

MvSSIM: A Quality Assessment Index for Hyperspectral Images

Rui Zhu^a, Fei Zhou^b, Jing-Hao Xue^{a,*}

^a*Department of Statistical Science, University College London, London WC1E 6BT, UK*

^b*The Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055, China*

Abstract

Quality assessment indexes play a fundamental role in the analysis of hyperspectral image (HSI) cubes. To assess the quality of an HSI cube, the structural similarity (SSIM) index has been widely applied in a band-by-band manner, as SSIM was originally designed for 2D images, and then the mean SSIM (MeanSSIM) index over all bands is adopted. MeanSSIM fails to accommodate the spectral structure which is a unique characteristic of HSI. Hence in this paper, we propose a new and simple multivariate SSIM (MvSSIM) index for HSI, by treating the pixel spectrum as a multivariate random vector. MvSSIM maintains SSIM's ability to assess the spatial structural similarity via correlation between two images of the same band; and adds an ability to assess the spectral structural similarity via covariance among different bands. MvSSIM is well founded on multivariate statistics and can be easily implemented through simple sample statistics involving mean vectors, covariance matrices and cross-covariance matrices. Experiments show

*Corresponding author. Tel.: +44-20-7679-1863; Fax: +44-20-3108-3105

Email addresses: r.zhu.12@ucl.ac.uk (Rui Zhu),
flyingzhou@sz.tsinghua.edu.cn (Fei Zhou), jinghao.xue@ucl.ac.uk (Jing-Hao Xue)

that MvSSIM is a proper quality assessment index for distorted HSIs with different kinds of degradations.

Keywords: Hyperspectral images, quality assessment, structural similarity (SSIM), spectral structure, spatial structure.

1. Introduction

Hyperspectral images (HSIs) are captured on 100s of narrow spectral bands ranging from 400 to 2400 nm, represented as a 3D data cube containing both the spatial structure in two dimensions and the spectral structure in the other dimension. Quality assessment plays a crucial role in evaluating the performance of many HSI preprocessing techniques, such as image restoration [1–4]. The quality of the preprocessed images is usually assessed by some quality assessment indexes. A good quality assessment index can identify well-preprocessed HSI and can thus assist the HSI analysis afterwards, such as classification, target detection and unmixing. Quality assessment for HSI has been discussed extensively in literature [5–7].

The structural similarity (SSIM) index has been widely used in the quality assessment of HSI [2–4, 8–10]. SSIM was originally designed for traditional 2D greyscale images to assess the image quality resembling human perception [11–14]. SSIM can evaluate the similarity in the spatial structure between two images (a reference image and a test image). Recently, many extensions of SSIM for 2D images have been proposed, such as multi-scale SSIM [15], complex wavelet SSIM [16], information content weighting SSIM [17] and intra-and-inter patch similarity [18], among others.

The literature of image quality assessment can be classified to three cat-

egories based on the availability of the reference image: full-reference assessment [15–19], reduced-reference assessment [20] and no-reference assessment [21]. As with the work on SSIM, in this paper we focus on the full reference assessment, i.e. a reference image (an HSI cube in our case) is provided.

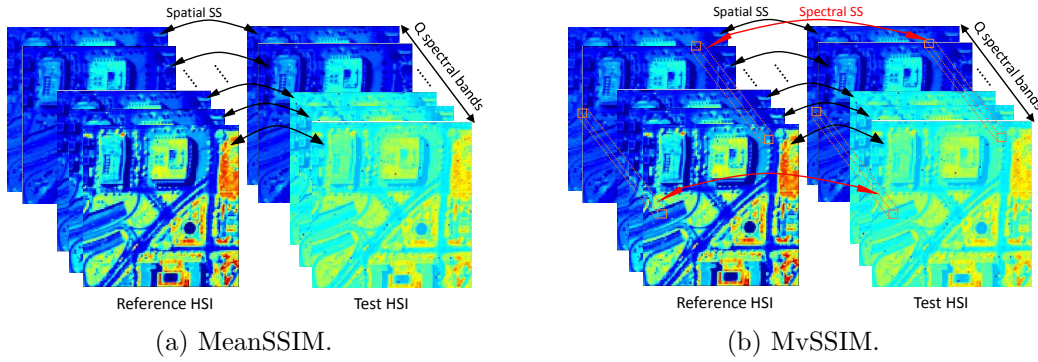


Figure 1: Illustration of MeanSSIM and MvSSIM (‘SS’ for structural similarity).

In the literature on using SSIM for HSI, usually a band-by-band manner is adopted for the 3D cube. The SSIM index for the image of each spectral band is calculated and then the mean of all these SSIM indexes (MeanSSIM) is taken as the quality measure of the whole HSI cube, as illustrated in Figure 1a. This simple strategy can compare the within-band spatial structure between each pair of images for the same band in the reference HSI and the test HSI. However, the similarity in the cross-band spectral structure, arising from the continuity property of the spectra, has been neglected, although such information is rich, unique and crucial in HSI. It is well known that both spatial and spectral structures are of great importance in the analysis of HSI and omitting the spectral structure is undesirable. Alparone et al. [6] and Garzelli and Nencini [5], extend SSIM to HSI by representing the pixel spec-

38 trum as a hypercomplex number. However, restricted by the properties of
 39 hypercomplex numbers, their index needs a recursive procedure to compute,
 40 making it not as popular as MeanSSIM in HSI restoration and denoising.

41 In this context, we propose in this paper a new and simple quality assess-
 42 ment index for HSI, termed multivariate SSIM (MvSSIM). In a 2D image
 43 a pixel is treated as a univariate random variable by SSIM; in contrast, in
 44 an HSI cube a pixel is in nature a multivariate random vector. To be more
 45 specific, each spectrum of a pixel in an HSI cube is represented as a multi-
 46 variate random vector, which contains the spectral information within each
 47 spectrum. Hence the cross-band spectral similarity between two HSI cubes
 48 can be naturally included in the index in this way. By replacing the univari-
 49 ate sampling statistics in SSIM with their multivariate versions, MvSSIM
 50 generalises SSIM to HSI.

51 Compared with MeanSSIM, MvSSIM can assess both the within-band
 52 spatial structural similarity, between images of the same band, and the cross-
 53 band spectral structural similarity, between spectra of the same pixel, as il-
 54 lustrated in Figure 1b between a reference cube and a test cube. MvSSIM is
 55 well founded on multivariate statistics and can be easily implemented through
 56 simple multivariate sample statistics involving mean vectors, covariance ma-
 57 trices and cross-covariance matrices. Experiments show that MvSSIM is a
 58 proper quality assessment index for distorted HSIs with different kinds of
 59 noises.

2. MvSSIM for hyperspectral images

2.1. SSIM

SSIM is a quality assessment index originally designed for 2D greyscale images. Suppose we have two images \mathbf{x} and \mathbf{y} , both containing $N = a \times b$ pixels: $\mathbf{x} = [x_1, \dots, x_N]^T \in \mathbb{R}^{N \times 1}$ and $\mathbf{y} = [y_1, \dots, y_N]^T \in \mathbb{R}^{N \times 1}$, aligned with each other. In SSIM, the N pixels of a 2D image are treated as N realisations of a univariate random variable: x_i and y_i ($i = 1, \dots, N$) are the realisations of random variables x and y , respectively.

SSIM consists of three comparisons between \mathbf{x} and \mathbf{y} : the similarity of luminance, $l(\mathbf{x}, \mathbf{y})$; the similarity of contrast, $c(\mathbf{x}, \mathbf{y})$; and the similarity of structure, $s(\mathbf{x}, \mathbf{y})$. It is defined as the product of the powers of these three similarities:

$$\text{SSIM}(\mathbf{x}, \mathbf{y}) = [l(\mathbf{x}, \mathbf{y})]^\alpha \times [c(\mathbf{x}, \mathbf{y})]^\beta \times [s(\mathbf{x}, \mathbf{y})]^\gamma, \quad (1)$$

where α , β and γ are three positive exponents adjusting the relative importance of the similarities and often all set to 1.

The three similarities are calculated by using the sample statistics of x and y . First, the similarity of luminance $l(\mathbf{x}, \mathbf{y})$ is obtained by comparing the sample means \bar{x} and \bar{y} :

$$l(\mathbf{x}, \mathbf{y}) = \frac{2\bar{x}\bar{y} + C_1}{\bar{x}^2 + \bar{y}^2 + C_1}, \quad (2)$$

where $\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$ and $\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i$, and C_1 is a constant that controls the stability of the fraction when $\bar{x}^2 + \bar{y}^2$ is close to zero. Constants C_2 and C_3

79 in the other two similarities play the same role as C_1 .

80 Second, the similarity of contrast $c(\mathbf{x}, \mathbf{y})$ is obtained by comparing the
81 sample standard deviations s_x and s_y :

$$c(\mathbf{x}, \mathbf{y}) = \frac{2s_x s_y + C_2}{s_x^2 + s_y^2 + C_2}, \quad (3)$$

82 where $s_x^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2$ and $s_y^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{y})^2$ are the sample variances.

83 Third, the similarity of structure $s(\mathbf{x}, \mathbf{y})$ is calculated as the sample cor-
84 relation coefficient of x and y :

$$s(\mathbf{x}, \mathbf{y}) = \frac{s_{xy}^2 + C_3}{s_x s_y + C_3}, \quad (4)$$

85 where $s_{xy}^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$ is the sample cross-variance. The sample
86 correlation coefficient measures the linear dependency between x and y , indi-
87 cating the similarity between two within-image spatial structures of the two
88 images, which were vectorised into a pair of two N -element vectors. Thus
89 $s(\mathbf{x}, \mathbf{y})$ is of great important in SSIM for assessing the spatial structural
90 similarity of two images.

91 SSIM possesses the following three good properties as a similarity index.
92 First, SSIM is symmetric, i.e. $\text{SSIM}(\mathbf{x}, \mathbf{y}) = \text{SSIM}(\mathbf{y}, \mathbf{x})$. Second, the value
93 of SSIM is bounded, i.e. $\text{SSIM}(\mathbf{x}, \mathbf{y}) \in [-1, 1]$. Third, SSIM has a unique
94 maximum, i.e. $\text{SSIM}(\mathbf{x}, \mathbf{y}) = 1$ if and only if $\mathbf{x} = \mathbf{y}$.

95 2.2. MeanSSIM

96 When SSIM is used in the quality assessment of HSI, it is commonly
97 applied in a band-by-band manner. That is, an SSIM index is obtained for a

pair of images of the same band, and then the mean index over bands is used as the quality measure of the test HSI cube against the reference cube, as illustrated in Figure 1a. We call this measure the mean SSIM (MeanSSIM) index.

Suppose we have two HSI cubes, $\mathbf{X}_H \in \mathbb{R}^{a \times b \times Q}$ and $\mathbf{Y}_H \in \mathbb{R}^{a \times b \times Q}$, where a and b represent the numbers of pixels in height and width, and Q is the number of spectral bands. \mathbf{X}_H and \mathbf{Y}_H can be rearranged as 2D matrices $\mathbf{X} = [\mathbf{x}_1^c, \mathbf{x}_2^c, \dots, \mathbf{x}_Q^c] \in \mathbb{R}^{N \times Q}$ and $\mathbf{Y} = [\mathbf{y}_1^c, \mathbf{y}_2^c, \dots, \mathbf{y}_Q^c] \in \mathbb{R}^{N \times Q}$, where $N = a \times b$ denotes the total number of pixels and $\mathbf{x}_q^c \in \mathbb{R}^{N \times 1}$ and $\mathbf{y}_q^c \in \mathbb{R}^{N \times 1}$ represent the image vectors of the q th spectral band of \mathbf{X}_H and \mathbf{Y}_H , respectively. The MeanSSIM index is calculated as

$$\text{MeanSSIM} = \frac{1}{Q} \sum_{q=1}^Q \text{SSIM}(\mathbf{x}_q^c, \mathbf{y}_q^c). \quad (5)$$

MeanSSIM can explore the similarity in spatial structure of each pair of band images. However, due to its band-by-band manner, it fails to adequately explore the cross-band spectral structure in HSI, while the spectrum of each pixel, i.e. each row of \mathbf{X} or \mathbf{Y} , contains crucial information like its chemical components. Thus, in addition to assessing the within-band spatial structural similarity between two images of the same band, assessing the cross-band spectral structural similarity between two spectra at the same spatial position should also be considered in the quality assessment of HSI.

2.3. *MvSSIM*

Since an HSI cube contains both spatial structure and spectral structure, its quality assessment should contain assessments for both structures.

120 Hence in this paper, we propose multivariate SSIM (MvSSIM) for the quality
 121 assessment of HSI, generalising SSIM via multivariate sample statistics.

122 In MvSSIM, the spectrum of each pixel of an HSI cube is treated as
 123 a realisation of a Q -dimensional random vector. To be more specific, we
 124 rewrite $\mathbf{X} \in \mathbb{R}^{N \times Q}$ and $\mathbf{Y} \in \mathbb{R}^{N \times Q}$ as $\mathbf{X} = [\mathbf{x}_1^r, \mathbf{x}_2^r, \dots, \mathbf{x}_N^r]^T$ and $\mathbf{Y} =$
 125 $[\mathbf{y}_1^r, \mathbf{y}_2^r, \dots, \mathbf{y}_N^r]^T$, where $\mathbf{x}_n^r \in \mathbb{R}^{Q \times 1}$ and $\mathbf{y}_n^r \in \mathbb{R}^{Q \times 1}$ represent the spectra of
 126 the n th pixel of \mathbf{X}_H and \mathbf{Y}_H , respectively. Here \mathbf{x}_n^r and \mathbf{y}_n^r are considered as
 127 the realisations of Q -dimensional random vectors $X \in \mathbb{R}^{Q \times 1}$ and $Y \in \mathbb{R}^{Q \times 1}$,
 128 respectively.

129 As an extension of SSIM, MvSSIM also consists of three similarity mea-
 130 surements between \mathbf{X} and \mathbf{Y} , i.e. $l(\mathbf{X}, \mathbf{Y})$, $c(\mathbf{X}, \mathbf{Y})$ and $s(\mathbf{X}, \mathbf{Y})$. These
 131 three similarities are defined on the following multivariate sample statistics
 132 of X and Y :

133 i) the sample means,

$$\bar{X} = \frac{1}{N} \sum_{n=1}^N \mathbf{x}_n^r \in \mathbb{R}^{Q \times 1}, \quad \bar{Y} = \frac{1}{N} \sum_{n=1}^N \mathbf{y}_n^r \in \mathbb{R}^{Q \times 1}; \quad (6)$$

ii) the sample covariance matrices,

$$\Sigma_X = \frac{1}{N-1} \sum_{n=1}^N (\mathbf{x}_n^r - \bar{X})(\mathbf{x}_n^r - \bar{X})^T \in \mathbb{R}^{Q \times Q}, \quad (7)$$

$$\Sigma_Y = \frac{1}{N-1} \sum_{n=1}^N (\mathbf{y}_n^r - \bar{Y})(\mathbf{y}_n^r - \bar{Y})^T \in \mathbb{R}^{Q \times Q}; \quad (8)$$

134 and iii) the sample cross-covariance matrix,

$$\Sigma_{XY} = \frac{1}{N-1} \sum_{n=1}^N (\mathbf{x}_n^r - \bar{X})(\mathbf{y}_n^r - \bar{Y})^T \in \mathbb{R}^{Q \times Q}. \quad (9)$$

135 Different from the univariate sample statistics in SSIM, the sample statis-
 136 tics in MvSSIM are vectors or matrices, rather than scalars. Thus the com-
 137 parisons between scalars in SSIM should be extended to comparisons be-
 138 tween vectors or matrices in MvSSIM. The extensions from $l(\mathbf{x}, \mathbf{y})$, $c(\mathbf{x}, \mathbf{y})$
 139 and $s(\mathbf{x}, \mathbf{y})$ to $l(\mathbf{X}, \mathbf{Y})$, $c(\mathbf{X}, \mathbf{Y})$ and $s(\mathbf{X}, \mathbf{Y})$ are described as follows.

140 2.3.1. From $l(\mathbf{x}, \mathbf{y})$ to $l(\mathbf{X}, \mathbf{Y})$

141 As with $l(\mathbf{x}, \mathbf{y})$, $l(\mathbf{X}, \mathbf{Y})$ measures the luminance similarity between im-
 142 ages by comparing the sample mean vectors, \bar{X} and \bar{Y} . Because $l(\mathbf{X}, \mathbf{Y})$
 143 compares the luminance similarity, the spectral structure is not included in
 144 this term and the inner products of vectors are used to make the numerator
 145 and denominator scalars. We define

$$l(\mathbf{X}, \mathbf{Y}) = \frac{2\langle \bar{X}, \bar{Y} \rangle + C_1}{\langle \bar{X}, \bar{X} \rangle + \langle \bar{Y}, \bar{Y} \rangle + C_1} = \frac{2 \sum_{q=1}^Q \bar{x}_q \bar{y}_q + C_1}{\sum_{q=1}^Q (\bar{x}_q^2 + \bar{y}_q^2) + C_1}, \quad (10)$$

146 where $\langle \cdot, \cdot \rangle$ denotes the inner product of two vectors, and \bar{x}_q and \bar{y}_q are the
 147 q th entries of \bar{X} and \bar{Y} , respectively.

148 It is easy to show that $l(\mathbf{X}, \mathbf{Y}) \in [0, 1]$ and $l(\mathbf{X}, \mathbf{Y}) = 1$ when $\mathbf{X} = \mathbf{Y}$. If
 149 $Q = 1$, i.e. the HSI becomes a 2-D image, (10) degenerates into (2) of SSIM.

150 2.3.2. From $c(\mathbf{x}, \mathbf{y})$ to $c(\mathbf{X}, \mathbf{Y})$

151 Similar to $c(\mathbf{x}, \mathbf{y})$, $c(\mathbf{X}, \mathbf{Y})$ compares the similarity between sample co-
 152 variance matrices Σ_X and Σ_Y . A sample covariance matrix (e.g. Σ_X) con-
 153 tains the variances within individual bands (of \mathbf{X}) in its diagonal entries, and
 154 the covariances between different spectral bands (of \mathbf{X}) in its off-diagonal en-
 155 tries. Hence when we compare \mathbf{X} and \mathbf{Y} through Σ_X and Σ_Y , we can achieve
 156 two comparisons simultaneously: comparing the contrasts of two images of
 157 the same band via the two standard deviations of this band, and comparing
 158 the contrasts of two spectra of the same spatial position via the covariances
 159 between different bands.

160 To make use of both the spatial and spectral information and to make
 161 the numerator and the denominator scalars, a natural choice is to use the
 162 nuclear norm to summarise the sample covariance matrix. Hence we define
 163 $c(\mathbf{X}, \mathbf{Y})$ as

$$c(\mathbf{X}, \mathbf{Y}) = \frac{2\|\Sigma_X\|_*^{\frac{1}{2}}\|\Sigma_Y\|_*^{\frac{1}{2}} + C_2}{\|\Sigma_X\|_* + \|\Sigma_Y\|_* + C_2} = \frac{2\sqrt{\lambda^{(s)}}\sqrt{d^{(s)}} + C_2}{\lambda^{(s)} + d^{(s)} + C_2}, \quad (11)$$

164 where $\|\cdot\|_*$ is the nuclear norm, $\lambda^{(s)} = \sum_{q=1}^Q \lambda_q$, $d^{(s)} = \sum_{q=1}^Q d_q$, λ_q 's are the
 165 singular values of Σ_X , and d_q 's are the singular values of Σ_Y .

166 The similarity $c(\mathbf{X}, \mathbf{Y})$ can take values in $[0, 1]$, and $c(\mathbf{X}, \mathbf{Y}) = 1$ when
 167 $\mathbf{X} = \mathbf{Y}$. If $Q = 1$, we treat the spectral norm of a scalar as itself and (11) is
 168 equivalent to (3) of SSIM.

169 2.3.3. From $s(\mathbf{x}, \mathbf{y})$ to $s(\mathbf{X}, \mathbf{Y})$

The term $s(\mathbf{x}, \mathbf{y})$ measures the spatial structural similarity between two images and is vital for SSIM resembling human perception. Preserving this good property of SSIM, we also adopt the correlation coefficient for MvSSIM. We define $s(\mathbf{X}, \mathbf{Y})$ as

$$\begin{aligned} s(\mathbf{X}, \mathbf{Y}) &= \frac{1}{Q} \text{trace}((\mathbf{\Sigma}_{\mathbf{XY}} + C_3 \mathbf{I}_Q)(\mathbf{\Gamma}_X^{\frac{1}{2}} \mathbf{\Gamma}_Y^{\frac{1}{2}} + C_3 \mathbf{I}_Q)^{-1}) \\ &= \frac{1}{Q} \sum_{q=1}^Q \frac{\sigma_{XYq}^2 + C_3}{\sigma_{Xq} \sigma_{Yq} + C_3}, \end{aligned} \quad (12)$$

170 where $\mathbf{\Gamma}_X$ and $\mathbf{\Gamma}_Y$ are diagonal matrices composed of the diagonal elements
171 of $\mathbf{\Sigma}_X$ and $\mathbf{\Sigma}_Y$, respectively; and σ_{XYq}^2 , σ_{Xq}^2 and σ_{Yq}^2 are the q th diagonal
172 entry of $\mathbf{\Sigma}_X$, $\mathbf{\Sigma}_Y$ and $\mathbf{\Sigma}_{\mathbf{XY}}$, respectively. It is obvious that $s(\mathbf{X}, \mathbf{Y})$ is the
173 mean of correlation coefficients of all spectral bands.

174 The similarity $s(\mathbf{X}, \mathbf{Y}) \in [-1, 1]$, and $s(\mathbf{X}, \mathbf{Y}) = 1$ when $\mathbf{X} = \mathbf{Y}$. If
175 $Q = 1$, (12) degenerates into (4) of SSIM.

176 2.3.4. MvSSIM

177 Combing the three similarity measurements defined above, the MvSSIM
178 index of \mathbf{X} and \mathbf{Y} can be written in a similar formulation to SSIM:

$$\text{MvSSIM}(\mathbf{X}, \mathbf{Y}) = [l(\mathbf{X}, \mathbf{Y})]^\alpha \times [c(\mathbf{X}, \mathbf{Y})]^\beta \times [s(\mathbf{X}, \mathbf{Y})]^\gamma, \quad (13)$$

179 where as with SSIM α , β and γ are three positive exponents that adjust the
180 relative importance of the components.

181 Among these three terms, $l(\mathbf{X}, \mathbf{Y})$ and $s(\mathbf{X}, \mathbf{Y})$ measure the similarity

182 between band images in luminance and spatial structure, while $c(\mathbf{X}, \mathbf{Y})$ mea-
 183 sures the similarity between both band images and pixel spectra. Thus in
 184 MvSSIM, both the within-band spatial structural similarity and the cross-
 185 band spectral structural similarity are assessed.

186 Moreover, comparing (1)-(4) with (10)-(13), we can find that MvSSIM is
 187 a natural generalisation of SSIM, and thus it can be readily embedded into
 188 other state-of-the-art SSIM-based quality assessment indexes such as [15–18].

189 3. Experiments

190 Besides MeanSSIM, MvSSIM is also compared with three other SSIM-
 191 based quality assessment indexes in literature, namely Q_λ , Q_m [7] and $Q2^n$ [5].

192 The index Q_λ measures the minimum SSIM between the pair of spectra of
 193 the same pixel among all pixels; Q_m is the product of Q_λ and the minimum
 194 SSIM between the pair of images of the same band among all bands; and
 195 $Q2^n$ is an extension of SSIM by expressing the spectrum as a hypercomplex
 196 number.

197 The five quality assessment indexes could be categorised into the following
 198 three groups: 1) Q_λ , which measures spectral similarities between spectra of
 199 the same pixel; 2) MeanSSIM, which measures spatial similarities between
 200 images of the same band; and 3) Q_m , $Q2^n$ and MvSSIM, which measure both
 201 spectral and spatial similarities.

202 3.1. Dataset

203 The Washington DC HSI is used for the synthetic experiments. The
 204 Washington DC HSI is a Hyperspectral Digital Imagery Collection Experi-
 205 ment (HYDICE) image of Washington DC Mall and can be downloaded from

206 <https://engineering.purdue.edu/~landgreb/Hyperspectral.Ex.html>. The
207 dataset is of size $1208 \times 1208 \times 191$, where 1208×1208 is the spatial size of
208 the HSI and 191 is the number of bands.

209 We extract a subcube from the whole Washington DC HSI cube for the
210 experiments following [1]. The subcube is of size $250 \times 250 \times 191$, where
211 250×250 is the size of the spatial size of the HSI and 191 is the number of
212 bands. The original HSI subcube serves as the reference cube while its noisy
213 version acts as a test cube.

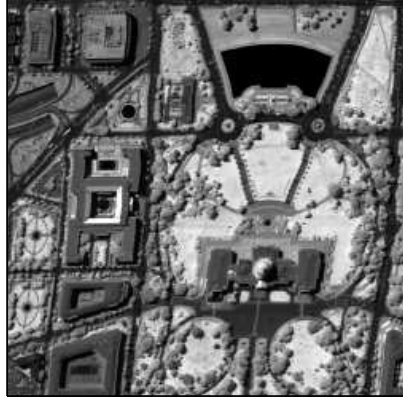
214 3.2. Experiment settings

215 MeanSSIM is computed using the MATLAB function ‘ssim’ with the de-
216 fault setting: window size is 11, $C_1 = 0.01$ and $C_2 = 0.03$. For MvSSIM, a
217 patch of size $5 \times 5 \times 191$ moves from pixel to pixel, the index of each patch is
218 calculated, and then the mean index of all the patches is taken as the index of
219 the whole HSI. We set constants C_i of MvSSIM to 0 and exponents α , β and
220 γ to 1 for simplicity. The index $Q2^n$ is calculated by using the pansharpening
221 toolbox of [22]. The block size is set to 32 and the block shift size is set to
222 32, as suggested in [5].

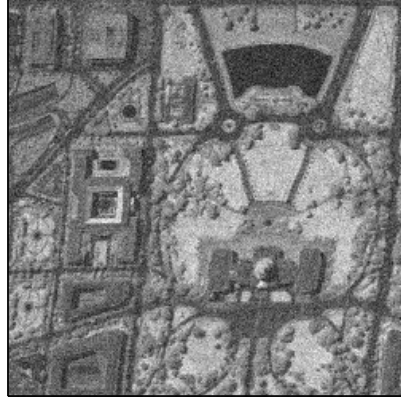
223 Following the experiments in [7], four typical degradations are applied to
224 the HSI to evaluate the quality assessment indexes: Gaussian white additive
225 noise, spatial smoothing, spectral smoothing and lossy compression. The
226 index values are calculated for different levels of degradations.

227 First, Gaussian white additive noises are added to 50 randomly-selected
228 bands of the spectra. We test 10 different variances: from 10 to 100 with a
229 step of 10, i.e. 10 different noisy HSIs are created with different variances.

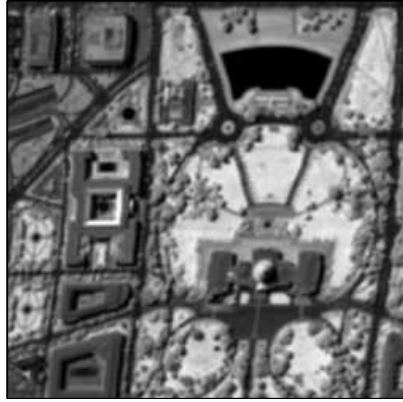
230 Second, Gaussian smoothing filters are applied to 50 randomly-selected



(a)



(b)



(c)



(d)



(e)

Figure 2: The reference image and noisy images of band 80. (a) Reference. (b) Gaussian white noise (variance 60). (c) Gaussian smoothing noise (standard deviation 1). (d) Savitzky-Golay smoothing noises (frame size 11). (e) JPEG2000 compression (compression ratio 30).

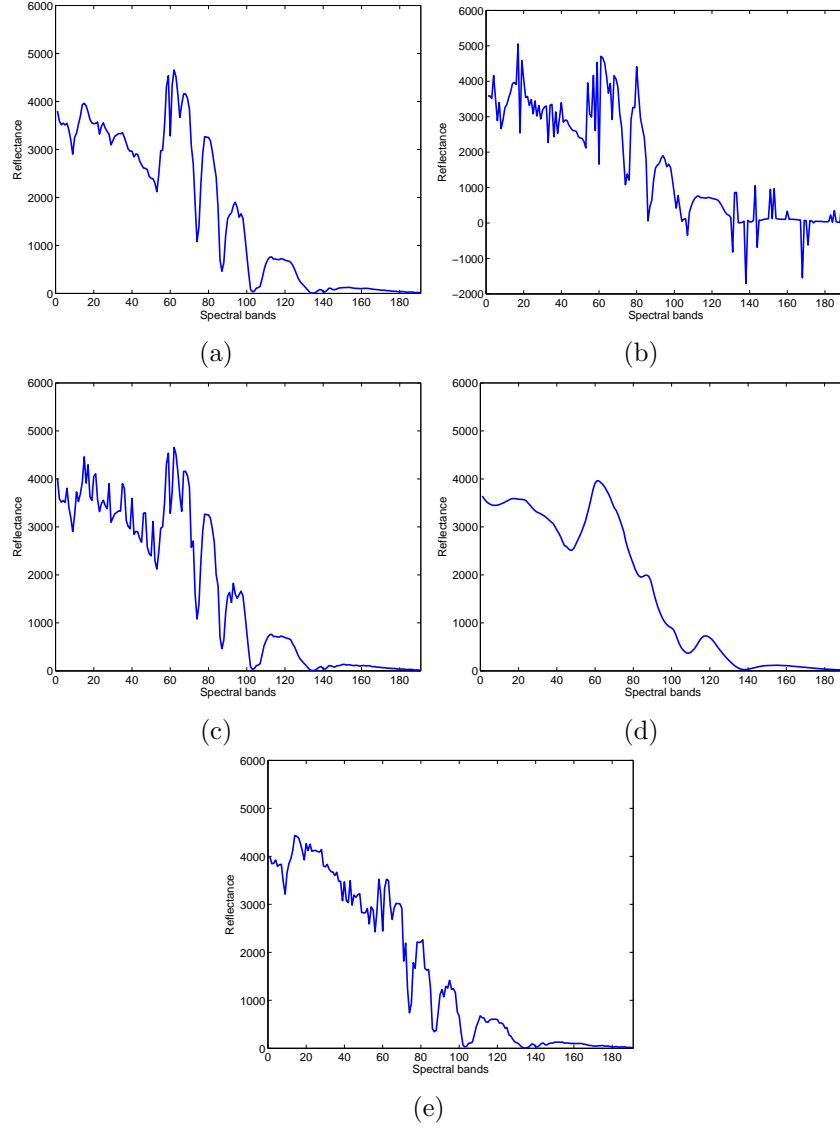


Figure 3: The reference spectrum and noisy spectra of the pixel at position (50, 50). (a) Reference. (b) Gaussian white noise (variance 60). (c) Gaussian smoothing noise (standard deviation 1). (d) Savitzky-Golay smoothing noises (frame size 11). (e) JPEG2000 compression (compression ratio 30).

bands to create spatially blurred band images, i.e. in the spatial dimensions of the HSI. Eight different standard deviations of the Gaussian smoothing kernels are tested: 0.1, 0.5, 1, 5, 10, 50, 100 and 500, i.e. eight different noisy HSIs are created with different standard deviations.

Third, Savitzky-Golay smoothing filter is applied to the spectra of all pixels to create smooth spectra, i.e. in the spectral dimension of the HSI. We test eight different frame sizes: 5, 11, 31, 71, 91, 131, 171 and 191, i.e. eight different noisy HSIs are created with different frame sizes..

Fourth, JPEG2000 compression is applied to the HSI in a band-by-band way. We test five different compression ratios: from 10 to 50 with a step of 10, i.e. five different noisy HSIs are created with different compression ratios.

The reference image and noisy images of band 80 and the reference spectrum and noisy spectra of pixel (50, 50) are shown in Figure 2 and Figure 3.

3.3. Results

3.3.1. Gaussian white additive noise

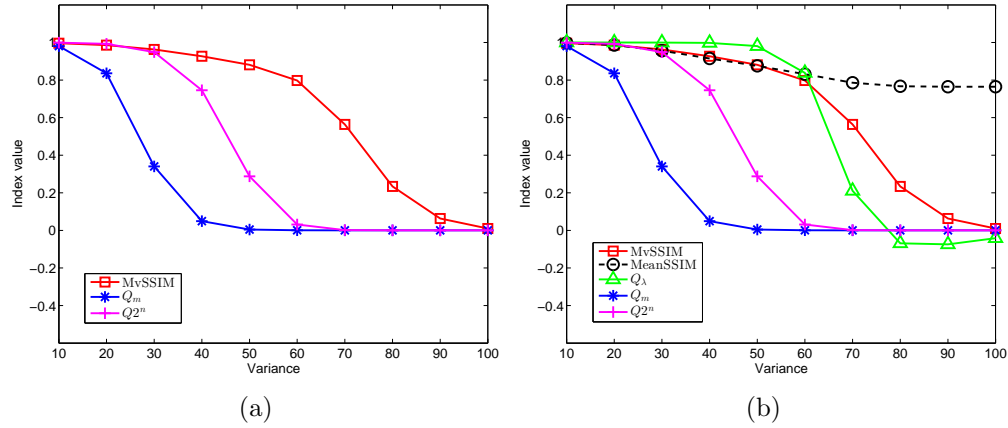


Figure 4: Assessments for the Gaussian white additive noise contaminated HSIs.

Figure 4 shows the assessments for the HSIs contaminated by the Gaussian white additive noises of different variances, which represent different degrees of contamination. The performances of the three indexes that measure both spectral and spatial similarities are shown in Figure 4a. It is obvious that Q_m is the most sensitive to the Gaussian white additive noise, Q_{2^n} is less sensitive, and MvSSIM is the least sensitive. However, sensitivity is not the only criterion to evaluate the performances of the indexes. The changes in the spatial structure and the spectral structure should also be considered when carrying out such evaluation.

We use MeanSSIM as a measurement for the spatial structural change and Q_λ as a measurement for the spectral structural change, and plot the performances of these two indexes in Figure 4b. In the plot, the value of Q_λ is high when the variance is less than 60 and drops fast when the variance becomes large; this indicates that the spectral structure changes little when the white noise is light but can change dramatically when the white noise is heavy. In the meantime, the figure shows that the value of MeanSSIM is relatively stable; this indicates that the spatial structure does not change much with the variance of white noise. This is because MeanSSIM averages out white noise over bands that the low similarities between contaminated band images are compensated by high similarities between other band images.

Considering the above behaviours of MeanSSIM and Q_λ , we prefer MvSSIM in the Gaussian white noise case even though it is the least sensitive index in Figure 4a. As shown in Figure 4b, it is clear that the values of Q_m and Q_{2^n} are close to zero even when the values of Q_λ are still close to one; this indicates that Q_m and Q_{2^n} fail to consider the high spectral structural

271 similarity in this case and are over-sensitive to the Gaussian white noise.
 272 In contrast, MvSSIM provides large values when the values of Q_λ are large.
 273 Also, compared with Q_λ , MvSSIM is more desired because it also reflects the
 274 spatial structural similarity, making it between MeanSSIM and Q_λ in the
 275 case of Gaussian white noise.

276 3.3.2. Gaussian smoothing noise

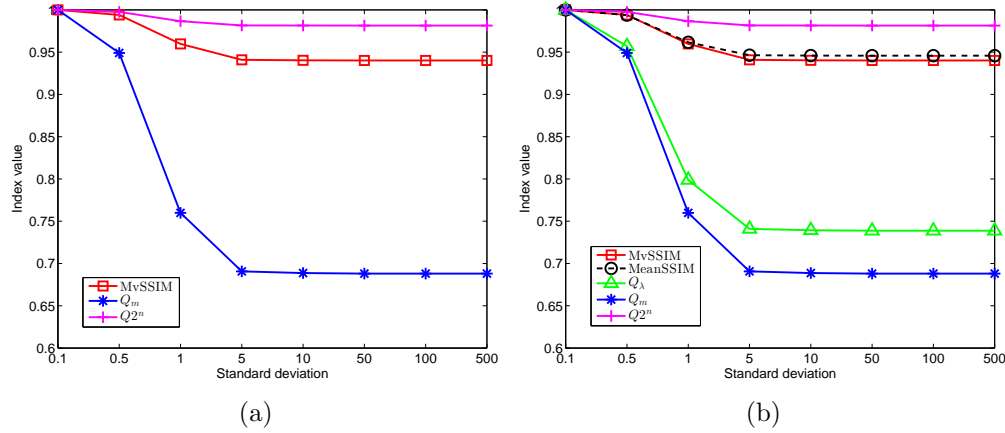


Figure 5: Assessments for the Gaussian smoothing noise contaminated HSIs.

277 Figure 5a shows the assessments for the HSIs contaminated by the Gaus-
 278 sian smoothing noise: Q_m is the most sensitive to the Gaussian smoothing
 279 noise, MvSSIM is less sensitive, and Q_{2^n} is the least sensitive.

280 Similarly to the case of Gaussian white noise, we use MeanSSIM to con-
 281 sider the spatial structural similarity and use Q_λ to consider the spectral
 282 structural similarity, as plotted in in Figure 5b to evaluate the relative per-
 283 formances of MvSSIM, Q_m and Q_{2^n} . The value of Q_λ drops quickly when the
 284 standard deviation of the Gaussian smooth noise is larger than one, while
 285 the value of MeanSSIM is less sensitive to the Gaussian smoothing noise

286 compared with that of Q_λ .

287 When Q_λ largely decreases due to the noise, $Q2^n$ remains relatively sta-
 288 ble; this indicates that $Q2^n$ fails to respond well to the decrease in the spec-
 289 tral structural similarity introduced by the Gaussian smoothing noise. In
 290 contrast, Q_m reflects well the changes in the spectral structural similarity.
 291 However, Q_m fails to consider the strong spatial structural similarity as indi-
 292 cated by the big values of MeanSSIM. Compared with $Q2^n$ and Q_m , MvSSIM
 293 is a more desired candidate to assess the Gaussian smoothing noise contam-
 294 inated HSIs. It is between MeanSSIM and Q_λ , demonstrating a reasonable
 295 compromise between the spatial structural similarity and the spectral struc-
 296 tural similarity.

297 3.3.3. Savitzky-Golay smoothing noise

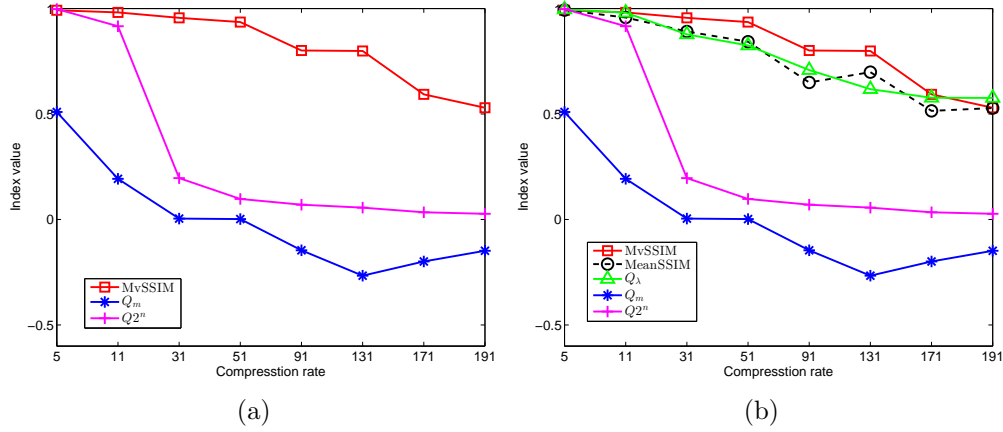


Figure 6: Assessments for the Savitzky-Golay smoothing noise contaminated HSIs.

298 Figure 6a shows the assessments for the HSIs contaminated by the Savitzky-
 299 Golay smoothing noise: Q_m is the most sensitive to the Savitzky-Golay
 300 smoothing noise, $Q2^n$ is less sensitive, and MvSSIM is the least sensitive.

301 Considering the behaviours of MeanSSIM and Q_λ in Figure 6b, the in-
 302 sensitive performance of MvSSIM is reasonable. It is obvious that Q_λ and
 303 MeanSSIM are not sensitive to the Savitzky-Golay spectral smoothing noise,
 304 i.e. neither the spatial and spectral structures are dramatically affected by
 305 the spectral smoothing noise. It makes sense that the spectral structural sim-
 306 ilarity is not largely affected by the Savitzky-Golay smoothing noise, because
 307 it is well known that the Savitzky-Golay filter can keep original signal struc-
 308 ture while removing noises with proper frame sizes [23]. Thus the large values
 309 of MvSSIM is reasonable as it assesses both spatial and spectral structural
 310 similarities. However, Q_m and Q^{2^n} provide small values when the values
 311 of MeanSSIM and Q_λ are still large, which indicates that Q_m and Q^{2^n} are
 312 over-sensitive to the spectral smoothing noise.

313 3.3.4. JPEG2000 compression noise

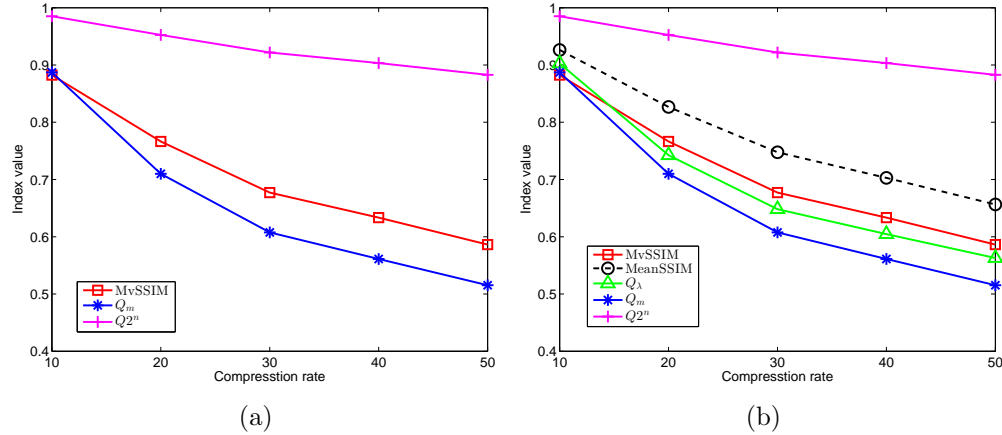


Figure 7: Assessments of the JPEG2000 compression noise contaminated HSIs.

314 Figure 7a shows the assessments of the HSIs contaminated by the JPEG2000
 315 compression noise: Q_m is the most sensitive to the JPEG2000 compression

316 noise, MvSSIM is less sensitive, and Q_{2n} is the least sensitive.

317 Considering the behaviours of MeanSSIM and Q_λ in Figure 7b, the com-
318 parative evaluation of MvSSIM, Q_m and Q_{2^n} is similar to that in 3.3.2:
319 Q_{2^n} does not manage to respond well to the spectral and spatial structural
320 changes; Q_m is over-sensitive to the JPEG2000 compression noise; and MvS-
321 SIM provides index values between Q_λ and MeanSSIM, which indicates that
322 MvSSIM more properly measures the influence of both spectral and spatial
323 structural similarities. Thus we can prefer MvSSIM for assessing the HSIs
324 contaminated by the JPEG2000 compression noise.

325 3.3.5. Summary

326 Two summaries could be made from these experiment results.

327 First, MvSSIM could provide appropriate assessments for noisy HSIs.

328 Second, as the indexes can perform differently for different kinds of noises,
329 by combining the performances of the indexes for a noisy HSI, we could
330 estimate the type of the noise added to the HSI based on the patterns of
331 the indexes, as suggested by [7]. For example, when MvSSIM is the least
332 sensitive to different levels of noises, there may be smoothing noise along the
333 spectral dimension.

334 4. Conclusion

335 In this paper, we proposed a new quality assessment method called MvS-
336 SIM for 3D HSI cubes. MvSSIM explores both spatial and spectral simi-
337 larities of HSI cubes. It can assess the similarities in both the within-band
338 spatial structure and the cross-band spectral structure, by treating each pixel
339 spectrum as a realisation of a multivariate random vector. The experiments

340 demonstrated that MvSSIM is a proper index of quality assessment for vari-
341 ous types of noises.

342 **Acknowledgements**

343 The authors are grateful to the anonymous reviewers for their constructive
344 comments.

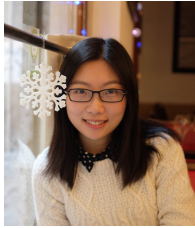
345 **References**

- 346 [1] Y. Yuan, X. Zheng, X. Lu, Spectral-spatial kernel regularized for hyper-
347 spectral image denoising, *IEEE Transactions on Geoscience and Remote*
348 *Sensing* 53 (7) (2015) 3815–3832.
- 349 [2] Q. Yuan, L. Zhang, H. Shen, Hyperspectral image denoising employing
350 a spectral-spatial adaptive total variation model, *IEEE Transactions on*
351 *Geoscience and Remote Sensing* 50 (10) (2012) 3660–3677.
- 352 [3] H. Zhang, W. He, L. Zhang, H. Shen, Q. Yuan, Hyperspectral image
353 restoration using low-rank matrix recovery, *IEEE Transactions on Geo-*
354 *science and Remote Sensing* 52 (8) (2014) 4729–4743.
- 355 [4] J. Li, Q. Yuan, H. Shen, L. Zhang, Hyperspectral image recovery em-
356 ploying a multidimensional nonlocal total variation model, *Signal Pro-*
357 *cessing* 111 (2015) 230–248.
- 358 [5] A. Garzelli, F. Nencini, Hypercomplex quality assessment of
359 multi/hyperspectral images, *IEEE Geoscience and Remote Sensing Let-*
360 *ters* 6 (4) (2009) 662–665.

- [6] L. Alparone, S. Baronti, A. Garzelli, F. Nencini, A global quality measurement of pan-sharpened multispectral imagery, *IEEE Geoscience and Remote Sensing Letters* 1 (4) (2004) 313–317.
- [7] E. Christophe, D. Léger, C. Mailhes, Quality criteria benchmark for hyperspectral imagery, *IEEE Transactions on Geoscience and Remote Sensing* 43 (9) (2005) 2103–2114.
- [8] J. Ren, J. Zabalza, S. Marshall, J. Zheng, Effective feature extraction and data reduction in remote sensing using hyperspectral imaging, *IEEE Signal Processing Magazine* 31 (4) (2014) 149–154.
- [9] Y.-Q. Zhao, J. Yang, Hyperspectral image denoising via sparse representation and low-rank constraint, *IEEE Transactions on Geoscience and Remote Sensing* 53 (1) (2015) 296–308.
- [10] M. Wang, J. Yu, J.-H. Xue, W. Sun, Denoising of hyperspectral images using group low-rank representation, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9 (9) (2016) 4420–4427.
- [11] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image quality assessment: from error visibility to structural similarity, *IEEE Transactions on Image Processing* 13 (4) (2004) 600–612.
- [12] N. Yun, Z. Feng, J. Yang, J. Lei, The objective quality assessment of stereo image, *Neurocomputing* 120 (2013) 121–129.
- [13] F. Zhou, Q. Liao, Single-frame image super-resolution inspired by perceptual criteria, *IET Image Processing* 9 (1) (2015) 1–11.

- 384 [14] W. Sun, F. Zhou, Q. Liao, MDID: A multiply distorted image database
385 for image quality assessment, *Pattern Recognition* 61 (2017) 153–168.
- 386 [15] Z. Wang, E. P. Simoncelli, A. C. Bovik, Multiscale structural similar-
387 ity for image quality assessment, in: *Signals, Systems and Computers*,
388 2004. Conference Record of the Thirty-Seventh Asilomar Conference on,
389 Vol. 2, IEEE, 2003, pp. 1398–1402.
- 390 [16] M. P. Sampat, Z. Wang, S. Gupta, A. C. Bovik, M. K. Markey, Com-
391 plex wavelet structural similarity: A new image similarity index, *IEEE*
392 *Transactions on Image Processing* 18 (11) (2009) 2385–2401.
- 393 [17] Z. Wang, Q. Li, Information content weighting for perceptual image
394 quality assessment, *IEEE Transactions on Image Processing* 20 (5)
395 (2011) 1185–1198.
- 396 [18] F. Zhou, Z. Lu, C. Wang, W. Sun, S.-T. Xia, Q. Liao, Image quality
397 assessment based on inter-patch and intra-patch similarity, *PloS one*
398 10 (3) (2015) e0116312.
- 399 [19] Y. Yuan, Q. Guo, X. Lu, Image quality assessment: A sparse learning
400 way, *Neurocomputing* 159 (2015) 227–241.
- 401 [20] R. Soundararajan, A. C. Bovik, Video quality assessment by reduced
402 reference spatio-temporal entropic differencing, *IEEE Transactions on*
403 *Circuits and Systems for Video Technology* 23 (4) (2013) 684–694.
- 404 [21] X. Li, Q. Guo, X. Lu, Spatiotemporal statistics for video quality assess-
405 ment, *IEEE Transactions on Image Processing* 25 (7) (2016) 3329–3342.

- 406 [22] G. Vivone, L. Alparone, J. Chanussot, M. Dalla Mura, A. Garzelli, G. A.
407 Licciardi, R. Restaino, L. Wald, A critical comparison among pansharp-
408 ening algorithms, *IEEE Transactions on Geoscience and Remote Sensing*
409 53 (5) (2015) 2565–2586.
- 410 [23] M. Browne, N. Mayer, T. R. Cutmore, A multiscale polynomial filter
411 for adaptive smoothing, *Digital Signal Processing* 17 (1) (2007) 69–75.



Rui Zhu received the B.S. degree in engineering from Xi-amen University in 2012 and the M.Sc. degree in statistics from University College London in 2013. She is currently working towards the Ph.D. degree in the Department of Statistical Science, University College London. Her research interests include spectral data analysis, hyperspectral image analysis, subspace-based classification methods and metric learning.



Fei Zhou received the B.Eng. degree in electronics and information engineering from the Huazhong University of Science and Technology, Wuhan, China in 2007, and the Ph.D. degree in electronics engineering from Tsinghua University, Beijing, China in 2013. Since 2013, he has been a Postdoctoral Fellow with the Graduate School at Shenzhen, Tsinghua University, Shenzhen, China. His research interests include applications of image processing and pattern recognition in video surveillance, image super-resolution, image interpolation, image quality assessment, and object tracking.



Jing-Hao Xue received the Dr.Eng. degree in signal and information processing from Tsinghua University in 1998 and the Ph.D. degree in statistics from the University of Glasgow in 2008. Since 2008, he has worked in the Department of Statistical Science at University College London as a Lecturer and Senior Lecturer. His current research interests include statistical classification, high-dimensional data analysis, pattern recognition and image analysis.