

Article

No-Reference Hyperspectral Image Quality Assessment via Quality-Sensitive Features Learning

Jingxiang Yang ^{1,2}, Yong-Qiang Zhao ^{1,*}, Chen Yi ¹ and Jonathan Cheung-Wai Chan ²

¹ School of Automation, Northwestern Polytechnical University, Xi'an 710072, China; yang123jx@mail.nwpu.edu.cn (J.Y.); chenyi0818@mail.nwpu.edu.cn (C.Y.)

² Department of Electronics and Informatics, Vrije Universiteit Brussel, Brussel 1050, Belgium; jcheungw@etrovub.be

* Correspondence: zhaoyq@nwpu.edu.cn

Academic Editors: Naoto Yokoya, Zhaoliang Li, Magaly Koch and Prasad S. Thenkabil

Received: 17 January 2017; Accepted: 20 March 2017; Published: 23 March 2017

Abstract: Assessing the quality of a reconstructed hyperspectral image (HSI) is of significance for restoration and super-resolution. Current image quality assessment methods such as peak signal-noise-ratio require the availability of pristine reference image, which is often not available in reality. In this paper, we propose a no-reference hyperspectral image quality assessment method based on quality-sensitive features extraction. Difference of statistical properties between pristine and distorted HSIs is analyzed in both spectral and spatial domains, then multiple statistics features that are sensitive to image quality are extracted. By combining all these statistics features, we learn a multivariate Gaussian (MVG) model as benchmark from the pristine hyperspectral datasets. In order to assess the quality of a reconstructed HSI, we partition it into different local blocks and fit a MVG model on each block. A modified Bhattacharyya distance between the MVG model of each reconstructed HSI block and the benchmark MVG model is computed to measure the quality. The final quality score is obtained by average pooling over all the blocks. We assess five state-of-the-art super-resolution methods on Airborne Visible Infrared Imaging Spectrometer (AVIRIS) and Hyperspec-VNIR-C (HyperspecVC) data using our proposed method. It is verified that the proposed quality score is consistent with current reference-based assessment indices, which demonstrates the effectiveness and potential of the proposed no-reference image quality assessment method.

Keywords: image assessment; no reference; hyperspectral; statistics feature; multivariate Gaussian

1. Introduction

Hyperspectral image (HSI) with rich spatial and spectral information of the scene is useful in many fields such as mineral exploitation, agriculture, and environment management [1–3]. To improve the quality of the acquired HSI due to limited spatial resolution, super-resolution is an important enhancement technique [4–12]. In order to evaluate the reconstructed high resolution HSI, conventional strategy is to degrade the original data into a coarser resolution by down-sampling. Then, the original data are used as reference image and compared with the reconstructed high resolution image. The disadvantage is that as the invariance of the super-resolution performance to scale changes cannot be guaranteed, the performance of super-resolution method on the original data may not be as good as on the down-sampled data [13,14]. While it is naturally better to assess the super-resolution method on the original data rather than on the down-sampled data, reference image is not available for assessment if the super-resolution is applied on the original data.

To our knowledge, there is no published work on no-reference quality assessment for the reconstruction of HSI. Alparone et al. proposed a no-reference pansharpening assessment method

in [14] where high resolution panchromatic image was needed to assess the reconstructed multispectral image. This method is not applicable to cases where the panchromatic image is not available. There are some other no-reference image assessment methods designed for color images [15–17], but they cannot be applied to hyperspectral image directly. These methods can only assess spatial quality and give quality scores which reflect human’s subjective visual sense. Furthermore, they cannot deal with spectral fidelity, which is important for the interpretation of HSI.

In this study, we propose a no-reference quality assessment method for HSI. HSI possesses some statistical properties that are sensitive to distortion, deviations of these statistics from their regular counterparts reflect the extent of distortion. These statistics can be extracted as quality-sensitive features. By analyzing the statistical properties of the pristine and distorted HSIs, we extract multiple quality-sensitive features in both spectral and spatial domains. After integrating all these features, we can learn a multivariate Gaussian model (MVG) of these features from the pristine hyperspectral training dataset. The learned MVG is treated as a benchmark to compare with the MVG model fitted on the reconstructed HSI. Distance between the two MVG models is computed as quality measure with high value representing low quality. To apply this method, we partition the reconstructed HSI into different local blocks, and measure the image quality for each of the local blocks. The final quality score of the reconstructed HSI is obtained by average pooling.

We consider four contributions in this paper. Firstly, we propose the first no-reference assessment method for hyperspectral image. Our method does not require any reference image or down-sampling the original image, which is well-suited for practical applications. Secondly, in order to exploit both the spectral and spatial information for quality assessment, other than the off-the-shelf spatial features, we analyze the statistical properties in the spectral domain, design quality-sensitive features for the spectral domain, and integrate them with the spatial features to form a joint spectral-spatial quality-sensitive feature vector. Thirdly, compared with current color image assessment methods, our method can also blindly assess the spectral fidelity. Finally, we verify the potential of our method as a HSI assessment tool by testing it on several real HSIs, which are reconstructed by state-of-the-art super-resolution methods.

The remainder of this paper is organized as follows. In Section 2, we analyze the statistical properties of HSI, and extract quality-sensitive features. The methodology of computing the quality score is given in Section 3. We present the experimental results and give discussions about the experimental results in Sections 4 and 5, respectively. We make the conclusions in Section 6.

2. Quality-Sensitive Statistics Features

An image possesses statistics that would deviate from their regular counterparts due to distortion, extracting these statistics as features and measuring their deviations makes it possible to assess HSI without any reference [17]. Previous quality-sensitive statistics features designed for color images mainly focus on the spatial domain [18–22]. In order to exploit the spectral correlation of a HSI, we also need to extract quality-sensitive features from the spectral domain. In this section, we firstly analyze the statistical properties in the spectral domain and design a quality-sensitive spectral feature extraction method. Then, we demonstrate that off-the-shelf spatial features are effective for HSI. By integrating our proposed spectral features and the spatial features, we form a joint spectral-spatial quality-sensitive feature vector.

2.1. Statistics Features in Spectral Domain

In this sub-section, spectral quality-sensitive features are proposed after analyzing the statistics in the spectral domain. We observe that locally normalized spectra of a pristine HSI would follow a Gaussian distribution, while those of distorted HSIs would deviate. Given a pristine HSI $\mathbf{I} \in \mathbb{R}^{M \times N \times L}$, we first apply local normalization to a spectrum s

$$\bar{s}(\lambda) = \frac{s(\lambda) - \mu(\lambda)}{\sigma(\lambda) + C}, \quad (1)$$

where $\lambda = 1, 2, \dots, L$ is the spectral coordinate, and C is a constant to stabilize the normalization when the denominator tends to zero. In our experiment, C is set to 1. $\mu(\lambda)$ and $\sigma(\lambda)$ are, respectively, local mean and standard variance:

$$\mu(\lambda) = \sum_{k=-K}^K w_k s(\lambda + k), \quad (2)$$

$$\sigma(\lambda) = \sqrt{\sum_{k=-K}^K w_k [s(\lambda + k) - \mu(\lambda)]^2}, \quad (3)$$

where $w = \{w_k | k = -K, -K + 1, \dots, K\}$ is a Gaussian weighting window. K determines the width of the window. The local normalization removes the local mean displacements and normalize the local variance, thus has a decorrelation effect. The locally normalized spectrum would be more homogeneous than the original spectrum. After the local normalization, the spectra of a pristine HSI would approximately have zero mean and unit variance.

We crop a sub-image from AVIRIS data [23], and then apply the above local normalization to the spectrum, as shown in Figure 1. Noise and blurring are common effects caused by distortion in HSI [24–26], so we add noise to the pristine HSI or blur it to simulate distorted HSIs. Figure 2 shows the sub-images added with different level of noise (Gaussian noise) and blurring (average filtering). We plot the histograms of all the spectra in the sub-image in Figure 3. It is observed that distribution of the locally normalized spectra of a pristine HSI follows a Gaussian distribution with zero mean, while the locally normalized spectra of distorted HSIs deviate. There are two interesting findings in Figure 3. Firstly, each type of distortion modifies the distribution in its own way. For example, with noise added, the distribution curve becomes flat and tends to be a uniform distribution. When the HSI is blurred, the distribution curve becomes thin and tends to be a Laplacian distribution. Secondly, heavier distortion causes greater modification of the distribution. Noise with standard variance $\sigma = 0.20$ makes the distribution curve much flatter than noise with $\sigma = 0.05$, and a 5×5 blurring kernel generates a narrower bell-shaped curve than a 3×3 kernel.

Therefore, some statistical properties in the spectral domain can be modified by the distortion, and measuring the changes of these statistics makes it possible to assess the spectral distortion. Generalized Gaussian distribution (GGD) can be used to capture the statistical changes between the pristine and distorted HSIs. The function of a GGD with zero mean is

$$f(x; \alpha, \beta, \sigma^2) = \frac{\alpha}{2\beta\Gamma(1/\alpha)} \exp\left[-\left(\frac{|x|}{\beta}\right)^\alpha\right], \quad (4)$$

where

$$\beta = \sigma \sqrt{\frac{\Gamma(1/\alpha)}{\Gamma(3/\alpha)}}, \quad (5)$$

$$\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt \quad a > 0, \quad (6)$$

where α and β represent shape parameter and scale parameter, respectively. σ is standard variance. The GGD model can describe broadly the statistics of multiple distributions. The GGD model reduces to a Laplacian distribution and a Gaussian distribution when $\alpha = 1$ and $\alpha = 2$. It tends to a uniform distribution when α approaches infinity. When distortion is introduced, the locally normalized spectra would deviate from Gaussian distribution and tend to a uniform-like or a Laplacian-like distribution, all of them can be captured by a GGD model. The statistics of a GGD model is described by its model parameters, so we select the parameters $[\alpha, \beta]$ as the spectral quality-sensitive features, which can be estimated using moment-matching algorithm [17].

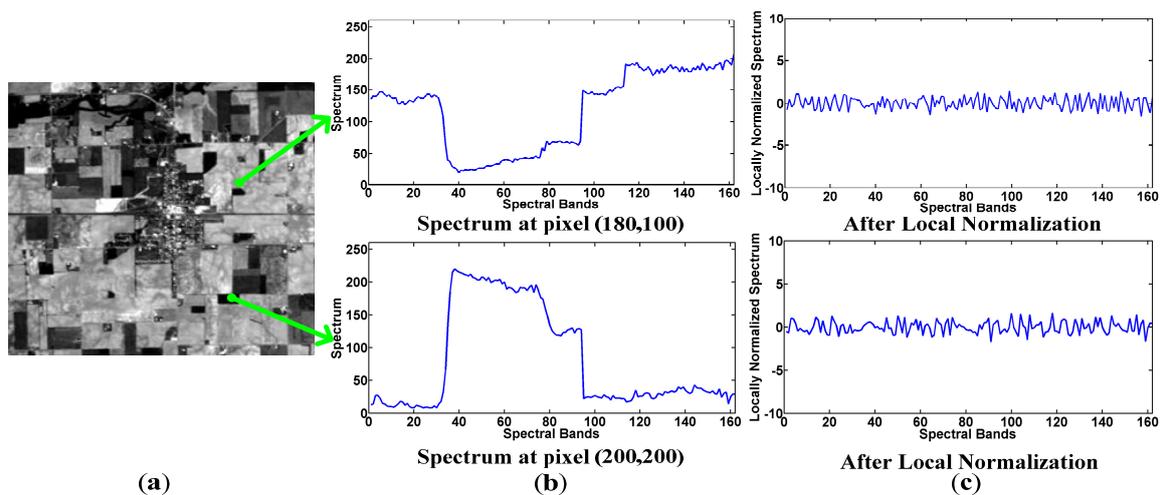


Figure 1. Illustration of the local normalization on spectrum: (a) the 20th band (589.31 nm) of a pristine sub-image cropped from AVIRIS data, the size is 256×256 ; (b) spectra curves selected from two pixels; and (c) the locally normalized spectra.

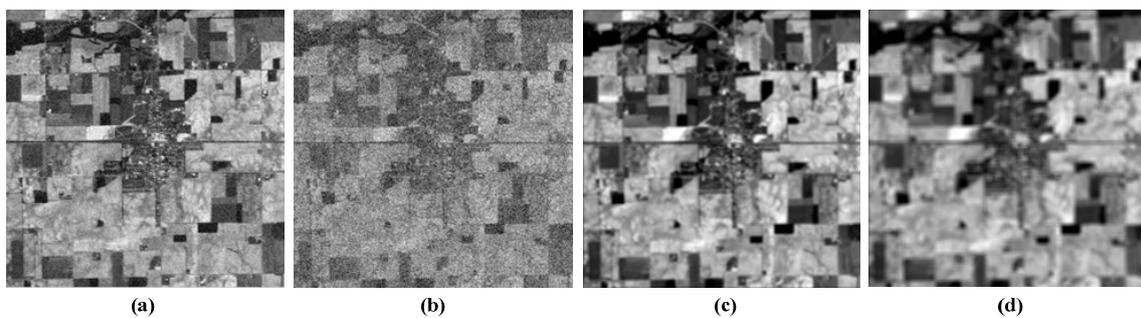


Figure 2. Distorted versions of the sub-image: (a) distorted by Gaussian noise with standard variance $\sigma = 0.05$; (b) distorted by Gaussian noise with standard variance $\sigma = 0.20$; (c) distorted by blurring with 3×3 average filtering kernel; and (d) distorted by blurring with 5×5 average filtering kernel.

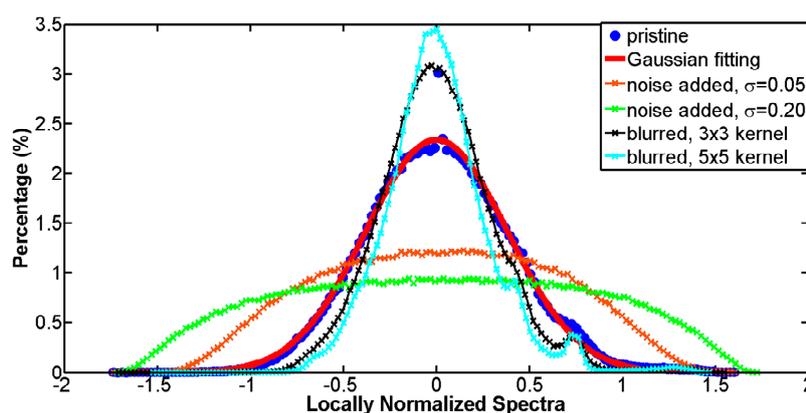


Figure 3. Histograms of locally normalized spectra of pristine hyperspectral image (HSI) and distorted HSIs.

To show that our extracted feature is sensitive to image quality, we randomly crop 200 pristine sub-images of size $64 \times 64 \times 224$ from AVIRIS dataset [23]. Then, we introduce different types of distortions to each sub-image. After applying the local normalization on the spectra of each sub-image, we fit the histogram of spectra in each sub-image with GGD model. The extracted features $[\alpha, \beta]$ are

plotted in Figure 4. As shown in Figure 4, features belonging to the same distortion form a cluster. It is easy to separate different distortions in the feature space, which demonstrates the sensitivity of the extracted feature to image quality.

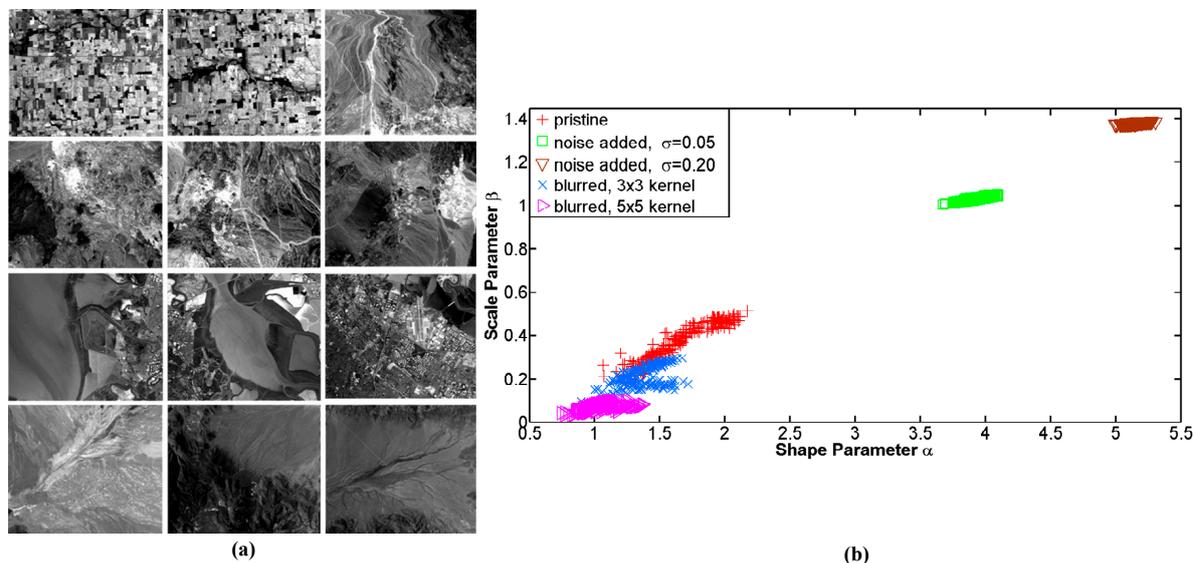


Figure 4. (a) The AVIRIS data of different scenes, 200 sub-images are randomly cropped from them; and (b) spectral quality-sensitive features visualization. Each point represents feature of a sub-image, each color represents a type of distortion.

2.2. Statistics Features in Spatial Domain

Image quality distortion can be reflected in local image structures [17], image gradient [19], and multi-scale and multi-orientation decomposition [20]. To exploit these information, we adopt multiple types of spatial features (originally proposed for color images [27]) and verify their effectiveness on HSI in this sub-section.

2.2.1. Statistics of Panchromatic Image

A hyperspectral image often contains large number of continuous spectral bands with narrow bandwidth, extracting the spatial features band-by-band would be time-consuming and result in huge number of redundant features. In order to extract features from the spatial domain in a fast and simple way, we analyze the statistics and extract the spatial features on a synthesized panchromatic image, which is simulated by [27]

$$P = w_r I_r + w_g I_g + w_b I_b, \quad (7)$$

where I_r , I_g , and I_b are spectral bands of the HSI with band centers corresponding to the red, green, and blue bands. In the experiment, the weights w_r , w_g , and w_b are set to 0.06, 0.63, and 0.27 as suggested in [27]. The simulated panchromatic image of the HSI in Figure 1 is shown in Figure 5. The structural and textural information contained in the panchromatic image would be exploited in extracting the spatial quality-sensitive features. Similar to the spectral domain, we apply the local normalization to the simulated panchromatic image

$$\bar{P}(i, j) = \frac{P(i, j) - \mu(i, j)}{\sigma(i, j) + C}, \quad (8)$$

where i and j are the spatial coordinates, and $\mu(i, j)$ and $\sigma(i, j)$ are local mean and standard variance, respectively, computed by [17]

$$\mu(i, j) = \sum_{s=-S}^S \sum_{t=-T}^T w_{s,t} P(i + s, j + t), \tag{9}$$

$$\sigma(i, j) = \sqrt{\sum_{s=-S}^S \sum_{t=-T}^T w_{s,t} [P(i + s, j + t) - \mu(i, j)]^2}, \tag{10}$$

where $w = \{w_{s,t} | s = -S, \dots, S, t = -T, \dots, T\}$ is the Gaussian weighting window, and the window size is determined by S and T . After local normalization, the value of most pixels would be decorrelated and close to zero, the locally normalized result exhibits a homogeneous appearance with a few residual edges, as shown in Figure 6a. In Figure 6b, we present the histograms of the locally normalized panchromatic images simulated from pristine and distorted HSIs. It has been observed that the locally normalized panchromatic image follows a Gaussian distribution with zero mean, while it deviates when distortion exists [17,27]. The pattern of curves in Figure 6b is similar to Figure 3, and the statistics of the panchromatic image is modified by distortions in a similar way as what has been discovered in the spectral domain. We also use the GGD model to measure the difference of statistics between the pristine and distorted HSIs. The shape parameter and scale parameter of GGD model are used as spatial quality-sensitive features.

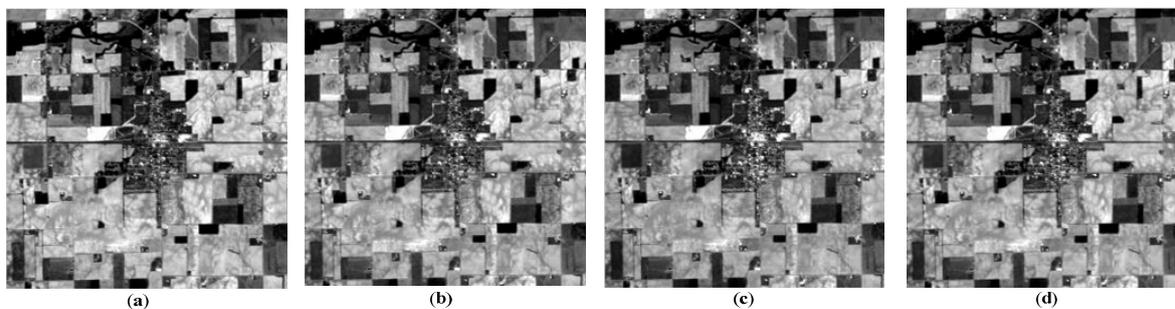


Figure 5. Simulated panchromatic image and the original images corresponding to red, green, and blue bands: (a) red band (665.59 nm); (b) green band (589.31 nm); (c) blue band (491.90 nm); and (d) the simulated panchromatic image.

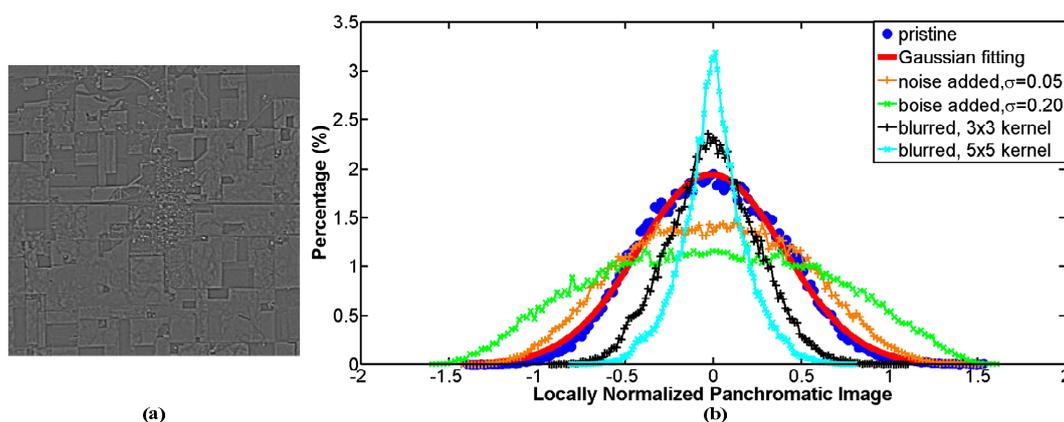


Figure 6. (a) The local normalization of pristine panchromatic image in Figure 5; and (b) histograms of locally normalized panchromatic images, under different kind of distortions.

2.2.2. Statistics of Texture

The quality of image can also be revealed by the quality of texture which should be exploited for the quality assessment. Log-Gabor filters decompose an image in multi-scales and multi-orientations,

thus can capture textual information. The textures of a HSI are captured in the panchromatic image, so we apply Log-Gabor filters to the simulated panchromatic image. The Log-Gabor filter is expressed as [27]

$$G(\omega, \theta) = \exp\left(-\frac{(\log(\omega/\omega_0))^2}{2\sigma_r^2}\right) \cdot \exp\left(-\frac{(\theta - \theta_j)^2}{2\sigma_\theta^2}\right), \quad (11)$$

where $\theta_j = j\pi/J, j = \{0, 1, \dots, J-1\}$ is orientation; J is the number of orientations; ω_0 is center frequency; and σ_r and σ_θ determine radial bandwidth and angular bandwidth of the filter, respectively. Applying Log-Gabor filters with N center frequencies and J orientations to the simulated panchromatic image would generate $2NJ$ response maps $\{(e_{n,j}, o_{n,j}) | n = 0, \dots, N-1, j = 0, \dots, J-1\}$, where $e_{n,j}$ and $o_{n,j}$ represent the real part and the imaginary part of the response, respectively.

In Figure 7a, we present a response map $o_{1,3}$ ($N = 3, J = 4$) as an example. It is shown that texture and edges of the panchromatic image are extracted by the Log-Gabor filter. In order to analyze the statistical difference of the Log-Gabor filtering response between the pristine and the distorted HSIs, we take the response map $o_{1,3}$ as an example and plot the histograms of $o_{1,3}$ under different distortions in Figure 7b. It is clear that different distortions lead to different distributions of the Log-Gabor filtering response, the distribution of Log-Gabor response can be used as an indicator of distortion. We also use the GGD model to describe the distribution of the Log-Gabor response $e_{n,j}$ and $o_{n,j}$, the shape parameter and scale parameter of the fitted GGD model form another type of spatial quality-sensitive features.

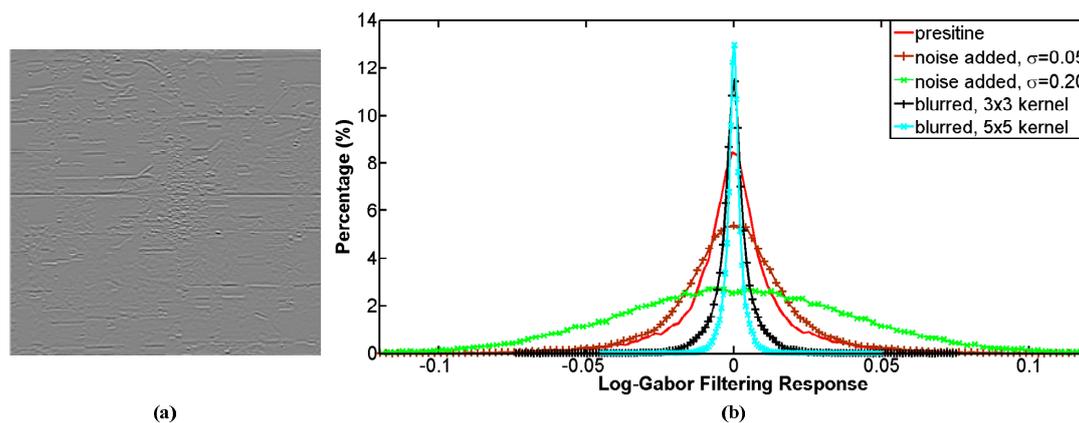


Figure 7. (a) Log-Gabor filtering response map $o_{1,3}$ of the pristine panchromatic image in Figure 5; and (b) histograms of Log-Gabor filtering response map $o_{1,3}$, under different kind of distortions.

In order to further exploit the texture information, we also analyze the statistics of directional gradient of the Log-Gabor filtering response map. The vertical gradient of $o_{1,3}$ is shown in Figure 8a. The histograms of the vertical gradient of $o_{1,3}$ under different distortions are given in Figure 8b. The distribution of the directional gradient is modified by distortion in a similar way to the Log-Gabor response map, therefore GGD model is used to describe the distribution of directional gradients (both horizontal and vertical) of $e_{n,j}$ and $o_{n,j}$ [19,27]; the shape parameter and scale parameter of the fitted GGD model are another spatial quality-sensitive features.

In addition to directional gradient, gradient magnitude of the Log-Gabor filtering response map is also analyzed. The gradient magnitude of $o_{1,3}$ is shown in Figure 9a. The histograms of gradient magnitude of $o_{1,3}$ under different distortions are presented in Figure 9b. The histogram follows Weibull distribution [27,28]

$$f(x; \lambda, k) = \begin{cases} \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} \exp\left(-\left(\frac{x}{\lambda}\right)^k\right) & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (12)$$

where λ is the scale parameter and k is the shape parameter of Weibull model. Since the distribution of the gradient magnitude can be fitted by the Weibull model, alterations of the Weibull model can be used as an indicator for the degree of distortion. Thus, the parameters λ and k of the fitted Weibull model can be used as quality-sensitive features.

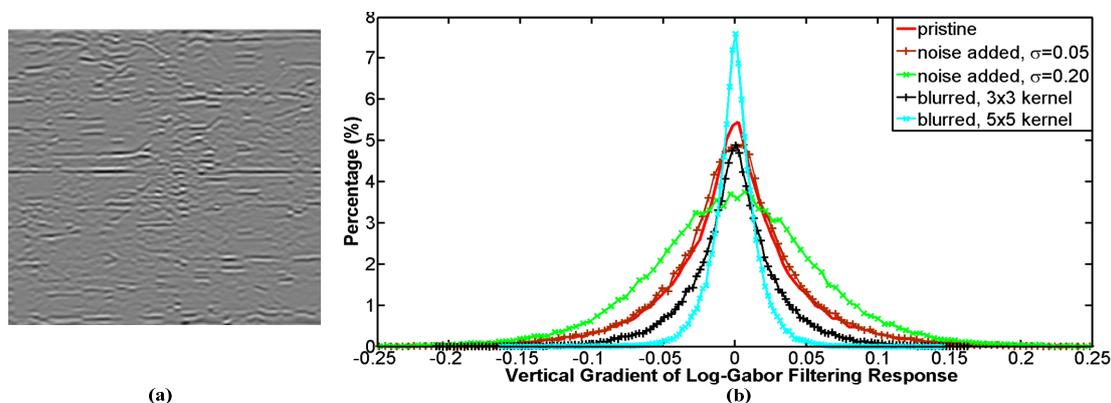


Figure 8. (a) Vertical Gradient of Log-Gabor response map $o_{1,3}$ of the pristine panchromatic image in Figure 5; and (b) histograms of Log-Gabor filtering response map $o_{1,3}$, under different kind of distortions.

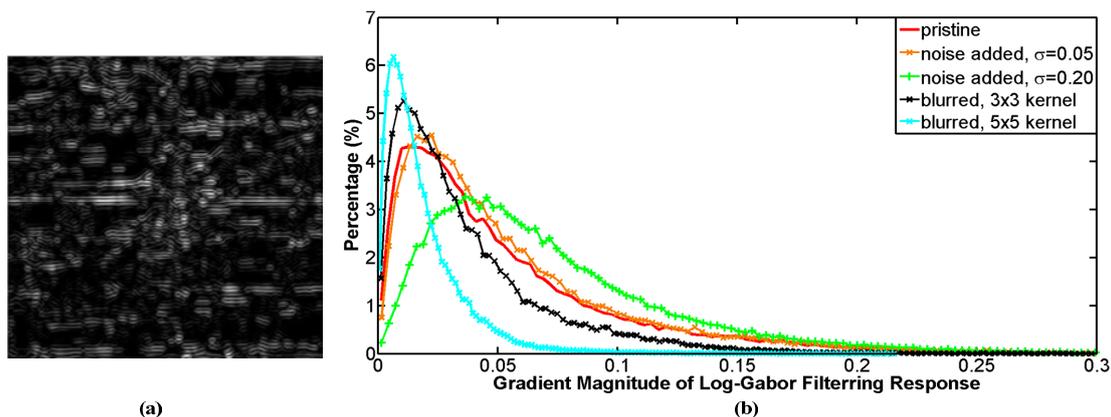


Figure 9. (a) Gradient magnitude of Log-Gabor response map $o_{1,3}$ of the pristine panchromatic image in Figure 5; and (b) histograms of gradient magnitude of $o_{1,3}$, under different kind of distortions.

To demonstrate that the extracted features above are sensitive to image quality, we visualize the features as in Figure 4. We randomly crop 200 pristine sub-images of size $64 \times 64 \times 224$ from the AVIRIS dataset, and introduce different kinds of distortions to each sub-image. We apply Log-Gabor filters on each sub-image, then we fit the histograms of $o_{1,3}$ and the vertical gradient of $o_{1,3}$ with GGD model. We fit the histogram of gradient magnitude of $o_{1,3}$ with Weibull model. The parameters of the fitted model are used as features. Feature of each sub-image is plotted as a point. As shown in Figures 10–12, even though there is some overlapping between different kinds of distortions, most features belonging to the same distortion tend to group into the same cluster. Different distortions occupy different regions in the feature space, which demonstrates the sensitivity of the extracted feature to image quality.

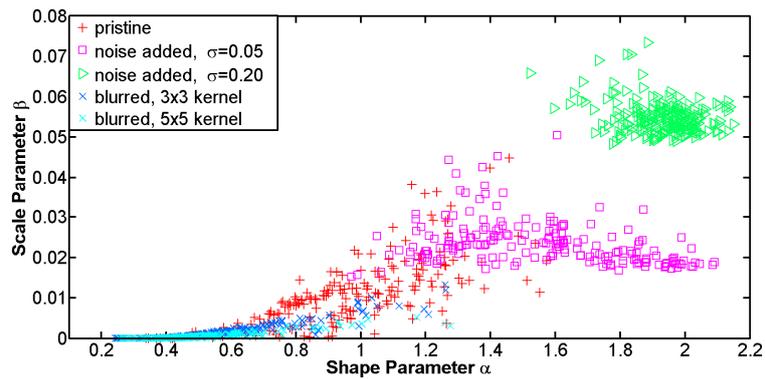


Figure 10. Visualization of spatial quality-sensitive features extracted from Log-Gabor response map $o_{1,3}$. Each point represents feature of a sub-image, each color represents a type of distortion.

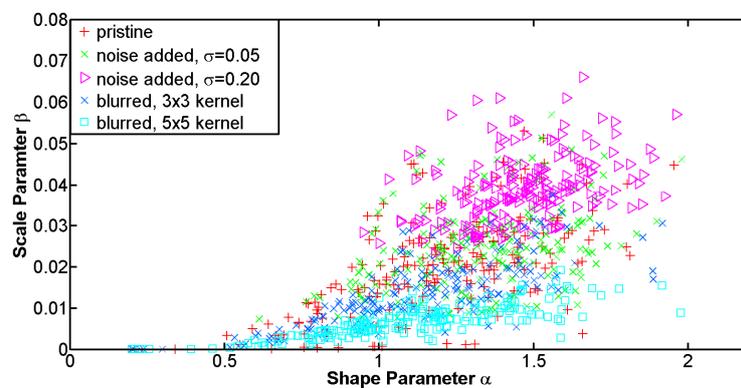


Figure 11. Visualization of spatial quality-sensitive features extracted from vertical gradient of Log-Gabor response map $o_{1,3}$. Each point represents feature of a sub-image, each color represents a type of distortion.

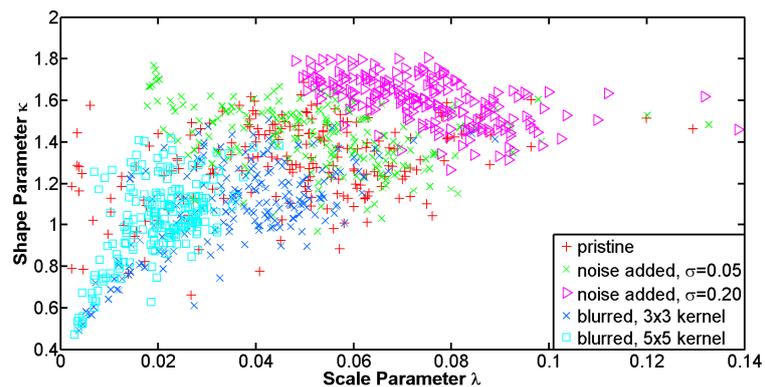


Figure 12. Visualization of spatial quality-sensitive features extracted from gradient magnitude of Log-Gabor response map $o_{1,3}$. Each point represents feature of a sub-image, each color represents a type of distortion.

In order to extract joint features that contain both structural and spectral information, we need to integrate the spatial features with the proposed spectral features. All the features extracted in the spatial domain are stacked, then they are concatenated with the spectral features, a joint spectral-spatial feature vector that is sensitive to image quality can be obtained, as shown in Figure 13.

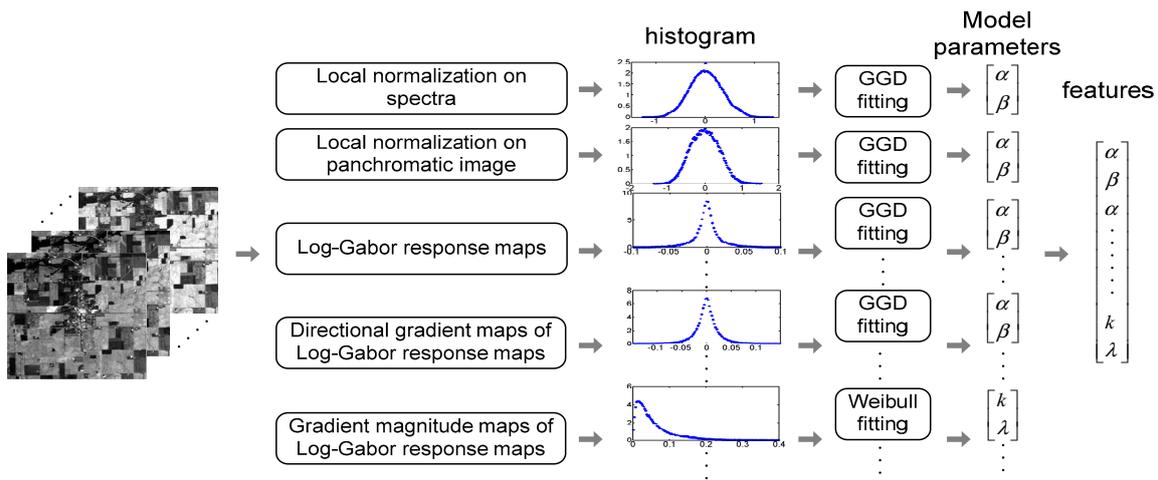


Figure 13. Flow chart of quality-sensitive features extraction for each HSI.

3. Quality Assessment: From Features to Score

If we can extract the spectral-spatial features from the pristine HSI training set and distorted HSIs using the method in Section 2, the distortion of the HSI could be quantified by computing the distance of the quality-sensitive features between the training set and the distorted HSI. In this work, we adopt the strategy of multivariate Gaussian (MVG) learning originally proposed in [18], the flow chart is in Figure 14. In the training stage, there are three main steps: collecting training hyperspectral data, extracting quality-sensitive features, and learning MVG distribution.

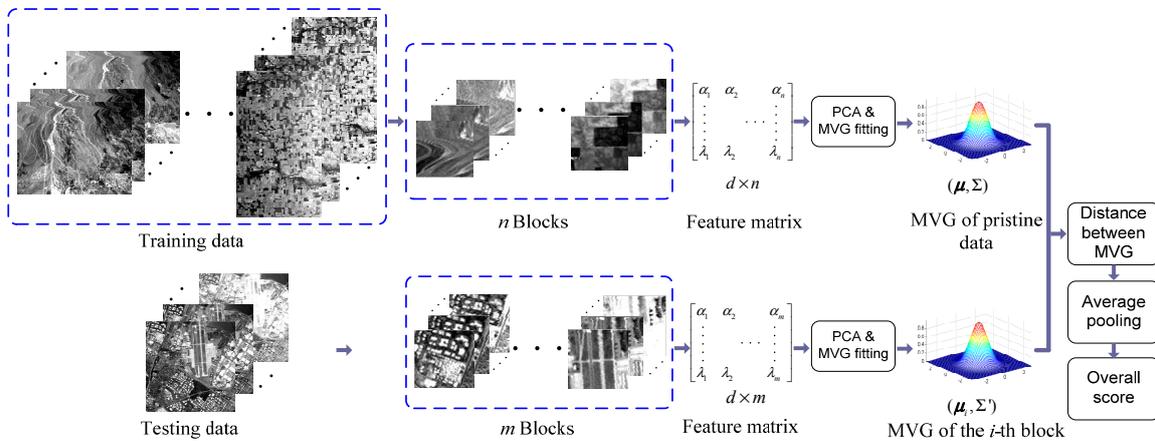


Figure 14. Flow chart of the proposed HSI assessment method.

A set of pristine HSI is firstly collected as training set. Noisy bands and water absorption bands are removed. Different local image regions contain different structures, and have different contributions to the overall image quality [18,27]. In order to exploit the local structural information of the image, we divide the HSI into non-overlapping local 3D blocks. Quality-sensitive features are extracted from each block. By stacking all the spectral and spatial quality-sensitive features, a feature vector $x \in \mathbb{R}^{d \times 1}$ would be extracted from each block. Suppose there are n blocks in the training set in total, a feature matrix $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{d \times n}$ would be obtained from the training set.

There are correlations among different kinds of features; for example, directional gradient and gradient magnitude are highly correlated. In order to remove the correlation and reduce the computation burden, PCA transform is applied to the feature matrix X , a projection matrix Φ and a dimension-reduced feature matrix can be obtained

$$\mathbf{X}' = \Phi \mathbf{X}, \quad (13)$$

where $\mathbf{X}' = [\mathbf{x}'_1, \mathbf{x}'_2, \dots, \mathbf{x}'_n] \in \mathbb{R}^{d' \times n}$ is the dimension-reduced feature matrix of the training data. Each feature vector in \mathbf{X}' is extracted from different blocks, and there is no overlapping among the blocks. Thus, the feature vectors can be assumed to be independent of each other and all the feature vectors should conform to a common multivariate Gaussian model [21,22]. The MVG model can be learned from \mathbf{X}' with the standard maximum likelihood estimation algorithm, the MVG model is

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{d'/2} |\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})\right], \quad (14)$$

where $\mathbf{x} \in \mathbb{R}^{d' \times 1}$ is the feature vector after dimension reduction, and $\boldsymbol{\mu}$ and Σ are mean vector and covariance matrix, respectively. Since there is no distortion in the training set, the normal distribution of the features is represented by the learned MVG model, which is a benchmark for assessing distorted image [18]. When distortion exists in HSI, the distribution of the feature vector would deviate from the learned MVG model. The deviation can be measured and quality score of a distorted image can be computed.

For each testing HSI, we divide it into several blocks, of which the size is the same as that of training data. After extracting quality-sensitive features and stacking them into a feature vector as in the training stage, we can obtain a feature matrix $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m] \in \mathbb{R}^{d' \times m}$, where m is the number of blocks in the testing image. With the pre-learned projection matrix Φ , dimension-reduced feature matrix is

$$\mathbf{Y}' = \Phi \mathbf{Y}, \quad (15)$$

where $\mathbf{Y}' = [\mathbf{y}'_1, \mathbf{y}'_2, \dots, \mathbf{y}'_m] \in \mathbb{R}^{d' \times m}$ is the dimension-reduced feature matrix of the testing image. Different blocks make different contribution to the quality of testing image, so we compute quality score on each local block. Every block should be fitted by a MVG model $(\boldsymbol{\mu}_i, \Sigma_i)$, and then compared with the learned benchmark MVG model $(\boldsymbol{\mu}, \Sigma)$.

It should be noted that MVG model of each block can be estimated from its neighboring blocks, but it is complex and time-costly. In this work, $\boldsymbol{\mu}_i$ and Σ_i of the i -th block's features are simply approximated by \mathbf{y}'_i and covariance matrix of \mathbf{Y}' , which is denoted as Σ' . A modified Bhattacharyya distance is used to compute the distance between the benchmark MVG and the fitted MVG of the i -th block [27]

$$dis_i = \sqrt{(\boldsymbol{\mu} - \mathbf{y}'_i)^T \left(\frac{\Sigma + \Sigma'}{2}\right)^{-1} (\boldsymbol{\mu} - \mathbf{y}'_i)}, \quad (16)$$

The distance measures disparity between statistics of the i -th block and the pristine training data, it is used as the measurement for image quality. The smaller the distance, the better the image quality is. Quality score of the whole image is computed by averaging the distances over all the blocks.

4. Experiment Design and Results

4.1. Experiment Setting and Data

To demonstrate the effectiveness of the proposed assessment method, we test if the proposed quality scores are consistent with other reference-based indices. We firstly apply five state-of-the-art super-resolution methods to the simulated and real HSIs, and then quality scores of the reconstructed HSIs are computed and compared with reference-based evaluation indices to see if there is consistency.

The following super-resolution methods are used to reconstruct HSIs; these methods are selected due to their good performance in both reconstruction accuracy and speed:

- Coupled negative matrix factorization based hyperspectral fusion (denoted as CNMF) [8];
- Sparse spatial-spectral representation based super-resolution (denoted as SSR) [9];

- Sparse image fusion algorithm (denoted as sparseFU) [10];
- Bayesian sparse representation based super-resolution (denoted as BayesSR) [11]; and
- Spectral unmixing based super-resolution (denoted as SUn) [12].

Two datasets are used in the experiment. The first dataset was acquired by Airborne Visible Infrared Imaging Spectrometer (AVIRIS) sensor [23], which consists of 224 spectral bands in the range of 400 ~2500 nm. This dataset includes four images collected over *Moffett Field*, *Cuprite*, *Lunar Lake*, and *Indian Pines* sites with dimensions 753×1923 , 614×2207 , 781×6955 , and 614×1087 , respectively. Spatial resolution of *Moffett Field*, *Cuprite* and *Lunar Lake* is 20 m, and spatial resolution of *Indian Pines* is 4 m. After discarding the water absorption bands and noisy bands, there are 162 bands remaining. The second dataset was acquired by airborne Headwall Hyperspec-VNIR-C (HyperspecVC) sensor over agricultural and urban areas in Chikusei, Ibaraki, Japan [29]. It was made public by Dr. Naoto Yokoya and Prof. Akira Iwasaki from the University of Tokyo. The dataset has 128 bands in the range of 363 ~1018 nm. Size of the data is 2517×2335 . Spatial resolution of the data is 2.5 m. After discarding noisy bands, 125 bands are used in the experiment.

We crop two sub-images from each dataset as testing images, the rest of the dataset is treated as pristine data and used in the training stage. Then we apply the five super-resolution methods on the testing images, and evaluate the enhanced sub-images using the proposed assessment method. The parameters in the algorithm are set as follows. The size in spatial domain of each block is 64×64 , the size in spectral domain is the number of bands. The window sizes for local normalization in spectral and spatial domains are set as $K = 3$ and $S = T = 2$, respectively. The dimension of features after PCA is determined by the number of high order Principal Components (PCs) which has preserved at least 90% information of the original input. The parameters related to Log-Gabor filtering are adopted from [27]: $N = 3$, $J = 4$, $\sigma_r = 0.60$, $\sigma_\theta = 0.71$, $\omega_0^1 = 0.417$, $\omega_0^2 = 0.318$, $\omega_0^3 = 0.243$, where ω_0^1 , ω_0^2 , and ω_0^3 are the center frequencies of Log-Gabor filters at three different scales. All the parameters of the super-resolution methods are tuned to achieve the best reconstruction results.

4.2. Reference-Based Evaluation Indices

Peak signal-noise-ratio (PSNR), structural similarity index measurement (SSIM) [21], feature similarity index measurement (FSIM) [22], and spectral angle mean (SAM) are representatives of popular quantitative measures for image quality and have been applied to evaluate enhancement methods. They are selected to compare with the proposed quality score. PSNR computes the mean square errors of the reconstructed HSI. SSIM and FSIM calculate the similarity between the reconstructed HSI and reference. SAM measures the spectral distortion. Mathematically, PSNR of the l -th band is computed as

$$MSE_l = \frac{1}{MN} \left\| \mathbf{I}_l^{ref} - \mathbf{I}_l^{rec} \right\|_2, \quad (17)$$

$$PSNR_l = 10 \log_{10} \left(\frac{I_{max,l}}{MSE_l} \right)^2, \quad (18)$$

where $I_{max,l}$ is the maximum of the image on the l -th band, \mathbf{I}_l^{ref} and \mathbf{I}_l^{rec} are the reference image and reconstructed image on the l -th band. M and N are number of rows and columns. SSIM of the l -th band is computed as [21]

$$SSIM_l = \frac{4\sigma_{\mathbf{I}_l^{rec}\mathbf{I}_l^{ref}}\bar{I}_l^{rec}\bar{I}_l^{ref}}{(\sigma_{\mathbf{I}_l^{rec}}^2 + \sigma_{\mathbf{I}_l^{ref}}^2)}, \quad (19)$$

where \bar{I}_l^{ref} and \bar{I}_l^{rec} are mean of the reference and reconstructed image; and $\sigma_{\mathbf{I}_l^{rec}\mathbf{I}_l^{ref}}$, $\sigma_{\mathbf{I}_l^{ref}}$ and $\sigma_{\mathbf{I}_l^{rec}}$ are covariance and standard variance. FSIM of the l -th band is [22]

$$FSIM_l = \frac{\sum_{z \in \Omega} S_L(z) \cdot PC_m(z)}{\sum_{z \in \Omega} PC_m(z)}, \quad (20)$$

where Ω is the whole spatial domain, and

$$S_{PC}(z) = \frac{2PC_l^{ref}(z) \cdot PC_l^{rec}(z) + C_1}{(PC_l^{ref}(z))^2 + (PC_l^{rec}(z))^2 + C_1}, \quad (21)$$

$$S_{GM}(z) = \frac{2GM_l^{ref}(z) \cdot GM_l^{rec}(z) + C_2}{(GM_l^{ref}(z))^2 + (GM_l^{rec}(z))^2 + C_2}, \quad (22)$$

$$S_L(z) = S_{PC}(z) \cdot S_{GM}(z), \quad (23)$$

$$PC_m(z) = \max(PC_l^{ref}(z), PC_l^{rec}(z)), \quad (24)$$

where $PC_l^{ref}(z)$ and $PC_l^{rec}(z)$ are phase congruency at pixel z of the reference and the reconstructed images. The PSNR, SSIM, and FSIM are computed by averaging over all the acquired bands of HSI. SAM at pixel z is computed as

$$SAM(z) = \arccos\left(\frac{\langle S^{ref}(z), S^{rec}(z) \rangle}{\|S^{ref}(z)\|_2 \cdot \|S^{rec}(z)\|_2}\right), \quad (25)$$

where $S^{ref}(z)$ and $S^{rec}(z)$ represent the spectrum at pixel z , SAM of HSI is computed by averaging over the entire spatial domain.

4.3. Comparison With Reference-Based Indices

We crop two sub-images from *Indian Pines* and *Moffett Field* of AVIRIS data, and two sub-images (denoted as *Chikusei-1* and *Chikusei-2*) from HyperspecVC data. After down-sampling the sub-images by a factor of two, we apply the super-resolution methods on them. It is noted that our goal here is to compare our assessment method with previous indices. Down-sampling is necessary to obtain reference for these indices. The indices are reported in Tables 1–4. In order to present the trend of different indices clearly, we also plot the curves of different indices in Figures 15–18. The reconstructed HSIs using all enhancement methods are shown in Figures 19–22.

Table 1. Comparison among peak signal-noise-ratio (PSNR), structural similarity index measurement (SSIM), feature similarity index measurement (FSIM), and our score on *Indian Pines* of Airborne Visible Infrared Imaging Spectrometer (AVIRIS) data.

	sparseFU	SUn	BayesSR	SSR	CNMF
PSNR	23.6208 dB	28.7583 dB	29.0122 dB	30.5461 dB	30.9304 dB
SSIM	0.8317	0.9514	0.9455	0.9513	0.9616
FSIM	0.9125	0.9634	0.9640	0.9683	0.9698
Our score	30.4231	26.6541	25.8696	25.7163	25.3713

Table 2. Comparison among PSNR, SSIM, FSIM, and our score on *Moffett Field* of AVIRIS data.

	sparseFU	SUn	BayesSR	SSR	CNMF
PSNR	23.4575 dB	30.0800 dB	30.3489 dB	30.6237 dB	30.7831 dB
SSIM	0.8152	0.9226	0.9301	0.9478	0.9525
FSIM	0.9117	0.9516	0.9571	0.9669	0.9647
Our score	31.4860	28.7071	27.3159	27.2858	26.0752

Table 3. Comparison among PSNR, SSIM, FSIM, and our score on *Chikusei-1* of HyperspecVC data.

	sparseFU	SSR	SUn	BayesSR	CNMF
PSNR	29.1765 dB	33.1108 dB	34.5367 dB	36.5812 dB	36.9954 dB
SSIM	0.9521	0.9714	0.9735	0.9650	0.9883
FSIM	0.9557	0.9769	0.9828	0.9823	0.9899
Our score	21.3899	15.4410	15.3373	14.1547	13.9024

Table 4. Comparison among PSNR, SSIM, FSIM, and our score on *Chikusei-2* of HyperspecVC data.

	sparseFU	SSR	SUn	BayesSR	CNMF
PSNR	29.3492 dB	30.8350 dB	35.4586 dB	37.4310 dB	37.4797 dB
SSIM	0.9419	0.9463	0.9618	0.9663	0.9840
FSIM	0.9514	0.9640	0.9793	0.9808	0.9894
Our score	30.7928	23.3912	23.7168	23.1677	23.1171

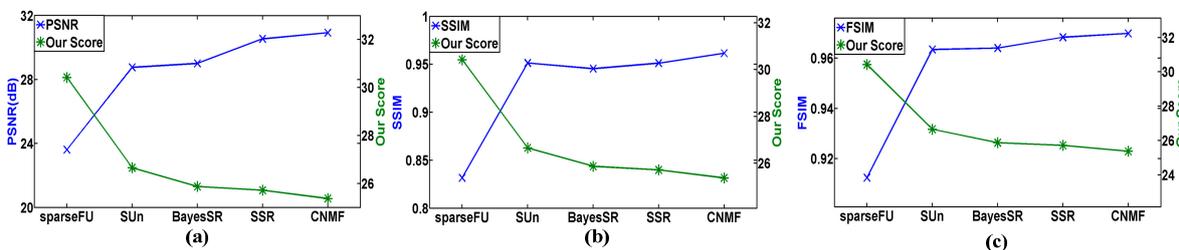


Figure 15. Consistency of our score and reference-based indices on *Indian Pines* of AVIRIS data: (a) our score and PSNR; (b) our score and SSIM; and (c) our score and FSIM.

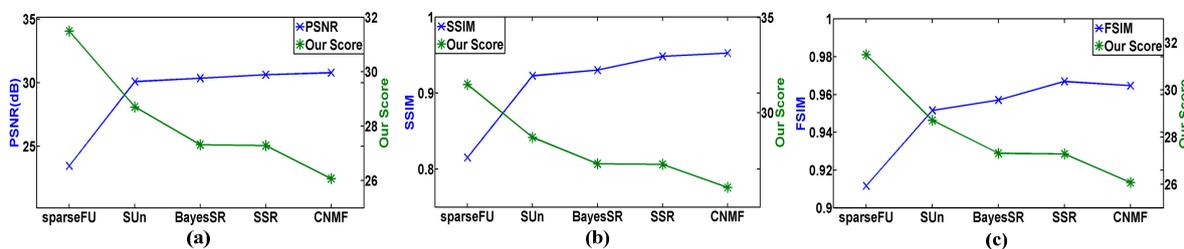


Figure 16. Consistency of our score and reference-based indices on *Moffett Field* of AVIRIS data: (a) our score and PSNR; (b) our score and SSIM; and (c) our score and FSIM.

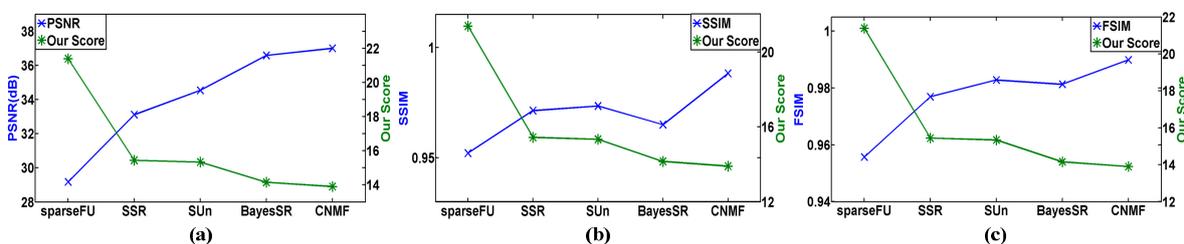


Figure 17. Consistency of our score and reference-based indices on *Chikusei-1* of HyperspecVC data: (a) our score and PSNR; (b) our score and SSIM; and (c) our score and FSIM.

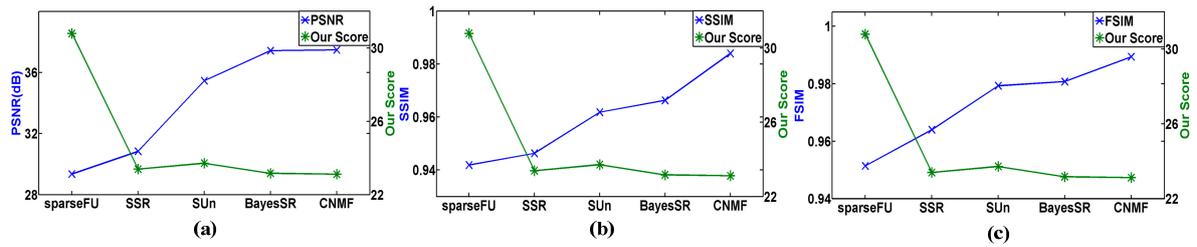


Figure 18. Consistency of our score and reference-based indices on *Chikusei-2* of HyperspecVC data: (a) our score and PSNR; (b) our score and SSIM; and (c) our score and FSIM.

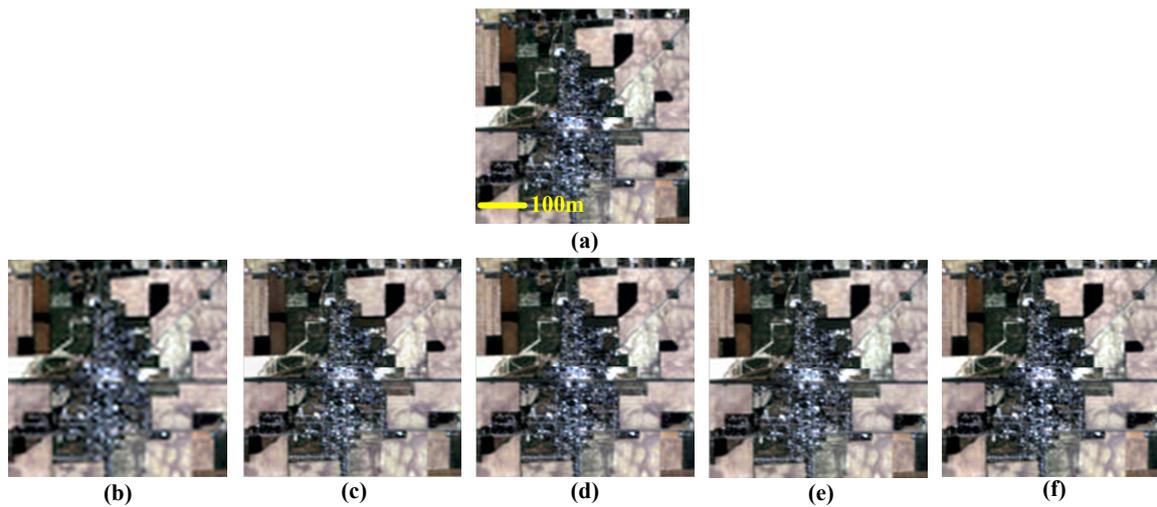


Figure 19. Reconstructed HSI of different super-resolution methods, the images are shown in RGB (band 35, 25, 15). The sub-image with size $128 \times 128 \times 162$ is cropped from *Indian Pines* of AVIRIS data: (a) original sub-image; (b) result of sparseFU; (c) result of SUn; (d) result of BayesSR; (e) result of SSR; and (f) result of CNMF.

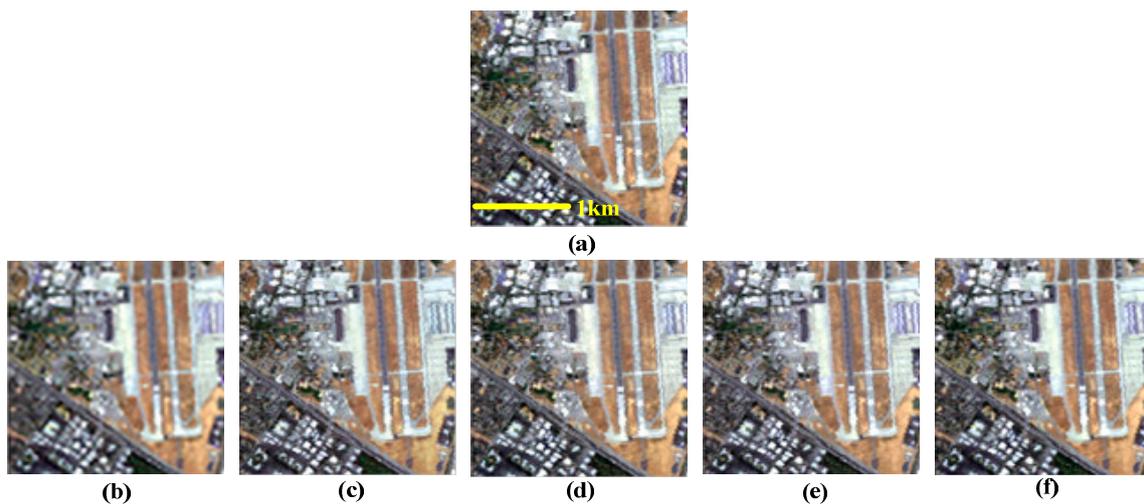


Figure 20. Reconstructed HSI of different super-resolution methods, the images are shown in RGB (band 35, 25, 15). The sub-image with size $128 \times 128 \times 162$ is cropped from *Moffett Field* of AVIRIS data: (a) original sub-image; (b) result of sparseFU; (c) result of SUn; (d) result of BayesSR; (e) result of SSR; and (f) result of CNMF.

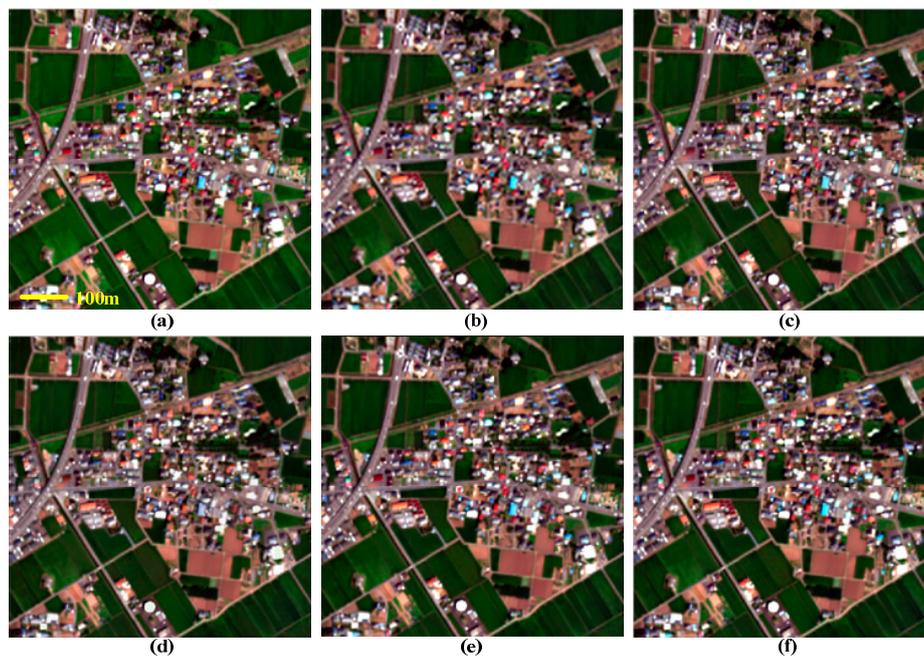


Figure 21. Reconstructed HSI of different super-resolution methods, the images are shown in RGB (band 56, 34, 19). The sub-image *Chikusei-1* with size $256 \times 256 \times 125$ is cropped from HyperspecVC data: (a) original sub-image; (b) result of sparseFU; (c) result of SSR; (d) result of SUn; (e) result of BayesSR; and (f) result of CNMF.

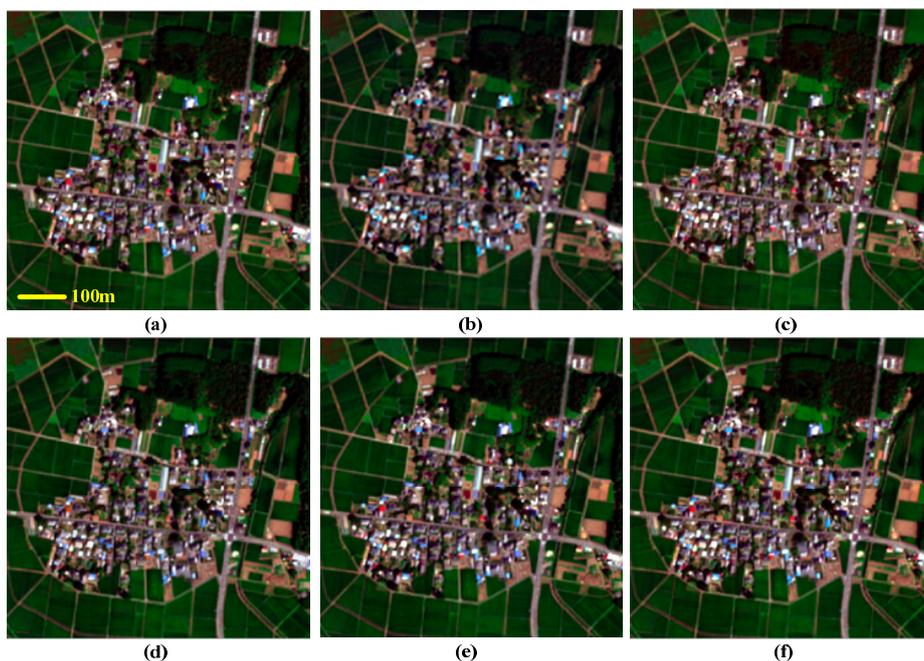


Figure 22. Reconstructed HSI of different super-resolution methods, the images are shown in RGB (band 56, 34, 19). The sub-image *Chikusei-2* with size $256 \times 256 \times 125$ is cropped from HyperspecVC data: (a) original sub-image; (b) result of sparseFU; (c) result of SSR; (d) result of SUn; (e) result of BayesSR; and (f) result of CNMF.

Our score measures the extent of distortion in the reconstructed HSI, with higher score representing lower quality, which should correspond to, e.g., a lower PSNR. In each table and figure, different methods are arranged in ascending order of PSNR from left to right. As shown in the tables

and figures, the corresponding scores of the proposed method are in descending order from left to right, which means that our no-reference score is consistent with PSNR in assessing the reconstructed HSI. We find that the no-reference score is not consistent with SSIM and FSIM of BayesSR, as shown in Figure 17b,c. This is caused by the inconsistency of SSIM and FSIM, both of them are not consistent with PSNR of BayesSR. Nevertheless, our score is consistent with SSIM and FSIM in most cases. PSNR, SSIM, and FSIM are the most common reference-based indices for image reconstruction evaluation, and the consistency between our measured scores and these three indices indicates that the proposed assessment method has potential to be implemented as a no-reference measure in evaluating spatially enhanced HSIs.

It should be noted that the result of SSR on *Chikusei-2* is inconsistent with other indices. In Table 4, PSNR values of SSR and SUn are 30.8350 dB and 35.4586 dB, showing that the quality of SSR is lower than that of the SUn. However, the proposed score method obtained the scores of 23.3912 for SSR and 23.7168 for SUn, showing the former has a better quality. If SSR can be evaluated correctly, its score should be higher than SUn and slightly lower than sparseFU. This inconsistency may be attributed to the limited number of training samples. AVIRIS dataset contains HSI acquired over multiple sites, more blocks can be extracted for training the benchmark MVG model, so it leads to great consistency on evaluations of *Indian Pines* and *Moffett Field*, as shown in Tables 1 and 2. However, the HyperspecVC data are taken only over *Chikusei*, the number of training blocks is smaller than those taken from the AVIRIS images, which may explain the failure of evaluating SSR in Table 4.

4.4. Spectral Distortion Assessment

Spectral fidelity is of high importance for the interpretation of HSI, so assessing spectral distortion is necessary for the reconstructed HSIs. Spectral angle mean (SAM), as a reference-based spectral assessment index, computes the disparity between the spectra of original HSI and reconstructed HSI. In this sub-section, we compute the spectral distortion without reference using the proposed method. Quality-sensitive features are extracted from both spectral and spatial domains in our method. If we extract quality-sensitive features only from the spectral domain and then train the benchmark MVG model, the quality score would measure spectral deviation of the reconstructed HSI from the pristine HSI, which can be treated as a measurement of spectral distortion. The spectral quality scores of the reconstructed HSIs are given in Tables 5–8. We also plot the spectral quality scores as curves in Figure 23.

Table 5. Comparison between SAM and spectral quality score on *Indian Pines* of AVIRIS data.

	sparseFU	SUn	SSR	CNMF	BayesSR
SAM	5.5003°	4.2875°	4.1631°	3.7864°	3.6997°
Spectral score	1.6776	1.6625	1.5130	1.0657	1.0469

Table 6. Comparison between SAM and spectral quality score on *Moffett Field* of AVIRIS data.

	sparseFU	BayesSR	SUn	SSR	CNMF
SAM	3.9216°	3.2950°	2.6870°	2.6214°	2.3456°
Spectral score	1.5505	1.3589	1.3252	1.2964	1.0489

Table 7. Comparison between SAM and spectral quality score on *Chikusei-1* of HyperspecVC data.

	BayesSR	SSR	sparseFU	SUn	CNMF
SAM	3.1975°	3.1424°	2.4779°	2.4702°	1.8175°
Spectral score	1.4390	1.4210	1.4322	1.3802	1.2435

Table 8. Comparison between SAM and spectral quality score on *Chikusei-2* of HyperspecVC data.

	SSR	BayesSR	SUn	sparseFU	CNMF
SAM	4.6458°	3.4304°	3.0957°	2.5936°	2.1912°
Spectral score	1.3016	1.4139	1.4024	1.3387	1.2043

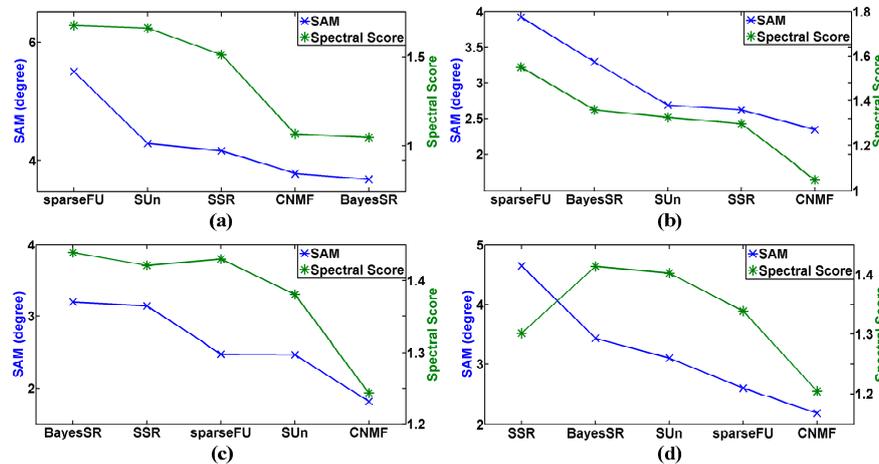


Figure 23. Comparison between SAM and spectral score: (a) on *Indian Pines*; (b) on *Moffett Field*; (c) on *Chikusei-1*; and (d) on *Chikusei-2*.

Different methods are arranged in descending order of SAM in the tables. As shown in Tables 5 and 6, the corresponding spectral quality scores are in descending order as well, which demonstrates that our no-reference spectral quality score is consistent with SAM on AVIRIS data. However, on *Chikusei-1*, SAM values of SSR and sparseFU are 3.1424° and 2.4779°, indicating that SSR has larger spectral distortion, while the spectral scores of SSR and sparseFU are 1.4210 and 1.4322, showing that sparseFU has larger distortion. Similarly, the spectral score of SSR is inconsistent with SAM on *Chikusei-2*. The fewer number of training samples that can be extracted from this dataset, the same reason suggested in Section 4.3, may have caused this inconsistency. However, most of our spectral quality scores are consistent with SAM on HyperspecVC data.

4.5. Analyzing Each Type of Spatial Features

There are four types of statistics features extracted from the spatial domain. They are based on histogram of the normalized panchromatic image, histograms of Log-Gabor filtering responses, histograms of directional gradient of Log-Gabor filtering responses, and histograms of gradient magnitude of Log-Gabor filtering responses. In order to analyze their contribution separately, we extract the spectral features and incorporate them with only one type of spatial features each time, then train the benchmark MVG and compute the quality score. We report the quality scores in Tables 9–12. We also plot the scores as curves in Figure 24.

Table 9. Comparison of each type of spatial features on *Indian Pines* of AVIRIS data.

	sparseFU	SUn	BayesSR	SSR	CNMF
PSNR	23.6208 dB	28.7583 dB	29.0122 dB	30.5461 dB	30.9304 dB
Norm. pan.	1.8954	3.2748	3.2500	3.1095	3.3361
Log-Gabor	12.8284	10.0834	10.0197	9.7432	9.8424
Log-Gabor grad.	14.8344	15.8623	15.2599	15.0333	14.4433
Log-Gabor grad. mag.	11.6690	11.6845	11.5943	11.3729	11.0730

Table 10. Comparison of each type of spatial features on *Moffett Field* of AVIRIS data.

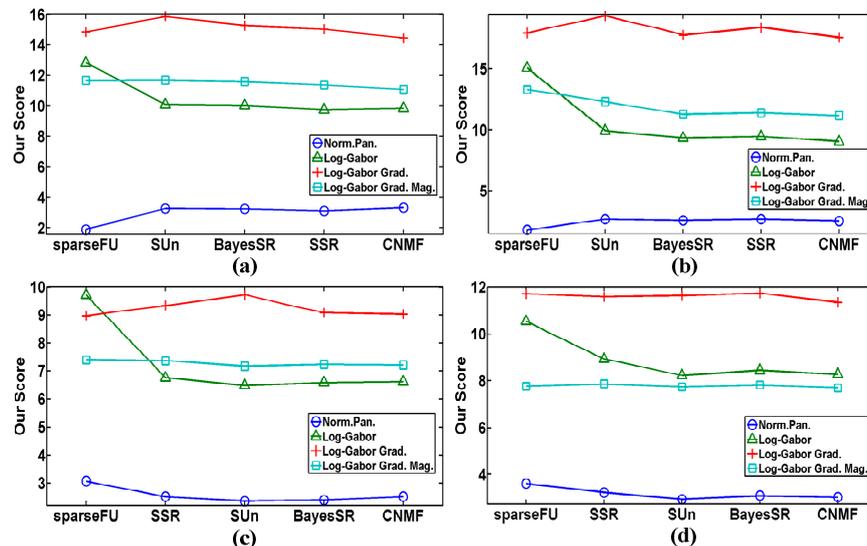
	sparseFU	SUn	BayesSR	SSR	CNMF
PSNR	23.4575 dB	30.0800 dB	30.3489 dB	30.6237 dB	30.7831 dB
Norm. pan.	1.8197	2.6929	2.6068	2.7086	2.5588
Log-Gabor	15.0683	9.9437	9.3638	9.4834	9.0729
Log-Gabor grad.	17.9726	19.3613	17.8142	18.4023	17.6334
Log-Gabor grad. mag.	13.2884	12.3348	11.2945	11.4125	11.1886

Table 11. Comparison of each type of spatial features on *Chikusei-1* of HyperspecVC data.

	sparseFU	SSR	SUn	BayesSR	CNMF
PSNR	29.1765 dB	33.1108 dB	34.5367 dB	36.5812 dB	36.9954 dB
Norm. pan.	3.0628	2.5048	2.3498	2.3832	2.5048
Log-Gabor	9.7003	6.7508	6.4789	6.5799	6.6106
Log-Gabor grad.	8.9788	9.3372	9.7097	9.0800	9.0335
Log-Gabor grad. mag.	7.3913	7.3545	7.1662	7.2224	7.2018

Table 12. Comparison of each type of spatial features on *Chikusei-2* of HyperspecVC data.

	sparseFU	SSR	SUn	BayesSR	CNMF
PSNR	29.3492 dB	30.8350 dB	35.4586 dB	37.4310 dB	37.4797 dB
Norm. pan.	3.6266	3.2416	2.9475	3.1052	3.0355
Log-Gabor	10.5328	8.9106	8.2312	8.4462	8.2740
Log-Gabor grad.	11.6873	11.5684	11.6145	11.7111	11.3460
Log-Gabor grad. mag.	7.7580	7.8534	7.7350	7.8038	7.6945

**Figure 24.** The curves of the quality scores with single type of spatial features used: (a) on *Indian Pines*; (b) on *Moffett Field*; (c) on *Chikusei-1*; and (d) on *Chikusei-2*.

We can make the following two conclusions from the results. Firstly, integrating multiple types of spatial features performs better than using a single type of spatial features. When only one type of spatial features is extracted, the curve is not monotonically descending, which means that some quality scores are not consistent with the reference-based indices, as shown in Figure 24. When all the spatial features are extracted, the scores are consistent with reference-based indices in most cases, as presented in Section 4.3. Secondly, among all these spatial features, the features based on Log-Gabor filtering are more efficient. As shown in Figure 24, the curves of Log-Gabor features are generally in descending

order. While other features, such as the features based on locally normalized panchromatic image, could not lead to a satisfactory assessment, as shown in the tables. This phenomenon is reasonable because Log-Gabor filters describe texture, edges, and details, which play a key role in reflecting the quality of image [19,20].

4.6. Robustness Analysis Over Training Data

To further investigate the robustness of our method, we design an experiment by varying the training data where the benchmark MVG model is trained on data from one sensor and used to evaluate enhanced data of another sensor. We train the benchmark MVG model on HyperspecVC data, then with the trained model, we evaluate the reconstructed images from AVIRIS data. We also compute the spectral distortion by training the benchmark MVG with only spectral features. The quality scores are presented in Tables 13 and 14, the spectral scores are presented in Tables 15 and 16, and the curves of scores are plotted in Figures 25 and 26.

The quality scores of SUn, BayesSR, SSR, and CNMF are consistent with PSNR on both *Indian Pines* and *Moffett Field*, but sparseFU cannot be assessed correctly. On *Indian Pines*, the spectral scores of SUn and SSR are not consistent with SAM. SAM of SUn and SSR are 4.2875° and 4.1631° respectively, showing that spectral distortion of SUn is larger than SSR. While our spectral score shows that spectral distortion of SSR is larger than SUn. On *Moffett Field*, except sparseFU, the spectral scores are consistent with SAM. The above inconsistency may be caused by the huge difference between the training datasets, as HyperspecVC data and AVIRIS data have big difference in spatial resolution and number of spectral bands. Although there is minor inconsistency between the scoring method and conventional reference-based indices, most of the super-resolution methods can still be assessed correctly, which demonstrates the robustness of the proposed method to some extent.

All the experiments are implemented on Matlab 2014a, with Intel PC Core 3.10 GHz, RAM of 12 GB. The training of our method takes about 30 min, while assessing the reconstructed HSI takes about 2 min. All the codes for super-resolution methods are provided by the authors.

Table 13. Performance on *Indian Pines* of AVIRIS data, trained on HyperspecVC data.

	sparseFU	SUn	BayesSR	SSR	CNMF
PSNR	23.6208 dB	28.7583 dB	29.0122 dB	30.5461 dB	30.9304 dB
Our score	72.3626	82.3826	80.3653	80.1138	79.9110

Table 14. Performance on *Moffett Field* of AVIRIS data, trained on HyperspecVC data.

	sparseFU	SUn	BayesSR	SSR	CNMF
PSNR	23.4575 dB	30.0800 dB	30.3489 dB	30.6237 dB	30.7831 dB
Our score	79.0081	89.3360	88.8949	87.9909	86.0247

Table 15. Spectral scores on *Indian Pines* of AVIRIS data, trained on HyperspecVC data.

	sparseFU	SUn	SSR	CNMF	BayesSR
SAM	5.5003°	4.2875°	4.1631°	3.7864°	3.6997°
Spectral score	1.7102	1.5146	1.5844	1.0776	0.9915

Table 16. Spectral scores on *Moffett Field* of AVIRIS data, trained on HyperspecVC data.

	sparseFU	BayesSR	SUn	SSR	CNMF
SAM	3.9216°	3.2950°	2.6870°	2.6214°	2.3456°
Spectral score	1.6405	1.7076	1.5679	1.4515	1.0755

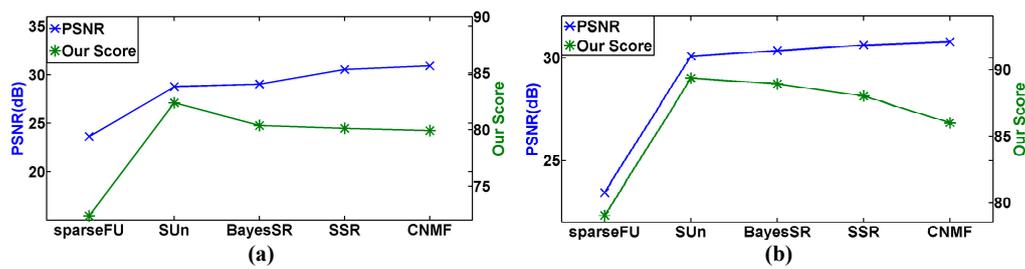


Figure 25. Consistency of our score and PSNR with HyperspecVC data used for training: (a) on *Indian Pines*; and (b) on *Moffett Field*.

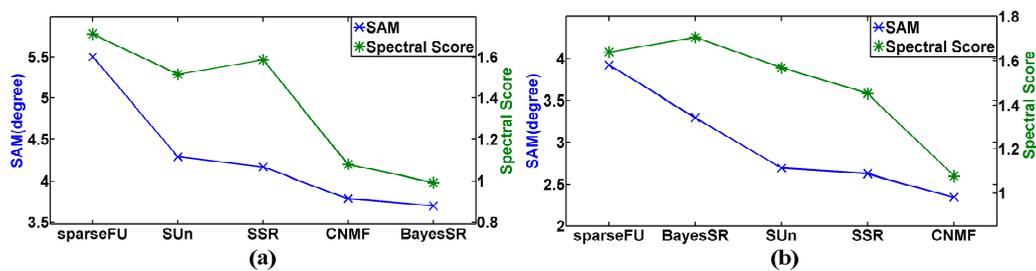


Figure 26. Consistency of our spectral score and SAM with HyperspecVC data used for training: (a) on *Indian Pines*; and (b) on *Moffett Field*.

5. Discussion

From the experiments, we can make the following interesting discussions:

1. The spectral features based on locally normalized spectra are more efficient than a single type of spatial features. If we only use the spectral features, the quality score reflects the spectral distortion of the reconstructed HSI and they are consistent with SAM, as shown in Section 4.4. However, if we use only a single type of spatial features, the quality scores of some reconstructed HSIs are not consistent with PSNR, as shown in Section 4.5. The efficiency of spectral features for distortion characterization can be further verified by comparing Figures 4b and 10, Figures 11 and 12 where spectral features belonging to the same distortion tend to form clusters which are more compact and more separable than that of spatial features.
2. Texture information is necessary in reflecting image quality, which is verified by the efficiency of features based on Log-Gabor responses and the gradients. We have tested several types of spatial features to characterize spatial quality of the reconstructed images. By comparing the performance among different types of spatial features, it is found that features based on statistics of Log-Gabor responses and the gradients often lead to better results than statistics of locally normalized panchromatic image, as shown in Section 4.4. It is worth noting that some other filters, such as wavelet and ridgelet [30,31], are also effective in texture analysis, extracting quality-sensitive features using these filters may lead to a better result.
3. Integrating multiple features is helpful for enhancing the performance. Multiple features are extracted from spectral and spatial domains and incorporated in the proposed method. By comparing the results in Sections 4.3 and 4.3, we find that if we only exploit a single type of features, some reconstructed HSI cannot be assessed correctly, while, if multiple features are exploited, most of the reconstructed HSI can be assessed correctly, which means that these features are complementary to each other in predicting image quality. Additional statistics features can also be integrated in our framework to obtain a better result.
4. The benchmark MVG is robust over the training data. In this study, the training data come from the same sensor with the testing data. When we train the benchmark MVG model on

HyperspecVC data and test it on AVIRIS data, it is observed that even though there is huge difference in spatial and spectral configuration of the two sensors, we obtained comparable results; most of the scores are consistent with PSNR and SAM on AVIRIS data. In real applications, if training data from the same sensor are not sufficient, training the benchmark MVG model with data from other sensors may be an alternative option.

5. The proposed method has potential to be applied in reality. The speed of our assessment method is fast: it only takes less than two minutes to evaluate the reconstructed HSI in the experiments. In addition, the proposed assessment method is fully blind, both the reference image and information related to the distortion type in HSI are not necessary to be known. All those characteristics make it possible to be applied in reality.

However, it should be noted that there are still some questions that need to be studied further in the future:

1. Research in models that are more efficient in representing the quality-sensitive features. In this study, we learn MVG models to represent the quality-sensitive features of pristine HSI and reconstructed HSI. MVG model is simple and fast to be implemented, but it may not be the most efficient one in feature representation. Some other advanced machine learning models, such as sparse representation [32], which were used in this work, could be more efficient. If we exploit these models to represent the quality-sensitive features, better performance may be obtained.
2. Determining the optimal number of features. According to our experiments, integrating multiple features is helpful. In this study, one type of spectral features and four types of spatial features are exploited. However, if more quality-sensitive features are exploited in the future, more training samples would be required and the computation burden would increase. In order to balance the computation burden and the performance, we need to determine the optimal number of features.

6. Conclusions

We propose a no-reference quality assessment method to assess reconstructed HSI. Image distortion can be characterized by statistics of HSI, measuring the deviation of these statistics makes it possible to assess the image quality of HSI. Based on this principle, statistical properties of pristine and distorted HSIs are analyzed, and then multiple statistics that are sensitive to image quality are extracted as features from both spectral and spatial domains. A MVG model is built for the features extracted from pristine training data and treated as benchmark. Reconstructed HSI is divided into several blocks, quality-sensitive features are extracted from each block, and a MVG model of the features is fitted for each block. Quality score of each block is computed by measuring the distance between the benchmark and the fitted MVG. Overall quality score is obtained by average pooling. We apply five state-of-the-art super-resolution methods on AVIRIS and HyperspecVC data, and then compute the quality scores of the reconstructed HSIs. Our quality scores have good consistency with PSNR, SSIM, FSIM, and SAM, which demonstrates the effectiveness and potential of the proposed no-reference assessment method.

Supplementary Materials: The following are available online at www.mdpi.com/2072-4292/9/4/305/s1. Supplementary experiments on testing sub-images cropped from *Cuprite* site of AVIRIS data.

Acknowledgments: The authors gratefully acknowledge Space Application Laboratory, Department of Advanced Interdisciplinary Studies, the University of Tokyo for providing the hyperspectral data. The authors also thank the anonymous reviewers for their helpful comments. This work is supported by the National Natural Science Foundation of China (61371152, 61071172, 61374162), the National Natural Science Foundation of China and South Korean National Research Foundation Joint Funded Cooperation Program (61511140292), New Century Excellent Talents Award Program from Ministry of Education of China (NCET-12-0464), the Ministry of Education Scientific Research Foundation for the Returned Overseas, the Fundamental Research Funds for the Central Universities (3102015ZY045), the China Scholarship Council for joint PhD students (201506290120), and the Innovation Foundation of Doctor Dissertation of Northwestern Polytechnical University (CX201621).

Author Contributions: Jingxiang Yang conceived the methodology, designed the experiments, and wrote the paper; Jonathan C.-W. Chan proposed the theme of research, assisted in experiment design and revised the manuscript; Yong-Qiang Zhao revised the manuscript; and Chen Yi helped with the experiment.

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

References

1. He, L.; Song, X.; Feng, W.; Guo, B.B.; Zhang, Y.S.; Wang, Y.H.; Wang, C.Y.; Guo, T.C. Improved remote sensing of leaf nitrogen concentration in winter wheat using multi-angular hyperspectral data. *Remote Sens. Environ.* **2016**, *174*, 122–133. [[CrossRef](#)]
2. Yokoya, N.; Chan, J.C.W.; Segl, K. Potential of resolution-enhanced hyperspectral data for mineral mapping using simulated EnMAP and Sentinel-2 images. *Remote Sens.* **2016**, *8*, 172. [[CrossRef](#)]
3. Pôças, I.; Rodrigues, A.; Gonçalves, S.; Costa, P.M.; Gonçalves, I.; Pereira, L.S.; Cunha, M. Predicting grapevine water status based on hyperspectral reflectance vegetation indices. *Remote Sens.* **2015**, *7*, 16460–16479. [[CrossRef](#)]
4. Loncan, L.; de Almeida, L.B.; Bioucas-Dias, J.M.; Briottet, X.; Chanussot, J.; Dobigeon, N.; Fabre, S.; Liao, W.; Licciardi, G.A.; Simões, M.; et al. Hyperspectral pansharpening: A review. *IEEE Geosci. Remote Sens. Mag.* **2015**, *3*, 27–46. [[CrossRef](#)]
5. Yue, L.; Shen, H.; Li, J.; Yuan, Q.; Zhang, H.; Zhang, L. Image super-resolution: The techniques, applications, and future. *Signal Process.* **2016**, *128*, 389–408.
6. 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](#)]
7. 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](#)]
8. 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](#)]
9. Akhtar, N.; Shafait, F.; Mian, A. Sparse spatio-spectral representation for hyperspectral image super-resolution. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014; pp. 63–78.
10. Zhu, X.X.; Bamler, R. A sparse image fusion algorithm with application to pan-sharpening. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 2827–2836. [[CrossRef](#)]
11. 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, Boston, MA, USA, 7–12 June 2015; pp. 3631–3640.
12. Lanaras, C.; Baltsavias, E.; Schindler, K. Hyperspectral Super-Resolution by Coupled Spectral Unmixing. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 3586–3594.
13. Aiazzi, B.; Alparone, L.; Baronti, S.; Garzelli, A.; Selva, M. MTF-tailored multiscale fusion of high-resolution MS and Pan imagery. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 591–596. [[CrossRef](#)]
14. 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](#)]
15. Hou, W.; Gao, X.; Tao, D.; Li, X. Blind image quality assessment via deep learning. *IEEE Trans. Neural Netw. Learn. Syst.* **2015**, *26*, 1275–1286. [[PubMed](#)]
16. Xue, W.; Zhang, L.; Mou, X. Learning without human scores for blind image quality assessment. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 23–28 June 2013; pp. 995–1002.
17. Mittal, A.; Moorthy, A.K.; Bovik, A.C. No-reference image quality assessment in the spatial domain. *IEEE Trans. Image Process.* **2012**, *21*, 4695–4708. [[CrossRef](#)] [[PubMed](#)]
18. Mittal, A.; Soundararajan, R.; Bovik, A.C. Making a “completely blind” image quality analyzer. *IEEE Signal Process. Lett.* **2013**, *20*, 209–212. [[CrossRef](#)]
19. Xue, W.; Mou, X.; Zhang, L.; Bovik, A.C.; Feng, X. Blind image quality assessment using joint statistics of gradient magnitude and Laplacian features. *IEEE Trans. Image Process.* **2014**, *23*, 4850–4862. [[CrossRef](#)] [[PubMed](#)]

20. Zhang, Y.; Chandler, D.M. No-reference image quality assessment based on log-derivative statistics of natural scenes. *J. Electron. Imaging* **2013**, *22*, 043025. [[CrossRef](#)]
21. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)] [[PubMed](#)]
22. Zhang, L.; Zhang, L.; Mou, X.; Zhang, D. FSIM: A feature similarity index for image quality assessment. *IEEE Trans. Image Process.* **2011**, *20*, 2378–2386. [[CrossRef](#)] [[PubMed](#)]
23. AVIRIS—Airborne Visible/Infrared Imaging Spectrometer—Data. Available online: http://aviris.jpl.nasa.gov/data/free_data.html (accessed on 30 October 2016).
24. Zhao, Y.Q.; Yang, J. Hyperspectral image denoising via sparse representation and low-rank constraint. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 296–308. [[CrossRef](#)]
25. Yang, J.; Zhao, Y.Q.; Chan, J.C.W.; Kong, S.G. Coupled sparse denoising and unmixing with low-rank constraint for hyperspectral image. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 1818–1833. [[CrossRef](#)]
26. Berisha, S.; Nagy, J.G.; Plemmons, R.J. Deblurring and sparse unmixing of hyperspectral images using multiple point spread functions. *SIAM J. Sci. Comput.* **2015**, *37*, S389–S406. [[CrossRef](#)]
27. Zhang, L.; Zhang, L.; Bovik, A.C. A feature-enriched completely blind image quality evaluator. *IEEE Trans. Image Process.* **2015**, *24*, 2579–2591.
28. Scholte, H.S.; Ghebreab, S.; Waldorp, L.; Smeulders, A.W.; Lamme, V.A. Brain responses strongly correlate with Weibull image statistics when processing natural images. *J. Vis.* **2009**, *9*, 29–29. [[CrossRef](#)] [[PubMed](#)]
29. Yokoya, N.; Iwasaki, A. *Airborne Hyperspectral Data over Chikusei*; Technical Report; Space Application Laboratory, The University of Tokyo: Tokyo, Japan, 2016.
30. Ramos, R.P.; do Nascimento, M.Z.; Pereira, D.C. Texture extraction: An evaluation of ridgelet, wavelet and co-occurrence based methods applied to mammograms. *Expert Syst. Appl.* **2012**, *39*, 11036–11047. [[CrossRef](#)]
31. Xu, Y.; Yang, X.; Ling, H.; Ji, H. A new texture descriptor using multifractal analysis in multi-orientation wavelet pyramid. In Proceedings of the IEEE Conference Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13–18 June 2010; pp. 161–168.
32. Rubinstein, R.; Bruckstein, A.M.; Elad, M. Dictionaries for sparse representation modeling. *Proc. IEEE* **2010**, *98*, 1045–1057. [[CrossRef](#)]



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).