified MSPs. Although some classification errors could be found in the land cover classification map, the fused images generated from the original IKONOS and WV-2 datasets by the proposed method had lower spectral and spatial distortion than the fused images generated by any of the HR, PS, GS and GSA fusion methods. However, improving the accuracy of the classification is a topic requiring further

research. In order to minimize spectral distortions caused by incorrect substitutions, the land cover category of pixel  $t$  (*i.e.*,  $M_{LC}(t)$ ) is checked with respect to the corresponding NDVI value (*i.e.*,  $I_{NDVI}(t)$ ), whereas the NDVI value of the substitution pixel  $n$  is strictly checked before the fusion is performed. In detail, the value of  $M_{LC}(t)$  is checked using the two conditions in Equation (9). If neither of the two conditions in Equation (9) is met, indicating that pixel  $t$  may be misclassified in  $M_{LC}$ , the un-mixing is not performed for pixel  $t$ , in order to avoid spectral distortion caused by an incorrect substitution pixel employed. The NDVI of the substitution pixel  $n$  ( $I_{NDVI}(n)$ ) is checked using Equation (10). In detail, if pixel  $t$  is a vegetation pixel in  $M_{LC}$ ,  $I_{NDVI}(n)$  should be higher than  $I_{NDVI}(t)$ , since the substitution pixel should correspond to a pure vegetation pixel in the up-sampled MS image, but  $t$  corresponds to a VNV mixed pixel in the up-sampled MS image. Similarly, if pixel  $t$  is a non-vegetation pixel in  $M_{LC}$ ,  $I_{NDVI}(n)$  should be lower than  $I_{NDVI}(t)$ . Although both pixels are of the same category (*i.e.*, vegetation or non-vegetation) in the PAN band, there is a spectral difference between the pixel  $t$  in the true high spatial resolution MS image and the fused pixel  $t$ , which is generated by the HR method with respect to the spectra of the substitution pixel. In addition, the MSPs that lie across VNV boundaries, which are mixed pixels of vegetation and non-vegetation in the reference MS image, are set to pure vegetation or non-vegetation pixels in the resultant fused images, which may introduce spectral distortions. The substitution strategy works fine, but needs further optimization. Although the HR method is employed in the proposed fusion scheme to calculate the fused versions of the sub-pixels, some other fusion methods based on the SDM injection model, such as the Brovey, SFIM and PS methods, could also be used in the proposed scheme to improve the quality of the fusion products.

## 5. Conclusions

Pansharpening algorithms can be summarized as the injection of the spatial details of a PAN image into an up-sampled MS image to generate a high resolution MS image. How to minimize spectral distortions while preserving the spatial details of the PAN image is a major topic in pansharpening. Most current methods focus on optimizing the spatial details derived from the PAN image and the weights of the spatial details during the injection. However, the mixed sub-pixels in the up-sampled MS image also contribute to the spectral and spatial distortions in the resulting fused image, and few fusion algorithms include the reduction of the spectral distortions caused by these mixed sub-pixels. To reduce the spectral distortions generated by currently-used fusion methods, an image fusion method that realizes the spectral un-mixing of mixed MS sub-pixels near VNV boundaries in the fusion process was proposed in this paper. In tests using high resolution WorldView-2 and IKONOS datasets over Beijing, the proposed method was compared to the PANSHARP, Gram–Schmidt (Modes 1 and 2), Adapted Gram–Schmidt and haze- and ratio-based fusion methods. The fused images generated by the proposed method produced lower spectral and spatial distortions and much sharper boundaries between vegetation and non-vegetation objects, indicating that un-mixing of the MSPs near VNV boundaries is effective at reducing spectral distortions. The land cover category (*i.e.*, vegetation or non-vegetation) of each of the MSPs is a prerequisite for the correct un-mixing of the identified MSPs. Improving the accuracy of the land cover classification and optimizing the spectral substitution strategy are the subjects of further studies. Although the HR method is employed by the proposed fusion method to generate fused sub-pixels, some other fusion methods employing the SDM injection model, such as the Brovey, SFIM and PS methods, can also be used in the proposed scheme to improve the quality of the fusion products.

**Acknowledgments:** The authors wish to thank Gemine Vivone for guiding the assessment of the quality of the fusion products using QNR and to acknowledge the anonymous reviewers for providing helpful suggestions that greatly improved the manuscript. This research was supported by the Major International Cooperation and Exchange Project of National Natural Science Foundation of China (Grant No. 41120114001), the One Hundred Person Project of the Chinese Academy of Sciences (Grant Nos. Y34005101A and Y2ZZ03101B), the National Science and Technology Support Program of China (Grant No. 2015BAB05B05-02), the China Geological Survey (Grant No. 12120113089200), the Chinese Academy of Sciences-Third World Academy of Sciences (CAS-TWAS)

Centre of Excellence on Space Technology for Disaster Mitigation (Grant No. Y3YI2702KB), and the National Natural Science Foundation of China (Grant No. 41201357).

**Author Contributions:** Hui Li drafted the manuscript and was responsible for the research design, experiments and the analysis. Linhai Jing provided technical guidance and reviewed the manuscript. Liming Wang provided a portion of the data and provided technical support, and Qiuming Cheng provided technical support.

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

## References

- Zhang, Y. System and Method for Image Fusion. U.S. Patent 7,340,099, 22 July 2004.
- Huang, B.; Zhang, H.; Song, H.; Wang, J.; Song, C. Unified fusion of remote-sensing imagery: Generating simultaneously high-resolution synthetic spatial-temporal-spectral earth observations. *Remote Sens. Lett.* **2013**, *4*, 561–569. [[CrossRef](#)]
- Xu, Q.; Zhang, Y.; Li, B. Recent advances in pansharpening and key problems in applications. *Int. J. Image Data Fusion* **2014**, *5*, 175–195. [[CrossRef](#)]
- Shahdoosti, H.R.; Ghassemian, H. Fusion of MS and Pan images preserving spectral quality. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 611–615. [[CrossRef](#)]
- Pohl, C.; van Genderen, J.L. Review article multisensor image fusion in remote sensing: Concepts, methods and applications. *Int. J. Remote Sens.* **1998**, *19*, 823–854. [[CrossRef](#)]
- Zhang, J. Multi-source remote sensing data fusion: Status and trends. *Int. J. Image Data Fusion* **2010**, *1*, 5–24. [[CrossRef](#)]
- Jing, L.; Cheng, Q. Spectral change directions of multispectral subpixels in image fusion. *Int. J. Remote Sens.* **2011**, *32*, 1695–1711. [[CrossRef](#)]
- 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](#)]
- 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.
- 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](#)]
- Welch, R.; Ehlers, M. Merging multiresolution SPOT HRV and Landsat TM data. *Photogramm. Eng. Remote Sens.* **1987**, *53*, 301–303.
- Laben, C.A.; Brower, B.V. Process for Enhancing the Spatial Resolution of Multispectral Imagery Using Pan-Sharpener. U.S. Patent 6,011,875, 4 January 2000.
- Aiazzi, B.; Baronti, S.; Selva, M.; Alparone, L. MS + PAN image fusion by an enhanced Gram-Schmidt spectral sharpening. In Proceedings of 26th Earsel Symposium on New Developments and Challenges in Remote Sensing, Warsaw, Poland, 29 May–2 June 2006; pp. 113–120.
- Pradines, D. Improving spot images size and multispectral resolution. *Proc. SPIE* **1986**. [[CrossRef](#)]
- Munehika, C.K.; Warnick, J.S.; Salvaggio, C.; Schott, J.R. Resolution enhancement of multispectral image data to improve classification accuracy. *Photogramm. Eng. Remote Sens.* **1993**, *59*, 67–72.
- Zhang, Y. A new merging method and its spectral and spatial effects. *Int. J. Remote Sens.* **1999**, *20*, 2003–2014. [[CrossRef](#)]
- Liu, J.G. Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details. *Int. J. Remote Sens.* **2000**, *21*, 3461–3472. [[CrossRef](#)]
- Wald, L.; Ranchin, T. Liu “smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details”. *Int. J. Remote Sens.* **2002**, *23*, 593–597. [[CrossRef](#)]
- Zhang, Y. A new automatic approach for effectively fusing Landsat 7 as well as IKONOS images. In Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS), Toronto, ON, Canada, 24–28 June 2002; Volume 4, pp. 2429–2431.
- Zhang, Y. Problems in the fusion of commercial high-resolution satellite as well as landsat 7 images and initial solutions. In Proceedings of the ISPRS Commission IV Symposium on Geospatial Theory, Processing and Applications, Ottawa, ON, Canada, 9–12 July 2002.

21. 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](#)]
22. Vicinanza, M.R.; Restaino, R.; Vivone, G.; Dalla Mura, M.; Chanussot, J. A pansharpening method based on the sparse representation of injected details. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 180–184. [[CrossRef](#)]
23. Shensa, M.J. The discrete wavelet transform—Wedding the a trous and mallat algorithms. *IEEE Trans. Signal Process.* **1992**, *40*, 2464–2482. [[CrossRef](#)]
24. Teggi, S.; Cecchi, R.; Serafini, F. TM and IRS-1C-PAN data fusion using multiresolution decomposition methods based on the “a trous” algorithm. *Int. J. Remote Sens.* **2003**, *24*, 1287–1301. [[CrossRef](#)]
25. Mallat, S.G. Multifrequency channel decompositions of images and wavelet models. *IEEE Trans. Acoust. Speech Signal Process.* **1989**, *37*, 2091–2110. [[CrossRef](#)]
26. Yocky, D.A. Multiresolution wavelet decomposition image merger of landsat thematic mapper and spot panchromatic data. *Photogramm. Eng. Remote Sens.* **1996**, *62*, 1067–1074.
27. 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](#)]
28. 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](#)]
29. Yang, X.H.; Jiao, L.C. Fusion algorithm for remote sensing images based on nonsubsampled contourlet transform. *Acta Autom. Sin.* **2008**, *34*, 274–281. [[CrossRef](#)]
30. 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](#)]
31. 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](#)]
32. 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](#)]
33. 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](#)]
34. Guo, Q.; Liu, S. Performance analysis of multi-spectral and panchromatic image fusion techniques based on two wavelet discrete approaches. *Opt. Int. J. Light Electron Opt.* **2011**, *122*, 811–819. [[CrossRef](#)]
35. 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](#)]
36. 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](#)]
37. Choi, J.; Yu, K.; Kim, Y. A new adaptive component-substitution-based satellite image fusion by using partial replacement. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 295–309. [[CrossRef](#)]
38. Zhang, H.; Huang, B. A new look at image fusion methods from a Bayesian perspective. *Remote Sens.* **2015**, *7*, 6828–6861. [[CrossRef](#)]
39. Adams, J.B.; Smith, M.O.; Johnson, P.E. Spectral mixture modeling: A new analysis of rock and soil types at the Viking Lander I site. *J. Geophys. Res. Solid Earth* **1986**, *91*, 8098–8112. [[CrossRef](#)]
40. Zhukov, B.; Oertel, D.; Lanzl, F.; Reinhäkel, G. Unmixing-based multisensor multiresolution image fusion. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 1212–1226. [[CrossRef](#)]
41. Marr, D.; Hildreth, E.C. Theory of edge detection. *Proc. R. Soc. Lond. Ser. B Biol. Sci. (1934–1990)* **1980**, *207*, 187–217. [[CrossRef](#)]
42. Otsu, N. A threshold selection method from gray level histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *SMC-9*, 62–66.
43. Chavez, P.S. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sens. Environ.* **1988**, *24*, 459–479. [[CrossRef](#)]
44. Chavez, P.S. Image-based atmospheric corrections revisited and improved. *Photogramm. Eng. Remote Sens.* **1996**, *62*, 1025–1036.
45. 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](#)]

46. 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](#)]
47. Wald, L.; Ranchin, T.; Mangolini, M. Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images. *Photogramm. Eng. Remote Sens.* **1997**, *63*, 691–699.
48. Wald, L. Quality of high resolution synthesised images: Is there a simple criterion? In Proceedings of International Conference on Fusion Earth Data, Sophia Antipolis, France, 26–28 January 2000; pp. 99–103.
49. Yuhas, R.; Goetz, A.; Boardman, J. Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm. In Proceedings of Summaries of the Third Annual JPL Airborne Geoscience Workshop, Pasadena, CA, USA, 15 June 1992; pp. 147–149.
50. Kruse, F.A.; Lefkoff, A.B.; Boardman, J.W.; Heidebrecht, K.B.; Shapiro, A.T.; Barloon, P.J.; Goetz, A.F.H. The spectral image-processing system (sips)—Interactive visualization and analysis of imaging spectrometer data. *Remote Sens. Environ.* **1993**, *44*, 145–163. [[CrossRef](#)]
51. 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](#)]
52. Garzelli, A.; Nencini, F. Hypercomplex quality assessment of multi/hyperspectral images. *IEEE Geosci. Remote Sens. Lett.* **2009**, *6*, 662–665. [[CrossRef](#)]
53. Alparone, L.; Aiazzi, B.; Baronti, S.; Garzelli, A.; Nencini, F.; Selva, M. Multispectral and panchromatic data fusion assessment without reference. *Photogramm. Eng. Remote Sens.* **2008**, *74*, 193–200. [[CrossRef](#)]
54. Wang, Z.; Bovik, A.C. A universal image quality index. *IEEE Signal Process. Lett.* **2002**, *9*, 81–84. [[CrossRef](#)]
55. Aiazzi, B.; Alparone, L.; Baronti, S.; Garzelli, A. Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 2300–2312. [[CrossRef](#)]
56. 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](#)]



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).