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)

Preprint submitted to Neurocomputing *February 18, 2017*

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

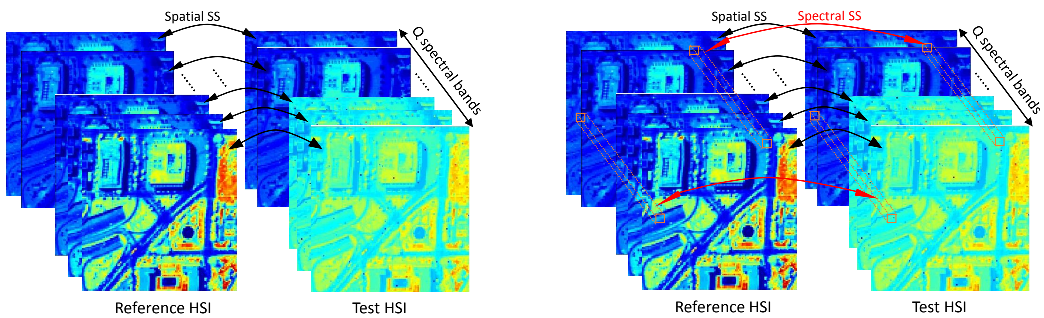
Keywords: Hyperspectral images, quality assessment, structural similarity (SSIM), spectral structure, spatial structure.

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 categories 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.



(a) MeanSSIM. (b) MvSSIM.

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

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

Compared with MeanSSIM, MvSSIM can assess both the within-band spatial structural similarity, between images of the same band, and the cross-band spectral structural similarity, between spectra of the same pixel, as illustrated in Figure 1b between a reference cube and a test cube. MvSSIM is well founded on multivariate statistics and can be easily implemented through simple multivariate sample statistics involving mean vectors, covariance matrices and cross-covariance matrices. Experiments show that MvSSIM is a proper quality assessment index for distorted HSIs with different kinds of 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 $\boldsymbol{x}$ and $\boldsymbol{y}$, both containing $N = a \times b$ pixels: $\boldsymbol{x} = [x_1, \ldots, x_N]^T \in \mathbb{R}^{N\times 1}$ and $\boldsymbol{y} = [y_1, \ldots, y_N]^T \in \mathbb{R}^{N\times 1}$, 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 $(i = 1, \ldots, N)$ are the realisations of random variables x and y, respectively.

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

$$\mathrm{SSIM}(\boldsymbol{x}, \boldsymbol{y}) = [l(\boldsymbol{x}, \boldsymbol{y})]^{\alpha} \times [c(\boldsymbol{x}, \boldsymbol{y})]^{\beta} \times [s(\boldsymbol{x}, \boldsymbol{y})]^{\gamma}, \tag{1}$$

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 $l(\boldsymbol{x}, \boldsymbol{y})$ is obtained by comparing the sample means $\bar{x}$ and $\bar{y}$:

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

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

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

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

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

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.

Third, the similarity of structure $s(\boldsymbol{x}, \boldsymbol{y})$ is calculated as the sample correlation coefficient of x and y:

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

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 correlation coefficient measures the linear dependency between x and y, indicating the similarity between two within-image spatial structures of the two images, which were vectorised into a pair of two N-element vectors. Thus $s(\boldsymbol{x}, \boldsymbol{y})$ is of great important in SSIM for assessing the spatial structural similarity of two images.

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

2.2. MeanSSIM

When SSIM is used in the quality assessment of HSI, it is commonly 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, $\boldsymbol{X}_H \in \mathbb{R}^{a\times b\times Q}$ and $\boldsymbol{Y}_H \in \mathbb{R}^{a\times b\times Q}$, where a and b represent the numbers of pixels in height and width, and Q is the number of spectral bands. $\boldsymbol{X}_H$ and $\boldsymbol{Y}_H$ can be rearranged as 2D matrices $\boldsymbol{X} = [\boldsymbol{x}_1^c, \boldsymbol{x}_2^c, \ldots, \boldsymbol{x}_Q^c] \in \mathbb{R}^{N\times Q}$ and $\boldsymbol{Y} = [\boldsymbol{y}_1^c, \boldsymbol{y}_2^c, \ldots, \boldsymbol{y}_Q^c] \in \mathbb{R}^{N\times Q}$, where $N = a \times b$ denotes the total number of pixels and $\boldsymbol{x}_q^c \in \mathbb{R}^{N\times 1}$ and $\boldsymbol{y}_q^c \in \mathbb{R}^{N\times 1}$ represent the image vectors of the qth spectral band of $\boldsymbol{X}_H$ and $\boldsymbol{Y}_H$, respectively. The MeanSSIM index is calculated as

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

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 $\boldsymbol{X}$ or $\boldsymbol{Y}$, 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.

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

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

As an extension of SSIM, MvSSIM also consists of three similarity measurements between $\boldsymbol{X}$ and $\boldsymbol{Y}$, i.e. $l(\boldsymbol{X}, \boldsymbol{Y})$, $c(\boldsymbol{X}, \boldsymbol{Y})$ and $s(\boldsymbol{X}, \boldsymbol{Y})$. These three similarities are defined on the following multivariate sample statistics of X and Y:

i) the sample means,

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

ii) the sample covariance matrices,

$$\boldsymbol{\Sigma}_X = \frac{1}{N-1}\sum_{n=1}^{N}(\boldsymbol{x}_n^r - \bar{X})(\boldsymbol{x}_n^r - \bar{X})^T \in \mathbb{R}^{Q\times Q}\ , \tag{7}$$

$$\boldsymbol{\Sigma}_Y = \frac{1}{N-1}\sum_{n=1}^{N}(\boldsymbol{y}_n^r - \bar{Y})(\boldsymbol{y}_n^r - \bar{Y})^T \in \mathbb{R}^{Q\times Q}; \tag{8}$$

and iii) the sample cross-covariance matrix,

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

Different from the univariate sample statistics in SSIM, the sample statistics in MvSSIM are vectors or matrices, rather than scalars. Thus the comparisons between scalars in SSIM should be extended to comparisons between vectors or matrices in MvSSIM. The extensions from $l(\boldsymbol{x},\boldsymbol{y})$, $c(\boldsymbol{x},\boldsymbol{y})$ and $s(\boldsymbol{x},\boldsymbol{y})$ to $l(\boldsymbol{X},\boldsymbol{Y})$, $c(\boldsymbol{X},\boldsymbol{Y})$ and $s(\boldsymbol{X},\boldsymbol{Y})$ are described as follows.

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

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

$$l(\boldsymbol{X},\boldsymbol{Y}) = \frac{2\langle\bar{X},\bar{Y}\rangle + C_1}{\langle\bar{X},\bar{X}\rangle + \langle\bar{Y},\bar{Y}\rangle + C_1} = \frac{2\sum\limits_{q=1}^{Q}\bar{x}_q\bar{y}_q + C_1}{\sum\limits_{q=1}^{Q}(\bar{x}_q^2 + \bar{y}_q^2) + C_1}, \tag{10}$$

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

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

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

Similar to $c(\boldsymbol{x}, \boldsymbol{y})$, $c(\boldsymbol{X}, \boldsymbol{Y})$ compares the similarity between sample covariance matrices $\boldsymbol{\Sigma}_X$ and $\boldsymbol{\Sigma}_Y$. A sample covariance matrix (e.g. $\boldsymbol{\Sigma}_X$) contains the variances within individual bands (of $\boldsymbol{X}$) in its diagonal entries, and the covariances between different spectral bands (of $\boldsymbol{X}$) in its off-diagonal entries. Hence when we compare $\boldsymbol{X}$ and $\boldsymbol{Y}$ through $\boldsymbol{\Sigma}_X$ and $\boldsymbol{\Sigma}_Y$, we can achieve two comparisons simultaneously: comparing the contrasts of two images of the same band via the two standard deviations of this band, and comparing the contrasts of two spectra of the same spatial position via the covariances between different bands.

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

$$c(\boldsymbol{X}, \boldsymbol{Y}) = \frac{2||\boldsymbol{\Sigma}_X||_*^{\frac{1}{2}}||\boldsymbol{\Sigma}_Y||_*^{\frac{1}{2}} + C_2}{||\boldsymbol{\Sigma}_X||_* + ||\boldsymbol{\Sigma}_Y||_* + C_2} = \frac{2\sqrt{\lambda^{(s)}}\sqrt{d^{(s)}} + C_2}{\lambda^{(s)} + d^{(s)} + C_2}, \tag{11}$$

where $||\ ||_*$ 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 singular values of $\boldsymbol{\Sigma}_X$, and d_q's are the singular values of $\boldsymbol{\Sigma}_Y$.

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

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

The term $s(\boldsymbol{x}, \boldsymbol{y})$ 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 $s(\boldsymbol{X}, \boldsymbol{Y})$ as

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

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

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

2.3.4. MvSSIM

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

$$\mathrm{MvSSIM}(\boldsymbol{X}, \boldsymbol{Y}) = [l(\boldsymbol{X}, \boldsymbol{Y})]^{\alpha} \times [c(\boldsymbol{X}, \boldsymbol{Y})]^{\beta} \times [s(\boldsymbol{X}, \boldsymbol{Y})]^{\gamma}, \tag{13}$$

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

Among these three terms, $l(\boldsymbol{X}, \boldsymbol{Y})$ and $s(\boldsymbol{X}, \boldsymbol{Y})$ measure the similarity between band images in luminance and spatial structure, while $c(\boldsymbol{X}, \boldsymbol{Y})$ measures the similarity between both band images and pixel spectra. Thus in MvSSIM, both the within-band spatial structural similarity and the cross-band spectral structural similarity are assessed.

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

3. Experiments

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

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

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

3.1. Dataset

The Washington DC HSI is used for the synthetic experiments. The Washington DC HSI is a Hyperspectral Digital Imagery Collection Experiment (HYDICE) image of Washington DC Mall and can be downloaded from https://engineering.purdue.edu/~landgreb/Hyperspectral.Ex.html. The dataset is of size $1208 \times 1208 \times 191$, where 1208×1208 is the spatial size of the HSI and 191 is the number of bands.

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

3.2. Experiment settings

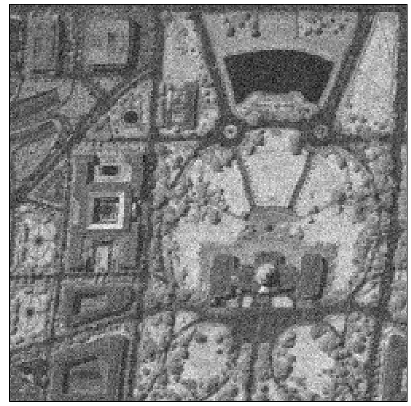
MeanSSIM is computed using the MATLAB function 'ssim' with the default setting: window size is 11, $C_1 = 0.01$ and $C_2 = 0.03$. For MvSSIM, a patch of size $5 \times 5 \times 191$ moves from pixel to pixel, the index of each patch is calculated, and then the mean index of all the patches is taken as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

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

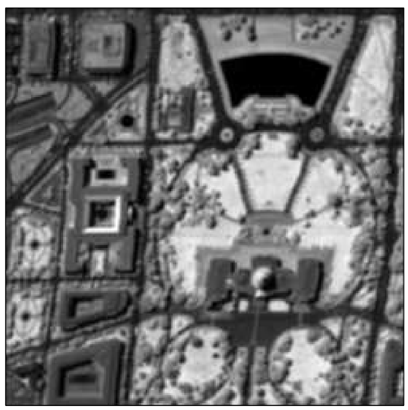
First, Gaussian white additive noises are added to 50 randomly-selected bands of the spectra. We test 10 different variances: from 10 to 100 with a step of 10, i.e. 10 different noisy HSIs are created with different variances.

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).

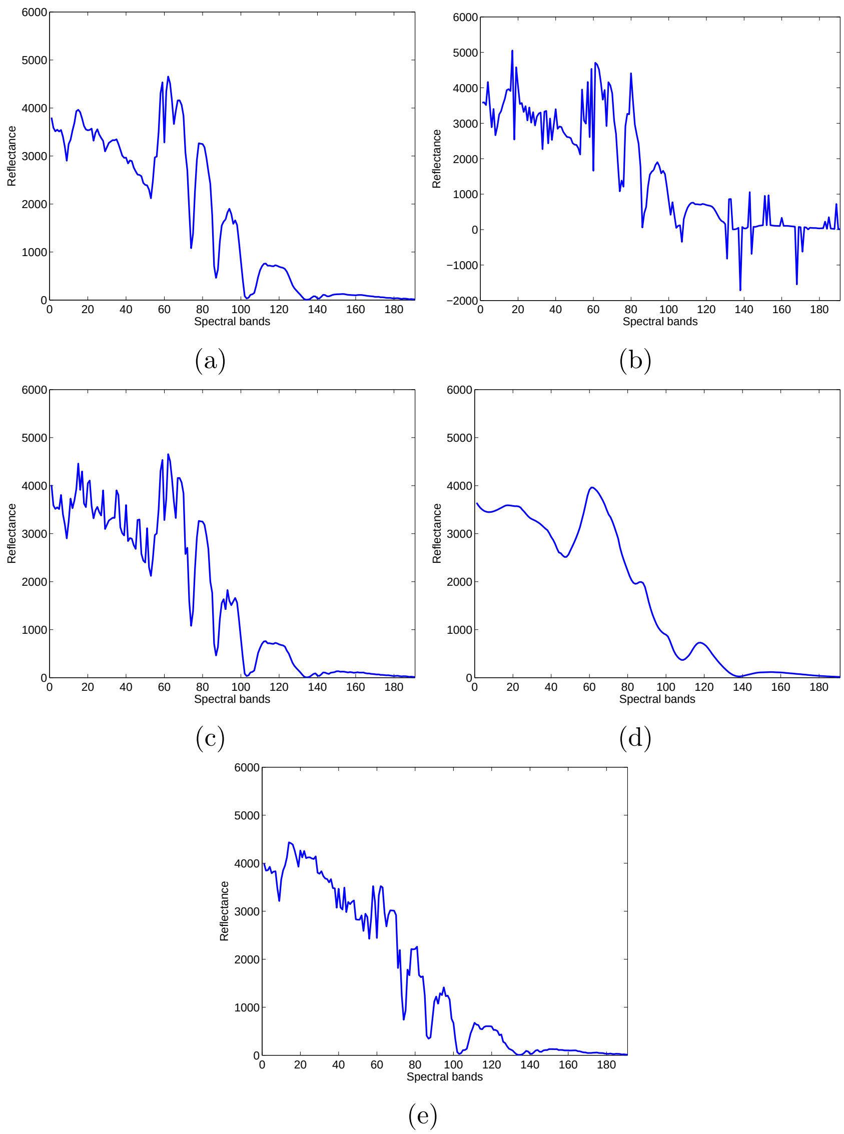


Figure 3: The reference spectrum and noisy spectra of the pixel at position (50, 50). (a) Reference. (b) Gaussian white noise (variance 60). (c) Gaussian smoothing noise (standard deviation 1). (d) Savitzky-Golay smoothing noises (frame size 11). (e) JPEG2000 compression (compression ratio 30).

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.

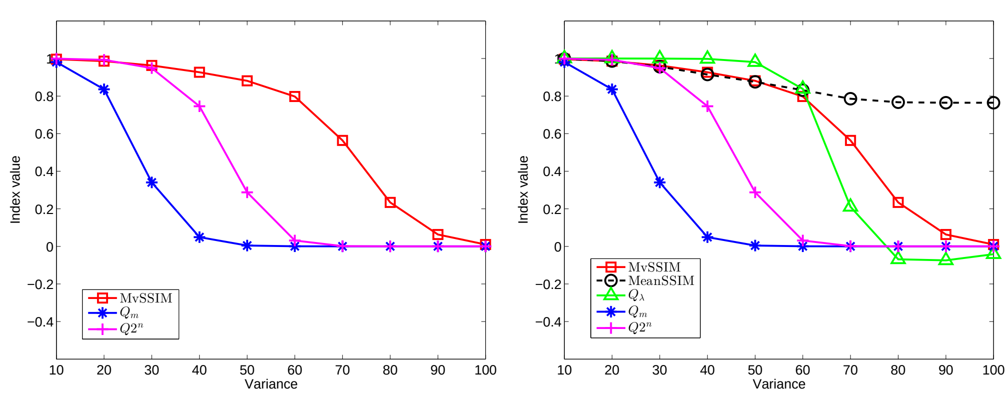
Third, Savitzky-Golay smoothing filter is applied to the spectra of all pixels to create smooth spectra, i.e. in the spectral dimension of the HSI. We test eight different frame sizes: 5, 11, 31, 71, 91, 131, 171 and 191, i.e. eight different noisy HSIs are created with different frame sizes..

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



(a) (b)

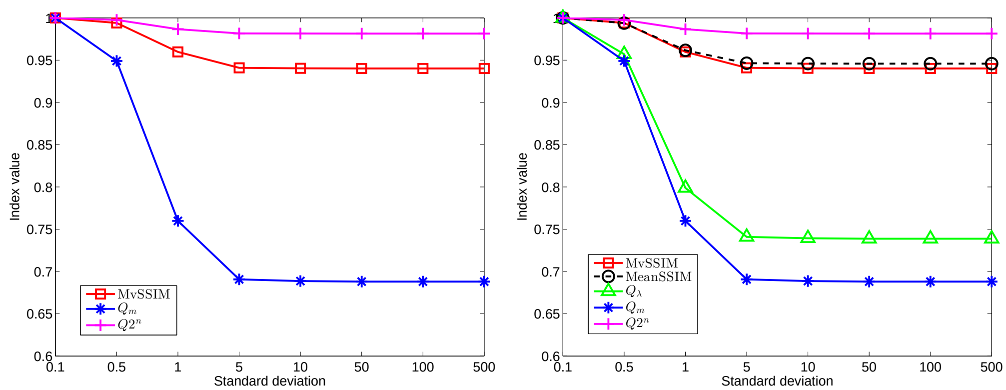
Figure 4: Assessments for the Gaussian white additive noise contaminated HSIs.

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, $Q2^n$ 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 $Q2^n$ are close to zero even when the values of Q_λ are still close to one; this indicates that Q_m and $Q2^n$ fail to consider the high spectral structural similarity in this case and are over-sensitive to the Gaussian white noise. In contrast, MvSSIM provides large values when the values of Q_λ are large. Also, compared with Q_λ, MvSSIM is more desired because it also reflects the spatial structural similarity, making it between MeanSSIM and Q_λ in the case of Gaussian white noise.

3.3.2. Gaussian smoothing noise



(a) (b)

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

Figure 5a shows the assessments for the HSIs contaminated by the Gaussian smoothing noise: Q_m is the most sensitive to the Gaussian smoothing noise, MvSSIM is less sensitive, and $Q2^n$ is the least sensitive.

Similarly to the case of Gaussian white noise, we use MeanSSIM to consider the spatial structural similarity and use Q_λ to consider the spectral structural similarity, as plotted in in Figure 5b to evaluate the relative performances of MvSSIM, Q_m and $Q2^n$. The value of Q_λ drops quickly when the standard deviation of the Gaussian smooth noise is larger than one, while the value of MeanSSIM is less sensitive to the Gaussian smoothing noise compared with that of Q_λ.

When Q_λ largely decreases due to the noise, $Q2^n$ remains relatively stable; this indicates that $Q2^n$ fails to respond well to the decrease in the spectral structural similarity introduced by the Gaussian smoothing noise. In contrast, Q_m reflects well the changes in the spectral structural similarity. However, Q_m fails to consider the strong spatial structural similarity as indicated by the big values of MeanSSIM. Compared with $Q2^n$ and Q_m, MvSSIM is a more desired candidate to assess the Gaussian smoothing noise contaminated HSIs. It is between MeanSSIM and Q_λ, demonstrating a reasonable compromise between the spatial structural similarity and the spectral structural similarity.

3.3.3. *Savitzky-Golay smoothing noise*

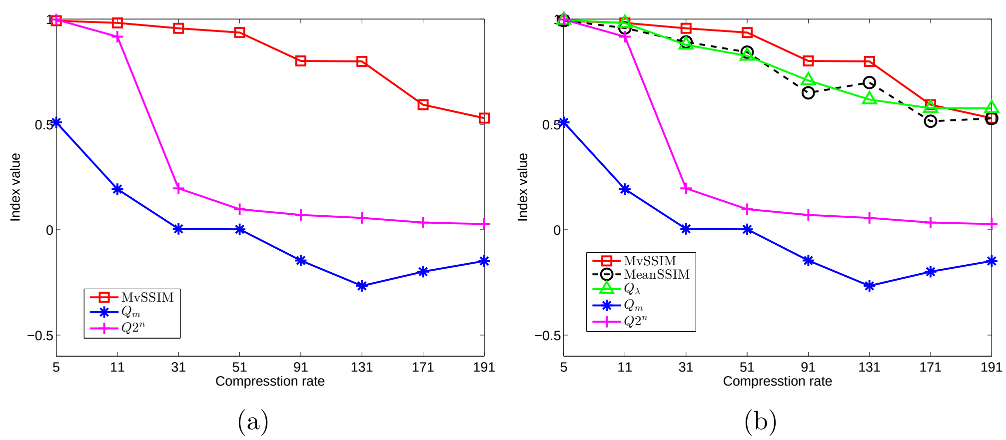
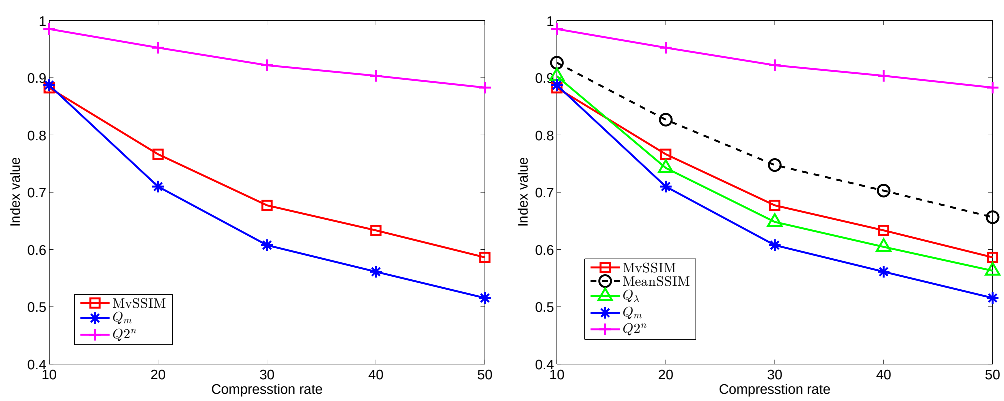


Figure 6: Assessments for the Savitzky-Golay smoothing noise contaminated HSIs.

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

Considering the behaviours of MeanSSIM and Q_λ in Figure 6b, the insensitive performance of MvSSIM is reasonable. It is obvious that Q_λ and MeanSSIM are not sensitive to the Savitzky-Golay spectral smoothing noise, i.e. neither the spatial and spectral structures are dramatically affected by the spectral smoothing noise. It makes sense that the spectral structural similarity is not largely affected by the Savitzky-Golay smoothing noise, because it is well known that the Savitzky-Golay filter can keep original signal structure while removing noises with proper frame sizes [23]. Thus the large values of MvSSIM is reasonable as it assesses both spatial and spectral structural similarities. However, Q_m and $Q2^n$ provide small values when the values of MeanSSIM and Q_λ are still large, which indicates that Q_m and $Q2^n$ are over-sensitive to the spectral smoothing noise.

3.3.4. JPEG2000 compression noise



(a) (b)

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

Figure 7a shows the assessments of the HSIs contaminated by the JPEG2000 compression noise: Q_m is the most sensitive to the JPEG2000 compression noise, MvSSIM is less sensitive, and $Q2n$ is the least sensitive.

Considering the behaviours of MeanSSIM and Q_λ in Figure 7b, the comparative evaluation of MvSSIM, Q_m and $Q2^n$ is similar to that in 3.3.2: $Q2^n$ does not manage to respond well to the spectral and spatial structural changes; Q_m is over-sensitive to the JPEG2000 compression noise; and MvSSIM provides index values between Q_λ and MeanSSIM, which indicates that MvSSIM more properly measures the influence of both spectral and spatial structural similarities. Thus we can prefer MvSSIM for assessing the HSIs contaminated by the JPEG2000 compression noise.

3.3.5. Summary

Two summaries could be made from these experiment results.

First, MvSSIM could provide appropriate assessments for noisy HSIs.

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

4. Conclusion

In this paper, we proposed a new quality assessment method called MvSSIM for 3D HSI cubes. MvSSIM explores both spatial and spectral similarities of HSI cubes. It can assess the similarities in both the within-band spatial structure and the cross-band spectral structure, by treating each pixel spectrum as a realisation of a multivariate random vector. The experiments demonstrated that MvSSIM is a proper index of quality assessment for various types of noises.

Acknowledgements

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

References

[1] Y. Yuan, X. Zheng, X. Lu, Spectral–spatial kernel regularized for hyperspectral image denoising, IEEE Transactions on Geoscience and Remote Sensing 53 (7) (2015) 3815–3832.

[2] Q. Yuan, L. Zhang, H. Shen, Hyperspectral image denoising employing a spectral–spatial adaptive total variation model, IEEE Transactions on Geoscience and Remote Sensing 50 (10) (2012) 3660–3677.

[3] H. Zhang, W. He, L. Zhang, H. Shen, Q. Yuan, Hyperspectral image restoration using low-rank matrix recovery, IEEE Transactions on Geoscience and Remote Sensing 52 (8) (2014) 4729–4743.

[4] J. Li, Q. Yuan, H. Shen, L. Zhang, Hyperspectral image recovery employing a multidimensional nonlocal total variation model, Signal Processing 111 (2015) 230–248.

[5] A. Garzelli, F. Nencini, Hypercomplex quality assessment of multi/hyperspectral images, IEEE Geoscience and Remote Sensing Letters 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.

[14] W. Sun, F. Zhou, Q. Liao, MDID: A multiply distorted image database for image quality assessment, Pattern Recognition 61 (2017) 153–168.

[15] Z. Wang, E. P. Simoncelli, A. C. Bovik, Multiscale structural similarity for image quality assessment, in: Signals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on, Vol. 2, IEEE, 2003, pp. 1398–1402.

[16] M. P. Sampat, Z. Wang, S. Gupta, A. C. Bovik, M. K. Markey, Complex wavelet structural similarity: A new image similarity index, IEEE Transactions on Image Processing 18 (11) (2009) 2385–2401.

[17] Z. Wang, Q. Li, Information content weighting for perceptual image quality assessment, IEEE Transactions on Image Processing 20 (5) (2011) 1185–1198.

[18] F. Zhou, Z. Lu, C. Wang, W. Sun, S.-T. Xia, Q. Liao, Image quality assessment based on inter-patch and intra-patch similarity, PloS one 10 (3) (2015) e0116312.

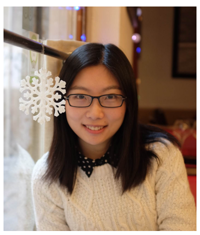
[19] Y. Yuan, Q. Guo, X. Lu, Image quality assessment: A sparse learning way, Neurocomputing 159 (2015) 227–241.

[20] R. Soundararajan, A. C. Bovik, Video quality assessment by reduced reference spatio-temporal entropic differencing, IEEE Transactions on Circuits and Systems for Video Technology 23 (4) (2013) 684–694.

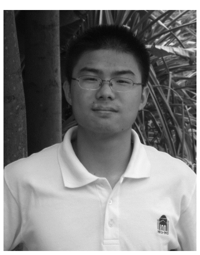
[21] X. Li, Q. Guo, X. Lu, Spatiotemporal statistics for video quality assessment, IEEE Transactions on Image Processing 25 (7) (2016) 3329–3342.

[22] G. Vivone, L. Alparone, J. Chanussot, M. Dalla Mura, A. Garzelli, G. A. Licciardi, R. Restaino, L. Wald, A critical comparison among pansharpening algorithms, IEEE Transactions on Geoscience and Remote Sensing 53 (5) (2015) 2565–2586.

[23] M. Browne, N. Mayer, T. R. Cutmore, A multiscale polynomial filter for adaptive smoothing, Digital Signal Processing 17 (1) (2007) 69–75.



Rui Zhu received the B.S. degree in engineering from Xiamen 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.