Self-Dictionary Regression for Hyperspectral Image Super-Resolution

Dongsheng Gao [1], Zhentao Hu [2] and Renzhen Ye [3,*] iD

[1] The State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China; gds881@163.com
[2] The College of Computer and Information Engineering, Henan University, Kaifeng 475004, China; hzt@henu.edu.cn
[3] The College of Sciences, Huazhong Agricultural University, Wuhan 430074, China
[*] Correspondence: yerenzhen@gmail.com

check for updates

Abstract: Due to sensor limitations, *hyperspectral images* (HSIs) are acquired by hyperspectral sensors with high-spectral-resolution but low-spatial-resolution. It is difficult for sensors to acquire images with high-spatial-resolution and high-spectral-resolution simultaneously. Hyperspectral image super-resolution tries to enhance the spatial resolution of HSI by software techniques. In recent years, various methods have been proposed to fuse HSI and *multispectral image* (MSI) from an unmixing or a spectral dictionary perspective. However, these methods extract the spectral information from each image individually, and therefore ignore the cross-correlation between the observed HSI and MSI. It is difficult to achieve high-spatial-resolution while preserving the spatial-spectral consistency between low-resolution HSI and high-resolution HSI. In this paper, a self-dictionary regression based method is proposed to utilize cross-correlation between the observed HSI and MSI. Both the observed low-resolution HSI and MSI are simultaneously considered to estimate the endmember dictionary and the abundance code. To preserve the spectral consistency, the endmember dictionary is extracted by performing a common sparse basis selection on the concatenation of observed HSI and MSI. Then, a consistent constraint is exploited to ensure the spatial consistency between the abundance code of low-resolution HSI and the abundance code of high-resolution HSI. Extensive experiments on three datasets demonstrate that the proposed method outperforms the state-of-the-art methods.

Keywords: hyperspectral image super-resolution; data fusion; self-dictionary regression

1. Introduction

With the developments of hyperspectral sensors, *hyperspectral images* (HSIs) have been widely used in numerous applications [1–3], such as remote sensing classification [4–6], change detection [7] and target detection [8]. A hyperspectral sensor captures high-spectral-resolution information by constructing a continuous radiance spectrum for every pixel in the HSIs. However, due to the instrument limitation, it is difficult for hyperspectral sensors to simultaneously acquire high-spatial-resolution HSI. Moreover, the low-spatial-resolution in HSI will result in mixed pixels and greatly degrade the further processing in the remote sensing applications [9,10]. Therefore, enhancing the spatial-resolution of HSI has become an important issue in the remote sensing community [11–13].

To mitigate this issue, HSI super-resolution [14,15] has been investigated to enhance the spatial-resolution of HSI. As a software technique, HSI super-resolution does not modify the sensor array or the imaging optics. HSI super-resolution is supposed as an inverse problem [11,16,17]: the original high-spatial-resolution HSI can be recovered from the low-resolution observations [14,18,19]. Fundamentally, the missing spatial information in low-resolution HSI can

be compensated by utilizing the prior knowledge in a high-resolution coincident image of the same scene [20], such as *panchromatic image* (PAN), RGB image and *multispectral image* (MSI).

Recently, many HSI super-resolution methods [12,21–23] have been proposed to fuse low-resolution HSI with a high-resolution coincident image. An overview of recent state-of-the-art hyperspectral and multispectral image fusion methods can be found in [23,24]. A general trend in the existing methods is to exploit the spatial information by spectral mixture analysis (hyperspectral unmixing) [25]. In spectral mixture analysis, the HSI can be described by a mixture of some "pure" spectral signatures (the so-called endmembers). These spectral signatures are the spectra of the underlying materials presented in the observed scene. Based on the spectral mixture analysis, the original high-resolution HSI can be unmixed into endmember dictionary (or endmember spectra) and abundance code (or abundance matrix). The endmember dictionary denotes the "pure" spectral signatures, while the abundance code indicates the proportions of endmember spectra within each pixel. The high-resolution HSI can be estimated by combining the endmember dictionary and the abundance code. In this way, HSI super-resolution is transferred to estimate the endmember dictionary and the abundance code. The main limitation of these super-resolution methods is that their performance largely depends on the accuracy of estimating endmember dictionary (or endmember spectra) and abundance (or code of the dictionary) [26,27].

In the past several decades, a lot of research has been made to develop the efficient estimation of the endmember dictionary and the abundance code. Because the MSI captures the same scene with HSI, the endmember dictionary or the abundance code should be the same [27,28]. The endmember dictionary is extracted from the low-resolution HSI and the abundance code is estimated by the spatial fractional abundances of MSI. For example, Yokoya et al. proposed *coupled nonnegative matrix factorization* (CNMF) to estimate the endmember dictionary and abundance code from HSI and MSI, respectively. To efficiently estimate the endmember dictionary and the abundance code, various constraints are considered such as spatial smoothness, the nonnegativity and sparsity constraints. Simoes et al. [26] proposed a convex subspace-based formulation by considering a total variation abundance regularization. Zou et al. [29] proposed a double regularization HSI super-resolution by introducing the spatial structure information and the nonnegative factorization. Veganzones et al. [27] proposed a local dictionary learning to exploit locally low rank property for HSI super-resolution. Zhao et al. [19] proposed a joint spatial and spectral regulation for HSI super-resolution. Dong et al. [30] proposed a non-negative structured sparse representation to exploit the spatial correlation among the learned sparse codes. However, most methods are inspired by a simple assumption [26,27,31]: *the spectral information extracted from one of the images should also be able to explain the other one*. The hyperspectral super-resolution image can be reconstructed by combining the endmember dictionary and the abundance code. In these conventional dictionary-based methods, the endmember dictionary is extracted from the low-resolution HSI and the abundance code is estimated by the spatial fractional abundances of MSI. Only the observed low-resolution HSI is used to estimate the endmember dictionary and the observed MSI is used to estimate the abundance code. These methods usually ignore the cross-correlation between the observed HSI and MSI, which is helpful in obtaining high quality reconstructed images.

In this paper, a *self-dictionary sparse regression* (SDSR) method is proposed to fuse the HSI and MSI. Both the observed low-resolution HSI and MSI are simultaneously considered to estimate the endmember dictionary and the abundance code. The proposed method first extracts the endmember dictionary by self-dictionary sparse regression, and then estimates the abundance code by constrained least squares and consistent constraint. (1) To extract the endmember dictionary, the proposed method performs a common sparse basis selection on the concatenation of observed HSI and MSI. The endmember dictionary is learned by finding a smallest (or sparse) subset of spectral signatures to represent the whole set of spectral signatures. In particular, the learned endmember dictionary can preserve the spectral consistency between the observed HSI and MSI, since HSI and MSI share the same sparse basis. (2) Although the abundance code can be estimated based on the learned endmember

dictionary, directly estimating the abundance code of the MSI is difficult. This is because the number of multispectral bands is usually lower than the number of induced endmembers, estimating the abundance of MSI is an ill-posed problem [27]. To improve the estimation of the abundance code, a consistent constraint is exploited to ensure the spatial consistency between the abundance code of low-resolution HSI and the abundance code of high-resolution HSI. This means that spatial fractional abundances in the corresponding spatial position should coincide.

Generally, the conventional dictionary-based HSI super-resolution method consists of four steps (see in Figure 1). Given the observed HSI, (1) the first step is to learn the endmember dictionary of HSI from the observed HSI. (2) Then, the endmember dictionary is projected onto the multispectral domain to estimate the endmember dictionary of MSI by a spectral response. (3) The projected endmember dictionary is used to estimate the abundance code from the observed MSI. (4) Finally, the endmember dictionary of HSI and the abundance code of MSI are combined to reconstruct the high-resolution HSI. In these conventional methods, the spectral responses of the sensors are assumed known. For example, Wei et al. [32] proposed HSI super-resolution by jointly estimating the endmember signatures and abundances from the observed HSIs. The spectral response of the multispectral sensor is assumed to be known and available. Ghasrodashti et al. [33] proposed an HSI super-resolution by combining spectral unmixing and Bayesian sparse. The spectral response is estimated from the observed images using the method presented in [26]. Guerra et al. [34] proposed a computationally efficient algorithm for fusing multispectral and hyperspectral images by incorporating the spatial details of the MSI into the HSI. It is also assumed that the spectral response of the multispectral sensor is known. The common point of these methods is to estimate the endmember dictionary of HSI and the endmember dictionary of MSI alternatively. In this paper, the endmember dictionaries are estimated by performing a common sparse basis selection on the concatenation of observed HSI and MSI. Both the endmember dictionary of HSI and the endmember dictionary of MSI are estimated jointly. In other words, the endmember dictionary of MSI can be estimated without requiring a spectral response.
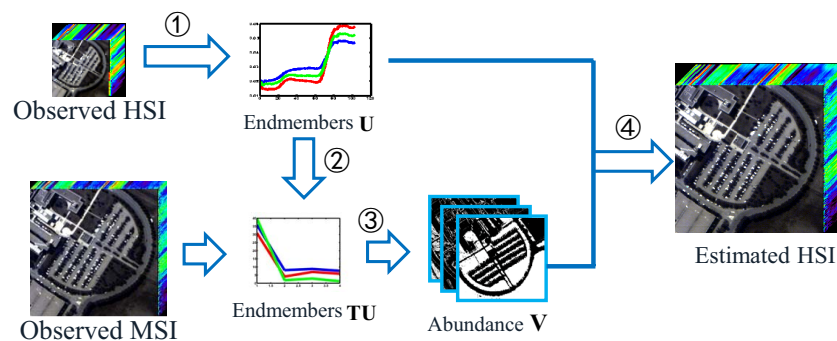
Figure 1. The conventional dictionary-based method for HSI super-resolution. (1) The first step is to learn the dictionary from the observed hyperspectral image. (2) Then, the dictionary is projected onto the multispectral domain by a spectral response. (3) The projected dictionary is used to estimate the code from the observed MSI. (4) Finally, the dictionary and the code are combined to reconstruct the hyperspectral super-resolution image.

Consequently, the proposed method includes two main contributions:

(1) A self-dictionary regression is proposed to identify the endmember dictionary on the concatenation of observed HSI and MSI. The learned endmember dictionary can preserve the spectral consistency between the observed HSI and MSI, since HSI and MSI share the same sparse basis.

(2) A consistent constraint is proposed to preserve the spatial consistency between low-resolution HSI and high-resolution HSI. The abundance code is estimated by constrained least squares and consistent constraint.

This paper is organized as follows: the HSI super-resolution problem is formulated in Section 2. Section 3 overviews the conventional HSI super-resolution methods. Then, Section 4 presents the proposed method. Section 5 details the experimental results. Section 6 concludes this paper.

2. Related Work

In recent years, many HSI super-resolution methods have been proposed to enhance the spatial-resolution of HSI. A general solution is to fuse the low-resolution HSI with a high-resolution coincident image of the same scene. According to the coincident images, most methods can be categorized into two groups: **the pan-sharpening methods** and **the HSI-MSI fusion methods**.

The pan-sharpening methods enhance the spatial-resolution of HSI by fusing the low-resolution HSI with a corresponding high-spatial resolution *panchromatic image* (PAN). Due to the high-spatial-resolution, PANs have been widely used to enhance the resolution of MSI, known as MSI pan-sharpening. With the increasing availability of hyperspectral sensors, PANs are now extending to enhance the resolution of HSI, known as HSI pan-sharpening [23]. Naturally, HSI pan-sharpening can be handled by the popular MSI pan-sharpening methods, such as component substitution (CS) [35] and multiresolution analysis (MRA) [36]. To obtain the high-resolution HSI, the CS methods substitute the spatial component of HSI with PAN. The MRA methods inject the spatial details of PAN into HSI. In the CS methods, the HSI is first separated into spatial and spectral component. Subsequently, the fused image is obtained by substituting the spatial component with the PAN. The MRA methods first generate the spatial details by a multiresolution decomposition of the PAN. Then, the generated details can be injected into the HSI. The CS methods may provide high spatial quality, but compromise spectral distortion. The MRA methods can preserve spectral consistency but deteriorate the spatial quality.

The HSI-MSI fusion methods obtain the high-resolution HSI by combining an HSI with a high-spatial-resolution MSI because the MSI provides both the spatial and the spectral information. The HSI-MSI fusion methods are significantly more difficult and advantageous than the pan-sharpening methods. To fuse the HSI and MSI, many methods have been proposed in the recent decades: frequency-based methods and dictionary-based methods (unmixing or spectral mixture analysis). The frequency-based methods transform the original images (HSI and MSI) into the frequency components, and combine the wavelet coefficients in the transform domain. More recently, the dictionary-based methods uses spectral mixture analysis to enhance the spatial resolution of HSI. Based on the spectral mixture analysis, the original high-resolution HSI can be unmixed into endmember dictionary (or endmember spectra) and abundance code (or abundance matrix). In these methods, the high-resolution HSI can be fused by combining the estimated endmember dictionary and abundance code. Much research has been conducted to develop the efficient estimation of the endmember dictionary and the abundance code. Furthermore, various constraints [19,27,29,30,37] are considered to efficiently estimate the endmember dictionary and the abundance code, such as spatial smoothness [26], the nonnegativity and sparsity constraints [37]. Simoes et al. [26] proposed a convex subspace-based method by considering a total variation abundance regularization. Akhtar et al. [37] presented a constrained sparse representation by imposing non-negativity and the spatial structure for HSI super-resolution. Zou et al. [29] proposed a double regularization HSI super-resolution by introducing spatial structure information and the nonnegative factorization. Lanaras et al. [38] proposed a hyperspectral super-resolution method by jointly unmixing the two input images into pure reflectance spectra of the observed materials. Veganzones et al. [27] proposed a local dictionary learning method to exploit the low rank property for HSI super-resolution. Zhao et al. [19] proposed a joint spatial-spectral regulation to utilize the nonlocal similarities for HSI super-resolution. Dong et al. [30] proposed a non-negative structured sparse representation to exploit the spatial correlation among the learned sparse codes. Mei et al. [21] proposed a novel three-dimensional full convolutional neural network for hyperspectral super-resolution. The main advantage of the dictionary-based methods is physically reasonable and effective for HSI-MSI fusion.

3. Problem Formulation

Let $\mathbf{X} \in \mathbb{R}^{L \times N}$ denote an HSI with L spectral bands (rows of $\mathbf{X}$) and N pixels (columns of $\mathbf{X}$). In this paper, the goal of HSI super-resolution is to recover the original high-resolution HSI $\mathbf{X} \in \mathbb{R}^{L \times N}$ from two degraded observations of $\mathbf{X}$: a low-resolution HSI $\mathbf{Y}_h \in \mathbb{R}^{L \times n}$ and a high-resolution MSI $\mathbf{Y}_m \in \mathbb{R}^{l \times N}$. For these observations, $n \ll N$, and $l \ll L$, which make the estimation of a super-resolution HSI $\mathbf{X}$ severely ill-posed. In particular, both observations $\mathbf{Y}_h$ and $\mathbf{Y}_m$ can be considered to be mapping of the original HSI:

$$\mathbf{Y}_h = \varphi_h\left(\mathbf{X}\right), \mathbf{Y}_m = \varphi_m\left(\mathbf{X}\right), \tag{1}$$

where $\varphi_h : \mathbb{R}^{L \times N} \to \mathbb{R}^{L \times n}$ and $\varphi_m : \mathbb{R}^{L \times N} \to \mathbb{R}^{l \times N}$. Generally, MSI Y_m can be approximated as:

$$\mathbf{Y}_m \approx \mathbf{TX}, \tag{2}$$

where $\mathbf{T} \in \mathbb{R}^{l \times L}$ is a transformation matrix or spectral response. This means the spectral quantization from the high-spectral-resolution image (HSI) to the low-spectral-resolution image (MSI). In experiments, the spectral response is often assumed to be known [26,31].

The HSI can be described by a mixture of some pure spectral signatures [8,9]. If the spectrum at each pixel is assumed to be a linear combination of several endmember spectra, the high-resolution HSI is formulated as:

$$\mathbf{X} = \mathbf{UV}, \tag{3}$$

where $\mathbf{U} \in \mathbb{R}^{L \times p}$ is the spectral dictionary and denotes the spectral signatures of the endmembers. $\mathbf{V} \in \mathbb{R}^{p \times N}$ is the abundance code and denotes the proportions of endmember spectra at each pixel. In addition, p represents the number of endmember spectra.

If $\mathbf{U}$ and $\mathbf{V}$ are known or estimated, an estimation of $\mathbf{X}$ can thereby be obtained by Equation (3). This reasoning led to the conventional dictionary-based method for HSI super-resolution [27,31] (see in Figure 1). Given the observed HSI $\mathbf{Y}_h$, (1) the first step is to learn the dictionary $\mathbf{U}$ from the observed HSI. (2) Then, the dictionary is projected onto the multispectral domain by a spectral response $\mathbf{T}$. (3) The projected dictionary $\mathbf{TU}$ is used to estimate the code $\mathbf{V}$ from the observed MSI $\mathbf{Y}_m$. (4) Finally, the dictionary $\mathbf{U}$ and the code $\mathbf{V}$ are combined to reconstruct the high-resolution HSI $\widehat{\mathbf{X}}$.

As shown in Figure 1, the success of the dictionary-based methods depends fundamentally on the first three procedures. Many recent efforts have been made in the past few years to develop the efficient estimation of $\mathbf{U}, \mathbf{T}$ and $\mathbf{V}$, related to the first three procedures—for example, Refs. [39,40] for $\mathbf{U}$, Refs. [26,27] for $\mathbf{T}$, and Refs. [27,30] for $\mathbf{V}$.

4. Proposed Method

The available dictionary-based HSI super-resolution methods usually obtain dictionary $\mathbf{U}$ from the observed low-resolution HSI and estimate the code $\mathbf{V}$ from the observed MSI. However, these methods usually ignore the cross-correlation between the observed HSI and MSI, which is helpful in obtaining high quality reconstructed images. In this section, a novel HSI super-resolution method is presented from the perspective of self-dictionary sparse regression [41]. A sparsity regularization is introduced to recover the spectral dictionary (or endmember signatures) from both observed low resolution HSI and MSI. In order to preserve the spectral consistency between the observed HSI and MSI, the spectral dictionary is estimated by performing a common sparse basis selection on the concatenation of observed HSI and MSI. Meanwhile, the dictionary of MSI can be also estimated without requiring a spectral response. Once the dictionary is acquired, the corresponding code can be recovered by standard non-blind unmixing methods (e.g., [42]). Here, the code of high-resolution HSI is estimated by ensuring the spatial consistency between the code of low-resolution HSI and the code of high-resolution HSI. Therefore, the proposed method contains two parts: (1) learning the dictionary with a common sparsity on the concatenation of observed HSI and MSI, (2) estimating the corresponding code by

exploiting a consistent constraint to preserve the spatial consistency between low-resolution HSI and high-resolution HSI. The flowchart of the proposed method is shown in Figure 2.
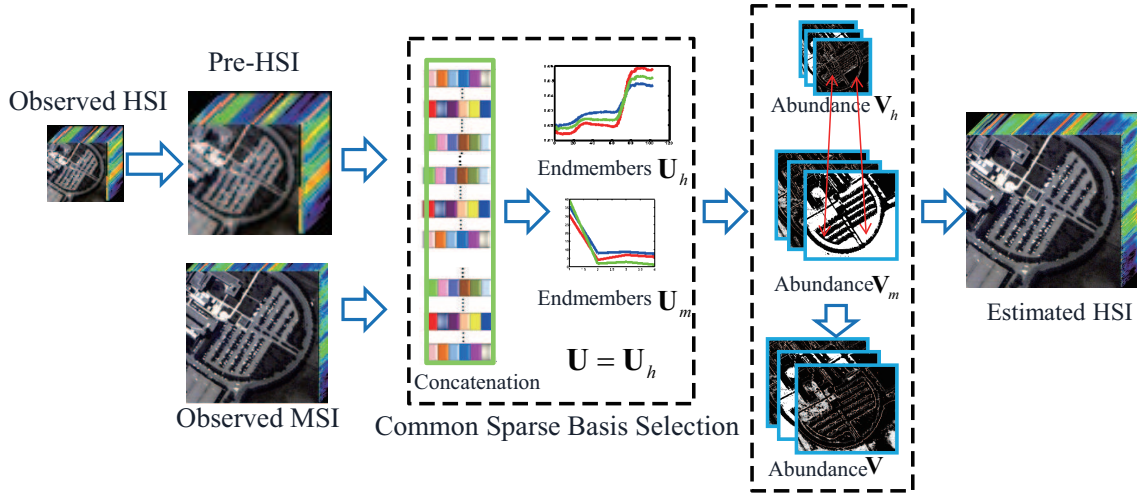
Figure 2. The proposed method first (1) learns the dictionary with a common sparsity on the concatenation of observed HSI and MSI, (2) then estimates the corresponding code by exploiting a consistent constraint to preserve the spatial consistency between low-resolution HSI and high-resolution HSI.

4.1. Self-Dictionary Sparse Regression

The performance of the super-resolution HSI is strongly influenced by the estimation of the spectral dictionary [27]. Since the low-resolution HSI and the corresponding high-resolution HSI capture the same scene, their endmembers should be the same [27,28]. In the general dictionary-based methods, the dictionary $\mathbf{U}_h \in \mathbb{R}^{L \times p}$ is learned from the observed low-resolution HSI itself and ignores the spectral consistency between the observed HSI and MSI. In particular, the elements of the dictionary $\mathbf{U}_h$ should be consistent with the dictionary $\mathbf{U}_m$ of MSI $\mathbf{Y}_m$. For this reason, a novel method is proposed to obtain the spectral dictionary entries from both observed HSI and MSI. The spectral dictionary (or underlying materials or endmembers) of high-resolution HSI $\mathbf{U} \in \mathbb{R}^{L \times p}$ is assumed to be same with the dictionary of low-resolution HSI $\mathbf{U}_h \in \mathbb{R}^{L \times p}$, $\mathbf{U} = \mathbf{U}_h$.

First, a spatially upscale version of the low-resolution HSI can be obtained by:

$$\mathbf{Y}_{f(h)} = f\left(\mathbf{Y}_h\right), \tag{4}$$

where $f\left(\right)$ is an upscaling function—for example, a bicubic interpolation. Then, the spectral dictionary is learned from the concatenation of pre-HSI and observed MSI by endmembers' induction algorithms:

$$\begin{bmatrix} \mathbf{Y}_{f(h)} \\ \mathbf{Y}_m \end{bmatrix} \approx \begin{bmatrix} \mathbf{U}_h \\ \mathbf{U}_m \end{bmatrix}, \mathbf{S} \tag{5}$$

where $\mathbf{U}_m \in \mathbb{R}^{l \times p}$ is the dictionary of MSI and $\mathbf{S} \in \mathbb{R}^{p \times N}$ is the coefficient matrix. In this paper, self-dictionary sparse regression [43,44] is used to extract the spectral dictionary (the spectra of the underlying endmembers). In this way, the dictionary is learned by finding a smallest subset of

measurement vectors to represent the whole set of measurement vectors. Therefore, the dictionary is learned by perform a common sparse basis selection:

$$\min_{\mathbf{C} \in \mathbb{R}^{N \times N}} \|\mathbf{C}\|_{row-0},$$
$$\text{s.t. } \begin{bmatrix} \mathbf{Y}_{f(h)} \\ \mathbf{Y}_m \end{bmatrix} = \begin{bmatrix} \mathbf{Y}_{f(h)} \\ \mathbf{Y}_m \end{bmatrix} \mathbf{C}, \mathbf{C} \geq 0, \mathbf{1}^T \mathbf{C} = \mathbf{1}^T, \tag{6}$$

where $\mathbf{C} \in \mathbb{R}^{N \times N}$ is the abundance matrix and two constraints are imposed: the *abundance non-negativity constraint* (ANC) $\mathbf{C} \geq 0$ and the *abundance sum-to-one constraint* (ASC) $\mathbf{1}^T \mathbf{C} = \mathbf{1}^T$. $\|\mathbf{C}\|_{row-0}$ denotes the number of nonzero rows of C.

Under the pure pixel assumption, Equation (6) may be formulated by identifying a complete pure pixel index set Λ [45]. Λ lists all indices of nonzero rows of C and contains a pure pixel index of every endmember:

$$\Lambda = \text{rowsupp}(\mathbf{C}), \tag{7}$$

$$|\Lambda| = p, \Lambda \subset \{1, \ldots, N\}. \tag{8}$$

In this paper, Equation (6) is tackled by greedy pursuit [41,46] under unknown number of endmembers. In this case, both pure pixels and the number of endmembers are identified simultaneously.

Once Λ is identified, we have

$$\begin{bmatrix} \mathbf{U}_h \\ \mathbf{U}_m \end{bmatrix} = \begin{bmatrix} \mathbf{Y}_{f(h)} \\ \mathbf{Y}_m \end{bmatrix}_{\Lambda}. \tag{9}$$

The intuition of the above observation is that, by solving Equation (6), we may recover a complete pure pixel index set Λ, and, consequently, identify the true endmember signature matrix $\mathbf{U}_h$ and $\mathbf{U}_m$.

4.2. Sparse Codes Estimation via Constrained Least Squares and Consistent Constraints

Once the spectral dictionaries $\mathbf{U}_h$ and $\mathbf{U}_m$ are estimated, the corresponding codes $\mathbf{V}_h$ and $\mathbf{V}_m$ are estimated by solving the following *constrained least squares* (CLS) optimization problem:

$$\mathbf{V}_h = \arg \min_{\mathbf{V}_h \geq 0} \|\mathbf{Y}_h - \mathbf{U}_h \mathbf{V}_h\|_F^2, \tag{10}$$

$$\mathbf{V}_m = \arg \min_{\mathbf{V}_m \geq 0} \|\mathbf{Y}_m - \mathbf{U}_m \mathbf{V}_m\|_F^2, \tag{11}$$

where codes $\mathbf{V}_h \in \mathbb{R}^{p \times n}$ and $\mathbf{V}_m \in \mathbb{R}^{p \times N}$ represent abundances of materials, and $\|\cdot\|_F$ denotes the Frobenius norm. The constraint $\mathbf{V}_h \geq 0, \mathbf{V}_m \geq 0$ means that the codes are element-wise non-negative [47]. It is possible to add the abundances sum-to-one constraint [47]: $\mathbf{V}_h^T \mathbf{1}_{p \times 1} = \mathbf{1}_{N \times 1}$, and $\mathbf{V}_m^T \mathbf{1}_{p \times 1} = \mathbf{1}_{n \times 1}$. However, this constraint is usually dropped due to possible scale model mismatches [48]. Multiplicative update rules have been developed to minimize the residual errors in Equations (10) and (11). The multiplicative update rules for Equations (10) and (11) are given by:

$$\mathbf{V}_h \leftarrow \mathbf{V}_h .* \left(\mathbf{U}_h^T \mathbf{Y}_h\right) ./ \left(\mathbf{U}_h^T \mathbf{U}_h \mathbf{V}_h\right), \tag{12}$$

$$\mathbf{V}_m \leftarrow \mathbf{V}_m .* \left(\mathbf{U}_m^T \mathbf{Y}_m\right) ./ \left(\mathbf{U}_m^T \mathbf{U}_m \mathbf{V}_m\right), \tag{13}$$

where $.*$ and $./$ denote element-wise multiplication and division, respectively. $(\mathbf{A})^T$ denotes the transposition of the matrix $\mathbf{A}$.

In the conventional dictionary-based methods, the code of desired high-resolution HSI $\mathbf{V}$ is approximated by estimating the code from the observed MSI $\mathbf{V}_m$: $\mathbf{V} \approx \mathbf{V}_m$. However, a main limitation

of these methods is their inability to estimate the code $\mathbf{V}_m$ from MSI [27]. Since the number of multispectral bands is usually lower than the number of entries in the spectral dictionary, estimating the abundance code of MSI is an ill-posed problem [27]. To address this issue, the spatial consistency between $\mathbf{V}_h$ and $\mathbf{V}_m$ is considered to improve the estimation of $\mathbf{V}$:

$$\mathbf{V} = \arg\min_{\mathbf{V}} \|\mathbf{V} - \mathbf{V}_m\|_F^2 + \lambda \|\mathbf{V}_h\mathbf{F}_L - \mathbf{V}\mathbf{F}_H\|_F^2, \tag{14}$$

where λ is a weighting factor to balance the importance of the two terms in Equation (14). When λ is too small ($\lambda = 0$), the second term $\lambda \|\mathbf{V}_h\mathbf{F}_L - \mathbf{V}\mathbf{F}_H\|_F^2$ will have less effect. In this case, $\mathbf{V}$ is only related to $\mathbf{V}_m$. When λ is too big ($\lambda = \infty$), the first term $\|\mathbf{V} - \mathbf{V}_m\|_F^2$ will have less effect. In this case, $\mathbf{V}$ is only related to $\mathbf{V}_h$. Therefore, the appropriate parameter λ will balance the contributions between the codes $\mathbf{V}_m$ and $\mathbf{V}_h$. $\mathbf{F}_L \in \mathbb{R}^{n \times q}$ and $\mathbf{F}_H \in \mathbb{R}^{N \times q}$ represent the corresponding spatial locations in low-resolution HSI $\mathbf{Y}_h$ and high-resolution HSI $\mathbf{X}$, respectively. $\mathbf{F}_H = [\mathbf{f}_{H1}, \ldots, \mathbf{f}_{Hq}]$ indicates the q samples on high-resolution HSI $\mathbf{X}$, and $\mathbf{F}_L = [\mathbf{f}_{L1}, \ldots, \mathbf{f}_{Lq}]$ indicates a set of corresponding samples on low-resolution HSI $\mathbf{Y}_h$. Each column of $\mathbf{F}_L$ and $\mathbf{F}_H$ is indicator functions that contain one at the selected samples and zeros elsewhere. In other words, 1 means that the pixel is selected, and 0 means that the pixel is not selected. In addition, there is a positional correspondence between these selected pixels on the low-resolution HSI $\mathbf{Y}_h$ and the high-resolution HSI $\mathbf{X}$. HSI super-resolution assumes that a pixel in the low-resolution HSI can be obtained by averaging the pixels of high-resolution HSI belonging to the same area. This assumption refers to pixel aggregation [32]. Given a set of q samples $\mathbf{F}_H = [\mathbf{f}_{H1}, \ldots, \mathbf{f}_{Hq}]$ on high-resolution HSI $\mathbf{X}$, and a set of corresponding samples $\mathbf{F}_L = [\mathbf{f}_{L1}, \ldots, \mathbf{f}_{Lq}]$ on low-resolution HSI $\mathbf{Y}_h$, their codes in the respective samples must coincide, $\mathbf{V}_h\mathbf{F}_L = \mathbf{V}\mathbf{F}_H$.

Taking derivative with regard to $\mathbf{V}$ and setting the resultant equation to zero yield, the solution of Equation (14) is given by:

$$\hat{\mathbf{V}} = \left(\mathbf{V}_m + \lambda\mathbf{V}_h F_L\mathbf{F}_H^T\right)\left(\mathbf{I} + \lambda\mathbf{F}_H\mathbf{F}_H^T\right)^{-1}, \tag{15}$$

where $\mathbf{I}$ is a $N \times N$ identity matrix. The desired high-resolution HSI can be recovered by Equation (3), after estimating the spectral dictionary U and the abundance code V. The overall description for estimating the high-resolution HSI is given as Algorithm 1.

Algorithm 1 The proposed method

Input:

    $\mathbf{Y}_h \in \mathbb{R}^{L \times n}$ - the observed hyperspectral image;
    $\mathbf{Y}_m \in \mathbb{R}^{l \times N}$ - the observed multispectral image;
    Parameter $\lambda = 1$, and $q = n$.

1. Identify a complete pure pixel index set Λ by Equation (6);
2. Estimate the spectral dictionaries $\mathbf{U}_h$ and $\mathbf{U}_m$ by Equation (9), calculate the spectral dictionary $\mathbf{U}$ by $\mathbf{U} = \mathbf{U}_h$;
3. Calculate the code of low-resolution HSI $\mathbf{V}_h$ by Equation (10);
4. Calculate the code of MSI $\mathbf{V}_m$ by Equation (11);
5. Estimate the code of high-resolution HSI $\mathbf{V}$ by Equation (15);
6. $\mathbf{X}$ is output by $\mathbf{X} = \mathbf{UV}$.

Output:

    $\mathbf{X}$ - the high-resolution hyperspectral image.

---

5. Experiments

To verify the performance of the proposed method, four subsections are presented in this section: (1) Section 5.1 details three public datasets; (2) Section 5.2 introduces the competing methods and evaluation indexes; (3) Section 5.3 describes the implementation details of the proposed method; and, (4) finally, Section 5.4 displays the experiments and comparisons.

5.1. Datasets

To evaluate the performance of the proposed method, extensive datasets are considered, including ground based and remote sensing imagery. These datasets are commonly used in HSI super-resolution. Through experiments on these databases, we can fairly evaluate the performance of the proposed method. We consider the HSIs in the datasets as the original high-resolution HSIs [37]. As common practice [23,37,49], the low-resolution HSI is simulated by first blurring the high-resolution HSI and then down-sampling the result by a factor of 3 in each direction. In order to facilitate the experiments, we use the $Norm(\cdot)$ defined in Equation (16) to normalize the HSIs in the datasets to the range $[0, 1]$ in advance, where **X** donates the HSIs in datasets. For evaluating the proposed method, we converted the resulting image into an 8-bit image:

$$Norm(\mathbf{X}) = \frac{\mathbf{X} - min(\mathbf{X})}{max(\mathbf{X}) - min(\mathbf{X})}. \tag{16}$$

Ground based dataset. The ground based dataset is CAVE (http://www.cs.columbia.edu/CAVE/databases/multispectral/). CAVE [37,38,49] comprises 32 HSI with $512 \times 512 \times 31$. Among them, 512×512 is the image size, and 31 is the number of bands. Each HSI is acquired at a wavelength interval of 10 nm in the range 400–700 nm. The high-spatial-resolution MSI is created by integrating a high-resolution HSI over the spectral dimension, using the Nikon D700 (Sendai, Japan) spectral response (https://www.maxmax.com/spectral_response.htm). Here, the HSI of "Balloons" is selected to give visual illustration, as shown in Figure 3. Three bands (3, 25 and 30) are selected as a pseudo color view.
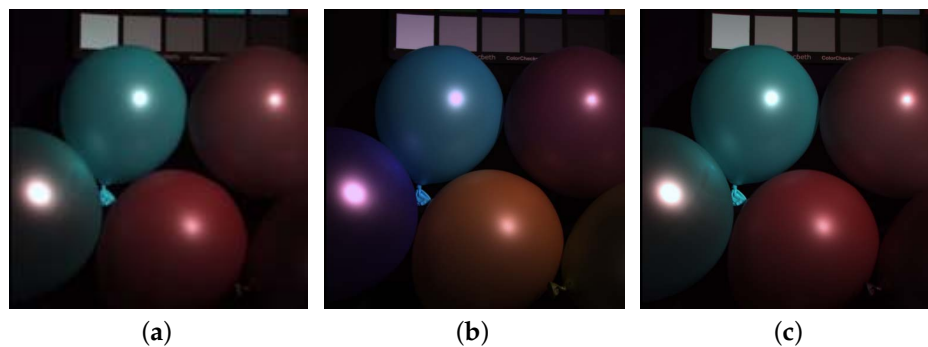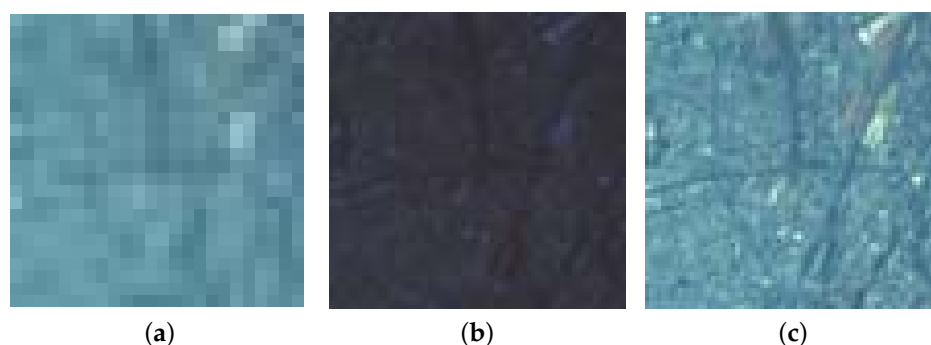

Figure 3. Balloons data from CAVE dataset (color image composed of bands 3, 25 and 30 for the red, green and blue channels, respectively). (**a**) low-resolution hyperspectral image; (**b**) multispectral image; (**c**) high-resolution hyperspectral image.

Remote sensing dataset. The first remote sensing dataset is Pavia University scene (http://www.ehu.eus/ccwintco/uploads/e/e3/Pavia.mat), which is widely used in hyperspectral classification with nine kinds of samples [26,28]. This image is acquired from the *reflective optics system imaging spectrometer* (ROSIS), a sensor of DLR that has 115 spectral bands. ROSIS hyperspectral remote sensing satellite captured the image at Pavia University in northern Italy in 2002. This HSI contains 610×340 pixels and 103 spectral bands, after removing water absorption bands. A part of size $200 \times 200 \times 93$ including abundant detailed information is selected from the original image, as shown in Figure 4. This image is used as the high-resolution HSI. In the HSI super-resolution problem, the spectral responses of the sensors is assumed known. The spectral response of the IKONOS satellite is used to create the observed MSI.
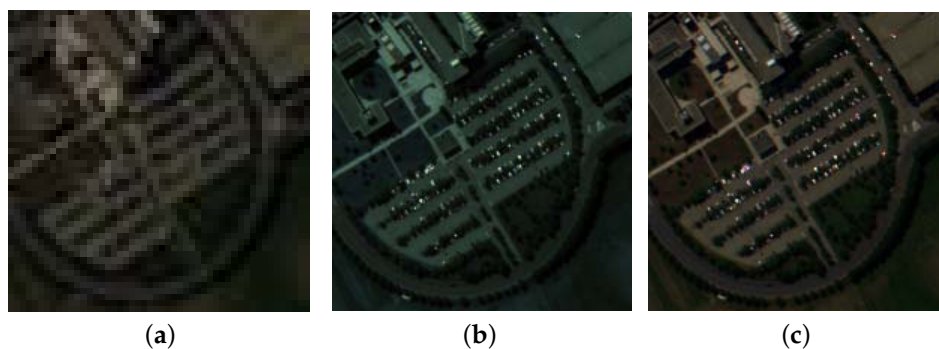
Figure 4. Pavia dataset and its auxiliary images (color image composed of bands 20, 16 and 5 for the red, green and blue channels, respectively). (**a**) low-resolution hyperspectral image; (**b**) multispectral image with four bands (blue, green, red, and near-infrared); (**c**) high-resolution hyperspectral image.

The second remote sensing dataset is Paris scene [26]. This dataset is taken above Paris and obtained by two instruments on board the *Earth Observing-1 Mission* (EO-1) satellite: the Hyperion instrument, and the *Advanced Land Imager* (ALI) instrument. The Hyperion instrument provides an HSI with a spatial resolution of 30 m, while the ALI instrument provides an MSI at a resolution of 30 m, as shown in Figure 5. In the experiments of HSI-MSI fusion, the observed HSI should have lower resolution than the observed MSI. Therefore, the low-resolution HSI is simulated by blurring and down-sampling the original high-resolution HSI. Furthermore, to further verify the effectiveness of the proposed method, Pavia dataset is used for hyperspectral classification.

Figure 5. Paris dataset and its auxiliary images (color image composed of bands 28, 13 and 3 for the red, green and blue channels, respectively). (**a**) low-resolution hyperspectral image; (**b**) multispectral image with four bands (blue, green, red, and near-infrared); (**c**) high-resolution hyperspectral image.

5.2. Competitors and Evaluation Indexes

Competing method. To demonstrate the superiority of the proposed method, six state-of-the-art methods are selected as competing methods. All of these competing methods can be obtained from the corresponding authors' MATLAB+MEX (R2014a, MathWorks, Natick, MA, USA) implementation. All the parameters involved in the competing methods are automatically chosen as described in the references. For a fair comparison, we use the same number of endmembers in all experiments. Furthermore, the spectral responses of the sensors can be estimated by HySure [26]. The estimated spectral responses are used as basis for all the competing methods:

- *Greedy simultaneous orthogonal matching pursuit* (GSOMP) (http://www.csse.uwa.edu.au/~ajmal/code/HSISuperRes.zip) [37] is a constrained sparse representation by imposing non-negativity and the spatial structure.

- *Coupled nonnegative matrix factorization* (CNMF) [31] achieves the high-resolution HSI by alternately unmixing the observed HSI and MSI. The endmember dictionary and abundance code are estimated from HSI and MSI, respectively.
- *Bayesian sparse representation* (BSR) (http://staffhome.ecm.uwa.edu.au/~00053650/code.html) [50] performs comparably to GSOMP because BSR exploits the non-parametric Bayesian framework without requiring explicit parameter tuning.
- *Hyperspectral SuperResolution* (HySure) (https://github.com/alfaiate/HySure) [26] formulates HSI super-resolution as a convex subspace-based optimization problem. A total variation abundance regularization is considered to promote piecewise-smooth. In addition, the spectral responses of the sensors are estimated by assuming relative smooth.
- *Convolutional neural network collaborative nonnegative matrix factorization* (CNNCNMF) [28] employs the convolutional neural network to learn the spatial mapping between the observed HSI and MSI. In addition, the collaborative nonnegative matrix factorization is introduced to explore the spectral characteristic between the observed HSI and MSI.
- *Proximal alternating linearized minimisation* (PALM) (https://github.com/lanha/SupResPALM) [49] obtains the high-resolution HSI by jointly unmixing the observed HSI and MSI.

Evaluation indexes. To assess the performance of all competing methods, six quantitative indices [26,50] are employed, including *root mean square error* (RMSE), *mean peak signal-to-noise ratio* (MPSNR), *mean structure similarity index* (MSSIM), *erreur relative globale adimensionnelle de synthèse* (ERGAS), *universal image quality index* (UIQI), and *spectral angle mapper* (SAM). The similarity between the target image and the reference image is evaluated by RMSE and MPSNR based on mean square error. The structural consistency between the target image and the reference image is evaluated by MSSIM. The fidelity is evaluated by ERGAS based on the weighted sum of MSE for each band. UIQI is a universal image quality index to evaluate image reconstruction quality. The fidelity of the spectral reflectance is described by SAM. Furthermore, the larger MPSNR, MSSIM, and UIQI are, the better image reconstruction quality is. The smaller RMSE, ERGAS, and SAM are, the better the image reconstruction quality is.

5.3. Parameter Determination

The parameters of the proposed method can be clearly delineated in Algorithm 1: the regularization parameter λ, and the number of consistency samples q. The impact of parameters is evaluated on Balloons data, Pavia dataset and Paris dataset. Experimental results for various λ and q are shown in Figure 6. In the CAVE dataset, the parameters are set to be $\lambda = 1$, $q = n$. In the Pavia dataset, the parameters are set to be $\lambda = 0.7$, $q = n$. In the Paris dataset, the parameters are set to be $\lambda = 10$, $q = n$.

The parameter λ is the regularization parameter which balances the contributions between the consistency constraint term and the fidelity term. With respect to λ, Figure 6a,d,g show the experimental results from $\lambda = 0$ to $\lambda = 10$. For both RMSE and MPSNR, the performance increases with a large λ. However, in the "Balloons" data, when the value of λ is larger than 1, the performance decreases. In the Pavia dataset, when the value of λ is larger than 0.7, the performance decreases.

Experimental results for various q/n (from 0 to 1) are shown in Figure 6b,e,h. It is observed that with the increase of the spatial bandwidth q/n, the performance increases slightly. When the number of consistency samples is equal to the number of pixels in the low-resolution HSI n, we achieve the best performance on RMSE and MPSNR. In this case, all the pixels in the low-resolution HSI are considered. $\mathbf{F}_L$ becomes a $n \times n$ identity matrix, $\mathbf{F}_H$ accounts for a uniform sub-sampling of the image, in order to yield the lower-spatial-resolution of the HSI.

Furthermore, we analyze the effect of the number of endmembers p on the performance of proposed method, as shown in Figure 6c,f,i. It is observed that a proper number of endmembers is important to different datasets. In the CAVE dataset, Pavia dataset and Paris dataset, the number of endmembers is set to 5, 10 and 20, respectively.
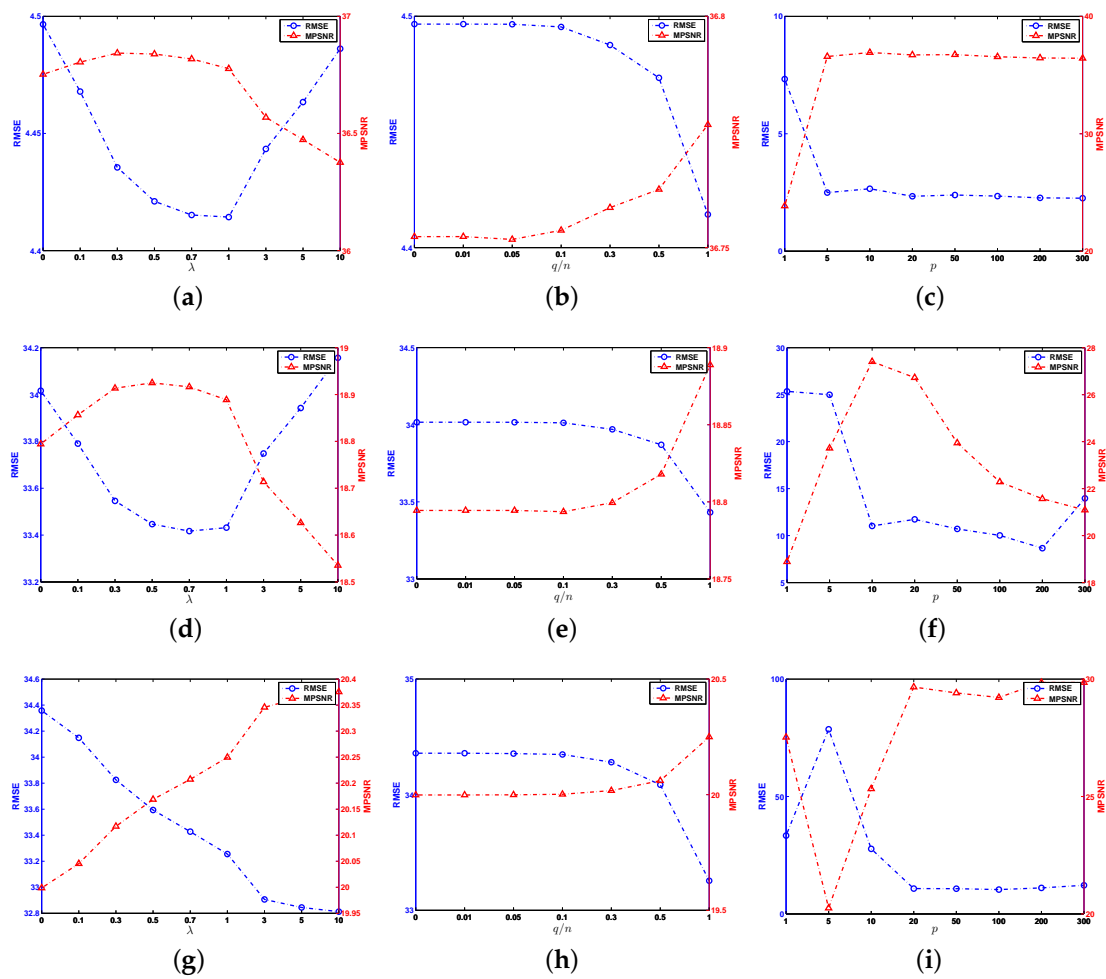
Figure 6. (**a**) evaluation of the parameters λ on Balloons data; (**b**) evaluation of the parameters q on Balloons data; (**c**) evaluation of the parameters p on Balloons data; (**d**) evaluation of the parameters λ on Pavia dataset; (**e**) evaluation of the parameters q on Pavia dataset; (**f**) evaluation of the parameters p on Pavia dataset; (**g**) evaluation of the parameters λ on Paris dataset; (**h**) evaluation of the parameters q on Paris dataset; (**i**) evaluation of the parameters p on Paris dataset.

5.4. Experimental Results

5.4.1. Comparison to the State-of-the-Art Methods

The experimental results on three datasets are discussed in this subsection. The proposed method is compared with six state-of-the-art methods in the experiments. These methods are CNMF, GSOMP, BSR, HySure, CNNCNMF and PALM. For the sake of comparison, we present three types of evaluation: (a) quantitative assessments, (b) visual results, and (c) spectral differences between the original and estimated HSIs. Tables 1–3 show HSI super-resolution performances of different methods on the CAVE, Pavia, and Paris dataset, respectively.

Table 1. Comparison with the competing methods on the CAVE dataset. The reported evaluation values are calculated in the 8-bit resulting images.

	Bicubic	CNMF	GSOMP	BSR	HySure	CNNCNMF	PALM	SDSR
RMSE	8.692	5.261	4.858	4.518	4.997	7.809	7.582	2.845
MPSNR	28.66	30.17	31.62	33.09	33.13	33.77	32.7	32.75
MSSIM	0.7123	0.8570	0.9142	0.9233	0.8509	0.9149	0.9062	0.9183
ERGAS	237.52	420.68	209.79	151.93	149.24	151.48	162.1	148.75
UIQI	0.5540	0.7001	0.7921	0.7607	0.6939	0.7962	0.7520	0.7637
SAM	0.2451	0.3809	0.3485	0.3184	0.2068	0.3013	0.2103	0.3723
Time	0.05	14.8	254.6	1650.0	142.0	139.4	62.7	20.8

Table 2. Comparison with the competing methods on the Pavia dataset. The reported evaluation values are calculated in the 8-bit resulting images.

	Bicubic	CNMF	GSOMP	BSR	HySure	CNNCNMF	PALM	SDSR
RMSE	13.019	4.799	4.342	5.090	7.569	4.622	3.898	2.349
MPSNR	25.96	34.76	35.77	34.58	33.99	35.32	36.54	42.04
MSSIM	0.7106	0.9683	0.9771	0.9656	0.9463	0.9679	0.9813	0.9908
ERGAS	195.20	74.01	68.73	80.42	97.27	72.00	60.89	33.36
UIQI	0.7905	0.9754	0.9791	0.9733	0.9563	0.9777	0.9825	0.9941
SAM	0.0960	0.0733	0.0537	0.0788	0.0877	0.0768	0.0492	0.0411
Time	0.02	3.5	27.2	249.9	35.8	34.8	19.6	58.4

Table 3. Comparison with the competing methods on the Paris dataset. The reported evaluation values are calculated in the 8-bit resulting images.

	Bicubic	CNMF	GSOMP	BSR	HySure	CNNCNMF	PALM	SDSR
RMSE	13.332	12.205	8.532	12.979	8.071	8.183	8.125	7.942
MPSNR	28.21	29.48	31.90	29.36	32.41	32.29	32.35	32.50
MSSIM	0.6671	0.7521	0.8438	0.7428	0.8381	0.8500	0.8472	0.8466
ERGAS	144.31	126.24	94.34	131.53	88.56	91.11	91.00	88.19
UIQI	0.6282	0.6565	0.7959	0.6448	0.8132	0.8185	0.8120	0.8209
SAM	0.0598	0.0929	0.0547	0.1009	0.0491	0.0513	0.0492	0.0482
Time	0.01	0.39	2.99	21.82	1.89	1.76	4.07	1.12

For qualitative analysis, Figures 7–9 show the pseudo color images in three datasets. It is clearly observed that the proposed method receives a better visual quality than other competing methods. As can be observed from Figures 7–9, the proposed method produces much sharper edges than other methods without any obvious artifacts across the image. It is very interesting to see that the proposed method can achieve the comparable visual performance. Without requiring the spectral response, the proposed method is suitable for complex real-word super resolution tasks.
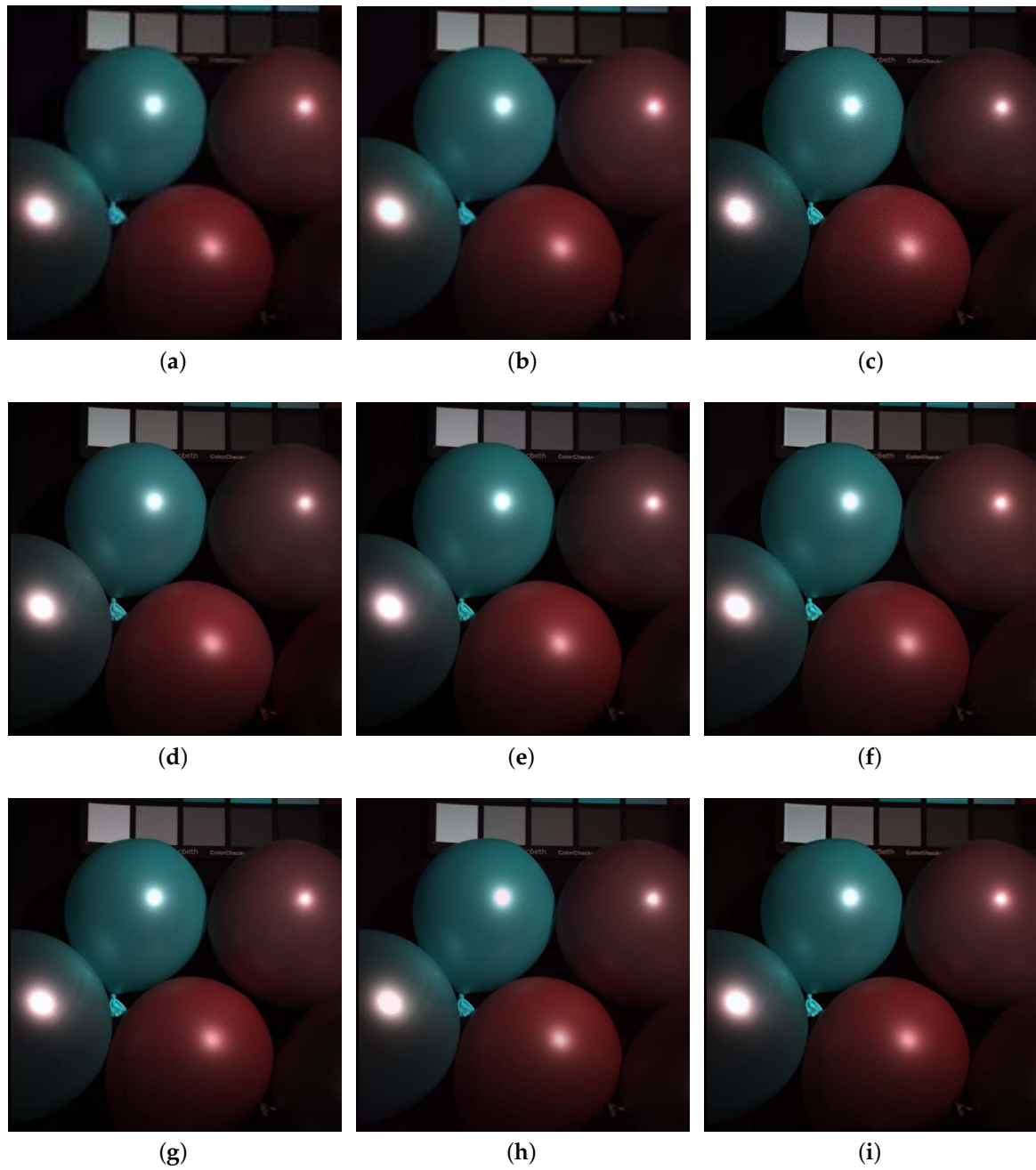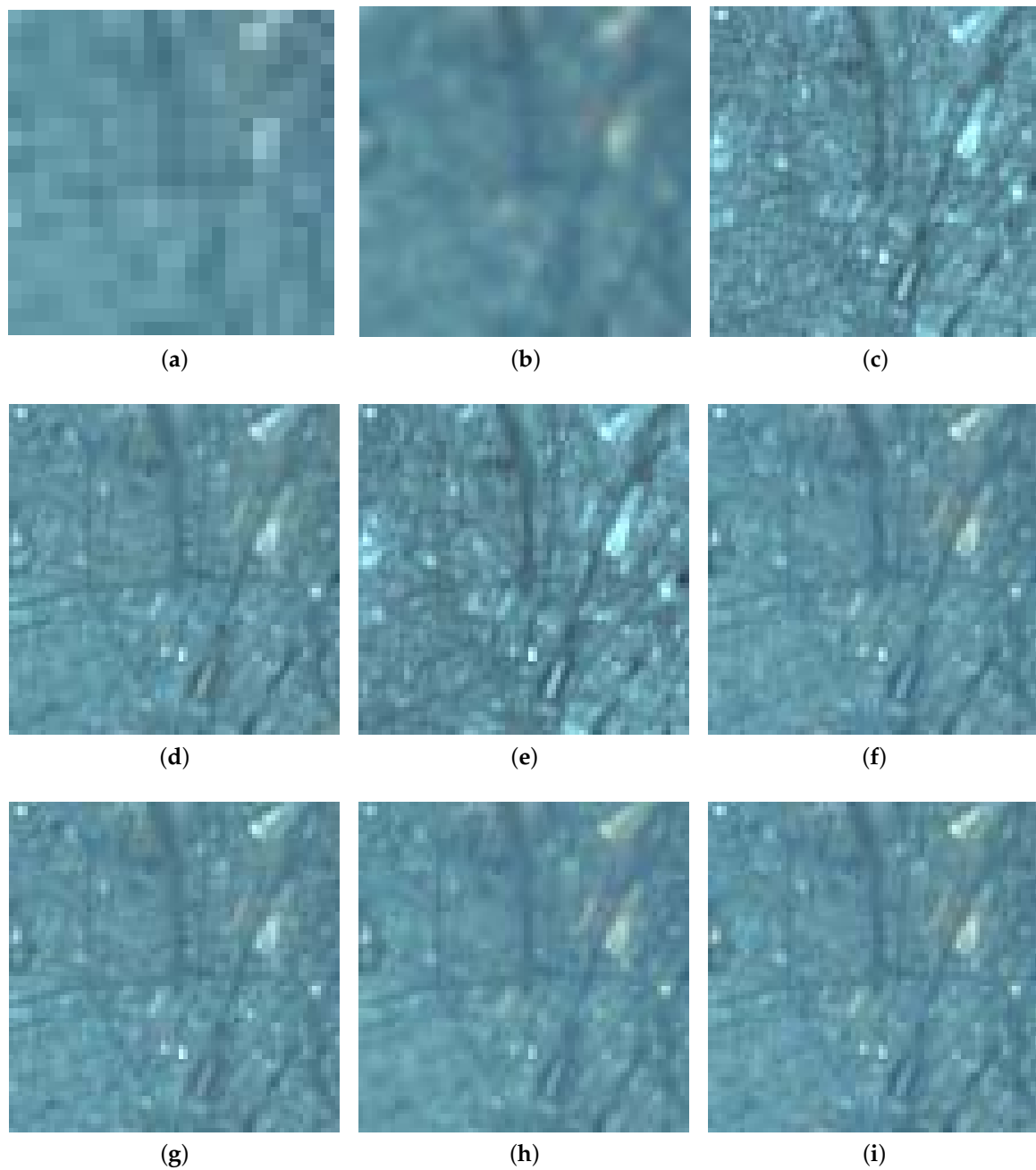
Figure 7. Super resolution image reconstruction on 'Balloons' data. (**a**) low-resolution HSI; (**b**) bicubic (RMSE: 6.784); (**c**) CNMF (RMSE: 6.448); (**d**) GSOMP (RMSE:4.267); (**e**) BSR (RMSE: 3.981); (**f**) HySure (RMSE: 3.428); (**g**) CNNCNMF (RMSE: 3.750); (**h**) PALM (RMSE: 3.691); (**i**) SDSR (RMSE: 3.055).
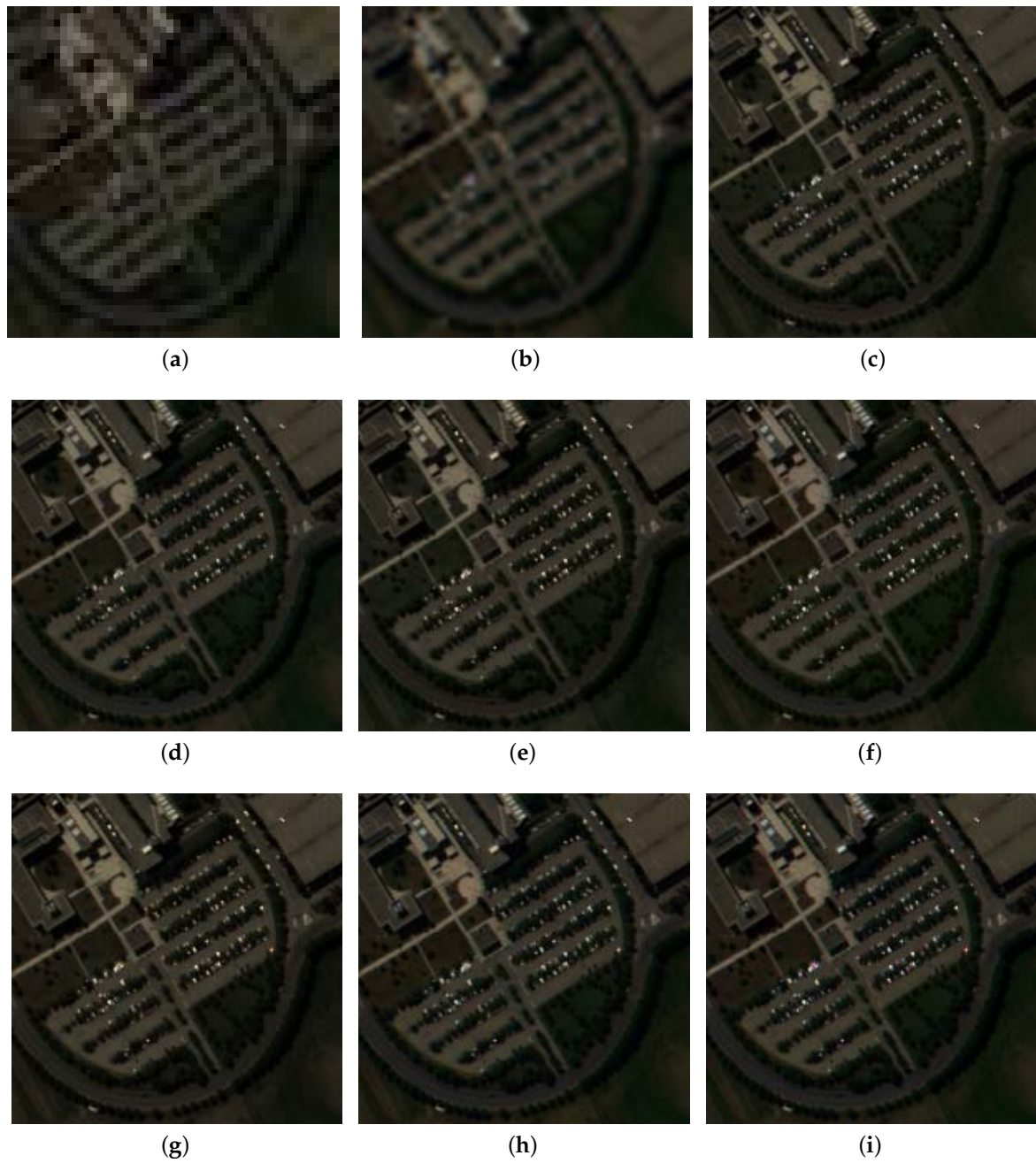
Figure 8. Super resolution image reconstruction on Pavia dataset. (**a**) low-resolution HSI; (**b**) bicubic (RMSE: 13.019); (**c**) CNMF (RMSE: 4.799); (**d**) GSOMP (RMSE: 4.342); (**e**) BSR (RMSE: 5.090); (**f**) HySure (RMSE: 7.569); (**g**) CNNCNMF (RMSE: 4.622); (**h**) PALM (RMSE: 3.898); (**i**) SDSR (RMSE: 2.349).

Figure 9. Super resolution image reconstruction on Paris data. (**a**) low-resolution HSI; (**b**) bicubic (RMSE: 13.332); (**c**) CNMF (RMSE: 12.205); (**d**) GSOMP (RMSE: 8.532); (**e**) BSR (RMSE: 12.979); (**f**) HySure (RMSE: 8.071); (**g**) CNNCNMF (RMSE: 8.183); (**h**) PALM (RMSE: 8.125); (**i**) SDSR (RMSE: 7.942).

To gain further intuition, the average squared error between the original spectra and obtained spectra are given. Figure 10 shows spectral residuals of several pixels after super resolution. In Figure 10, six typical objects are selected in the datasets such as Pixel (50,50) in Balloons data, Pixel (300,300) in Balloons data, Pixel (50,50) in Pavia dataset, Pixel (100,100) in Pavia dataset, Pixel (50,50) in Paris dataset and Pixel (70,70) in Paris dataset. Compared with other competing methods, Figure 10 illustrates that the proposed method can still preserve the useful spectral information of the original HSI. Furthermore, the average squared errors between the obtained and ground truth spectra are given in Table 4.
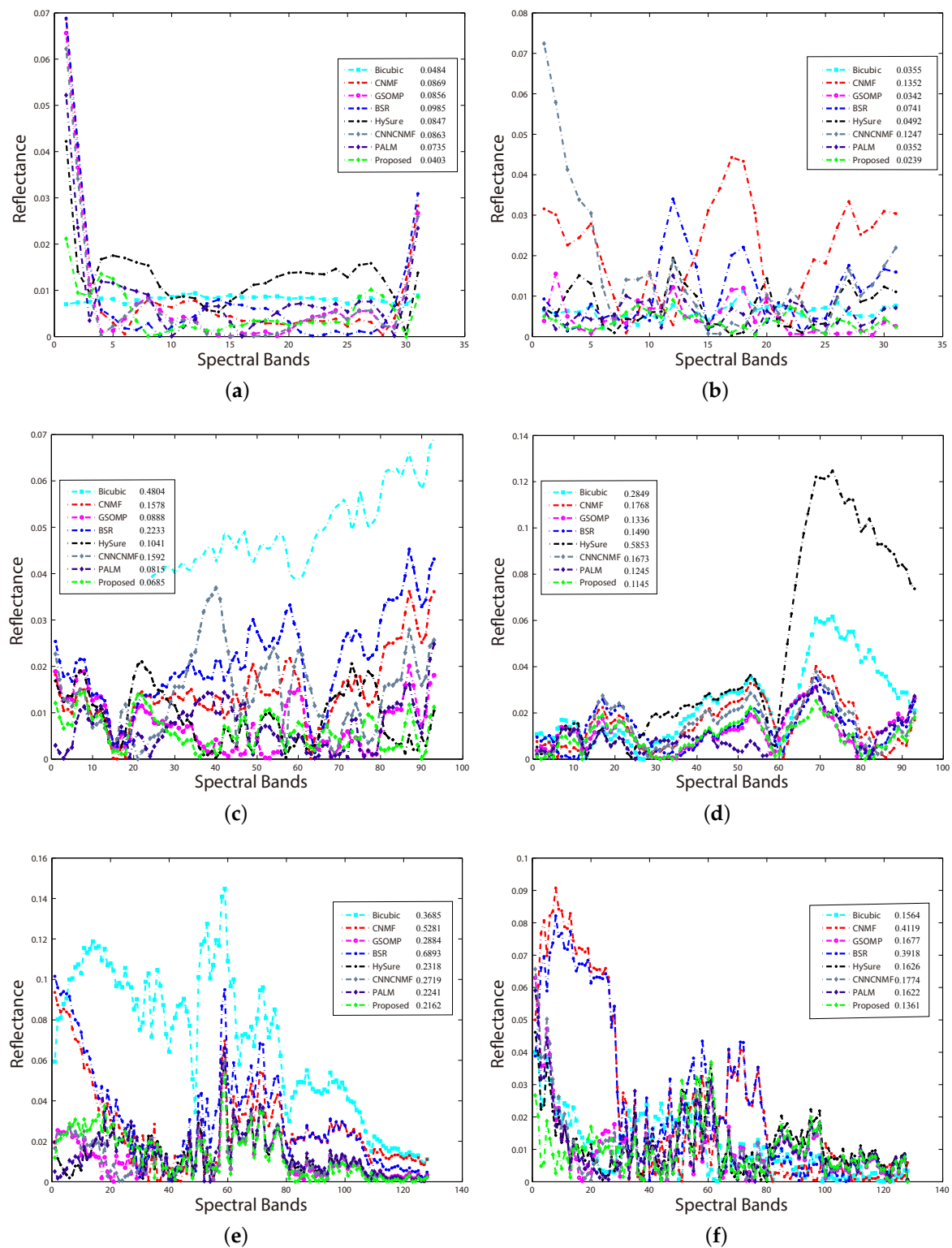
Figure 10. Spectral squared residuals of several pixels in the datasets. (**a**) Pixel (50,50) in Balloons data; (**b**) Pixel (300,300) in Balloons data; (**c**) Pixel (50,50) in Pavia dataset; (**d**) Pixel (100,100) in Pavia dataset; (**e**) Pixel (50,50) in Paris dataset; (**f**) Pixel (70,70) in Paris dataset.

Table 4. Average squared error between obtained and ground truth spectra of some pixels.

Pixels	Bicubic	CNMF	GSOMP	BSR	HySure	CNNCNMF	PALM	SDSR
Balloons (50,50)	0.0484	0.0869	0.0856	0.0985	0.0847	0.0863	0.0735	0.0403
Balloons (300,300)	0.0355	0.1352	0.0342	0.0741	0.0492	0.1247	0.0352	0.0239
Pavia (50,50)	0.4804	0.1578	0.0888	0.2233	0.1041	0.1592	0.0815	0.0685
Pavia (100,100)	0.2849	0.1768	0.1336	0.1490	0.5853	0.1673	0.1245	0.1145
Paris (50,50)	0.3685	0.5281	0.2884	0.6893	0.2318	0.2719	0.2241	0.2162
Paris (70,70)	0.1564	0.4119	0.1677	0.3918	0.1626	0.1774	0.1622	0.1361

5.4.2. Effectiveness of Concatenation in the Proposed SDSR Method

In the proposed method, the cross-correlation between the observed HSI and MSI is utilized. The endmember dictionaries are learned by performing a common sparse basis selection on the concatenation of observed HSI and MSI. Both the endmember dictionary $\mathbf{U}_h$ and $\mathbf{U}_m$ can be estimated simultaneously. We call the proposed method with concatenation operation the concatenation proposed method. In contrast, there are still two ways to estimate the endmember dictionary without the concatenation of observed HSI and MSI: coupling proposed and uncoupling proposed method.

- The concatenation proposed method simultaneously estimate the endmember dictionary $\mathbf{U}_h$ and $\mathbf{U}_m$ from the concatenation of observed HSI and MSI.
- The coupling proposed method refers to first estimating the endmember dictionary $\mathbf{U}_h$ from the observed HSI, and then utilizing the estimated $\mathbf{U}_h$ and spectral response to calculate the $\mathbf{U}_m$.
- The uncoupling proposed method refers to estimating the endmember dictionary $\mathbf{U}_h$ and $\mathbf{U}_m$ from the observed HSI and MSI, respectively.

A comparison of the proposed SDSR methods is shown in Table 5. Among the proposed methods, the concatenation proposed method achieves impressive performance. This is because the concatenation operation can exploit the cross-correlation between the observed HSI and MSI. Furthermore, the coupling proposed method can obtain similar performance with the concatenation proposed method, but require a spectral response. In the coupling proposed method, a spectral response is assumed known and provides spectral information between HSI and MSI. In the uncoupling proposed method, the endmember dictionary $\mathbf{U}_m$ is estimated from MSI, which is an ill-posed problem. The number of spectral bands in MSI is usually lower than the number of induced endmembers, which leads to the worst performance.

Table 5. The effectiveness of concatenation in the proposed method. The reported RMSE is calculated in the 8-bit resulting images.

	Coupling SDSR	Uncoupling SDSR	Concatenation SDSR
CAVE	4.1175	10.1147	2.2835
Pavia	2.6015	31.604	2.3488
Paris	8.3705	19.7726	8.0685

5.4.3. Computational Complexity

Based on the multiplicative updates, the overall cost for NMF is $O(tNLp)$. The computational complexity for the proposed method is $O(tN(L+l)p)$, where t is the number of iterations in Algorithm 1. It can be seen that the computational complexity depends linearly on the HSI resolution. This means that the proposed method enhances resolution in an acceptable computational time.

Additionally, the running time is shown in Tables 1–3. In the proposed method, the experiments are implemented using Matlab R2014a, running on 3.4 GHz Intel CPU (Santa Clara, CA, USA) with 64 GB memory. In the experiments, the running time of the proposed method is comparable to the competing methods except for Bicubic. Bicubic interpolation is frequently used for super-resolution

with low computational complexity Both the proposed method and CNMF have the same running time, the running time is almost linear in the image size. Although the proposed method cannot outperform all the competing methods, the proposed method can obtain the high-resolution HSI in acceptable computational time.

5.4.4. Hyperspectral Classification

Hyperspectral classification is further presented on Pavia dataset to verify the effectiveness of the proposed method. The estimated super-resolution HSI is classified with *support vector machine* (SVM). The ground truth used in classification is shown in Figure 11. To obtain a reliable evaluation, we estimate the classification *overall accuracy* (OA) with the SVM using ten-fold cross-validation. The dataset is randomly divided into ten equal parts.

In this experiment, two types of evaluation are presented. (1) The average cross-validation classification accuracies are presented in Table 6; (2) The classification maps are presented in Figure 12. Specifically, *overall accuracy* (OA) is selected as the classification assessment index. It is clearly seen from Table 6 and Figure 12 that the classification results have been improved by several HSI super-resolution methods: Bicubic, HySure, and Proposed SDSR method. Bicubic and HySure obtain the better index than the proposed method. This is because that Bicubic and HySure can obtain a smooth HSI. Bicubic interpolates the data points on a two-dimensional regular grid. Images obtained by Bicubic interpolation are smoother and have fewer interpolation artifacts. In HySure, total variation is used as a regularization. Total variation imposes sparsity in the distribution of the absolute gradient of an image. In this case, the transitions between the pixels should be smooth in the spatial dimension. Compared with Bicbic and HySure, the proposed method gives the compared classification result without smoothing the HSI. The reasons are mainly as follows: (1) The endmember in HSI are learned by performing a common sparse basis selection on the concatenation of observed HSI and MSI, which can preserve the spectral consistency between the observed HSI and MSI; and (2) the spatial consistency between low-resolution HSI and high-resolution HSI is considered.

(a)   **(b)**
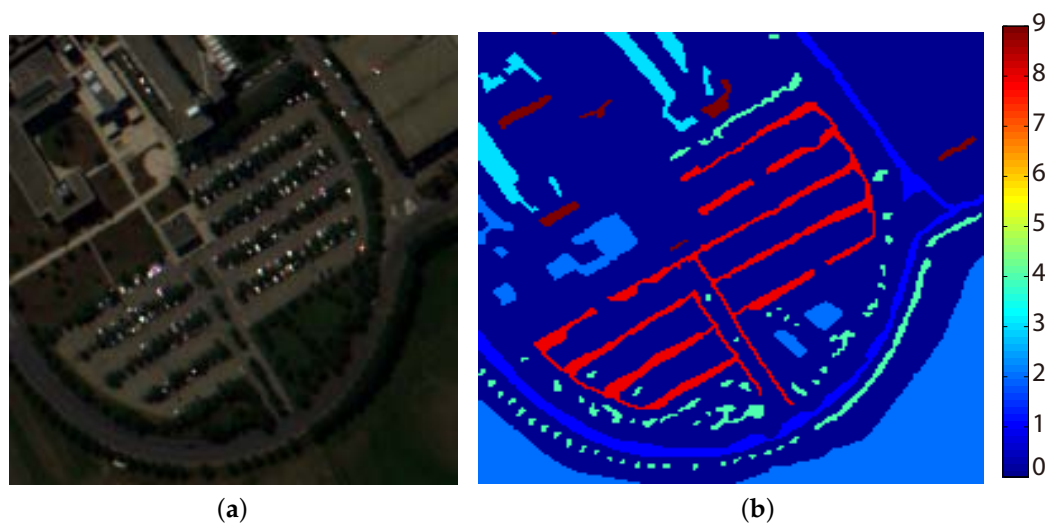
Figure 11. Pavia dataset and its Ground truth in classification. (**a**) high-resolution hyperspectral image; (**b**) ground truth in classification.

Table 6. Classification assessment index on the Pavia dataset.

	Original	Bicubic	CNMF	GSOMP	BSR	HySure	CNNCNMF	PALM	SDSR
OA	95.51%	99.87%	87.80%	90.20%	87.58%	99.52%	87.98%	97.40%	99.11%
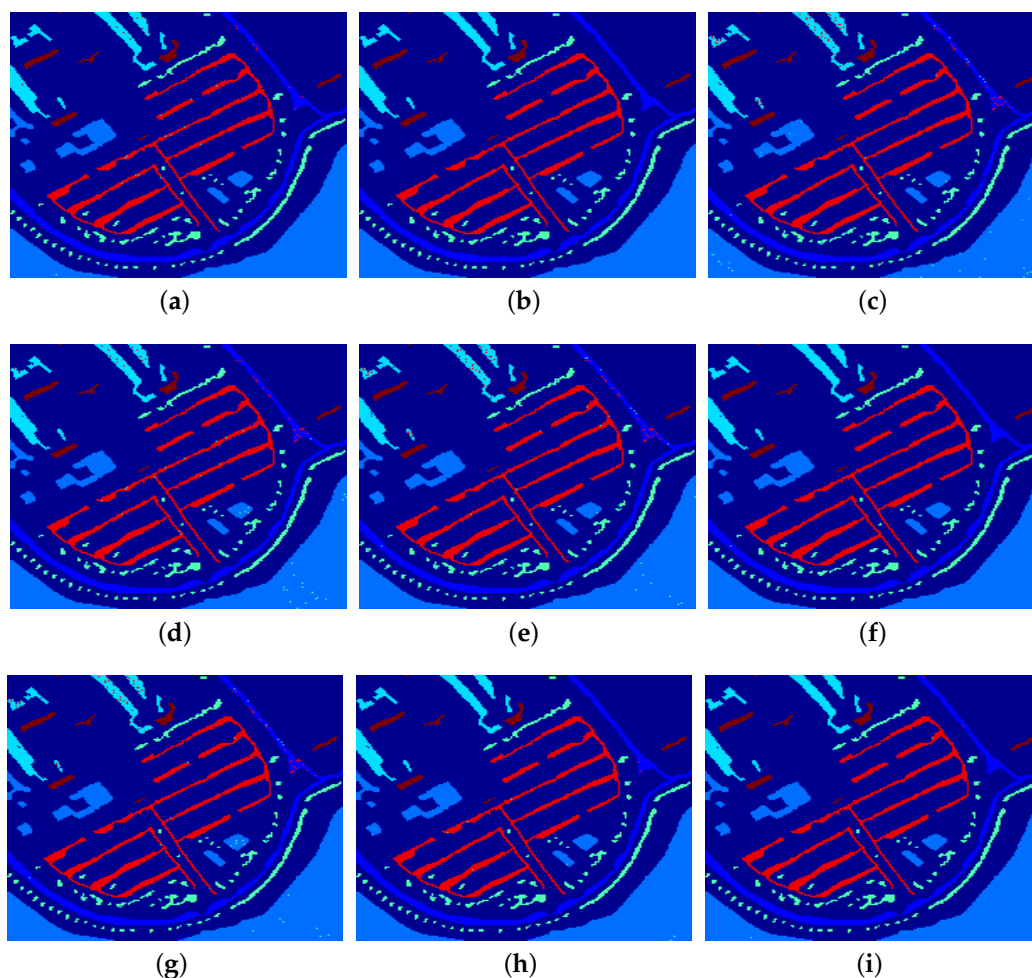
Figure 12. Classification maps of the Pavia dataset (the colors in this figure are consistent with those in Figure 11). (**a**) original (OA = 95.82%); (**b**) bicubic (OA = 100%); (**c**) CNMF (OA = 86.38%); (**d**) GSOMP (OA = 89.88%); (**e**) BSR (OA = 87.74%); (**f**) HySure (OA = 99.42%); (**g**) CNNCNMF (OA = 88.04%); (**h**) PALM (OA = 98.44%); (**i**) SDSR (OA = 99.03%).

6. Conclusions

In this paper, we propose a self-dictionary sparse regression to enhance the spatial-resolution of HSI. In the proposed method, the spatial-spectral consistency is considered to fuse the HSI and MSI. A notable difference between the existing HSI super-resolution and the proposed method is that both observed HSI and MSI are simultaneously considered to estimate the endmember dictionary and the abundance code. Specifically, the endmember dictionary is extracted by performing self-dictionary sparse regression on the concatenation of observed HSI and MSI. Then, a consistent constraint between the low-resolution HSI and the high-resolution HSI is exploited to improve the estimation of the abundance code. Experimental results on three datasets validate that the proposed method outperforms the conventional HSI super-resolution methods.

References

1. Lu, X.; Wang, B.; Zheng, X.; Li, X. Exploring Models and Data for Remote Sensing Image Caption Generation. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 2183–2195. [CrossRef]
2. Yuan, Y.; Zheng, X.; Lu, X. Spectral–Spatial Kernel Regularized for Hyperspectral Image Denoising. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 3815–3832. [CrossRef]
3. Tan, K.; Li, E.; Du, Q.; Du, P. Hyperspectral Image Classification Using Band Selection and Morphological Profiles. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2014**, *7*, 40–48. [CrossRef]
4. Lu, X.; Zheng, X.; Yuan, Y. Remote Sensing Scene Classification by Unsupervised Representation Learning. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 5148–5157. [CrossRef]
5. Yuan, Y.; Zheng, X.; Lu, X. Discovering Diverse Subset for Unsupervised Hyperspectral Band Selection. *IEEE Trans. Image Process.* **2017**, *26*, 51–64. [CrossRef] [PubMed]
6. Bioucas-Dias, J.M.; Plaza, A.; Camps-Valls, G.; Scheunders, P.; Nasrabadi, N.M.; Chanussot, J. Hyperspectral remote sensing data analysis and future challenges. *IEEE Geosci. Remote Sens. Mag.* **2013**, *1*, 6–36. [CrossRef]
7. Lu, X.; Yuan, Y.; Zheng, X. Joint dictionary learning for multispectral change detection. *IEEE Trans. Cybern.* **2017**, *47*, 884–897. [CrossRef] [PubMed]
8. Zheng, X.; Yuan, Y.; Lu, X. A target detection method for hyperspectral image based on mixture noise model. *Neurocomputing* **2016**, *216*, 331–341. [CrossRef]
9. Lu, X.; Wu, H.; Yuan, Y. Double constrained NMF for hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 2746–2758. [CrossRef]
10. Zheng, X.; Yuan, Y.; Lu, X. Single Image Super-Resolution Restoration Algorithm from External Example to Internal Self-Similarity. *Acta Opt. Sin.* **2017**, *37*, 0318006. [CrossRef]
11. Zheng, X.; Yuan, Y.; Lu, X. Dimensionality Reduction by Spatial-Spectral Preservation in Selected Bands. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 5185–5197. [CrossRef]
12. Mianji, F.A.; Zhang, Y.; Gu, Y.; Babakhani, A. Spatial-spectral data fusion for resolution enhancement of hyperspectral imagery. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Cape Town, South Africa, 12–17 July 2009; Volume 3, pp. III–1011.
13. Tang, Y.; Li, X. Set-based similarity learning in subspace for agricultural remote sensing classification. *Neurocomputing* **2016**, *173*, 332–338. [CrossRef]
14. Akgun, T.; Altunbasak, Y.; Mersereau, R.M. Super-resolution reconstruction of hyperspectral images. *IEEE Trans. Image Process.* **2005**, *14*, 1860–1875. [CrossRef] [PubMed]
15. Li, Y.; Hu, J.; Zhao, X.; Xie, W.; Li, J. Hyperspectral image super-resolution using deep convolutional neural network. *Neurocomputing* **2017**, *266*, 29–41. [CrossRef]
16. Tang, Y.; Chen, H.; Liu, Z.; Song, B.; Wang, Q. Example-based super-resolution via social images. *Neurocomputing* **2016**, *172*, 38–47. [CrossRef]
17. Tang, Y.; Shao, L. Pairwise operator learning for patch-based single-image super-resolution. *IEEE Trans. Image Process.* **2017**, *26*, 994–1003. [CrossRef] [PubMed]
18. Zhao, Y.; Yang, J.; Zhang, Q.; Song, L.; Cheng, Y.; Pan, Q. Hyperspectral imagery super-resolution by sparse representation and spectral regularization. *Eurasip J. Adv. Signal Process.* **2011**, *2011*, 87. [CrossRef]
19. Zhao, Y.; Yang, J.; Chan, J.C.W. Hyperspectral imagery super-resolution by spatial–spectral joint nonlocal similarity. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2014**, *7*, 2671–2679. [CrossRef]
20. Patel, R.C.; Joshi, M.V. Super-resolution of hyperspectral images: Use of optimum wavelet filter coefficients and sparsity regularization. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 1728–1736. [CrossRef]
21. Mei, S.; Yuan, X.; Ji, J.; Zhang, Y.; Wan, S.; Du, Q. Hyperspectral Image Spatial Super-Resolution via 3D Full Convolutional Neural Network. *Remote Sens.* **2017**, *9*, 1139. [CrossRef]
22. Hu, J.; Li, Y.; Xie, W. Hyperspectral Image Super-Resolution by Spectral Difference Learning and Spatial Error Correction. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1825–1829. [CrossRef]
23. Loncan, L.; de Almeida, L.B.; Bioucas-Dias, J.M.; Briottet, X.; Chanussot, J.; Dobigeon, N.; Fabre, S.; Liao, W.; Licciardi, G.A.; Simoes, M.; et al. Hyperspectral pansharpening: A review. *IEEE Geosci. Remote Sens. Mag.* **2015**, *3*, 27–46. [CrossRef]
24. Yokoya, N.; Grohnfeldt, C.; Chanussot, J. Hyperspectral and Multispectral Data Fusion: A comparative review of the recent literature. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 29–56. [CrossRef]

25. Li, J.; Yuan, Q.; Shen, H.; Meng, X.; Zhang, L. Hyperspectral image super-resolution by spectral mixture analysis and spatial–spectral group sparsity. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 1250–1254. [CrossRef]
26. Simões, M.; Bioucas-Dias, J.; Almeida, L.B.; Chanussot, J. A convex formulation for hyperspectral image superresolution via subspace-based regularization. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 3373–3388. [CrossRef]
27. Veganzones, M.A.; Simoes, M.; Licciardi, G.; Yokoya, N.; Bioucas-Dias, J.M.; Chanussot, J. Hyperspectral super-resolution of locally low rank images from complementary multisource data. *IEEE Trans. Image Process.* **2016**, *25*, 274–288. [CrossRef] [PubMed]
28. Yuan, Y.; Zheng, X.; Lu, X. Hyperspectral Image Superresolution by Transfer Learning. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2017**, *10*, 1963–1974. [CrossRef]
29. Zou, C.; Xia, Y. Hyperspectral Image Superresolution Based on Double Regularization Unmixing. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1022–1026. [CrossRef]
30. Dong, W.; Fu, F.; Shi, G.; Cao, X.; Wu, J.; Li, G.; Li, X. Hyperspectral Image Super-Resolution via Non-Negative Structured Sparse Representation. *IEEE Trans. Image Process.* **2016**, *25*, 2337–2352. [CrossRef] [PubMed]
31. Yokoya, N.; Yairi, T.; Iwasaki, A. Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 528–537. [CrossRef]
32. Wei, Q.; Bioucas-Dias, J.; Dobigeon, N.; Tourneret, J.; Chen, M.; Godsill, S. Multiband Image Fusion Based on Spectral Unmixing. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 7236–7249. [CrossRef]
33. Ghasrodashti, E.; Karami, A.; Heylen, R.; Scheunders, P. Spatial Resolution Enhancement of Hyperspectral Images Using Spectral Unmixing and Bayesian Sparse Representation. *Remote Sens.* **2017**, *9*, 541. [CrossRef]
34. Guerra, R.; Lopez, S.; Sarmiento, R. A Computationally Efficient Algorithm for Fusing Multispectral and Hyperspectral Images. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 5712–5728. [CrossRef]
35. 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]
36. Vivone, G.; Restaino, R.; Licciardi, G.; Mura, M.D.; Chanussot, J. MultiResolution Analysis and Component Substitution techniques for hyperspectral Pansharpening. In Proceedings of the IEEE Geoscience and Remote Sensing Symposium, Quebec City, QC, Canada, 13–18 July 2014; pp. 2649–2652.
37. Akhtar, N.; Shafait, F.; Mian, A. Sparse spatio-spectral representation for hyperspectral image super-resolution. In Proceedings of the 13th European Conference on Computer Vision (ECCV 2014), Zurich, Switzerland, 6–12 September 2014; pp. 63–78.
38. Lanaras, C.; Baltsavias, E.; Schindler, K. Hyperspectral Super-Resolution with Spectral Unmixing Constraints. *Remote Sens.* **2017**, *9*, 1196. [CrossRef]
39. Huang, W.; Xiao, L.; Liu, H.; Wei, Z. Hyperspectral Imagery Super-Resolution by Compressive Sensing Inspired Dictionary Learning and Spatial-Spectral Regularization. *Sensors* **2015**, *15*, 2041–2058. [CrossRef] [PubMed]
40. Huang, B.; Song, H.; Cui, H.; Peng, J.; Xu, Z. Spatial and spectral image fusion using sparse matrix factorization. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 1693–1704. [CrossRef]
41. Fu, X.; Ma, W.K.; Chan, T.H.; Bioucas-Dias, J.M. Self-dictionary sparse regression for hyperspectral unmixing: Greedy pursuit and pure pixel search are related. *IEEE J. Sel. Top. Signal Process.* **2015**, *9*, 1128–1141. [CrossRef]
42. Heinz, D.C. Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 529–545. [CrossRef]
43. Esser, E.; Moller, M.; Osher, S.; Sapiro, G.; Xin, J. A convex model for nonnegative matrix factorization and dimensionality reduction on physical space. *IEEE Trans. Image Process.* **2012**, *21*, 3239–3252. [CrossRef] [PubMed]
44. Elhamifar, E.; Sapiro, G.; Vidal, R. See all by looking at a few: Sparse modeling for finding representative objects. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2012), Providence, RI, USA, 16–21 June 2012; pp. 1600–1607.
45. Chan, T.H.; Ma, W.K.; Ambikapathi, A.; Chi, C.Y. A simplex volume maximization framework for hyperspectral endmember extraction. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 4177–4193. [CrossRef]
46. Chen, J.; Huo, X. Theoretical results on sparse representations of multiple-measurement vectors. *IEEE Trans. Signal Process.* **2006**, *54*, 4634–4643. [CrossRef]

47. Ma, W.K.; Bioucas-Dias, J.M.; Chan, T.H.; Gillis, N.; Gader, P.; Plaza, A.J.; Ambikapathi, A.; Chi, C.Y. A signal processing perspective on hyperspectral unmixing: Insights from remote sensing. *IEEE Signal Process. Mag.* **2014**, *31*, 67–81. [CrossRef]

48. Bioucas-Dias, J.M.; Plaza, A.; Dobigeon, N.; Parente, M.; Du, Q.; Gader, P.; Chanussot, J. Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2012**, *5*, 354–379. [CrossRef]

49. Lanaras, C.; Baltsavias, E.; Schindler, K. Hyperspectral Super-Resolution by Coupled Spectral Unmixing. In Proceedings of the IEEE International Conference on Computer Vision (ICCV 2015), Santiago, Chile, 12–16 December 2015; pp. 3586–3594.

50. Akhtar, N.; Shafait, F.; Mian, A. Bayesian sparse representation for hyperspectral image super resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2015), Boston, MA, USA. 7–12 June 2015; pp. 3631–3640.