ISSN 23499842(Online), Volume 1, Special Issue 1(NCFCSIT 15), March 2015
International Journal of Innovative Trends and Emerging Technologies

A HYBRID TECHNIQUE FOR DENOISING IN HYPERSPECTRAL IMAGES

A.AFREEN HABIBA[1], DR.B. RAGHU[2]
[1]Ph.D Scholar, CSE Department, Bharath University, Chennai
[2]Professor, CSE Department, Ramanujar Engineering College, Chennai
[1]*habiba.afreen@gmail.com*, [2]*raghu.balraj@gmail.com*

***Abstract*—In this paper we propose a hybrid technique for denoising in HSI. The HSI data cube is able to equally treat both spatial and spectral modes since they are measured as three order tensor. 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. To identify and distinguish spectrally unique materials, the HSI are spectrally determined, to provide sufficient spectral information. A noise-reduced hyperspectral image is then obtained by combining the rank-1 tensors using an Eigen value intensity sorting and reconstruction technique in case of single noise and K-SVD (Singular Value Decomposition) algorithm in case of Multiple Noise. By incorporating sparse regularization of small image patches, the proposed method can efficiently remove a variety of mixed or single noise while preserving the image textures well. The learned dictionary used clearly helps in removing the noise. This minimizing model removes the following mixed noise such as Gaussian-Gaussian mixture, Impulse noise and Gaussian-impulse noise from the HSI data. The weighted rank-one approximation problem arisen from the proposed model is solved by a new iterative scheme and the low rank approximation can be obtained by Singular Value Decomposition (SVD). The weighting function in the model can be determined by the algorithm itself, and it plays a role of noise detection in terms of the different estimated noise parameters. The proposed method proves the best results compared to the existing methods.**

***Keywords*— HSI (Hyperspectral Image), R1TD, K-SVD algorithm, learned dictionary, Gaussian noise, Impulse noise, Mixed noise.**

I. INTRODUCTION

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 [3]. 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 [1]. Acito [2] 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 [5] 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) and Wiener and wavelet filters some algorithms have been proposed to combine the spatial and spectral information for HSI noise reduction. Othman and Qian [15] proposed a hybrid spatial– spectral derivative-domain wavelet shrinkage noise reduction (HSSNR) approach. Chen and Qian proposed to simultaneously reduce the dimensionality and noise of HSI by the use of bivariate wavelet shrinkage. Yuan [18] 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 [17]. The multidimensional Wiener filtering (MWF) algorithm [15] 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. The MWF algorithm 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.

Disadvantages of Existing System: Existing algorithms may lead to a loss of the inter-dimensional information since the correlation between the spatial and spectral bands is not simultaneously considered. The application of a core tensor and n-mode tensor product may lead to information compression and loss of spatial detail.

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 Eigen-values generated by tensor decomposition. Here to distinguish the signal and noise profiles, that defines the contribution of the rank-1 profiles to the reconstructed signal-dominant components. 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, a series of the rank-1 tensors by R1TD should be estimated so that

the restored tensor b is as close as possible to the noise-free S 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.

The HSI are spectrally over determined, which means it provides ample spectral information to identify and distinguish spectrally unique materials. This entire denoising process is based on the K-SVD denoising algorithm. By incorporating sparse regularization of small image patches, the proposed method can efficiently remove a variety of mixed or single noise while preserving the image textures well. The learned dictionary used clearly helps in removing the noise. Our work involved in minimizing model to remove mixed noise such as Gaussian-Gaussian mixture, impulse noise and Gaussian-impulse noise from the HSI data. To solve the weighted rank-one approximation problem arisen from the proposed model, a new iterative scheme is given and the low rank approximation can be obtained by singular value decomposition(SVD) and we present a new weighting data fidelity function, which has the same minimize as the original likelihood functional but is much easier to optimize. The weighting function in the model can be determined by the algorithm itself, and it plays a role of noise detection in terms of the different estimated noise parameters.

*Advantages of Proposed System:*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.
K-SVD algorithm removes method can efficiently remove a variety of mixed or single noise while preserving the image textures well.

II. HSI AND TENSORS

A. *Analysis on Hyperspectral Images*

To understand the advantages of hyperspectral imagery, it may help to first review some basic spectral remote sensing concepts. You may recall that each photon of light has a wavelength determined by its energy level. Light and other forms of electromagnetic radiation are commonly described in terms of their wavelengths. For example, visible light has wavelengths between 0.4 and 0.7 microns, while radio waves have wavelengths greater than about 30 cm (Fig. 1).Reflectance is the percentage of the light hitting a material that is then reflected by that material (as opposed to being absorbed or transmitted). Some materials will reflect certain wavelengths of light, while other materials will absorb the same wavelengths. These patterns of reflectance and absorption across wavelengths can uniquely identify certain materials.

B. *Types of Imaging Techniques*

Depending on the number of spectral bands and wavelengths measured, an image is classified as a multispectral image when several wavelengths are measured and a hyperspectral image when a complete wavelength region, i.e., the whole spectrum, is measured

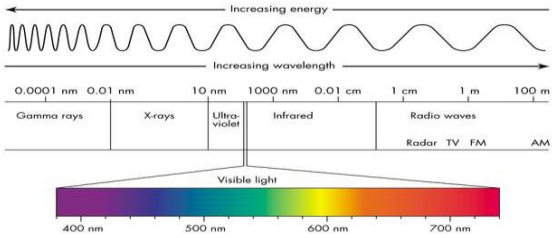

Fig 1 Electromagnetic Spectrum

for each spatial point. For example, a RGB image from a typical digital camera is a type of multispectral image that uses the light intensity at three specific wavelengths: red, green, and blue, to create an image in the visible region. The Fig 2 compares the optical information obtained by monochrome cameras, RGB cameras, and hyperspectral cameras.

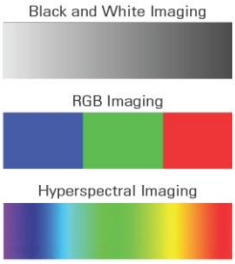

Fig 2 Differences in imaging

C. *Imaging spectrometer*

Hyperspectral images are produced by instruments called imaging spectrometers. The development of these complex sensors has involved the convergence of two related but distinct technologies: spectroscopy and the remote imaging of Earth and planetary surfaces.

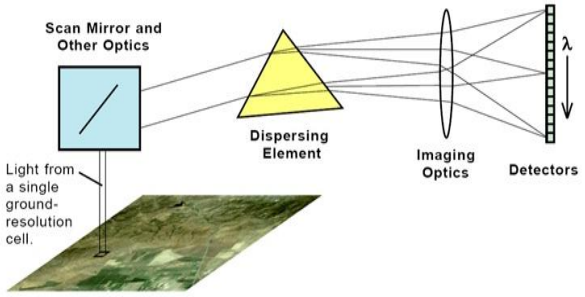

Fig 3 Imaging Spectrometer

Spectroscopy is the study of light that is emitted by or reflected from materials and its variation in energy with wavelength. As applied to the field of optical remote sensing, spectroscopy deals with the spectrum of sunlight that is diffusely reflected (scattered) by materials at the Earth's surface. Instruments called spectrometers (or spectro radiometers) are used to make ground-based or laboratory measurements of the light reflected from a test material. An optical dispersing element such as a grating or prism in the spectrometer splits this light into many narrow, adjacent wavelength bands and the energy in each band is measured by a separate detector. By using hundreds or even thousands of detectors, spectrometers

can make spectral measurements of bands as narrow as 0.01 micrometers over a wide wavelength range, typically at least 0.4 to 2.4 micrometers (visible through middle infrared wavelength ranges). Remote imagers are designed to focus and measure the light reflected from many adjacent areas on the Earth's surface. In many digital images, sequential measurements of small areas are made in a consistent geometric pattern as the sensor platform moves and subsequent processing is required to assemble them into an image.

D. Tensor

A tensor, represented as $A \in R^{L_1 \times L_2 \times \ldots\ldots \times L_N}$ is defined as a multidimensional array which is the higher-order equivalent of the vector (one-order tensor) and a matrix (two-order tensor). 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 modes1 and 2 represent the spatial modes and mode 3 denotes the spectral mode. Taking each vector to be in different mode, we can visualize the outer product of three vectors as follows,

Mathematically, we can write the outer product of three vectors a; b; c as follows,

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \circ \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \circ \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{matrix} a_1b_1c_1 & a_1b_2c_1 \\ a_2b_1c_1 & a_2b_2c_1 \end{matrix} \quad \begin{matrix} a_1b_1c_2 & a_1b_2c_2 \\ a_2b_1c_2 & a_2b_2c_2 \end{matrix}$$

We can see that the indexes of the entries in the resulting tensor.

Tensor matricization reorders the elements of an N-order tensor into a matrix from a given mode. The n-mode matricization of X belongs to $R^{L1 \times L2 \times \ldots \times LN}$ is matnX belongs to $R^{Li \times (L1L2 \ldots Ln-1\ Ln+1 \ldots LN)}$, which is the ensemble of vectors in the n-mode obtained by keeping index Li fixed and varying the other indices. A visual illustration of tensor matricization is shown in Fig 5.

III. STEPS TO REDUCE THE MIXED NOISE IN HSI

The HIS data is taken as input to the system. This HSI image is read and displayed. Then the HSI image processed in the R1TD algorithm to provide the Rank-1 Tensor profiles. With these profiles, we perform the Alternative Least Square Algorithm to optimize the tensors. Then we sort the tensors of higher order and reconstruct the noise free image by combining signal dominant components.

A. HSI Image Reader

The Hyperspectral imaging (HSI) collects and process information from across the electromagnetic spectrum. Much as the human eye sees visible light in three bands (red, blue, green), spectral imaging divides the spectrum into many more bands. This technique of dividing images into bands can be extended before can be extended beyond the visibility. Hyperspectral sensors collect information as a set of 'images'. Each image represents a range of the electromagnetic spectrum and is also known as a spectral band. These 'images' are then combined and form a three dimensional hyperspectral data cube for processing and analysis. This module is designed to read and visualize the HSI images.

The HSI data is considered as multiple images combined as a cube. Thus we have view each image in a well furnished manner. Each image ahs slice of images of different colors. This slice of image is not taken as a single color image for the calculation instead it is taken as whole cube called tensors.

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,

$$O=S+N \qquad (1)$$

In this model, the noise is assumed to be white, Gaussian and independent from signal.

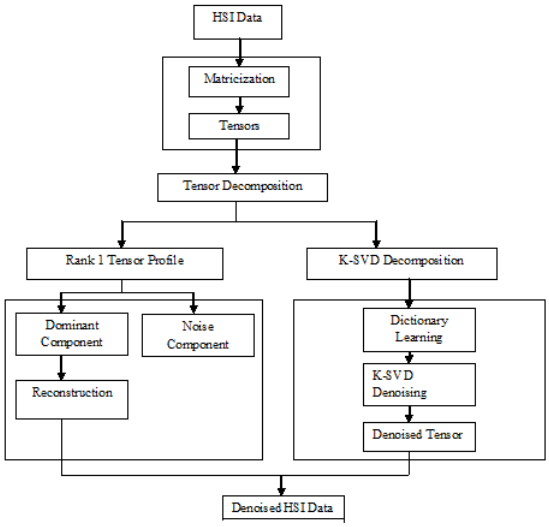

Fig 4 Architecture for Mixed Noise Reduction in HSI

B. Rank 1- Tensor Decomposition

The R1TD algorithm splits the signal into two components Signal Dominant Component and Noise Component. Finally the Signal Dominant Component alone is taken and the Reconstruction of image occurs to produce the Denoised HSI data. We develop a new tensor decomposition which jointly treats both the spatial and spectral modes. The R1TD algorithm is applied to the tensor data input which takes into account both the spatial and spectral information of the hyperspectral data cube. The tensor decomposition is of the form CANDECOMP/PARAFAC decomposition (Canonical decomposition and parallel factor decomposition).

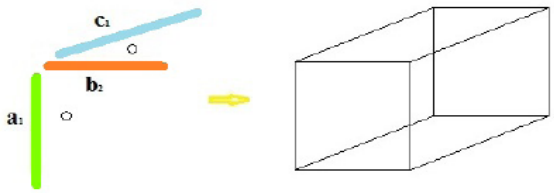

Fig 5 Tensor as the outer product of three vectors

The tensor decomposition was first attempted by Hitchcock in 1927 and Eckart and Young in 1936. However it was not fully introduced until 1970 with the

work of Harshman about the PARAFAC decomposition of Carroll and Chang about CANDECOMP. Both paper appeared in Psychometrika and explained the same decomposition. The CANDECMOP/PARAFAC is based on the fact that tensors can be rewritten as the sum of the several other tensors. Since the outer product of the three vectors gives a tensor as a result. We shall denote this tensor to be of rank 1 and we will use the term "rank 1 tensor" to denote tensors that can be written as the outer product of the vector triple. The CANDECOMP/PARAFAC decomposition rewrites a given tensor as a sum of several rank 1 tensors. Following the argument above, we define a tensor to be rank 2 if it can be expressed as the sum of two rank 1 tensors. Similarly, we define a tensor to be of rank 3 if it can be expressed as the sum of three rank 1 tensors. Thus the definition of a rank of a tensor T is the minimal number of rank 1 tensors that yield T as a linear combination.

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:

$$O = \sum_{r=1}^{M} \lambda_r U_r \circ V_r \circ W_r \quad (2)$$

where $U_r \in R^{L_1}, V_r \in R^{L_2}$ and $W_r \in R^{L_3}$ are vectors (rank-1 tensor 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, the above implies that the HSI data is a linear combination of a sequence of rank-1 tensors. However, there is currently no straight forward 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 is equivalent to the minimal number of triads necessary to describe the tensor.

C. Dictionary Learning for K-SVD

The four sub-minimization problems are solved as follows:

1) *Sparse Coding and Dictionary Learning:* The first minimization problem is

$$(\alpha^{v+1}, D^{v+1}) = \arg\min_{\alpha,D} J(\alpha, D, f_v, u_v, \Theta_v) \quad (3)$$

Applying the alternating algorithm again to this sub problem, this problem can be split into two convex sub problems corresponding to the so-called sparse coding step and the dictionary learning step, respectively. Let $\nu 1$ be an inner iteration number, then $\alpha_{v+1,}$ and D_{v+1} can be obtained by solving the following two minimization problems iteratively:

Sparse Coding (Conjugated OMP)

$$\alpha^{v_1+1} = \arg\min_{\alpha} J(\alpha, D, f_v, u_v, \Theta_v)$$

$$= arg \min_{\alpha} \left\{ \frac{\lambda}{2} \sum_{i=1}^{N} \left\| W_i D^{v_1} \alpha_{,i} - W_i R_i f^{v} \right\|_2^2 + \sum_{i=1}^{N} \mu_i \left\| \alpha_{,i} \right\|_0 \right\} \quad (4)$$

In the above, W_i is a diagonal matrix whose diagonal elements are $R_{i\omega}$.

Dictionary Learning: (Modified K-SVD)

The linear structure of K-SVD is significantly changed by the non-uniform weights. We denote

$$W = (R_{1\omega} \dots R_{N\omega}), X = (R_1 f \dots R_N f) \quad (5)$$

Then

$$D^{v_1+1} = \arg\min_{D, \|d_k\|_2=1} \{ \| W \circ (D\alpha^{v_1+1} - X^{v}) \|_F^2 \} \quad (6)$$

Similar to the K-SVD learning algorithm of [13], a natural approach to minimize each atom d_k from following energy:

$$d_k^{v_1+1} = \arg\min_{\|d_k\|_2=1} \left\| W \circ (E^k - d_k \alpha_{k,}^{v_1+1}) \right\|_F^2 \quad (7)$$

In the above, $E^k \triangleq X^v - \sum_{l=1, l \neq k}^{K} d_l^{v_1} \alpha_{l,}^{v_1+1}$. This problem is known as weighted approximation. An iterative algorithm [22] to address this difficulty is as follows

$$d_k^{v_1+1} = \arg\min_{\|d_k\|_2=1} \left\| W \circ \left(E^k - d_k \alpha_{k,}^{v_1+1}\right) + d_k^{v_1} \alpha_{k,}^{v_1+1} - d_k \alpha_{k,} \right\|_F^2 \quad (8)$$

via SVD. This algorithm cannot be used for the unweighted case. Thus we solve the minimization problem was:

$$d_k^{v_1+1} = \arg\min_{\|d_k\|_2=1} \left\| W \circ \left(E^k - d_k \alpha_{k,}^{v_1+1}\right) + \tau_k d_k^{v_1} \alpha_{k,}^{v_1+1} - \tau_k d_k \alpha_{k,} \right\|_F^2 \quad (9)$$

to update the atoms, where $\tau_k = (d_k^{v_1})^T \left(\frac{\sum_{i=1}^{N} W_i}{N} \right) d_k^{v_1}$.

Thus the modified scheme reduces the original K-SVD algorithm when all weights are the same.

Incorporating the sparse constraint, we get our modified K-SVD algorithm for weighted norm as follows:

1) Select the index set of patches S_k that use atom d_k

$$S_k = \{ i : \alpha_{k,i}^{v_1+1} \neq 0, 1 \leq i \leq N \}. \quad (10)$$

2) Let $\tau_k = (d_k^{v_1})^T \left(\frac{\sum_{i=1}^{N} W_i}{N} \right) d_k^{v_1}$, for each image patch with index i∈ S_k calculate the residual

$$e_i^{\sim k} = W_i \left(R_i f^{v} - D^{v_1} \alpha_{,i}^{v_1+1} \right) + \tau_k d_k^{v_1} \alpha_{k,i}^{v_1+1} \quad (11)$$

3) Set $\tilde{E}^k \in \mathbb{R}^{n_1 n_2 \times |S_k|}$ with its columns being the $e_i^{\sim k}$ and update $d_k^{v_1+1}$ by minimizing

$$(d_k^{v_1+1}, \beta^* = arg \min_{\|d_k\|_2=1, \beta} \left\| \tilde{E}^k - \tau_k d_k \beta^T \right\|_F^2 \quad (12)$$

where $\beta \in \mathbb{R}^{|S_k|}$. This rank-one approximation can be solved using SVD decomposition of $\tilde{E}^k$

4) Replace $\alpha_{k,i}^{v_1+1}, i \in S_k$ by relevant elements of β^*.

In our experiment, we choose the inner iteration number $v_1 = 10$.

2) *Reconstruction:* The minimization problem we have to solve is as follows, since J is quadratic with respect to f, thus

$$f^{v+1} = \left(d(\omega \circ \omega) + \lambda \sum_{i=1}^{N} R_i^T d((R_i \omega) \circ (R_i \omega)) R_i \right)^{-1} \times \left(d(\omega \circ \omega) g + \lambda (\sum_{i=1}^{N} R_i^T d((R_i \omega) \circ (R_i \omega)) R_i) \times D^{v+1} \alpha_{,i}^{v+1} \right) \quad (13)$$

Where $d(\omega \circ \omega)$ represents $diag(\omega \circ \omega)$ and R_i is a diagonal matrix. Thus the inverse matrix can be directly obtained.

3) *Noise Clustering* (Exception Step): The minimization problem we need to solve is u^{v+1} and it can be computed by

$$u_{.l}^{v+1} = \frac{\frac{r_l^v}{\sigma_l^v}\exp(-\tau_l)}{\sum_{s=1}^{M}\frac{r_s^v}{\sigma_s^v}\exp(-\tau_s)} \quad (14)$$

4) Parameter Estimation: The minimization for this step is

$$\Theta^{v+1} = \underset{\Theta,\sum r_l=1}{\arg\min}\, \mathcal{J}(\alpha^{v+1}, \mathbf{D}^{v+1}, f^{v+1}, \mathbf{u}^{v+1}, \Theta) = \underset{\Theta,\sum r_l=1}{\arg\min} \left\{ \begin{array}{l} \frac{\lambda}{2}\sum_{i=1}^{N} ||\mathbf{R}_i w \circ (\mathbf{D}^{v+1}\alpha_{.,i}^{v+1} - \mathbf{R}_i f^{v+1})||_2^2 \\ + < \mathbf{u}^{v+1} + \lambda \sum_{i=1}^{N} \mathbf{R}_i \mathbf{u}^{v+1}, \mathbf{1} > \ln \frac{\sigma_l}{r_l} \\ + \frac{1}{2}||w \circ (g - f^{v+1})||_2^2 \end{array} \right\} \quad (15)$$

From equation $\frac{\partial \mathcal{J}}{\partial \Theta} = 0$, we get the closed-form solution of Θ^{v+1}:

$$r_l^{v+1} = \frac{<u_{.,l}^{v+1}, \mathbf{1}> + \lambda <\mathbf{M} u_{.,l}^{v+1}, \mathbf{1}>}{<\mathbf{1}, \mathbf{1}> + \lambda <\mathbf{M}\mathbf{1}, \mathbf{1}>}$$

$$(\sigma_l^2)^{v+1} = \frac{<u_{.,l}^{v+1}, (g - f^{v+1}) \circ (g - f^{v+1})>}{<u_{.,l}^{v+1}, \mathbf{1}> + \lambda \sum_{i=1}^{N} <\mathbf{R}_i u_{.,l}^{v+1}, \mathbf{R}_i \mathbf{1}>} + \frac{\lambda \sum_{i=1}^{N} <\mathbf{R}_i u_{.,l}^{v+1}, (\mathbf{D}^{v+1}\alpha_{.,i}^{v+1} - \mathbf{R}_i f^{v+1}) \circ (\mathbf{D}^{v+1}\alpha_{.,i}^{v+1} - \mathbf{R}_i f^{v+1})>}{<u_{.,l}^{v+1}, \mathbf{1}> + \lambda \sum_{i=1}^{N} <\mathbf{R}_i u_{.,l}^{v+1}, \mathbf{R}_i \mathbf{1}>}. \quad (16)$$

D. Signal Dominant Component

Before going into the optimization we have to know about the operation performed with the tensors.

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

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1k}B \\ \vdots & \ddots & \vdots \\ a_{j1}B & \cdots & a_{jk}B \end{bmatrix} \in R^{(JM) \times (KN)} \quad (17)$$

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 $A \odot B$:

$$A \odot B = [a_1 \otimes b_1, a_2 \otimes b_2, \ldots, a_K \otimes b_K] \in R^{IJ \times K} \quad (18)$$

Figure shows how the Khatri–Rao product works.

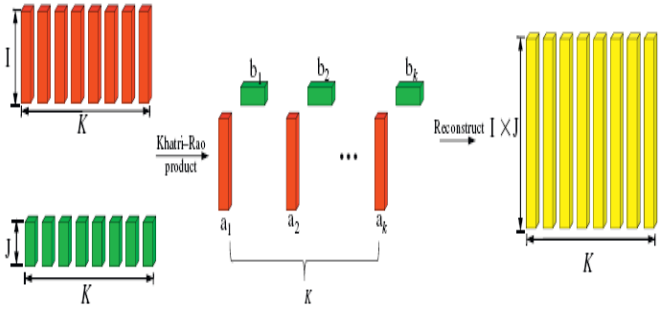

Fig 6 Visual illustration of the Khatri-Rao product

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,

$$\hat{S} = \sum_{r=1}^{k} \lambda_r U_r \circ V_r \circ W_r \quad (19)$$

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.

A series of the rank-1 tensors by R1TD should be estimated so that the restored tensor b is as close as possible to the noise-free S 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: 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.

The Optimization of each tensor is performed by the Alternative Least Square (ALS) Algorithm.

A series of the rank-1 tensors by R1TD should be estimated so that the restored tensor $\hat{S}$ 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:

$$\min_{\lambda, Ur, Vr, Wr} \left\| S - \sum_{r=1}^{k} \lambda_r U_r \circ V_r \circ W_r \right\|^2 \quad (20)$$

In this study, we combine the vectors of each rank-1 tensor in each mode into a factor matrix, i.e., U = [u1, u2, . . . , uk], V = [v1, v2, . . . , vk], and W = [w1, w2, . . . , wk]. In addition, we also denote Λ as a matrix form, i.e., Λ = diag (k1; k2; . . . ; kk). According to tensor matricization, the mode-1 flattening of Eq 5.3.3 should be expressed as:

$$mat_1 \hat{s} = U\Lambda(\mathrm{W} \odot \mathrm{V})^{\mathrm{T}} \quad (21)$$

In the alternating optimization, the solution of Eq 19 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:

$$\min_{\theta} \| mat_1 S - U^*(\mathrm{W} \odot \mathrm{V})^{\mathrm{T}} \|^2 \quad (22)$$

where U* is the weighted factor matrix computed by U* = U . Λ or $U^*_{(1)} = U_{(i)}.\Lambda_{(i,i)}$, and i = 1, 2, . . . , k since Λ is a diagonal matrix. The minimization Eq 22 is a linear least-squares problem, and its solution is written as follows:

$$U^* = mat_1 S.((\mathrm{W} \odot \mathrm{V})^{\mathrm{T}})^{-1} \quad (23)$$

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, i.e $\|U_{(i)}\| = \|V_{(i)}\| = \|W_{(i)}\| = 1$ for i = 1, 2, . . . , k. Thus, the solution for the mode-1 factor matrix [7] should be:

$$\Lambda_{(i,i)} = \|U^*_{(i)}\|, U_{(i)} = \frac{U^*_{(i)}}{\Lambda_{(i,i)}}, \quad i = 1,2,\dots\dots,k \qquad (24)$$

The objective function of Eq 22 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 t^{th} round of iteration, S is replaced by the estimated signal tensor $\hat{S}$ in the $(t-1)^{th}$ round of iteration. The Algorithm converges when the error of the estimated signal tensor $\hat{S}$ between two iterations decreases to a small value. Following algorithm summarizes the proposed ALS algorithm for HSI noise reduction.

ALS Algorithm

Input: Input HSI tensor $O \epsilon R^{L_1 \times L_2 \times L_3}$, decomposition level K, maximum number of iteration ITER

Initialization: Set U, V, W to the identity matrix, $So = 0$
Step 1: For t=1 to ITER{
Step 2: Calculate

$$U^* = mat_1 S^t.((W \odot V)^T)^{-1}, \quad \Lambda_{(i,j)} = \|U_{(1)}\|, \quad U_{(i)} = U^*_1/\Lambda_{(i,j)}$$
$$V = mat_1 S^t.((W \odot U)^T)^{-1}, \quad \Lambda_{(i,j)} = \|V_{(1)}\|, \quad V_{(i)} = V^*_1/\Lambda_{(i,j)}$$
$$W^* = mat_1 S^t.((U \odot V)^T)^{-1}, \quad \Lambda_{(i,j)} = \|W_{(1)}\|, \quad W_{(i)} = W^*_1/\Lambda_{(i,j)}$$

Step 3: Reconstruct estimated noise-free tensor by
Step 4: Check for convergence, if:

$$Err(t) = \|S^t - S^{t-1}\| < \varepsilon\}//\text{for loop in step}$$

Output: The denoised tensor $\hat{S} \in R^{L_1 \times L_2 \times L_3}$, weight matrix Λ, and factor matrices U,V,and W
After optimizing the tensors the Signal dominant tensor are chosen. The lowest value say (0.0001) and those having values less than this is taken as signal dominant components and all the other components are treated as noise components.

E. Denoising

The signal dominant components are combined leaving the noise tensors to the noise free image. After the noise components are removed, the signal-dominant components are obtained by reconstructing the remaining noise free tensor.

The tensors are reconstructed to form the noise free HSI data by the formula

$$\hat{S} = \sum_{r=1}^{k} \lambda_r U_r \circ V_r \circ W_r \qquad (25)$$

The value of K is the number of signal dominant tensors.

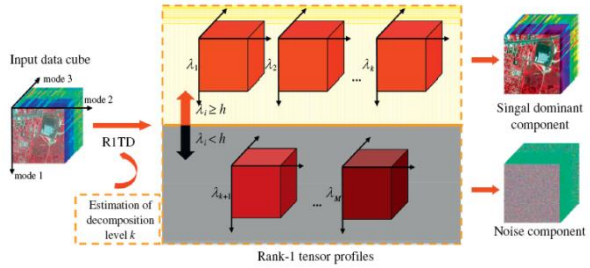

Fig 7 Main flow chart of the R1TD algorithm for HSI noise reduction

IV. RESULTS

The proposed algorithm is applied in 3 set of HSI data.
The HIS cannot be taken as an image itself. The values are to plotted as an image for our visualization. Thus a set of values of the received image is plotted as an image for our visualization. The values are plotted as an image for the original data and for the Denoised data. The original values are not plotted fully, only certain area shown for the visualization for a clear idea of the HSI image

To verify the effectiveness of the proposed algorithm, the proposed model is compared with several competitive methods: Spectral-Spatial Adaptive Total Variation (SSAHTV), Multidimensional Wiener Filtering (MWF) and Rank-1 Tensor Decomposition (R1TD). The algorithm SSAHTV [15] and MWF [13] may lead to loss of inter-dimensional information since the correlation between the spatial and spectral bands are not simultaneously considered. The application of a core tensor and n-mode tensor product may lead to information compression and loss of spatial details.
The R1TD [24] provides clear view than that of the other two but it deals with only Additive white and Gaussian noise.
The proposed algorithm deals mixed noise like impulse, Gaussian-Gaussian, Gaussian-impulse. Also it provides a higher PSNR than that of the existing system.

The PSNR is an engineering term for ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. Because many signals have a very wide dynamic range, PSNR is usually expressed in terms of logarithmic decibel scale.

The PSNR for the Existing System is compared with the proposed algorithm in the Table I. This comparison confirms that proposed method is has higher values than that of the existing systems. Also the existing system deals only with single noise whereas this deals with mixed noise.

TABLE I

PSNR COMPARISON FOR THE PROPOSED R1TD ALGORITHM AND THE EXISTING SYSTEMS

	Existing Method-PSNR		**Proposed Method-PSNR**
Band Number	**MWF**	**SSAHTV**	**R1TD**
1	24.47984	28.39629	30.3891
2	24.66283	26.44052	30.4351
3	25.48411	28.43831	30.4328
4	24.3141	28.2111	30.1965
5	24.73444	29.01946	29.9972
6	23.73444	27.5634	29.5229

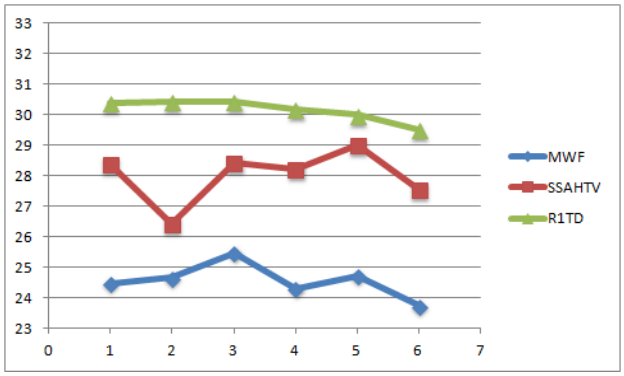

Fig 8 Graph for PSNR Comparison of The Proposed R1TD Algorithm and The Existing Systems

TABLE III

PSNR COMPARISON FOR THE PROPOSED R1TD ALGORITHM AND K-SVD ALGORITHM WITH THE EXISTING SYSTEMS

	Existing Method-PSNR		Proposed Method-PSNR	
Band Number	**MWF**	**SSAHTV**	**R1TD**	**Improved K-SVD**
1	24.48	28.4	30.39	41.8
2	24.66	26.44	30.43	40.23
3	25.48	28.43	30.43	37.2
4	24.31	28.21	30.2	37.1
5	24.73	29.02	29.99	36.45
6	23.73	27.56	29.52	36.25

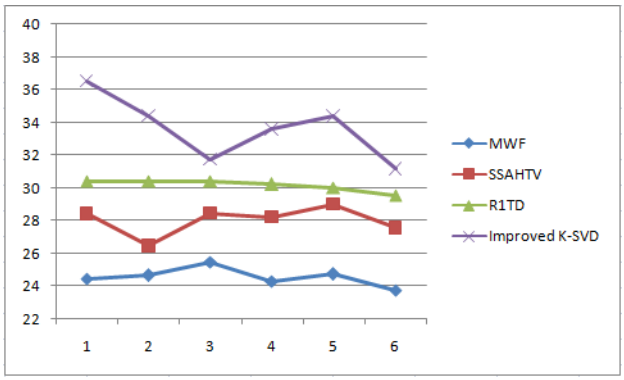

Fig 9 Graph for PSNR Comparison of The Proposed R1TD Algorithm and K-SVD Algorithms and The Existing Systems

V. CONCLUSION

In this study, the high-order rank-1 tensor decomposition (R1TD) model for single noise and the K-SVD model for mixed noise is investigated to develop a new noise removal algorithm for hyperspectral image pre-processing. The main advantage of the R1TD algorithm and the K-SVD 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 and the K-SVD 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 and mixed 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 is 39.578 as an average. Due to the utilization of tensor representation, images generated with both fixed noise intensity and random noise intensity was effectively processed in the synthetic scenarios.

References

[1] 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.
[2] 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.
[3] 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.
[4] Ana Rovi., 2010. Analysis of 2× 2 × 2 Tensors. master's thesis
[5] M. Elad and M. Aharon, "Image denoising via learned dictionaries and sparse representation," in Proc. IEEE Comput. Vis. Pattern Recognit.,Jun. 2006, pp. 895–900.
[6] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," IEEE Trans. Image Process.,vol. 15, no. 12, pp. 3736–3745, Dec. 2006.
[7] M. Aharon, M. Elad, and A. Bruckstein, "The K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representations,"IEEE Trans. Image Process., vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
[8] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," IEEE Trans. Image Process., vol. 17, no. 1, pp. 53–69, Jan. 2008.
[9] 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.
[10] Brett W. Bader and Tamara G. Kolda MATLAB Tensor Classes for Fast Algorithm Prototyping
[11] Bro, R., Kiers, H.A.L., 2003. A new efficient method for determining the number of components in PARAFAC models. J. Chemometr. 17 (5), 274–286Karami, 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.
[12] Landgrebe, D., 2002. Hyperspectral image data analysis. IEEE Signal Process. Mag. 19 (1), 17–28.
[13] Letexier, D., Bourennane, S., 2008. Noise removal from hyperspectral images by multidimensional filtering. IEEE Trans. Geosci. Remote Sens. 46 (7), 2061–2069.
[14] 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.
[15] 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.
[16] H. Wang and R. Haddad, "Adaptive median filters: New algorithms and results," IEEE Trans. Image Process., vol. 4, no. 4, pp. 499–502, Apr.1995.
[17] Y. Xiao, T. Zeng, J. Yu, and M. K. Ng, "Restoration of images corrupted by mixed Gaussian-impulse noise via l1-l0 minimization,"Pattern Recognit., vol. 44, no. 8, pp. 1708–1720, Aug. 2011.

[18] E. Lopez-Rubio, "Restoration of images corrupted by Gaussian
[19] and uniform impulsive noise," Pattern Recognit., vol. 43, no. 5,pp. 1835–1846, 2010.
[20] J. Liu, H. Huang, Z. Huan, and H. Zhang, "Adaptive variational method for restoring color images with high density impulse noise," Int. J.Comput. Vis., vol. 90, no. 2, pp. 131–149, 2010
[21] J. Bilmes. (1997). A Gentle Tutorial on the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.28.613
[22] N. Srebro and T. Jaakkola, "Weighted low-rank approximations," in Proc. 20th Int. Conf. Mach. Learn., 2003, pp. 720–727.
[23] S. Ko and Y. Lee, "Center weighted median filters and their applications to image enhancement," IEEE Trans. Circuits Syst., vol. 38, no. 9, pp.984–993, Sep. 1991.
[24] Xian Guo a, Xin Huang a,⇑, Liangpei Zhang a, Lefei Zhang"Hyperspectral image noise reduction based on rank-1 tensor Decomposition", ISPRS Journal of Photogrammetry and Remote Sensing 83 (2013) 50–63