Hyperspectral image noise reduction based on rank-1 tensor decomposition

Xian Guo [a], Xin Huang [a,*], Liangpei Zhang [a], Lefei Zhang [b]

The State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS), Wuhan University, Wuhan, Hubei 430079, PR China
Computer School, Wuhan University, Wuhan, Hubei 430079, PR China

ARTICLE INFO

ABSTRACT

In this study, a novel noise reduction algorithm for hyperspectral imagery (HSI) is proposed based on high-order rank-1 tensor decomposition. The hyperspectral data cube is considered as a three-order tensor that is able to jointly treat both the spatial and spectral modes. Subsequently, the rank-1 tensor decomposition (R1TD) algorithm is applied to the tensor data, which takes into account both the spatial and spectral information of the hyperspectral data cube. A noise-reduced hyperspectral image is then obtained by combining the rank-1 tensors using an eigenvalue intensity sorting and reconstruction technique. Compared with the existing noise reduction methods such as the conventional channel-by-channel approaches and the recently developed multidimensional filter, the spatial–spectral adaptive total variation filter, experiments with both synthetic noisy data and real HSI data reveal that the proposed R1TD algorithm significantly improves the HSI data quality in terms of both visual inspection and image quality indices. The subsequent image classification results further validate the effectiveness of the proposed HSI noise reduction algorithm.

1. Introduction

The development of hyperspectral remote sensing technology makes it possible to provide a large amount of spatial and spectral information for image analysis applications such as classification, unmixing, subpixel mapping, and target detection (Chang, 2003; Landgrebe, 2002). However, the acquired hyperspectral images (HSI) are often disturbed by radiometric noise such as sensor noise, photon (or shot) noise, calibration error, atmospheric scattering and absorption (Kerekes and Baum, 2005), which not only degrades the visual quality of the HSI data but also limits the precision of the subsequent image interpretation and analysis (Matteoli et al., 2011), for example, in classification, target detection, subpixel mapping, etc. Therefore, it is critical to remove the noise and retain the signal component before the subsequent process. The noise in hyperspectral imagery (HSI) can generally be categorized into two classes: random noise and fixed-pattern noise. Fixed-pattern noise like striping, generated during the calibration process, can be mitigated by a suitable model (Acito et al., 2011b). In contrast, random noise cannot be removed entirely, due to its stochastic nature. One widely used random noise model in HSI is the additive model, which is assumed to be white,

Gaussian, and independent-from-signal. However, with the improvement in the sensitivity of hyperspectral sensors, in some cases, the dominant noise source is no longer determined by signal-independent additive noise, but a mixture of signal-independent noise, signal-dependent noise, and fixed-pattern noise (Acito et al., 2010). Acito et al. (2011a) investigated the random noise estimation problem for HSI. Their newly developed model takes into account the signal-dependent noise contribution and is suitable for noise characterization in data where the signal-independent noise is not dominant. Bioucas-Dias and Figueiredo (2010) described a new approach to solve the optimization problem resulting from a variational estimation of images observed under multiplicative noise models. Total variation (TV) regularization was used as the prior and an augmented Lagrangian method was applied to the constrained problem.

Since the additive noise model is the situation generally found in HSI, many algorithms have been derived that are based on this model. The traditional methods employ denoising algorithms such as singular value decomposition (SVD) (Andrews and Patterson, 1976) and Wiener and wavelet filters, channel-by-channel (Banham and Katsaggelos, 1997). However, these algorithms may lead to a loss of the inter-dimensional information since the correlation between the spatial and spectral bands is not simultaneously considered. In recent years, some algorithms have been proposed to combine the spatial and spectral information for HSI noise

* Corresponding author.
   E-mail address: huang_whu@163.com (X. Huang).

reduction. Othman and Qian (2006) proposed a hybrid spatial–spectral derivative-domain wavelet shrinkage noise reduction (HSSNR) approach. Chen and Qian (2008) proposed to simultaneously reduce the dimensionality and noise of HSI by the use of bivariate wavelet shrinkage. Yuan et al. (2012) also presented a spectral-spatial adaptive total variation model for hyperspectral image denoising. Another type of HSI noise reduction algorithm for removing striping artifacts is based on wavelet transform and adaptive frequency domain filtering (Pande-Chhetri and Abd-Elrahman, 2011).

In multilinear algebra, the hyperspectral image data cube can be considered as a three-order tensor in which the spatial and spectral information is completely preserved (Zhang et al., 2008). Based on the tensor model and multilinear algebra, it has been shown that tensor representation can simultaneously deal with the two spatial dimensions and one spectral dimension of HSI to achieve a satisfying noise reduction performance. Examples of such approaches include multidimensional filtering based on Tucker tensor decomposition (Bourennane et al., 2011; Muti and Bourennane, 2005; Renard and Bourennane, 2008; Renard et al., 2008), and the use of the kernel trick in Tucker decomposition (Karami et al., 2011). The multidimensional Wiener filtering (MWF) algorithm is one of these Tucker-based noise reduction algorithms which jointly takes into account the spatial–spectral information and achieves a simultaneous improvement in image quality and classification accuracy. However, the application of a core tensor and n-mode tensor product may lead to information compression and loss of spatial detail (Letexier and Bourennane, 2008).

In this study, we develop a new tensor decomposition which represents a noisy image by a series of separate signal and noise profiles, and we estimate the noise-free image by extracting the signal-dominant component. Specifically, we present an HSI noise reduction algorithm based on rank-1 tensor decomposition (R1TD). As mentioned above, the HSI data cube is considered as a three-order tensor that is able to jointly treat both the spatial and spectral modes. The R1TD algorithm is then applied to the tensor data input, which takes into account both the spatial and spectral information of the hyperspectral data cube. Finally, the noise-reduced HSI is obtained by combining the rank-1 tensors by the use of an eigenvalue intensity sorting and reconstruction technique.

The remainder of this study is organized as follows. Section 2 introduces the definitions of tensor and multilinear algebra. Section 3 addresses the rank-1 tensor decomposition algorithm and decomposition-level estimator. Experiments on two widely used hyperspectral data sets are implemented in Section 4. Finally, Section 5 concludes this study.

2. Background

Tensor (Lathauwer, 1997) and multilinear algebra have been receiving more and more attention and have recently been applied to both computer vision and pattern recognition (Lu et al., 2008; Shashua and Levin, 2001; Tao et al., 2007; Vasilescu and Terzopoulos, 2003). A tensor, represented as $A \in R^{L_1 \times L_2 \times \cdots \times L_N}$, is defined as a multidimensional array, which is the higher-order equivalent of a vector (one-order tensor) and a matrix (two-order tensor). According to the above definition, tensor A's order is N, and each order is called the ith mode. An arbitrary element of A is a scalar denoted by $a_{l_1, l_2, \ldots, l_i, \ldots, l_N}$, where $1 \leqslant l_i \leqslant L_i$ and $1 \leqslant i \leqslant N$, with l_i being the location of this element in the ith mode. Specifically $A \in R^{L_1 \times L_2}$ is a two-mode tensor or matrix. Each column is defined as a_i, and an arbitrary element is denoted by a_{ij}, where $1 \leqslant i \leqslant L_1, 1 \leqslant j \leqslant L_2$. In this study, the HSI data cube is regarded as a three-order tensor $A \in R^{L_1 \times L_2 \times L_3}$, in which modes 1 and 2 represent the spatial modes, and mode 3 denotes the spectral mode. In this subsection, we give a brief review

of the relevant concepts for the matrix and tensor in multilinear algebra (Kolda and Bader, 2009).

Definition 1. Rank-1 tensor: Define the outer product of a tensor $A \in R^{I_1 \times I_2 \times \cdots \times I_p}$ and another tensor $B \in R^{J_1 \times J_2 \times \cdots \times J_q}$ as $A \circ B$. All the values of the indices are denoted by

$$(A \circ B)_{i_1 i_2, \ldots, i_p j_1 j_2, \ldots, j_q} = a_{i_1 i_2, \ldots, i_p} b_{j_1 j_2, \ldots, j_q} \tag{1}$$

Let $V_1, V_2, \ldots, V_N$ be N vectors in the Euclidean space with finite dimensions $I_1, I_2, \ldots, I_N$. Consider N vectors $u_1 \in V_1, u_2 \in V_2, \ldots, u_N \in V_N$. An N-mode $X \in R^{I_1 \times I_2 \times \cdots \times I_N}$ tensor is rank-1 if it can be written as the outer product of N vectors:.

$$X = u_1 \circ u_2 \circ \cdots \circ u_N \tag{2}$$

Definition 2. Tensor matricization. Also known as n-mode flattening or unfolding, tensor matricization reorders the elements of an N-order tensor into a matrix from a given mode. The n-mode matricization of $X \in R^{L_1 \times L_2 \times \cdots \times L_N}$ is $mat_n X \in R^{L_i \times (L_1 L_2 \ldots L_{n-1} L_{n+1} \ldots L_N)}$, which is the ensemble of vectors in the n-mode obtained by keeping index L_i fixed and varying the other indices. A visual illustration of tensor matricization is shown in Fig. 1.

Definition 3. The Kronecker product. The Kronecker product of matrices $A \in R^{J \times K}$ and $B \in R^{M \times N}$ is denoted by $A \otimes B$:

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1K}B \\ a_{21}B & a_{22}B & \cdots & a_{2K}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{J1}B & a_{J2}B & \cdots & a_{JK}B \end{bmatrix} \in R^{(JM) \times (KN)} \tag{3}$$

Definition 4. The Khatri–Rao product. It can be regarded as the column-wise Kronecker product. The Khatri–Rao product of matrices $A = [a_1, a_2, \ldots, a_K] \in R^{I \times K}$ and $B = [b_1, b_2, \ldots, b_K] \in R^{J \times K}$ is determined by $A \odot B$:

$$A \odot B = [a_1 \otimes b_1, a_2 \otimes b_2, \ldots, a_K \otimes b_K] \in R^{IJ \times K} \tag{4}$$

Fig. 2 shows how the Khatri–Rao product works.

3. Methodology: R1TD noise reduction model

The flowchart of the proposed R1TD algorithm for HSI noise reduction is shown in Fig. 3. The input HSI data cube is considered as a three-order tensor. Subsequently, the rank-1 tensor decomposition (R1TD) algorithm is used to extract the signal-dominant component from the observed HSI data cube by sorting the eigenvalues generated by tensor decomposition. Finally, the denoised HSI is obtained by combining the signal rank-1 profiles, shown as the red cube in Fig. 3. It is worth noting that the size of each red and dark red cube is equivalent to that of the input data. As the signal-dominant component and the noise component are defined as the sum of the red cubes and dark red cubes, respectively, each component cube size is also equal to the size of the input data.

3.1. Derivation of the R1TD noise reduction model

The HSI data can be regarded as a third-order tensor that completely preserves the spatial and spectral information. Here, we denote O as the observed HSI data cube consisting of the signal-dominant component S and the additive noise component N. By extending the classic two-dimensional additive noise model, the tensorial formulation is:
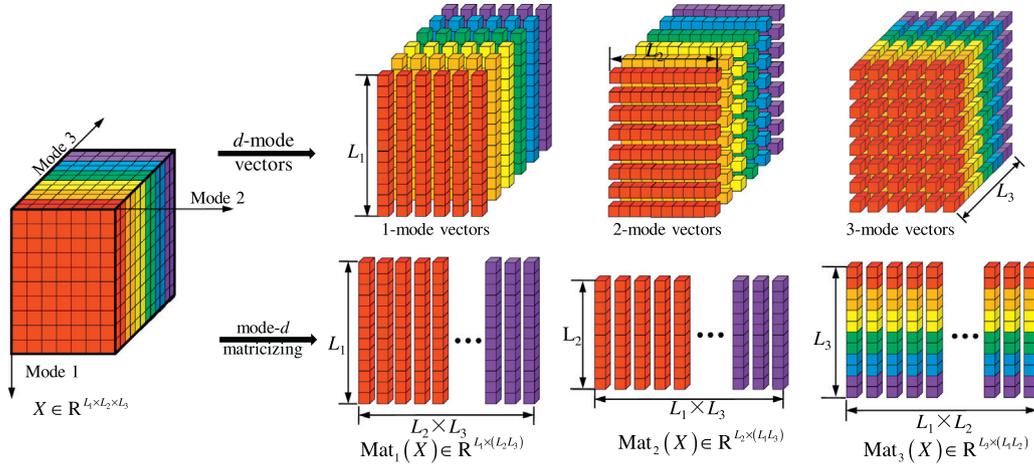
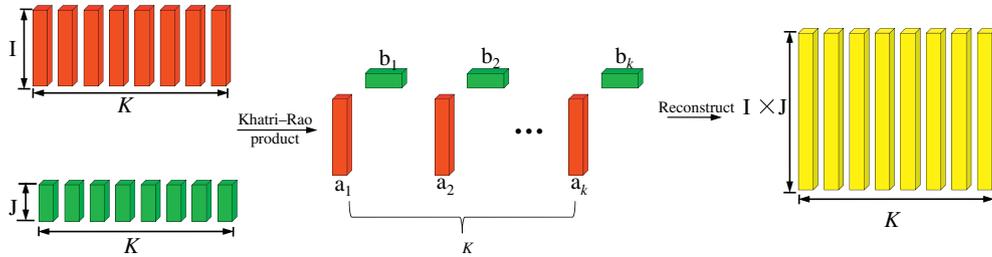**Fig. 1.** Illustration of tensor matricization in three modes.

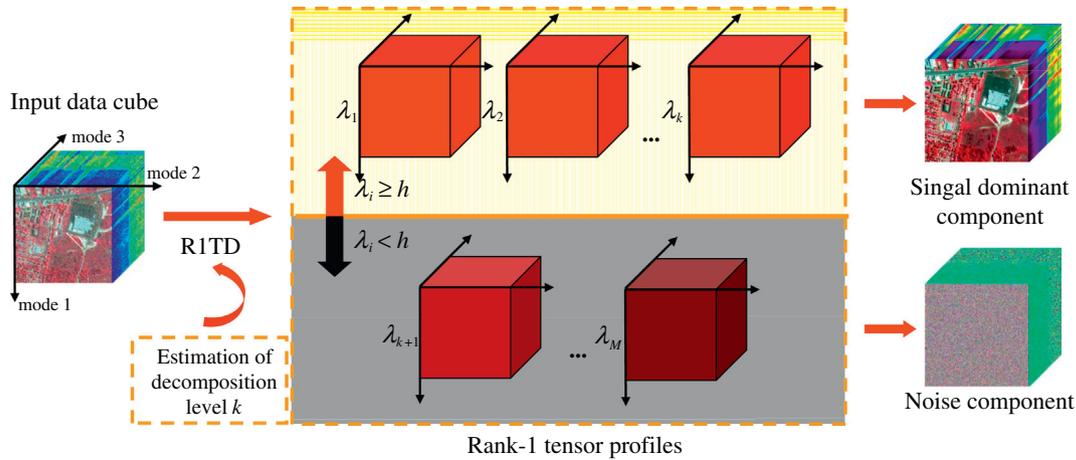**Fig. 2.** Visual illustration of the Khatri–Rao product.

Fig. 3. Main flowchart of the R1TD algorithm for HSI noise reduction.

$$O = S + N \qquad (5)$$

In this model, the noise is assumed to be white, Gaussian, and independent-from-signal. Based on the definitions of the rank-1 tensor and vector outer product, tensor $O \in R^{L_1 \times L_2 \times L_3}$ can be represented with the rank-1 tensor decomposition model:

$$O = \sum_{r=1}^{M} \lambda_r u_r \circ v_r \circ w_r, \qquad (6)$$

where $u_r \in R^{L_1}$, $v_r \in R^{L_2}$, and $w_r \in R^{L_3}$ ($r = 1, 2, \ldots, M$) are vectors (rank-1 tensors in this model) on three modes, and M is the number of rank-1 tensors used to restore the whole tensor O. Considering λ_r as the weight value, Eq. (6) implies that the HSI data is a linear combination of a sequence of rank-1 tensors (Bro and Kiers,

2003). However, there is currently no straightforward solution to M or the so-called tensor rank. The rationale of this problem is explained as follows: The rank of a three-order tensor (M in this model) is equivalent to the minimal number of triads necessary to describe the tensor. However, due to the special structure of the multilinear model compared to the bilinear one, there are no explicit rules for determining the tensor rank in general, except for the two-order case and some simple three-order tensors. Refer to Harshman (1970) and Kruskal (1976) for further details.

A common assumption for additive noise in subspace analysis is that the useful signals in HSI are highly correlated between the spectral channels, and the noise is accordingly less correlated because of its random distribution. R1TD weights can be used to indicate the correlation between each rank-1 profiles and the signals,

hence, are used here to distinguish the signal and noise profiles, which define the contribution of the rank-1 profiles to the reconstructed signal-dominant component. We therefore propose to extract the signal-dominant component from the observed data cube by sorting the weights of the rank-1 tensors, rather than finding the tensor rank of the noisy data. After the noise component of the input data cube is removed, the signal-dominant component is obtained by reconstructing the remaining rank-1 tensors. Based on this idea, the estimated signal-dominant component can be written as Eq. (7), leading to a linear combination of a series of the rank-1 tensors:

$$\widehat{S} = \sum_{r=1}^{k} \lambda_r u_r \circ v_r \circ w_r \tag{7}$$

where k is the decomposition level in this study, and it refers to the number of rank-1 tensors corresponding to the signal-dominant component, and is smaller than the value of M. Although the tensor rank M is difficult to calculate, it is unimportant in the R1TD model, in which the signal-dominant component is reconstructed from the k rank-1 profiles. Consequently, as long as an appropriate estimation of the decomposition rank k is performed, the denoising will be completed.

3.2. Solution for the rank-1 tensor in the R1TD procedure

A series of the rank-1 tensors by R1TD should be estimated so that the restored tensor $\widehat{S}$ is as close as possible to the noise-free tensor (or signal-dominant component) S, i.e., we minimize the mean squared error (MSE) between the ideal signal tensor and the reconstructed signal tensor:

$$\min_{\lambda,u_r,v_r,w_r} \left\| S - \sum_{r=1}^{k} \lambda_r u_r \circ v_r \circ w_r \right\|^2 \tag{8}$$

In this study, we combine the vectors of each rank-1 tensor in each mode into a factor matrix, i.e., $U = [u_1, u_2, \ldots, u_k]$, $V = [v_1, v_2, \ldots, v_k]$, and $W = [w_1, w_2, \ldots, w_k]$. In addition, we also denote Λ as a matrix form, i.e., $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_k)$. According to tensor matricization, the mode-1 flattening of Eq. (7) should be expressed as:

$$mat_1\widehat{S} = U\Lambda(W \odot V)^T \tag{9}$$

In the alternating optimization, the solution of (8) is optimized in each mode. Each time, only one factor matrix is optimized by the other fixed factor matrices. Here, we give the derivation to optimize U by fixing V and W in mode 1 as:

$$\min_{U} \|mat_1 S - U^*(W \odot V)^T\|^2 \tag{10}$$

where U^* is the weighted factor matrix computed by $U^* = U \cdot \Lambda$ or $U^*_{(1)} = U_{(i)} \cdot \Lambda_{(i,i)}$, and $i = 1, 2, \ldots, k$ since Λ is a diagonal matrix. The minimization (10) is a linear least-squares problem, and its solution is written as follows:

$$U^* = mat_1 S \cdot ((W \odot V)^T)^{-1}$$
$$= mat_1 S \cdot (W^T W \cdot V^T V)^{-1} (W \odot V)^T \tag{11}$$

where U^* is the weighted version of the mode-1 factor matrix U. In order to achieve a unique solution for the factor matrices, it is assumed that the columns of U, V, and W are normalized to length one (Sidiropoulos and Bro, 2000), i.e. $\|U_{(i)}\| = \|V_{(i)}\| = \|W_{(i)}\| = 1$ for $i = 1, 2, \ldots, k$. Thus, the solution for the mode-1 factor matrix (Kolda and Bader, 2009) should be:

$$\Lambda_{(i,i)} = \|U^*_{(i)}\|, \quad U_{(i)} = U^*_{(i)}/\Lambda_{(i,i)}, \quad i = 1, 2, \ldots, k \tag{12}$$

The objective function of (8) can therefore be solved by iteratively optimizing each factor matrix while keeping the other matrices fixed until the convergence criteria is met. In this study, since the ideal noise-free tensor S may not be identified in practice, the R1TD algorithm utilizes the input tensor O as the initialization value of S. In the tth round of iteration, S is replaced by the estimated signal tensor $\widehat{S}$ in the $(t-1)$th round of iteration. The algorithm converges when the error of the estimated signal tensor $\widehat{S}$ between two iterations decreases to a small value. Table 1 summarizes the proposed R1TD algorithm for HSI noise reduction.

3.3. Estimation of the n-mode rank and the decomposition-level selection

The decomposition level k is correlated with the tensor rank, which is the minimum number of rank-1 tensors necessary to describe a tensor (Kolda and Bader, 2009). Due to the special structure of the multilinear model compared to the bilinear one, it is difficult to find an accurate method in the literature to perform rank estimation for multilinear data. In this study, a new criterion is proposed for the decomposition-level selection. The n-mode rank criterion is formulated by:

$$k = \text{SNR} \cdot \prod_{i=1}^{n} I_i \bigg/ \prod_{i=1}^{n} K_i \tag{13}$$

where SNR is the signal-to-noise ratio defined by dividing the quantity of the power of the signal by that of the noise: $P_{\text{signal}}/P_{\text{noise}}$. I_i ($i = 1, 2, \ldots, n$) is the dimension of the ith mode, with K_i being the corresponding i-mode rank (Lathauwer, 1997). However, there are two questions: The first question is how to estimate the SNR value from the given image, and the second is how to choose the n-mode rank on each mode properly.

The SNR can be obtained directly when a noise-free image is available. In some situations, it is difficult to obtain a noise-free image as a priori knowledge. That is to say, the SNR estimation should be performed based only on the information provided by the input image. Since the noise is white and Gaussian (Pauluzzi and Beaulieu, 2000), the quantity of the signal and noise power can be evaluated by the signal and noise variances, respectively. In this part, a statistic-based algorithm is utilized to give a scalar measurement of the noise variance and the signal variance. This SNR estimator consists of the following three steps:

Step 1. Within a moving window of a given size, e.g., 5×5, the variance of the partial image can be calculated. Find the minimum variance value and denote it as the estimation of the noise variance $\hat{\sigma}^2_{\text{noise}}$.

Table 1
Alternating least squares optimization for R1TD noise reduction.

Input: Input HSI tensor $O \in R^{L_1 \times L_2 \times L_3}$, decomposition level k, and maximum number of iterations *ITER*
Initialization: Set U, V and W to the identity matrix, $S^0 = O$
Step 1. For $t = 1$ to *ITER* {
Step 2. Calculate
$\quad U^* = mat_1 S^t \cdot ((W \odot V)^T)^{-1}, \Lambda_{(i,j)} = \|U_{(1)}\|, \quad U_{(i)} = U^*_1/\Lambda_{(i,i)}$
$\quad V^* = mat_2 S^t \cdot ((W \odot V)^T)^{-1}, \Lambda_{(i,j)} = \|V_{(1)}\|, \quad V_{(i)} = V^*_1/\Lambda_{(i,i)}$
$\quad W^* = mat_3 S^t \cdot ((U \odot V)^T)^{-1}, \Lambda_{(i,j)} = \|W_{(1)}\|, \quad W_{(i)} = W^*_1/\Lambda_{(i,i)}$
Step 3. Reconstruct the estimated noise-free tensor by
$\quad S^t = \sum_{r=1}^{k} \lambda_r u_r \circ v_r \circ w_r$
Step 4. Check the convergence, if: $Err(t) = \|S^t - S^{t-1}\| \leqslant \varepsilon.$}// For loop in Step 2
Output: The denoised tensor $\widehat{S} \in R^{L_1 \times L_2 \times L_3}$, weight matrix Λ, and factor matrices U, V, and W

Step 2. Calculate the variance of the input image using (14).

$$\sigma^2_{signal} \approx \frac{\sum_{i=1}^{I_1}\sum_{j=1}^{I_2}\sum_{k=1}^{I_3} O^2_{i,j,k}}{I_1 \times I_2 \times I_3} - \hat{\sigma}^2_{noise} \tag{14}$$

Step 3. Perform the SNR prediction of the given image with the following equation:

$$SNR = \frac{P_{signal}}{P_{noise}} = \frac{\sigma^2_{signal}}{\sigma^2_{noise}} \tag{15}$$

Many algorithms with high estimation precision have been discussed in Pauluzzi and Beaulieu (2000). However, as the SNR estimate is not the main concern in this paper, the above method has been chosen and is adopted in this work as it is both convenient and rapid to carry out. The results of this estimation are close to the real SNR value, and the errors are within an acceptable range for the parameter estimation.

Differing from the tensor rank, the n-mode rank is denoted as the dimension of the vector space generated by unfolding the input tensor on the nth-mode (Lathauwer, 1997). From the perspective of hyperspectral image noise reduction, the n-mode rank can be regarded as the dimension of the signal subspace. As to the problem of n-mode rank selection, two information-based rank estimation criteria are taken into consideration in this study, namely, the Akaike information criterion (AIC) (Akaike, 1974) and the minimum description length (MDL) (Schwarz, 1978).

3.3.1. AIC (Akaike information criterion)

The AIC is introduced as an approach to select an appropriate model for a parameterized family of probability density functions used to best fit the sensor array data in passive sensor array
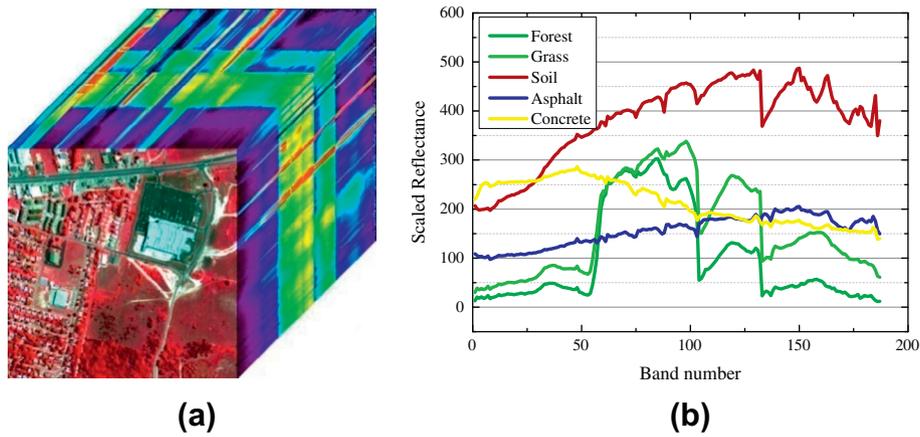

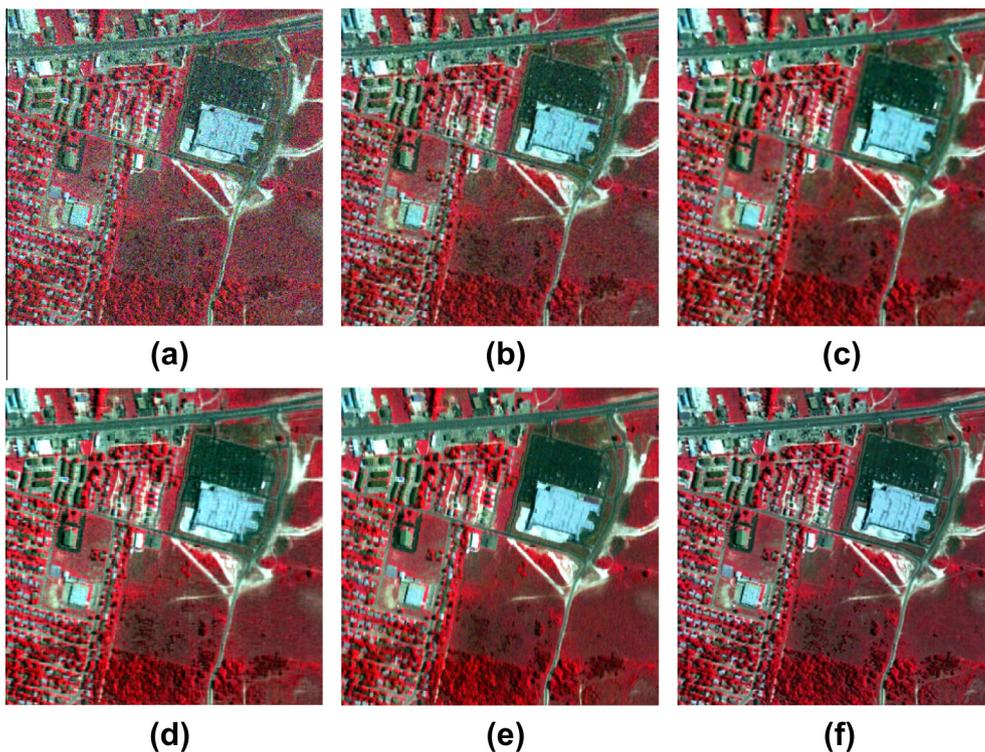**Fig. 4.** (a) The noiseless HYDICE data cube and (b) the scaled reflectance on each band.

Fig. 5. Noise reduction results: (a) synthetic noisy image (SNR = 15 dB), (b) CCSVD, (c) CCWF, (d) MWF, (e) SSAHTV, and (f) R1TD.

processing. Since this theoretical information criterion does not require any subjective threshold setting, the AIC model has been proven to be effective in a variety of problems, such as detecting the number of sources in signal processing (Chang and Du, 2004). For signal number estimation in a multichannel time-series, the criterion of the AIC model derived by Wax and Kailath (1985) is defined as:

$$\text{AIC}(r) = -2 \log \left(\frac{\prod_{i=r+1}^{L} \lambda_i^{\frac{1}{L-r}}}{\frac{1}{L-r} \sum_{i=r+1}^{L} \lambda_i} \right)^{(L-r)N} + 2r(2L - r) \qquad (16)$$

where r is the number of dominant eigenvalues that contain most of the information, and $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_L$ are eigenvalues generated by unfolding the input HSI tensor $O \in R^{L_1 \times L_2 \times L_3}$ in each mode. L and N are the height and width of the n-mode flattened matrix $mat_nO \in R^{L_n \times (L_1 L_2 \ldots L_{n-1} L_{n+1} \ldots L_3)}$. Subsequently, using Eq. (16), the rank of the corresponding mode is obtained:

$$(\text{Rank}_n O)_{\text{AIC}} = \arg\{\min_r \text{AIC}(r)\} \qquad (17)$$

3.3.2. MDL (minimum description length)

Another commonly used criterion for dimension selection is the MDL. This method was suggested by Schwarz (1978) for determining the appropriate number of factors that will fit a given set of observations in statistics. This model is an extension of maximum likelihood and gives a mathematical formulation of the principle of parsimony in model building. Results obtained by MDL are supposed to lean towards lower-dimensional models for large numbers of observations, when compared to AIC (Schwarz, 1978). Since the problem of determining the number of factors in the observation data set is similar to that of estimating the n-mode rank, in this work, the MDL algorithm is applied to hyperspectral image analysis for dimension selection by implementing the formula below:
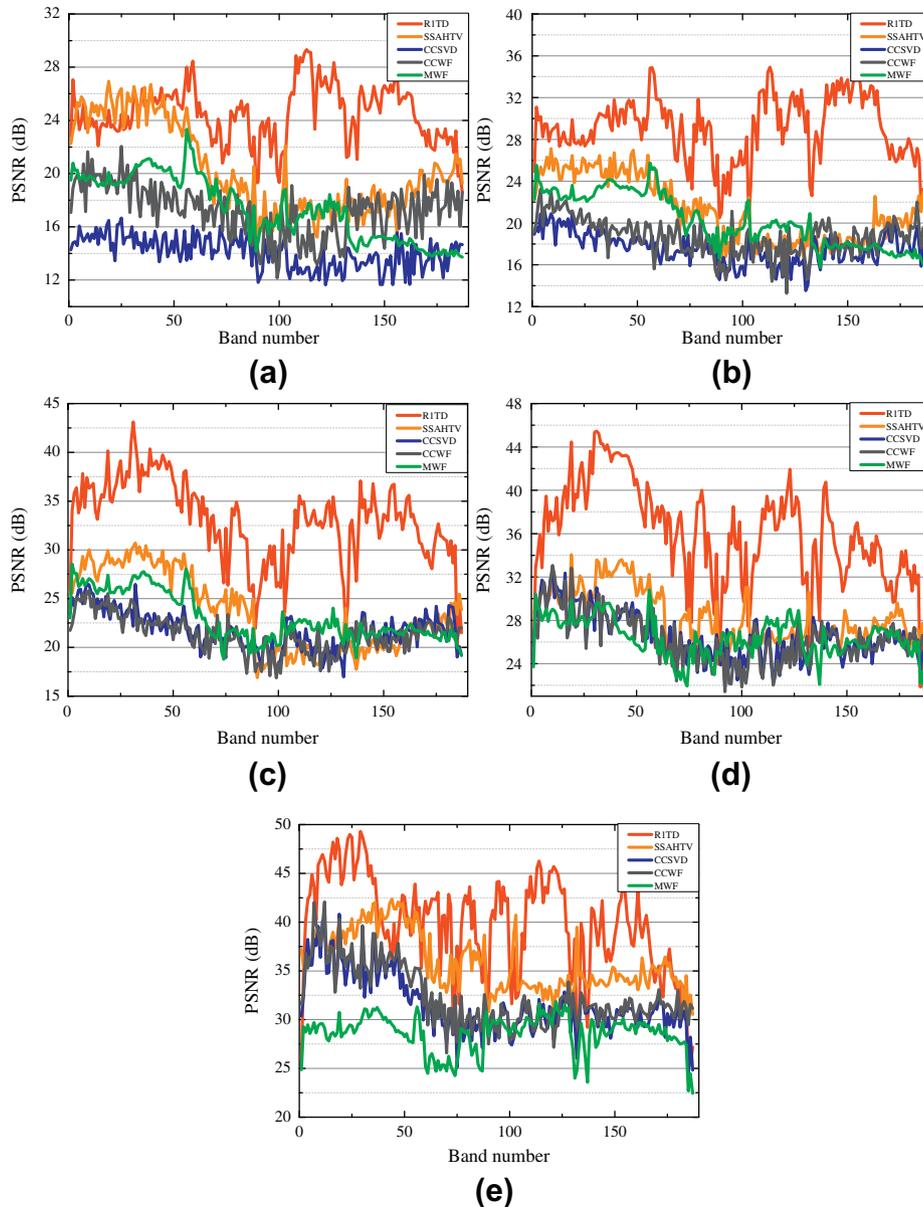

Fig. 6. PSNR values of the different denoising algorithms in each band for the synthetic scenarios: (a) SNR = 5 dB, (b) SNR = 10 dB, (c) SNR = 15 dB, (d) SNR = 20 dB, and (e) SNR = 25 dB.

$$\text{MDL}(r) = -\log\left(\frac{\prod_{i=r+1}^{L}\lambda_i^{\frac{1}{L-r}}}{\frac{1}{L-r}\sum_{i=r+1}^{L}\lambda_i}\right)^{(L-r)N} + \frac{1}{2}r(2L-r)\log N \qquad (18)$$

Similarly, r in Eq. (18) is the number of dominant eigenvalues, and $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_L$ are the eigenvalues generated by unfolding the input HSI tensor $O \in R^{L_1 \times L_2 \times L_3}$ in each mode. L stands for the height and N is the width of the n-mode matrix unfolding $mat_n O \in R^{L_n \times (L_1 L_2 \ldots L_{n-1} L_{n+1} \ldots L_3)}$. Subsequently, the n-mode rank can be calculated via:

$$(\text{Rank}_n O)_{\text{MDL}} = \arg\{\min_r \text{MDL}(r)\} \qquad (19)$$

To summarize this section, the procedure for selecting the decomposition level k involves the following steps. Firstly, two theoretical information criteria based estimators are suggested to compute the n-mode rank of each mode. Secondly, SNR estimation of the input image is conducted, and we can subsequently calculate the decomposition level k with Eq. (13).

4. Imagery and experiments

4.1. Data set 1

For the synthetic experiments, we use the Hyperspectral Digital Imagery Collection Experiment (HYDICE) urban image presented in Fig. 4 as the test HSI data set to verify the performance of the proposed algorithm. This urban data set was acquired by the HYDICE sensor system in October 1995. The data size is 307×307 pixels, with 210 spectral bands, and the spectral and spatial resolutions are 10 nm and 2 m. The image area is located at Copperas Cove near Fort Hood, Texas, US Before the denoising process, the water vapor absorption bands and the low signal/high noise bands (104–109, 138–152, and 203–210) were removed. For the purpose of comparing the noise reduction results at different noise levels, we generate the experimental images by adding synthetic noise with different SNR values ranging from 5 dB to 25 dB. Since the SNR in real data varies between different spectral bands, we
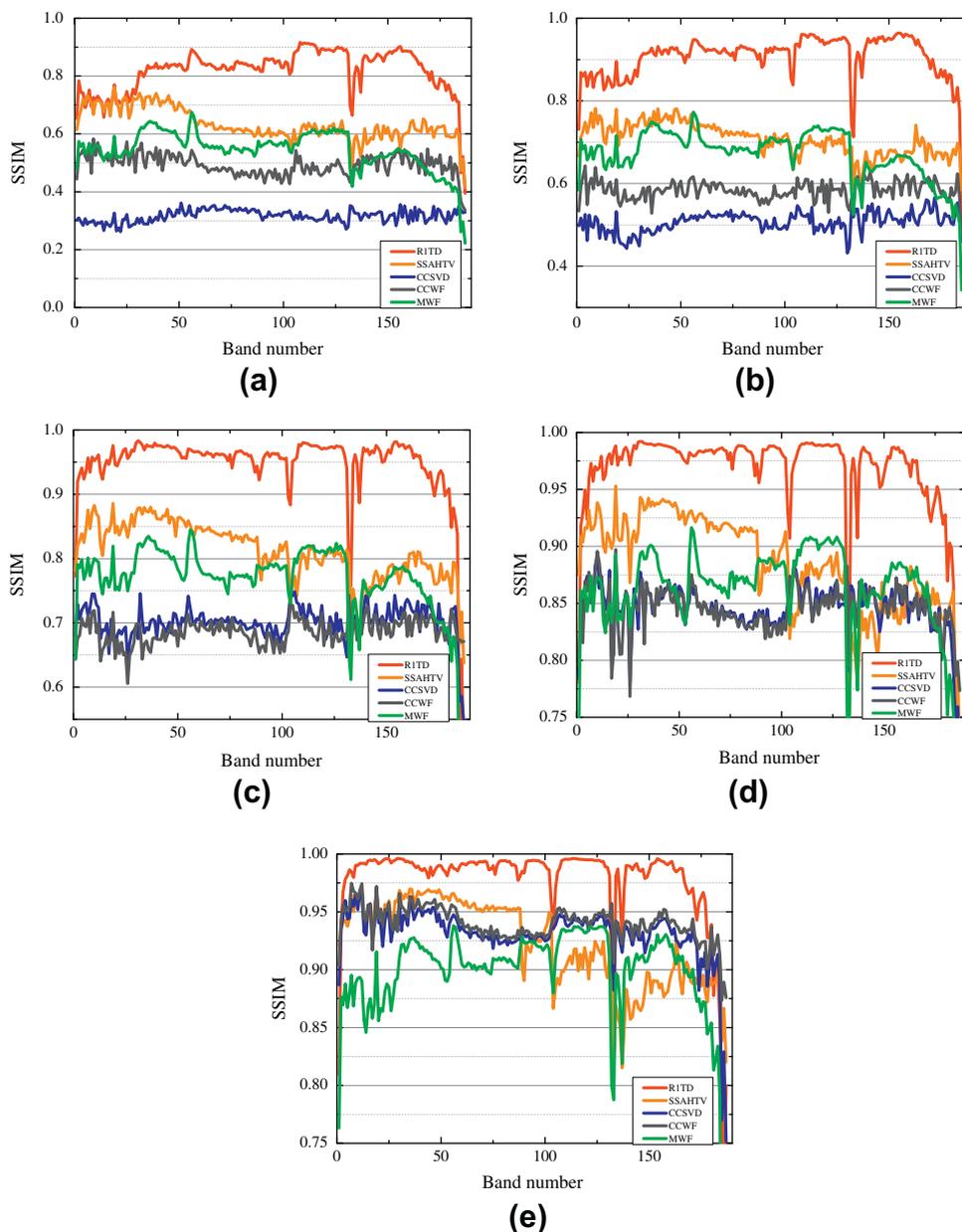

Fig. 7. SSIM values of the different denoising algorithms in each band for the synthetic scenarios: (a) SNR = 5 dB, (b) SNR = 10 dB, (c) SNR = 15 dB, (d) SNR = 20 dB, and (e) SNR = 25 dB.

randomly add different SNR noise in each band, and generate the denoising results using the R1TD method under this condition.

To verify the effectiveness of the proposed algorithm, the proposed R1TD tensor model is compared with several competitive methods: the channel-by-channel locally adaptive Wiener filter (CCWF) (Banham and Katsaggelos, 1997), channel-by-channel SVD (CCSVD) (Andrews and Patterson, 1976), SSAHTV (Yuan et al., 2012), and the MWF algorithm (Muti and Bourennane, 2005). Recently, several human visual perception indices have been proposed for the estimation of image quality based on a reference noise-free image. In addition to the visual interpretation, the following image quality evaluation indices are taken into account for the noise reduction performance comparison: (1) the peak signal-to-noise ratio (PSNR); (2) the structural similarity (SSIM) measure (Wang et al., 2004), which measures the spatial similarity between the output HSI and the reference data; (3) the spectral difference (SD); and (4) the spectral angle (SA) (Othman and Qian, 2006), which determines the spectral similarity between the processed HSI and the reference data.

From the denoising results shown in Fig. 5, it can be observed that the CCWF algorithm (Fig. 5c) is not effective in preserving details, and the result of CCSVD (Fig. 5b) is disturbed by residual noise. The MWF algorithm (Fig. 5d) is expected to generate noiseless images; however, some blurring is introduced after the denoising process. SSAHTV effectively removes the random noise and gives a clear view of the input image. However, some tiny details are over-smoothed in Fig. 5e. The denoising result of the R1TD algorithm is presented in Fig. 5f, where most of the details (e.g., the edge of the path, and individual trees) are visually better reconstructed, without artifacts, when compared with the results of the other methods.

PSNR and SSIM are employed to give a quantitative assessment of the denoising results from the perspectives of the spatial and structural information. In the synthetic experiments, the PSNR and SSIM values are computed between each clear band and the denoised band. The values of PSNR and SSIM for the different denoising strategies in the different bands of the five simulated cases are presented in Figs. 6 and 7.

The mean values of PSNR and SSIM are reported in Table 2 and 3. MWF, SSAHTV, and R1TD outperform the other two algorithms, because they simultaneously take both the spatial and spectral information into consideration. However, the MWF algorithm is mainly based on Tucker tensor decomposition and a generalized Wiener filtering. As a result, although the application of tensor representation helps to suppress the blurring brought about by the Wiener filter, the denoising results after MWF are often contaminated by artifacts. SSAHTV employs a spectral-spatial adaptive total variation (TV) model, in which the differences for both the noise and the spatial information between different bands are considered. However, the algorithm over-smoothes the details of the noisy image while removing the noise. Meanwhile, R1TD estimates the noisy image by a series of rank-1 tensors, and then removes the noise component to reconstruct the noise-free image. Since there are no artifacts introduced in the R1TD process, the corresponding denoising results are better than MWF and SSAHTV for the preservation of spatial information. It can also be inferred that as the SNR value increases, the proposed rank-1 tensor-based method achieves the best performance in terms of both the visual inspection and the quantitative results.

In addition, Fig. 8 shows the spectral difference between the spectra of the reference image and those of the denoised images obtained by the above noise reduction approaches (SNR = 15 dB). These pixels are chosen from five classes: grass, forest, asphalt, concrete, and soil. It can be observed that the results of R1TD shows a smaller spectral difference, compared with MWF, which has been proven to have good ability in keeping spectral informa-

Table 2
mPSNR values of the five noise reduction approaches applied to the synthetic data set.

mPSNR	CCSVD	CCWF	MWF	SSAHTV	R1TD
SNR = 5 dB	14.087	17.252	17.406	20.312	**24.509**
SNR = 10 dB	17.607	18.686	20.215	21.237	**29.603**
SNR = 15 dB	21.786	21.384	23.034	23.645	**33.011**
SNR = 20 dB	26.519	26.382	26.403	25.679	**36.038**
SNR = 25 dB	31.632	32.271	28.604	27.267	**39.825**
Random SNR	16.771	19.987	20.992	23.382	**24.668**

Table 3
mSSIM values of the five noise reduction approaches applied to the synthetic data set.

mSSIM	CCSVD	CCWF	MWF	SSAHTV	R1TD
SNR = 5 dB	0.329	0.493	0.545	0.634	**0.818**
SNR = 10 dB	0.516	0.658	0.668	0.706	**0.902**
SNR = 15 dB	0.692	0.686	0.768	0.812	**0.946**
SNR = 20 dB	0.846	0.845	0.860	0.886	**0.965**
SNR = 25 dB	0.931	0.939	0.897	0.929	**0.978**
Random noise	0.451	0.586	0.634	0.698	**0.816**

tion. It can be seen that the channel-by-channel denoising methods (CCSVD and CCWF) are worse than the tensor-based algorithm in keeping spectral consistency. Fig. 9 shows the spectral angle between the noise-free spectra and those obtained by the different noise reduction methods at the five specific pixels mentioned above. The results of the spectral angle in Fig. 9 also confirm the superiority of the proposed R1TD model.

A further comparison of the classification was also performed. The support vector machine (SVM) algorithm (Mountrakis et al., 2011) was used as the classifier, and the number of training and test pixels used in the classification of the HYDICE data set are shown in Table 4.

Classification maps of the HYDICE data set are presented in Fig. 10 for visual inspection. The accuracies of the spectral classification for CCSVD, CCWF, MWF, SSAHTV, and R1TD are listed and compared in Table 5, covering all six synthetic cases.

The computation of R1TD involves the selection of the decomposition level k, which determines the image restoration quality. It is the only free parameter of the proposed algorithm. The impact of the decomposition level k value on noise reduction is investigated by implementing the denoising process with various k values in synthetic experiments. In this comparative study, we employ the relative difference as a further measurement of the denoising performance. The relative difference value is computed with the denoised image $\widehat{S}$ and a reference noise-free image S, and is derived as: $RD = \|S - \widehat{S}\|^2/\|S\|^2$. Fig. 11 shows the relative difference value with respect to the variation of the decomposition level k. Here, we treat the optimal decomposition level as the k value that makes the reconstructed image most close to the ideal noise-free image, and minimizes the relative difference. Given different images, the optimal decomposition level varies correspondingly. It can be inferred from Fig. 11 that when the value of k is much smaller than that required for the optimal decomposition level, the number of signal rank-1 tensors used to estimate the noise-free image is not enough, and thus leads to a loss of detail in the restored image. When the value of k is much larger than that required for the optimal decomposition level, some noise rank-1 profiles are included during the reconstruction and will also result in a degradation of the reconstructed image. Although the optimal decomposition level cannot be identified by visual inspection or a priori knowledge in most cases, Fig. 11 indicates that the decomposition levels in a certain data range will yield denoising results that are equivalent to the optimal decomposition level. Here, we define this data range as
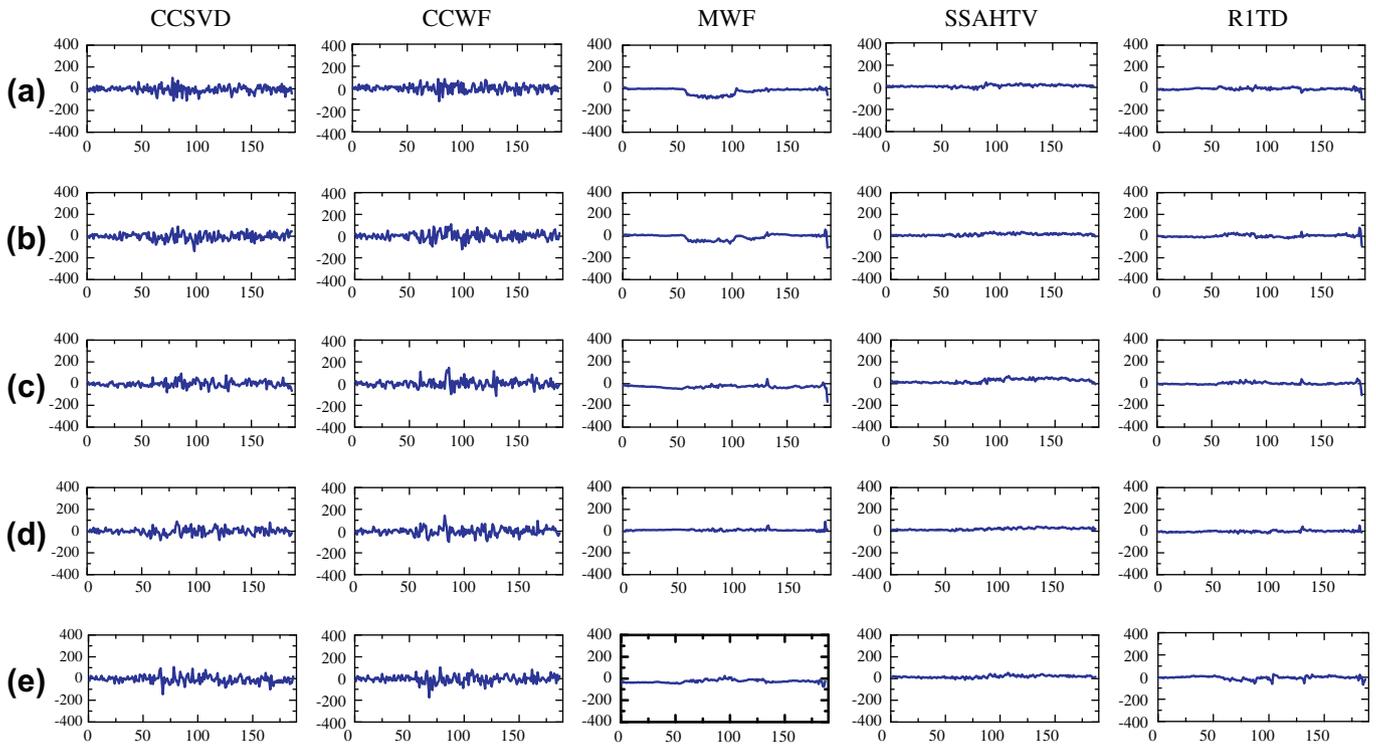
Fig. 8. Spectral differences between the noise-free spectra and the denoising results of the four denoising methods for the five classes (SNR = 15 dB). (a–e) are the pixels of grass, forest, asphalt, concrete, and soil, respectively.
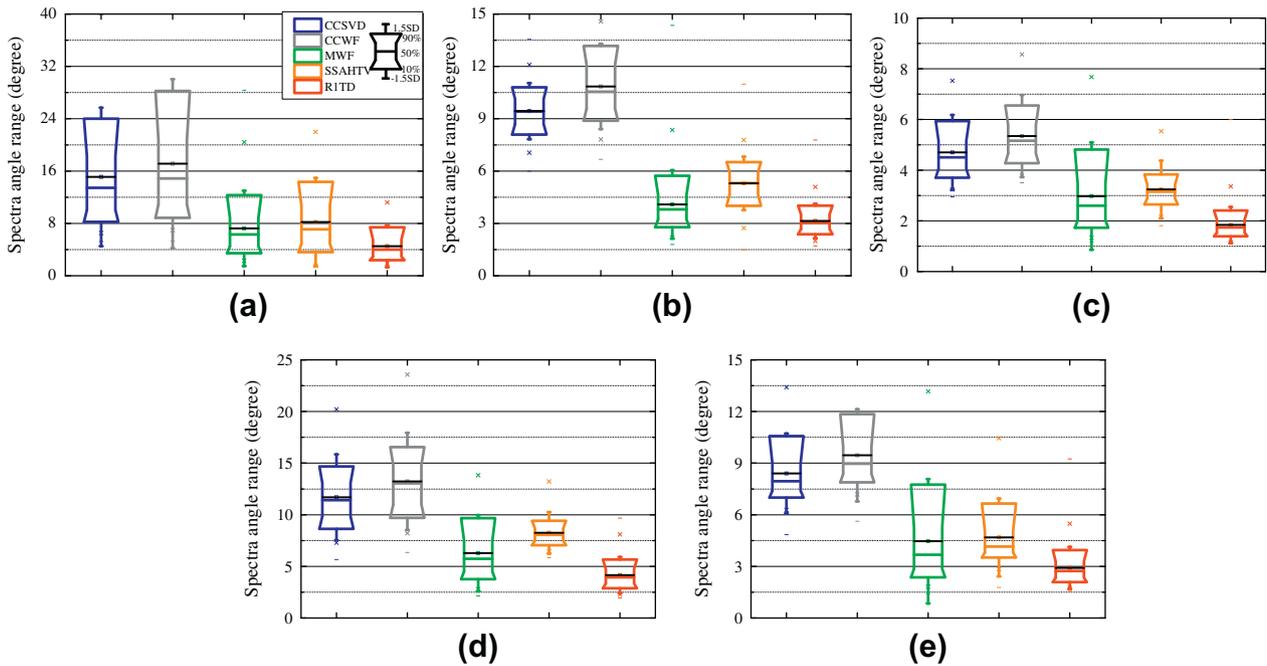

Fig. 9. Spectral angles between the noise-free spectra and the denoising results for the five pixels chosen from each class: (a–e) are the pixels of grass, forest, asphalt, concrete, and soil.

the acceptable R1TD range. Practically, we manually tune the acceptable R1TD range in each synthetic case, which is shown as the light-green[1] data set range in Fig. 11. The AIC-based and MDL-based decomposition-level assessments are performed with respect to various SNR values for all six synthetic experiments. Fig. 11a–f

demonstrate that all the estimated k values fall in the acceptable R1TD range, which helps R1TD to achieve satisfying denoising results. Therefore, we draw the conclusion that the proposed decomposition-level criterion works and fits well in the synthetic study.

Fig. 12 shows a convergence plot of R1TD for the different synthetic scenarios. It can be seen that the optimization procedure converges within 10 iterations, at the chosen decomposition level.

---

[1] For interpretation of color in Fig. 11, the reader is referred to the web version of this article.

Table 4
Number of training and test samples used in the classification of the HYDICE data set.

Name of class	Training samples	Test samples
Forest	25	2437
Grass	25	2191
Soil	25	1378
Asphalt	25	2379
Concrete	25	1923

To summarize the synthetic experiments, due to the utilization of the tensor representation, the proposed R1TD model is able to process the noisy image as an entity, which preserves the inter-dimensional relationships. The above synthetic experiments reveal that images generated with both a fixed noise intensity and a random noise intensity can achieve satisfactory denoising results.

4.2. Data set 2

The image data used in this real-world data experiment refers to an agricultural area of Indiana Pines in northern Indiana, US. The image size is 128×128 pixels and was acquired by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor in June 1992, with 220 spectral bands (from 0.4 to 2.5 μm). It should be noted that some bands of this AVIRIS image are corrupted by striping noise and other kinds of noise, as well as additive noise. The hyperspectral cube and scaled reflectance plot are presented in Fig. 13.

The conventional channel-to-channel denoising methods (CCSVD and CCWF), the spectral-spatial adaptive TV-based approach (SSAHTV) and the tensor-based noise removal algorithms (MWF and R1TD) are applied to this data set. According to the decomposition-level estimator, the estimation of SNR is 8.9 dB
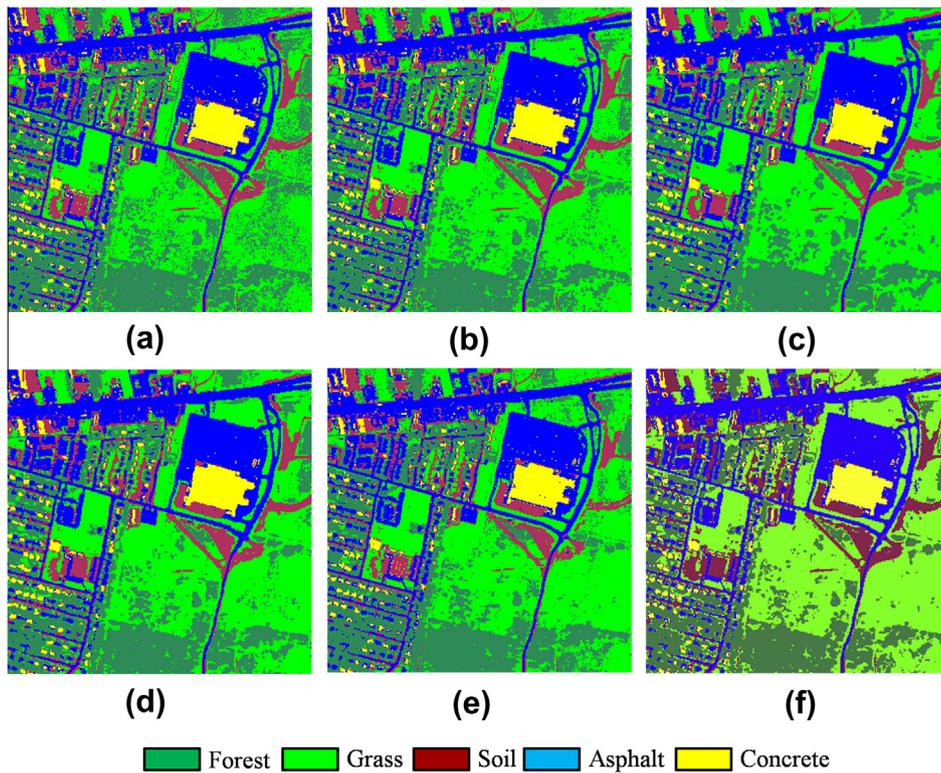

Fig. 10. Classification maps of the HYDICE data set: (a) spectral classification, (b) CCSVD, (c) CCWF, (d) MWF, (e) SSAHTV, and (e) R1TD.

Table 5
SVM Classification accuracy comparison of the R1TD-based and the state-of-the-art noise reduction algorithms for the HYDICE data set.

| Accuracy assessment		Input image	CCSVD	CCWF	MWF	SSAHTV	R1TD
Reference image	OA (%)	**98.14**	97.30	97.43	97.96	97.35	**98.09**
	Kappa	**0.976**	0.966	0.968	0.974	0.967	**0.976**
5 dB	OA (%)	93.78	95.03	95.71	95.17	95.89	**96.59**
	Kappa	0.922	0.937	0.946	0.939	0.951	**0.957**
10 dB	OA (%)	95.30	96.21	96.54	96.32	97.07	**97.40**
	Kappa	0.941	0.952	0.956	0.954	0.963	**0.967**
15 dB	OA (%)	96.32	96.67	96.32	97.11	97.03	**97.44**
	Kappa	0.954	0.958	0.963	0.963	0.963	**0.968**
20 dB	OA (%)	97.27	97.28	97.27	97.46	96.98	**97.86**
	Kappa	0.964	0.965	0.965	0.967	0.962	**0.972**
25 dB	OA (%)	97.48	97.44	97.50	97.67	96.66	**97.77**
	Kappa	0.968	0.968	0.968	0.969	0.958	**0.972**
Random SNR	OA (%)	90.82	93.28	95.27	96.09	96.22	97.32
	Kappa	0.884	0.915	0.943	0.953	0.955	0.968
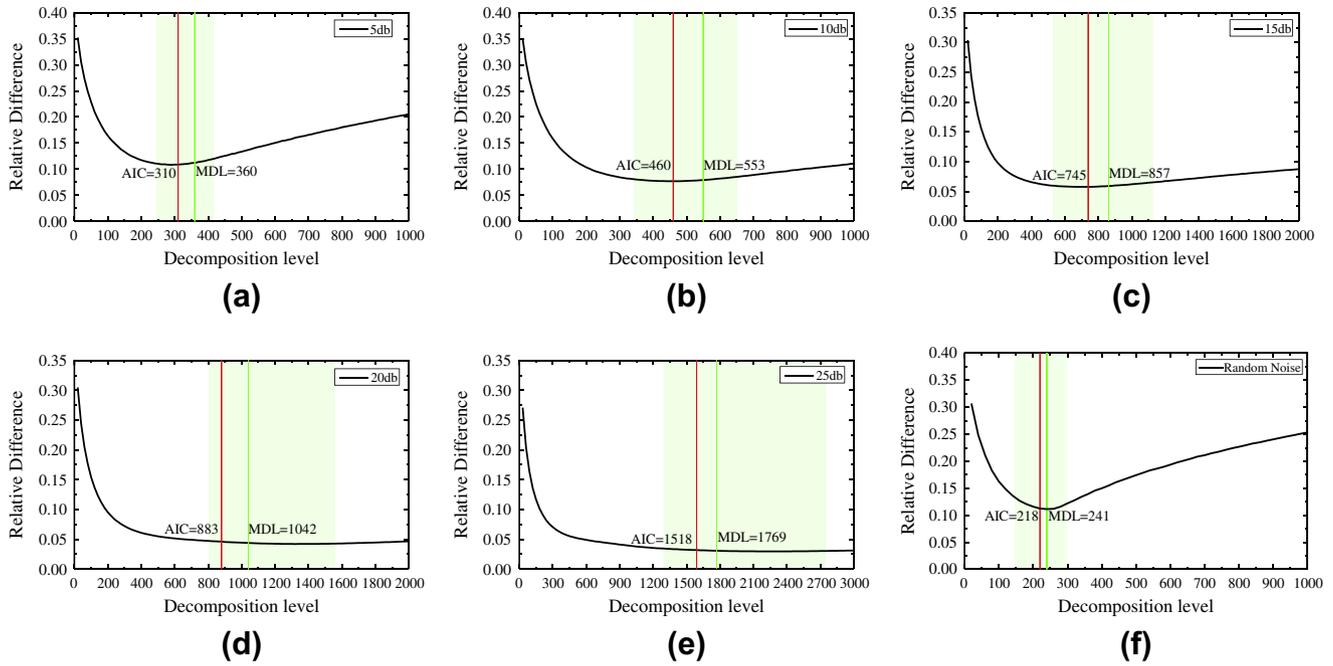
Fig. 11. The impact of different decomposition levels on the relative difference between the denoised image and the reference image in the synthetic scenarios, with the R1TD denoising approach.
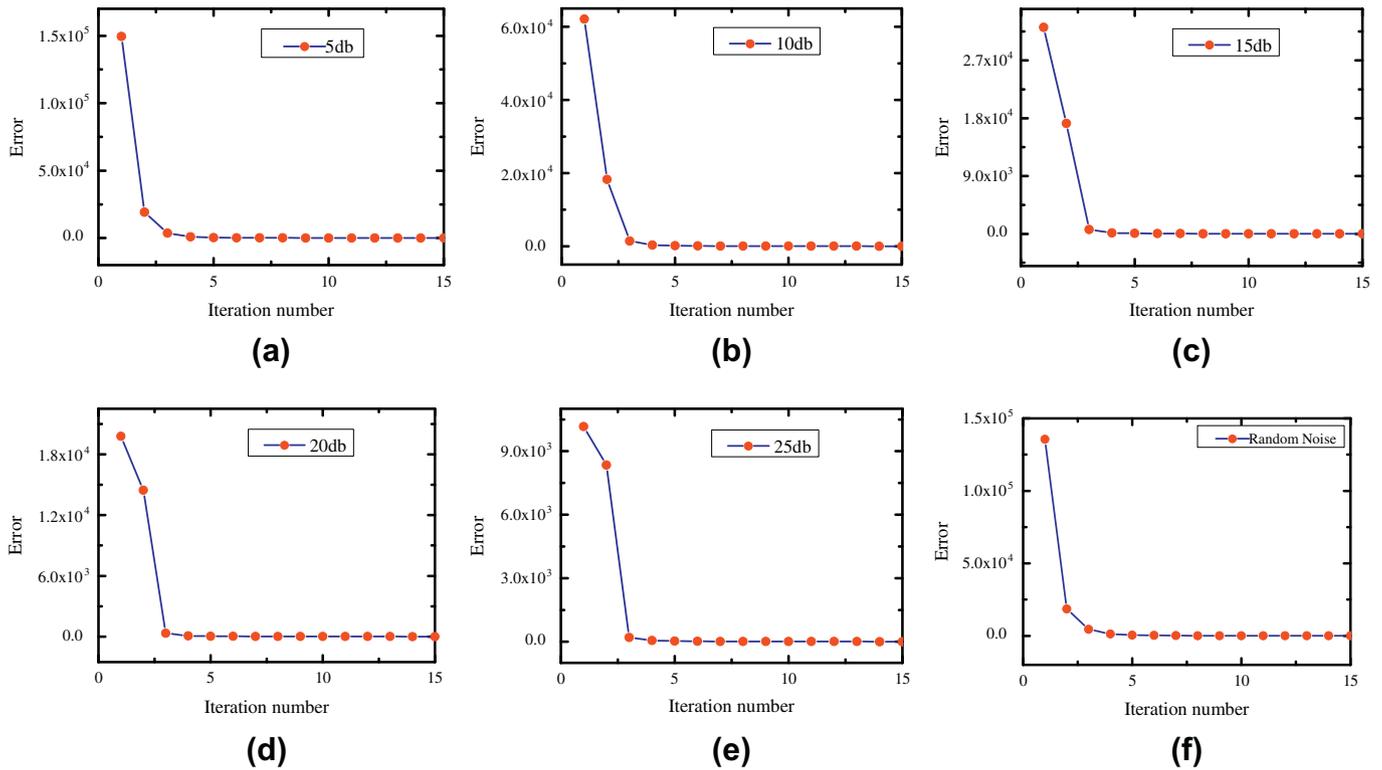

Fig. 12. Convergence plot of R1TD for the different synthetic scenarios. (a): 5 dB fixed additive noise, (b) 10 dB,(c) 15 dB, (d) 20 dB, (e) 25 dB, (f) random noise with mean SNR = 7.9 dB.

for the real-world data, and the n-mode rank values (K_1, K_2 and K_3) obtained using AIC and MDL are (45, 45, and 40) and (43, 43, and 42). Since the n-mode rank values generated by the AIC and MDL criteria for the decomposition-level estimation are identical, either

of these two criteria is suitable for the proposed algorithm. For this real data experiment, the decomposition level is set to 330, with respect to the AIC estimation. The parameters for the other three algorithms are as follows: CCSVD is conducted with a 100 decom-
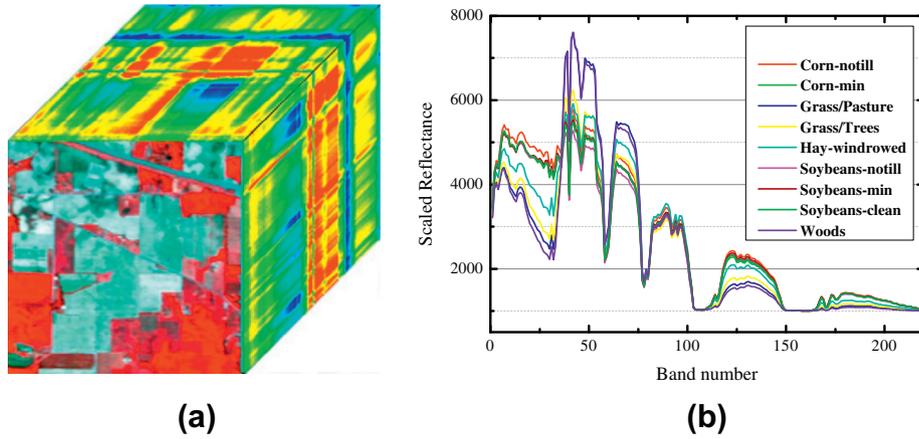
Fig. 13. AVIRIS Indian Pines data set: (a) 3-D hyperspectral cube and (b) the scaled reflectance plot.
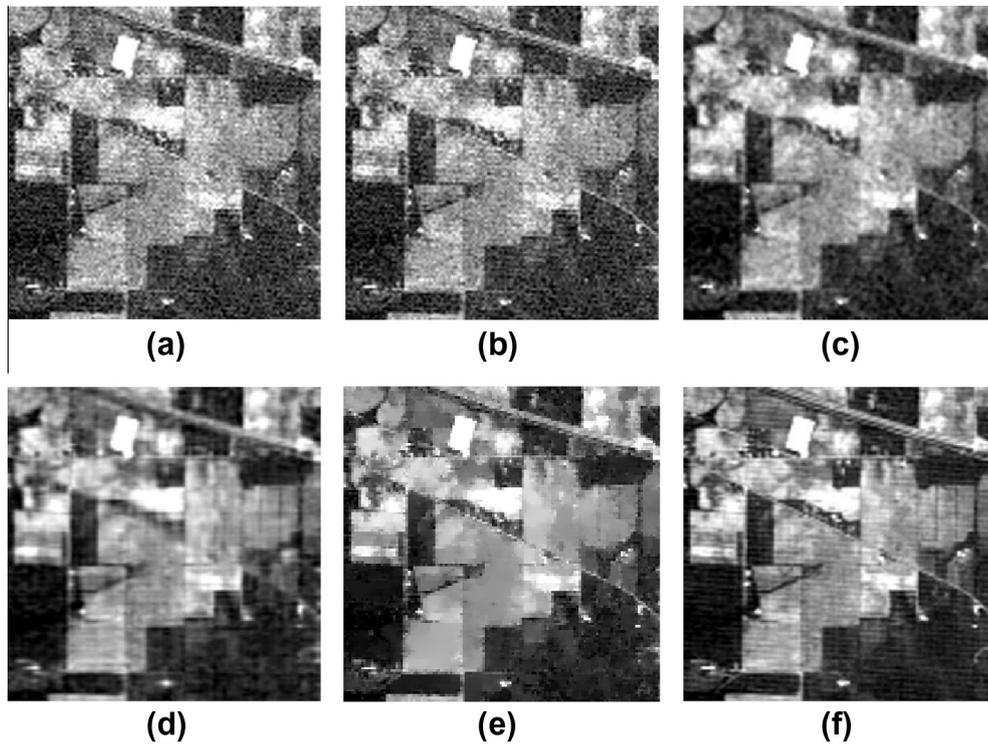

Fig. 14. Noise reduction results for the AVIRIS Indian Pines data set: (a) original noise band 3, (b) CCSVD, (c) CCWF, (d) MWF, (e) SSAHTV and (e) R1TD.

Table 6
Number of training and test samples used in the classification of data set 2.

Name of class	Training samples	Test samples
Corn-notill	50	1697
Corn-min	50	661
Grass/pasture	50	727
Grass/trees	50	712
Hay-windrowed	50	217
Soybeans-notill	50	844
Soybeans–min	50	2559
Soybeans–clean	50	490
Woods	50	763

position level; CCWF is performed with a 4×4 window; MWF is implemented with an n-mode rank (K_1, K_2, and K_3) = (45, 45, and 40), which was also estimated by the AIC model; and in SSAHTV

the lambda that controls the strength of the noise reduction is set to 2.2.

Fig. 14 gives the noise reduction results of the AVIRIS data set on the 3rd band. From a visual interpretation, R1TD is more effective in removing additive noise in this real-world test data set than the other denoising strategies. There are, however, some undesired horizontal stripes in the R1TD results. The main reason for this is that the R1TD approach is based on the assumption of an additive blend of noise and signal, and hence it is expected to be effective for the AWGN (Additive White and Gaussian Noise) scenario. However, striping-noise and mixed-noise bands are in fact found in the AVIRIS data set. The stripes can be explained by the striping noise that is contained in the original AVIRIS data set, which become clear after the removal of the additive noise. Further efforts should be devoted to improving the proposed algorithm for suppressing multiplicative and striping noise.
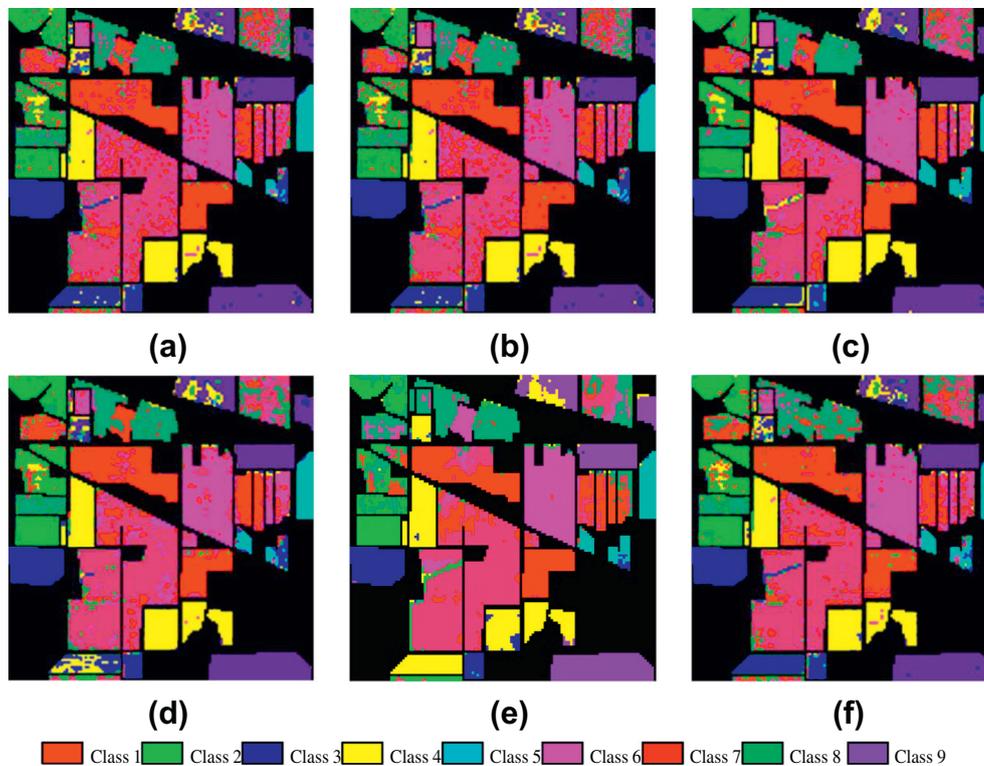
Class 1 ▮ Class 2 ▮ Class 3 ▮ Class 4 ▮ Class 5 ▮ Class 6 ▮ Class 7 ▮ Class 8 ▮ Class 9

Fig. 15. Classification map obtained before and after denoising of the Indian Pines data set: (a) original image, (b) CCSVD, (c) CCWF, (d) MWF, (e) SSAHTV, and (f) R1TD.

Table 7
Class-specific accuracies comparison of the R1TD-based and the state-of-the-art noise reduction algorithms for the Indiana Pines data set.

Class	1	2	3	4	5	6	7	8	9	OA	Kappa
Origin	73.7	79.6	94.5	95.7	94.3	63.5	83.3	69.7	99.5	81.1	0.777
CCSVD	73.2	80.7	94.4	95.9	93.1	65.1	84.3	68.9	99.6	81.4	0.780
CCWF	83.3	86.7	97.0	90.9	77.2	68.7	89.1	72.4	98.9	85.1	0.824
MWF	84.1	81.2	**97.1**	96.3	**97.3**	66.4	86.7	**73.3**	**99.8**	85.3	0.826
SSAHTV	**86.4**	**87.4**	94.2	92.0	83.1	73.2	90.2	70.1	98.5	86.1	0.837
R1TD	80.1	82.7	94.8	**96.3**	82.1	**81.7**	**93.7**	63.3	99.3	**86.8**	**0.843**

It is worth noting that, for the Indian Pines data set, the reference noise-free image is not available. Thus, to validate the noise removal performance, classification is implemented on each denoised image to evaluate the denoising algorithms. Nine land-cover classes are of interest: corn-notill, corn-min, grass/pasture, grass/trees, hay-windrowed, soybeans-notill, soybeans-min, soybeans-clean, and woods. Classification is performed via the SVM algorithm, and the number of training and test pixels for the nine classes are generated randomly from the ground truth reference (Table 6).

Fig. 15 shows the classification maps. It is clear that the noise level is different in each band of the real data cube, and the conventional channel-by-channel image denoising methods tend to remove the detailed features and destroy the spectral coherence of the image after the denoising process. The accuracy assessments given in Table 7 indicate that the classification accuracy using R1TD is higher than those of CCSVD and CCWF. In summary, the proposed algorithm is better for detail preservation and is more effective than MWF and SSAHTV, in the classification case.

5. Conclusion

In this study, the high-order rank-1 tensor decomposition (R1TD) model is investigated to develop a new noise removal

algorithm for hyperspectral image pre-processing. The main advantage of the R1TD algorithm is that it treats the HSI data as a cube and, hence, is able to simultaneously extract tensor features in both the spectral and spatial modes. Unlike the state-of-the-art Tucker model based denoising methods, the proposed R1TD algorithm considers the fact that the different parts of HSI can be represented by a sequence of rank-1 tensors. Then, in the additive noise condition, noise-free HSI can be obtained once the noise component is removed. However, the determination of the decomposition level in rank-1 tensor decomposition is a difficult and challenging problem. In this study, we present an n-mode rank based decomposition-level estimator, which performs the decomposition-level estimation with the signal-to-noise ratio (SNR), dimension, and n-mode rank of the input image. The experimental results revealed that, for synthetic data, the image quality was improved while the spectral information was well preserved. The PSNR and SSIM indices reached 39.578 and 0.978, respectively, in the 25 dB synthetic case. Due to the utilization of tensor representation, images generated with both a fixed noise intensity and a random noise intensity were effectively processed in the synthetic scenarios. Meanwhile, the proposed decomposition-level estimator was validated by the quantitative results, and the estimated values of k were within the acceptable range in all cases. Meanwhile, for real-world data, R1TD was effective for additive noise removal,

and the classification result was improved as the overall accuracy increased from 81.1% to 86.8% with the estimated k value.

As a concluding remark, compared with the conventional denoising methods, the proposed R1TD algorithm achieves a significant improvement in HSI data quality for both synthetic noisy data and real HSI data. However, the R1TD approach is based on the assumption of an additive blend of noise and signal. It is therefore effective for the additive white and Gaussian noise (AWGN) scenario. However, real-world HSI contains different kinds of noise in addition to the AWGN, and the proposed algorithm should be improved to deal with multiplicative and striping noise.

Acknowledgements

References

Acito, N., Diani, M., Corsini, G., 2010. Hyperspectral signal subspace identification in the presence of rare signal components. IEEE Trans. Geosci. Remote Sens. 48 (4), 1940–1954.

Acito, N., Diani, M., Corsini, G., 2011a. Signal-dependent noise modeling and model parameter estimation in hyperspectral images. IEEE Trans. Geosci. Remote Sens. 49 (8), 2957–2971.

Acito, N., Diani, M., Corsini, G., 2011b. Subspace-based striping noise reduction in hyperspectral images. IEEE Trans. Geosci. Remote Sens. 49 (4), 1325–1342.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Autom. Control 19 (6), 716–723.

Andrews, H.C., Patterson, C.L., 1976. Singular value decompositions and digital image processing. IEEE Trans. Acoust. Speech Signal Process. 24 (1), 26–53.

Banham, M.R., Katsaggelos, A.K., 1997. Digital image restoration. IEEE Signal Process Mag. 14 (2), 24–41.

Bioucas-Dias, J.M., Figueiredo, M.A.T., 2010. Multiplicative noise removal using variable splitting and constrained optimization. IEEE Trans. Image Process. 19 (7), 1720–1730.

Bourennane, S., Fossati, C., Cailly, A., 2011. Improvement of target-detection algorithms based on adaptive three-dimensional filtering. IEEE Trans. Geosci. Remote Sens. 49 (4), 1383–1395.

Bro, R., Kiers, H.A.L., 2003. A new efficient method for determining the number of components in PARAFAC models. J. Chemometr. 17 (5), 274–286.

Chang, C.-I., 2003. Hyperspectral Imaging: Techniques for Spectral Detection and Classification. Kluwer Academic/Plenum Publishers, New York.

Chang, C.-I., Du, Q., 2004. Estimation of number of spectrally distinct signal sources in hyperspectral imagery. IEEE Trans. Geosci. Remote Sens. 42 (3), 608–619.

Chen, G., Qian, S.-E., 2008. Simultaneous dimensionality reduction and denoising of hyperspectral imagery using bivariate wavelet shrinking and principal component analysis. Can. J. Remote Sens. 34 (5), 447–454.

Harshman, R.A., 1970. Foundations of the PARAFAC procedure: models and conditions for an "Explanatory" multimodal factor analysis. UCLA Work. Pap. Phonet. 16, 1–84.

Karami, A., Yazdi, M., Asli, A.Z., 2011. Noise reduction of hyperspectral images using kernel non-negative tucker decomposition. IEEE J. Sel. Top. Signal Process. 5 (3), 487–493.

Kerekes, J.P., Baum, J.E., 2005. Full-spectrum spectral imaging system analytical model. IEEE Trans. Geosci. Remote Sens. 43 (3), 571–580.

Kolda, T.G., Bader, B.W., 2009. Tensor decompositions and applications. SIAM Rev. 51 (3), 455–500.

Kruskal, J.B., 1976. More factors than subjects, tests and treatments: an indeterminacy theorem for canonical decomposition and individual differences scaling. Psychometrika 41 (3), 281–293.

Landgrebe, D., 2002. Hyperspectral image data analysis. IEEE Signal Process. Mag. 19 (1), 17–28.

Lathauwer, L.D., 1997. Signal Processing Based on Multilinear Algebra. Katholike University Leuven, Leuven, Belgium.

Letexier, D., Bourennane, S., 2008. Noise removal from hyperspectral images by multidimensional filtering. IEEE Trans. Geosci. Remote Sens. 46 (7), 2061–2069.

Lu, H., Plataniotis, K.N., Venetsanopoulos, A.N., 2008. MPCA: multilinear principal component analysis of tensor objects. IEEE Trans. Neural Netw. 19 (1), 18–39.

Matteoli, S., Acito, N., Diani, M., Corsini, G., 2011. An automatic approach to adaptive local background estimation and suppression in hyperspectral target detection. IEEE Trans. Geosci. Remote Sens. 49 (2), 790–800.

Mountrakis, G., Im, J., Ogole, C., 2011. Support vector machines in remote sensing: a review. ISPRS J. Photogramm. 66 (3), 247–259.

Muti, D., Bourennane, S., 2005. Multidimensional filtering based on a tensor approach. Signal Process. 85 (12), 2338–2353.

Othman, H., Qian, S.-E., 2006. Noise reduction of hyperspectral imagery using hybrid spatial–spectral derivative-domain wavelet shrinkage. IEEE Trans. Geosci. Remote Sens. 44 (2), 397–408.

Pande-Chhetri, R., Abd-Elrahman, A., 2011. De-striping hyperspectral imagery using wavelet transform and adaptive frequency domain filtering. ISPRS J. Photogramm. 66 (5), 620–636.

Pauluzzi, D.R., Beaulieu, N.C., 2000. A comparison of SNR estimation techniques for the AWGN channel. IEEE Trans. Commun. 48 (10), 1681–1691.

Renard, N., Bourennane, S., 2008. Improvement of target detection methods by multiway filtering. IEEE Trans. Geosci. Remote Sens. 46 (8), 2407–2417.

Renard, N., Bourennane, S., Blanc-Talon, J., 2008. Denoising and dimensionality reduction using multilinear tools for hyperspectral images. IEEE Geosci. Remote Sens. Lett. 5 (2), 138–142.

Schwarz, G., 1978. Estimating the dimension of a model. Annals Stat. 6 (2), 461–464.

Shashua, A., Levin, A., 2001. Linear image coding for regression and classification using the tensor-rank principle. In: IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, USA, pp. 42–49.

Sidiropoulos, N.D., Bro, R., 2000. On the Uniqueness of Multilinear Decomposition of N-Way Arrays. J. Chemom. 14, 229–239.

Tao, D., Li, X., Wu, X., Hu, W., Maybank, S.J., 2007. Supervised tensor learning. Knowl. Inf. Syst. 13 (1), 1–42.

Vasilescu, M.A.O., Terzopoulos, D., 2003. Multilinear subspace analysis of image ensembles. In: IEEE Conference on Computer Vision and Pattern Recognition, Madison, Wisconsin, USA, pp. 93–99.

Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. 13 (4), 600–612.

Wax, M., Kailath, T., 1985. Detection of signals by information theoretic criteria. IEEE Trans. Acoust. Speech Signal Process. 33 (2), 387–392.

Yuan, Q., Zhang, L., Shen, H., 2012. Hyperspectral image denoising employing a spectral-spatial adaptive total variation model. IEEE Trans. Geosci. Remote Sens. 50 (10), 3660–3677.

Zhang, Q., Wang, H., Plemmons, R.J., Pauca, V.P., 2008. Tensor methods for hyperspectral data analysis: a space object material identification study. J. Opt. Soc. Am. A 25 (12), 3001–3012.