Article

Efficient Blind Hyperspectral Unmixing Framework Based on CUR Decomposition (CUR-HU)

Muhammad A. A. Abdelgawad *, Ray C. C. Cheung and Hong Yan

Department of Electrical Engineering and Centre for Intelligent Multidimensional Data Analysis, City University of Hong Kong, Hong Kong; r.cheung@cityu.edu.hk (R.C.C.C.); h.yan@cityu.edu.hk (H.Y.)
* Correspondence: mabdelgaw2-c@my.cityu.edu.hk

Abstract: Hyperspectral imaging captures detailed spectral data for remote sensing. However, due to the limited spatial resolution of hyperspectral sensors, each pixel of a hyperspectral image (HSI) may contain information from multiple materials. Although the hyperspectral unmixing (HU) process involves estimating endmembers, identifying pure spectral components, and estimating pixel abundances, existing algorithms mostly focus on just one or two tasks. Blind source separation (BSS) based on nonnegative matrix factorization (NMF) algorithms identify endmembers and their abundances at each pixel of HSI simultaneously. Although they perform well, the factorization results are unstable, require high computational costs, and are difficult to interpret from the original HSI. CUR matrix decomposition selects specific columns and rows from a dataset to represent it as a product of three small submatrices, resulting in interpretable low-rank factorization. In this paper, we propose a new blind HU framework based on CUR factorization called CUR-HU that performs the entire HU process by exploiting the low-rank structure of given HSIs. CUR-HU incorporates several techniques to perform the HU process with a performance comparable to state-of-the-art methods but with higher computational efficiency. We adopt a deterministic sampling method to select the most informative pixels and spectrum components in HSIs. We use an incremental QR decomposition method to reduce computation complexity and estimate the number of endmembers. Various experiments on synthetic and real HSIs are conducted to evaluate the performance of CUR-HU. CUR-HU performs comparably to state-of-the-art methods for estimating the number of endmembers and abundance maps, but it outperforms other methods for estimating the endmembers and the computational efficiency. It has a 9.4 to 249.5 times speedup over different methods for different real HSIs.

Keywords: hyperspectral unmixing (HU); low-rank approximation; nonnegative matrix factorization (NMF); CUR matrix decomposition

1. Introduction

Hyperspectral imaging technique has attracted much attention in remote sensing, which collects reflected light of hundreds of contiguous and narrow wavelengths, having rich spectral information about the various materials in the scene. It is a powerful tool in different application fields, such as mineralogy, agriculture, surveillance, space exploration, environmental monitoring, and disease diagnosis [1,2]. There is a tradeoff between the spectral and spatial resolutions, which leads the observed signature of each pixel to be a mixture of several materials [3]. HU techniques are source separation methods that decompose each mixed pixel in HSI into a set of pure spectral signatures called endmembers and their corresponding fractions called abundance maps. More specifically, the HU process is composed of three tasks: (1) estimating the optimal number of endmembers, (2) extracting the endmember signatures, and (3) estimating the abundance maps [4]. Estimating the optimal number of endmembers, virtual dimensionality (VD) [5], is crucial because of the performance of extracting endmembers and estimating the abundance maps depending

on it. Most of HU algorithms mainly focus on the last two tasks of the HU process, i.e., estimating the endmembers and abundance maps sequentially or simultaneously [6,7], by assuming that the number of endmembers is available in advance.

The hyperspectral mixing and unmixing algorithms can be classified into linear and nonlinear models according to the assumptions of the interaction of light with the observed field of view [4,8,9]. Most existing HU algorithms are based on the linear mixing model (LMM), a straightforward and simple method for representing the mixing process. Furthermore, two physical abundance constraints are required while considering linear unmixing: the nonnegative abundance constraint (ANC) and the abundance sum-to-one constraint (ASC) [10]. The endmember extraction and abundance maps estimation of the HU process can be solved separately. Estimating the abundance maps is an ill-posed linear inversion problem, and we can solve it using the least-squares-based methods, such as fully constrained least squares (FCLS) [7]. To reduce the accumulative errors in two-stage-based HU algorithms, the statistical-based HU has been proposed, where the HU is considered a BSS problem, so the endmembers and abundance maps are estimated simultaneously [11].

Nonnegative matrix factorization (NMF) [12] is the popular method for solving BSS problems, which has been utilized in the HU process. NMF decomposes the nonnegative data matrix into a product of two nonnegative matrices: a source matrix and a mixing matrix. Although standard and constrained NMF methods have been proposed for blind HU, they do not guarantee to provide a unique solution, the decomposition results depend on the initialization, and they require high computational costs to obtain the solution because they are iterative methods. In addition, the interpretation and analysis of the resultant matrices from NMF are complex.

Because HSIs consist of relatively few prominent features (i.e., endmembers), we can approximate HSIs using the low-rank approximation method to explore those prominent features and their fraction at each pixel. Many low-rank decomposition methods exist, such as Singular Value Decomposition (SVD), QR factorization, NMF, Interpolative Decomposition (ID), CUR matrix decomposition, and Principal Component Analysis (PCA) [13]. In order to demonstrate that the HSI data are almost low-rank or can be well approximated as low-rank data, we consider a real HSI dataset known as the Samson HSI dataset [14]. It comprises 95×95 pixels with 156 spectral bands and has three reference endmembers (i.e., soil, tree, and water). The Samson data cub can be reshaped into an HSI data matrix of size 156×9025, whose rows represent the spectral bands, and its columns represent the pixels. Figure 1a shows the Samson data matrix, which clearly can be well approximated as low-rank data and constructed from a few components. Figure 1b shows the ground truth of normalized endmembers and their abundance maps of the Samson HSI dataset. Moreover, we selected 15 pixels randomly from the HSI data matrix and visualized their reflectance along the spectral bands, as shown in Figure 1c. As can be seen, the 15 signatures can be represented within three groups (i.e., solid, dash, and dash-dot lines), where each member in a group is a linear combination of the group members. In other words, each pixel of the HSI dataset can be represented as linear combinations of the main component of each group. Thus, the HSI is almost a low-rank and well-approximated by a low rank, and the rank here represents the number of endmembers.

Inspired by the CUR matrix factorization methods [15,16] for finding the low-rank approximation, interpreting the resulting matrices from the decomposition, and their performance in various application areas, we propose a novel framework for performing the main three tasks of the HU process based on CUR factorization called CUR-HU. It is a new blind unmixing method incorporating CUR matrix decomposition based on a deterministic sampling method for selecting the columns and rows from the observed mixed HSIs that have the maximum independence and provide a unique solution. CUR-HU automatically estimates the endmembers and their abundance maps without requiring prior knowledge about the scene. Unlike NMF, CUR-HU does not depend on the initialization method, and the resulting matrices are easy to interpret as they contain the same data element as the original given HSIs. To decrease the complexity of computing the SVD of an HSI, which

the sampling method requires computing the selected indices of rows and columns, the QR factorization based on an incremental approach is utilized to approximate the SVD and estimate the number of endmembers in a given HSI. Figure 2 shows an overview of the proposed framework (i.e., CUR-HU) for performing the whole HU process. The performance of the proposed method in each stage of the complete HU process is evaluated using synthetic and real standard HSI datasets, with a comparison with several selected state-of-the-art HU baselines. The experiments demonstrate that CUR-HU outperforms other state-of-the-art blind HU algorithms, especially for extracting the endmember spectra and the computational efficiency. CUR-HU has a 9.4 to 249.5 times speedup over different state-of-the-art methods for estimating the endmembers and their abundance maps of different real HSI datasets.
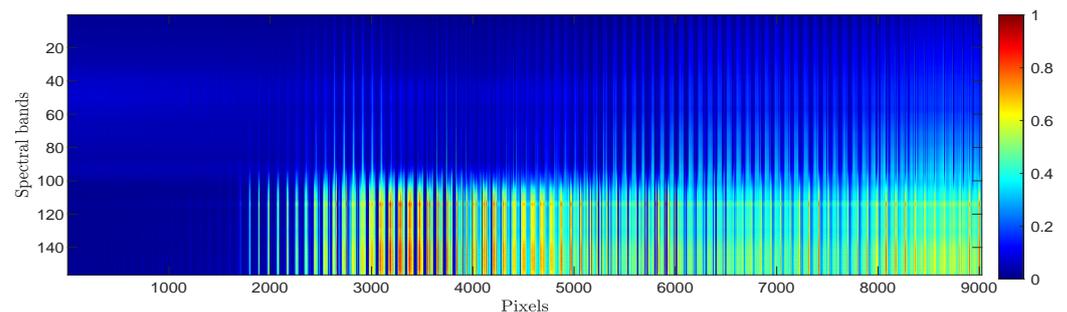
(**a**) A reshaped Samson HSI cube into HSI data matrix of size 156×9025, which can be well approximated as low-rank data and constructed from a few components..
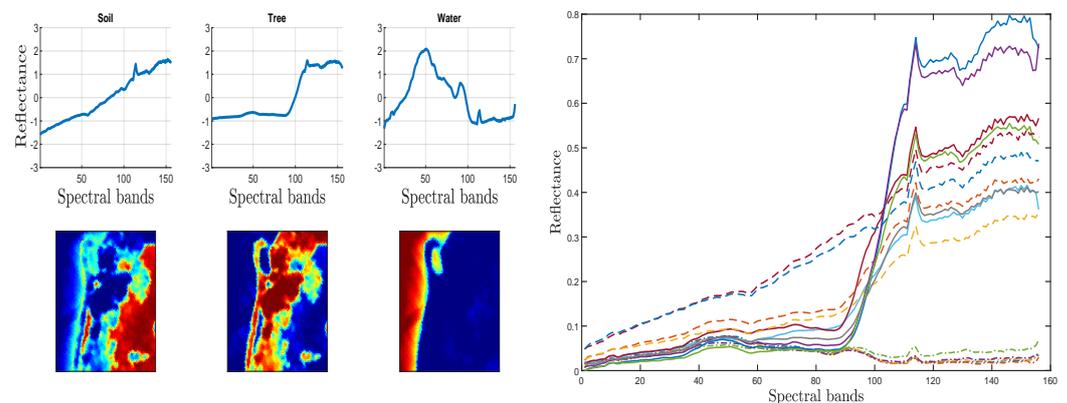
(**b**) Ground truth of normalized endmembers and their abundance map of the Samson HSI dataset.

(**c**) The signatures of 15 random pixels selected from the HSI data matrix of the Samson dataset.

Figure 1. Samson HSI dataset from the view of low-rank representation.

The contributions of this paper are summarized as follows:

- We propose a new framework based on CUR matrix factorization called CUR-HU for performing the complete HU process. To the best of our knowledge, this is the first time that CUR decomposition has been applied to solve the HU problem.
- CUR-HU incorporates several techniques to solve the blind HU problem with unmixing performance comparable to state-of-the-art methods but with much higher computational efficiency. It does not require initialization or prior information about the scene.
- Various experiments on synthetic and real HSIs were conducted, and the results show that CUR-HU performs comparably to state-of-the-art methods for estimating the number of endmembers, extracting the abundance maps, and robustness to spectral variability. It outperforms other methods for estimating the endmembers and the computational efficiency.

- CUR-HU can be used to initialize the existing HU algorithms, especially iterative methods that are sensitive to the initialization step. Furthermore, it can provide supervised information for methods based on deep learning techniques.
- CUR-HU has a 9.4 to 249.5 times speedup over different state-of-the-art methods for estimating the endmembers and their abundance maps of different real HSI datasets.

The rest of the paper is organized as follows. Section 2 presents the related work for solving the HU problem. Section 3 briefly introduces the concepts and methods used in our proposed method, including CUR decomposition, selection method, NMF, and the linear mixture model (LMM). Section 4 presents the proposed HU method, which is called CUR-HU. Section 5 provides extensive experiments and performance comparison with the state-of-the-art methods using both simulated and real HSIs. Finally, the conclusion and discussion of our future work are presented in Section 6.
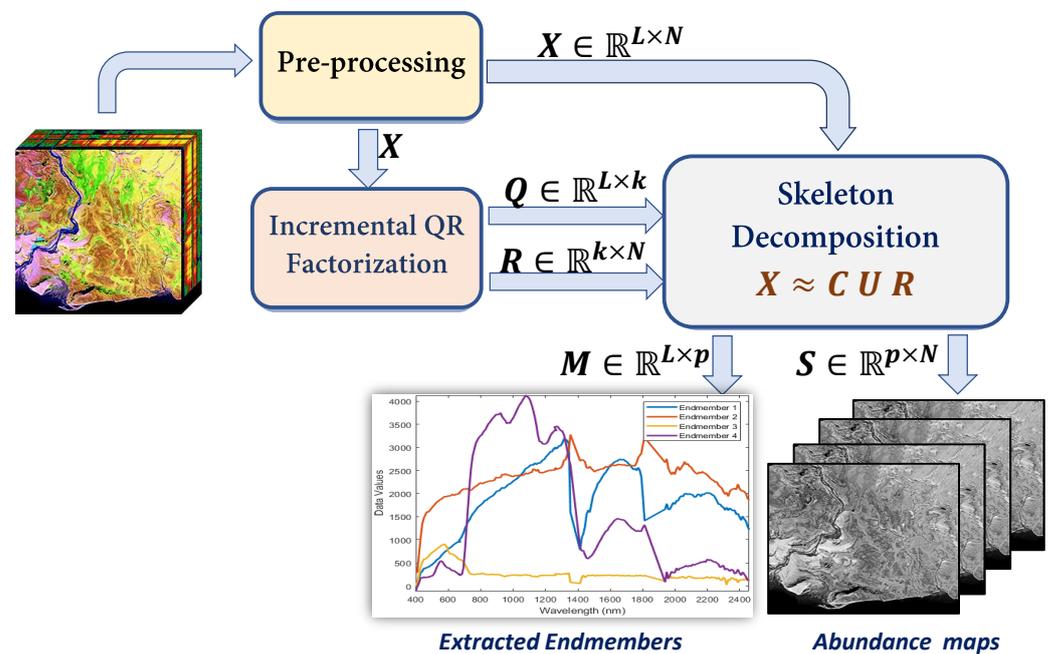
Figure 2. An overview of the proposed framework (i.e., CUR-HU) for solving the blind hyperspectral unmixing problem. CUR-HU considers the main three stages of the whole HU process.

Notation

In this paper, the lower and upper case bold are used for representing vectors and matrices, respectively, and x_{ij} is the element in i-th row and j-th column of matrix $\mathbf{X}$. We follow Matlab style notation, i.e., $X(I, J)$ is the submatrix which contains the rows and columns of X indexed by I and J. The SVD of $X = W\Sigma V^T$ in which $W \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ are the left and right singular matrices, and $\Sigma \in \mathbb{R}^{m \times n}$ contains the singular values of X. The Frobenius norm of a matrix is $\|X\|_F$ and the Moore–Penrose inverse of X is $\mathbf{X}^\dagger$. The identity matrix of size $m \times m$ is represented by $\mathbf{I}_m$.

2. Related Work

Various methods have been proposed for estimating the number of endmembers. They can be categorized into two groups [17]: (1) The eigen analysis-based methods, e.g., HFC, NWHFC [5], and HySime [18]. The HySime method tries to find the optimal signal subspace by minimizing the mean-square error (MSE) between the projected data onto the subspace and the original data, where the number of endmembers is equal to the optimal dimensionality of the subspace. NWHFC and HFC methods aim to find the most significant eigenvalue gap between the eigenvalues of the sample correlation and covariance matrices, which determines the appropriate number of eigenvectors that can represent the HSIs data. (2) The information-theoretic criteria-based algorithms, e.g.,

the Akaike information criterion (AIC) and the minimum description length (MDL) [19]. Many algorithms have been proposed to solve the HU problem, and they can be divided into three classes according to the computational methods: (1) geometric algorithms, (2) sparse regression-based algorithms [4], and (3) statistical algorithms. Geometrical methods utilize the orientation of HSIs by assuming that the endmembers of HSI are the vertices of a simplex with the minimum or maximum volume enclosing the data set or contained in the convex hull of the data set, respectively. The classical geometrical algorithms include N-FINDR [20], the pixel purity index (PPI) [21], the vertex component analysis (VCA) [6], the simplex growing algorithm (SGA) [22], and simplex identification via split augmented Lagrangian [23]. VCA and N-FINDR are iteratively searching about the simplex with maximum volume, whose vertices are considered endmembers. VCA is an iterative approach that projects data in the orthogonal direction to the subspace spanned by the extracted endmembers in the previous steps, while N-FINDR randomly selects some pixels as the initial endmembers matrix, then changes those pixels by other pixels to find a set of pixels with the maximum volume to extract endmembers. Although the geometrical unmixing methods have worked well for extracting endmembers, they may fail to extract endmembers from the highly mixed HSI data, especially if the pure spectral signatures do not exist. Sparse regression-based methods use large known libraries to formulate the HU problem as a sparse linear regression problem [24,25]. Extracting realistic endmembers directly from the HSIs data is a big challenge without prior knowledge. Because more large spectral libraries have become openly available, many algorithms have been proposed to select the best subset of endmembers signatures from those libraries to represent each pixel by considering the spectral library as a priori knowledge [26]. This approach aims to use only a few spectra in a given spectral library to model each mixed pixel in the HSI. It is called sparse unmixing because the number of endmembers in a given HSI is much smaller than the available spectra in the library, which leads to a sparse solution. The authors in [26] proposed a sparse unmixing using spectral a priori information unmixing method that uses the spectral a priori information in the HSI to address the high mutual coherence problem of the spectral library. However, extracting the endmembers from spectral libraries is usually inconsistent with HSI pixel signatures because of acquisition conditions, affecting the subsequent endmember extraction step, i.e., abundance estimation.

Statistical methods of HU use statistical representation to interpret the mixed pixels. They are powerful techniques for decreasing the accumulative errors in the two-step-based HU methods because they consider the HU as a BSS problem to estimate the endmember signatures and extract the abundance maps simultaneously.

Independent component analysis (ICA) [27,28], Bayesian self-organizing maps [29], independent factor analysis (IFA) [30], and nonnegative matrix factorization (NMF) [12] are some of the common statistical methods utilized for HU. HU based on ICA and IFA methods has been proposed based on two assumptions: the observed spectrum vector should be a linear of the endmember signatures weighted by the correspondent sources (abundance fractions) and the sources should be statistically independent. Thus, the HU itself is not a strict ICA or IFA problem. The authors in [30] claim that the ICA does not play a competitive role in HSI because the statistical independence of the sources, i.e., assumed by ICA and IFA methods, is violated in the HU, compromising the performance of ICA and IFA methods for solving HU. As a result, nonparametric statistical unmixing methods based on the dependent component analysis are proposed in [31–33]. It models the abundance maps as mixtures of Dirichlet densities, which enforces the non-negativity and constant sum constraints on the abundance fractions. The mixing matrix is estimated using a generalized expectation–maximization (GEM)-type algorithm. Although this method provides a reasonable approach to avoid the independence assumption of ICA, the Dirichlet densities of the abundances need to be acquired in advance. Moreover, a matrix inversion operation is required in each iteration step, which may lead to performance degradation. Compared with the sparse regression-based methods and geometrical algorithms, the

NMF-based HU algorithms are powerful and widely used to extract the endmembers and their abundance maps simultaneously [34].

NMF-based algorithms work well for solving HU, but they are ill-posed geometric algorithms, meaning that the solution may not be unique. Various auxiliary regularizers have been introduced to add to conventional NMF to improve the uniqueness of the HU solution, e.g., minimum volume rank deficient NMF (Min-vol NMF) [35], $l_{1/2}$-sparsity constrained NMF ($l_{1/2}$-NMF) [36], manifold regularized sparse NMF [37], spatial group sparsity regularized NMF (SGSNMF) [38], robust collaborative NMF (R-CoNMF) [39], graph regularized NMF [40], and subspace structure regularized NMF (SSRNMF) [41]. The SGSNMF method incorporates a spatial group sparsity regularizer into the NMF-based HU method to explore the sparse structure and the use of spatial information of HSIs. It exploits the idea that the pixels within a local spatial are expected to have the same pattern in the low-rank matrix of the abundance map. For $l_{1/2}$-NMF, the $l_{1/2}$-norm is used instead of the l_1-norm to obtain a sparse solution. Because the $l_{1/2}$ regularization provides high unmixing performance, it is used in different NMF-based HU methods, such as structure-constrained sparse NMF [42] and graph regularized L1/2-NMF [37]. Recently, nonnegative tensor factorizations have been introduced to solve the blind HU problem, e.g., Matrix-Vector Nonnegative Tensor Factorization for Blind HU (MVNTF) [43], which preserves the spatial and spectral information to perform the factorization on the hyperspectral alternative 2D image. Moreover, HU methods based on deep learning (DL) have been proposed to achieve more competitive unmixing performance [44–46]. However, there are still several drawbacks, i.e., requiring many training samples and network parameters to achieve satisfactory performance, high computational complexity, and difficulty interpreting the results from the network [47,48]. Most existing HU methods consider only one or two tasks of the HU process, i.e., estimating the end members and their abundance and assuming that the number of endmembers is available, which is uncommon in real scenarios. In this paper, we consider the whole HU process.

3. Preliminaries

This section introduces the primary methods we build on, the Linear Mixture Model (LMM), NMF, and CUR or Skeleton matrix factorization.

3.1. Linear Mixture Model (LMM)

In this paper, we consider the LMM model [3], in which each light ray interacts with only an endmember in the scene before reaching the imaging sensor. The spectrum at each pixel in HSI is modeled as a linear combination of all endmembers. Assume that endmembers exist in the HSI dataset and the observed spectral data vectors are given by

$$\mathbf{y} = \mathbf{x} + \mathbf{e} \tag{1}$$

where $\mathbf{x} \in \mathbb{R}^{L \times 1}$ and $\mathbf{e} \in \mathbb{R}^{L \times 1}$ are the signal and additive noise, respectively, and L is the number of spectral bands. The noise here includes endmember variability, model inadequacies, and sensor noise [49,50]. The signal vectors lie on an unknown p dimension subspace and are given by

$$\mathbf{x} = \sum_{k=1}^{p} m_k s_k = \mathbf{Ms} \tag{2}$$

where $\mathbf{M} \in \mathbb{R}^{L \times p}$ is a full rank matrix with a rank equal to p, where p is the number of endmembers in HSI, and $p < L$. Each column of $\mathbf{M}$ contains an endmember signature of HSI, and $\mathbf{s} \in \mathbb{R}^{p \times 1}$ is the abundance vector. Let $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_N] \in \mathbb{R}^{p \times N}$ represent abundance maps matrix. The LMM model can be formulated in a matrix form as

$$\mathbf{Y} = \mathbf{MS} + \mathbf{E} \tag{3}$$

where $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N] \in \mathbb{R}^{L \times N}$ is the HSI data matrix, and $\mathbf{E} \in \mathbb{R}^{L \times N}$ is the additive noise matrix, where N is the number of pixels. The abundance maps matrix should obey positivity and sum-to-one constraints to be physically meaningful.

3.2. Nonnegative Matrix Factorization

The LMM formulation of the HU problem can be considered a BSS problem to extract the endmember spectra and the abundance maps simultaneously. NMF can be used to solve the blind HU problem as it decomposes a given nonnegative data matrix $\mathbf{V} \in \mathbb{R}_{+}^{n \times m}$ into two other nonnegative matrices $\mathbf{W} \in \mathbb{R}_{+}^{n \times r}$ (the source matrix) and $\mathbf{H} \in \mathbb{R}_{+}^{r \times m}$ (the mixing matrix) which satisfy the following approximation.

$$\mathbf{V} \approx \mathbf{WH} \tag{4}$$

An infinite number of $\mathbf{H}$ and $\mathbf{W}$ pairs satisfy the above factorization. For example, $\mathbf{WH} = \left(\mathbf{W}\mathbf{\Gamma}^{-1}\right)(\mathbf{\Gamma}\mathbf{H})$ for any invertible $\mathbf{\Gamma} \in \mathbb{R}_{+}^{r \times r}$ matrix. The conventional procedure of NMF is to formulate the factorization using an objective function that quantifies the quality of approximation between $\mathbf{V}$ and $\mathbf{WH}$, and then, an optimization algorithm is used to minimize or maximize the defined objective function concerning $\mathbf{H}$ and $\mathbf{W}$. One of the widely utilized objective functions for NMF is the square of the Frobenius norm between the given data matrix and decomposed matrices as follows,

$$\|\mathbf{V} - \mathbf{WH}\|_F^2 = \sum_{ij} \left(\mathbf{V}_{ij} - (\mathbf{WH})_{ij}\right)^2. \tag{5}$$

Although the above equation is convex in $\mathbf{H}$ and $\mathbf{W}$ alone, it is not convex in $\mathbf{H}$ and $\mathbf{W}$ together [12]. Therefore, finding the global minima of this function w.r.t. $\mathbf{H}$ and $\mathbf{W}$ analytically is difficult. It is possible to find local minima numerically using optimization methods.

The HU problem can be represented as a conventional NMF form [12,39,51]. Let the HSI matrix $\mathbf{Y}$ be $\mathbf{V}$, the endmember matrix $\mathbf{M}$ is the source matrix, and the abundance matrix $\mathbf{S}$ is the mixing matrix of NMF. Then, the HU problem can be formulated as

$$\underset{\mathbf{M,S}}{\arg\min} \|\mathbf{Y} - \mathbf{MS}\|_F^2, \quad \text{s.t. } \mathbf{M}, \mathbf{S} \succeq 0 \tag{6}$$

Although various auxiliary regularizers on $\mathbf{M}$ and $\mathbf{S}$ have been proposed to improve the uniqueness of the solution, the NMF-based HU methods still have some drawbacks, such as the initialization methods, the difficulty of driving the objective function, requiring extensive computational time to solve the optimization problem, the resultant solution is not unique, and the difficulty of interpreting the resultant matrices concerning the original HSIs.
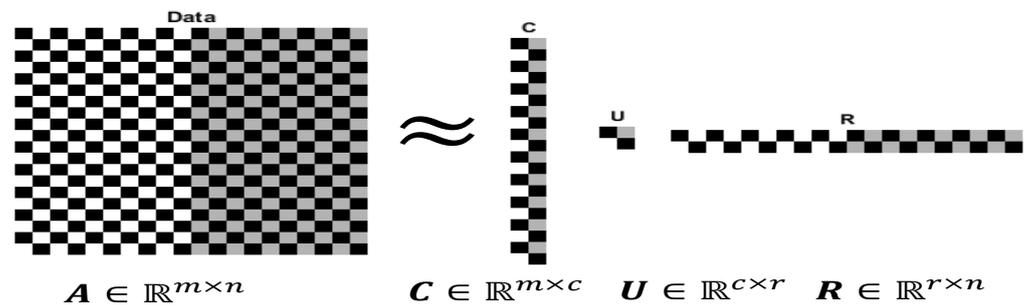
3.3. CUR Matrix Factorization

Low-rank matrix approximations based on the selected columns and rows from a given data matrix naturally provide interpretable results. They have been successfully applied in various application areas. CUR matrix factorization or CUR decomposition is a widely-used technique representing the data matrix as a product of three small submatrices of original data [15,16]. A visual representation of CUR matrix decomposition is shown in Figure 3. CUR of a matrix $A \in \mathbb{R}^{m \times n}$ is a low-rank approximation that can be represented as

$$A \approx CUR \tag{7}$$

where $\mathbf{C} \in \mathbb{R}^{m \times c}$ and $\mathbf{R} \in \mathbb{R}^{r \times n}$ matrices are constructed by extracting c columns and r rows of A, respectively. The matrix U should be formed so that the CUR approximation error $\|A - CUR\|_F$ is very small. CUR matrix decompositions give more interpretable results because they represent data in lower dimensional spaces using actual data points as

basis vectors that maintain sparseness or nonnegativity properties, unlike the other matrix factorizations. It originated with a pseudoskeleton and truncated QR approximation [52,53].

Figure 3. An illustration of CUR matrix decomposition, **C** and **R** matrices are subsets of columns and rows of a data matrix **A**, respectively.

4. CUR Factorization for Hyperspectral Unmixing (CUR-HU)

This section presents the proposed framework (i.e., CUR-HU) for performing the main three tasks of the HU process. A general assumption is that a low-dimensional subspace is hidden in the HSI dataset, as demonstrated in Figure 1. The endmember signatures of the given HSI represent the macroscopic objects of the HSI scene, e.g., water, soil, and vegetation. HU algorithms attempt to extract those macroscopic objects from the observed mixed signals. As those endmember signatures are statistically independent, extracting them from the mixed observed signals can be achieved by finding a group of pixel spectra with maximum independence between them. In other words, the extracted endmember signatures should be more mutually independent than the mixed pixel spectra. Therefore, such a methodology would be a progression toward extracting more realistic endmember spectra.

According to the LMM, the spectrum of each pixel of HSI is modeled as a linear combination of all endmembers, so the HSI data matrix has a low-rank structure with a rank almost equal to the number of endmembers. To explore the low-rank pattern of HSIs, we can project the pixels onto low dimensional subspace as follows. Let the HSI matrix be $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_N] \in \mathbb{R}^{L \times N}$ with a span dimension r. Then the basis of the reduced dimension $k < r$ is a set of basis vectors $\{\phi_i\}_{i=1}^{k} \subset \mathbb{R}^L$, which provides the best approximation to the original span of $\mathbf{Y}$. The set of basis vectors $\{\phi\}_{i=1}^{k}$ is constructed to minimize error as follows;

$$\min_{\{\phi_i\}_{i=1}^{k}} \sum_{j=1}^{N} \left\| \mathbf{y}_j - \sum_{i=1}^{k} \left(\mathbf{y}_j^T \phi_i \right) \phi_i \right\|_2^2, \text{ with } \phi_i^T \phi_t = \begin{cases} 1 & \text{if } i = t, \quad i, t = 1, \ldots, k. \\ 0 & \text{if } i \neq t, \end{cases} \tag{8}$$

The set of the left singular vectors of $\mathbf{Y}$ gives the optimal solution to this minimization problem. Let the SVD of the HSI data be $\mathbf{Y} = \mathbf{W} \Sigma \mathbf{V}^T$, where $\mathbf{W} \in \mathbb{R}^{L \times r}$ and $\mathbf{V} \in \mathbb{R}^{N \times r}$ are left and right singular vectors, respectively, and $\Sigma \in \mathbb{R}^{r \times r}$, with order as $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$. So the optimal basis vectors for solving are $\{\mathbf{w}_i\}_{i=1}^{k}$. Although this is an optimal solution, it is limited to approximating the linear or low-order polynomial nonlinearity functions. It is computationally expensive and difficult to physically interpret the resultant singular vectors of a given HSI dataset.

We propose to utilize the CUR matrix decomposition to find the hidden structure of HSI datasets. Therefore, HSI can be represented using CUR matrix decomposition, such that for a given HSI matrix $\mathbf{Y} \in \mathbb{R}_+^{L \times N}$, as $\mathbf{Y} \approx \mathbf{CUR}$, where $\mathbf{C} \in \mathbb{R}^{L \times p}$ contains the most informative p columns that can represent the columns space of the HSI matrix $\mathbf{Y}$ (i.e., a subset of p pixels that can represent the whole N pixels of HSI), $\mathbf{R} \in \mathbb{R}^{p \times N}$ contains the most informative rows from $\mathbf{Y}$, and $\mathbf{U} \in \mathbb{R}^{p \times p}$ which makes the difference between data matrix $\mathbf{Y}$ and low-rank approximation matrix $\mathbf{CUR}$ as small as possible. Analog to

Equation ([6]) of NMF-based HU methods, the HU problem is formulated based on CUR decomposition as follows.

$$\arg\min_{\mathbf{C},\mathbf{U},\mathbf{R}} \|\mathbf{Y} - \mathbf{CUR}\|_F^2 \tag{9}$$

The $\mathbf{C}$ matrix contains the endmembers of HSI, and the $\mathbf{UR}$ matrix represents the abundance matrix whose columns represent the fractional compositions at each of the N pixels. To construct $\mathbf{C}$ and $\mathbf{R}$ matrices, randomized and deterministic sampling approaches can be used. Randomized approaches are based on sampling columns and rows based on uniform or non-uniform distributions [15,54]. Deterministic approaches consider the matrix approximation as an optimization problem and try to find the best decomposition that minimizes the approximation error [55]. They provide superior accuracy when reconstructing the matrix but have higher computational costs. In this paper, we utilized a recent deterministic selection method, the discrete empirical interpolation method (DEIM), because it gives high approximation accuracy, provides the most informative columns and rows approximating the whole data matrix, and approximates the nonlinear part by considering the continuous domain as a finite set of discrete points in the continuous domain [55,56]. For more details about using the DEIM method to approximate the nonlinear function, please refer to Appendix [A.2].

The DEIM method is completely deterministic and involves few parameters to provide the selected rows and columns indices of the HSI matrix $\mathbf{Y}$ based on the knowledge of the SVD. Suppose the truncated SVD $\mathbf{Y} \approx \mathbf{W}_r \boldsymbol{\Sigma}_r \mathbf{V}_r^T$ is computed, then the DEIM algorithm selects the best rows and columns indices $\mathbf{q}$ and $\mathbf{p}$ by processing the basis vectors $\mathbf{W}$ and $\mathbf{V}$, respectively, to construct the $\mathbf{R}$ and $\mathbf{C}$ matrices. Given an orthonormal matrix $\mathbf{W} \in \mathbb{R}^{L \times r}$ and a set of distinct indices $\mathbf{p} \in \mathbb{R}^r$, then the interpolatory projector for indices $\mathbf{p}$ onto the column span of $\mathbf{W}$ is

$$\mathcal{P} \equiv \mathbf{W}\left(\mathbf{P}^T\mathbf{W}\right)^{-1}\mathbf{P}^T \tag{10}$$

where $\mathbf{P} = \mathbf{I}_L(:, \mathbf{p}) \in \mathbb{R}^{L \times r}$, and the $\mathbf{P}^T\mathbf{W} \in \mathbb{R}^{r \times r}$ is invertible. This projector $\mathcal{P}$ has a critical characteristic not generally available by orthogonal projectors: for any vector $\mathbf{s} \in \mathbb{R}^L$,

$$(\mathcal{P}\mathbf{s})(\mathbf{p}) = \mathbf{P}^T\mathcal{P}\mathbf{s} = \mathbf{P}^T\mathbf{W}\left(\mathbf{P}^T\mathbf{W}\right)^{-1}\mathbf{P}^T\mathbf{s} = \mathbf{s}(\mathbf{p}) \tag{11}$$

The original vector $\mathbf{s}$ and its projected vector $\mathcal{P}\mathbf{s}$ match in the $\mathbf{p}$ elements, so this projector is called an interpolatory projector. DEIM processes each vector of the left singular vectors $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_r]$. Starting from the first singular vector $\mathbf{w}_1$, which corresponds to the largest singular value. The first index p_1 is selected as the index of the largest magnitude element in $\mathbf{w}_1$, as

$$|\mathbf{w}_1(p_1)| = \|\mathbf{w}_1\|_\infty \tag{12}$$

Now, $\mathbf{p}$ contains only the first selected index p_1 as $\mathbf{p} \equiv [p_1]$. The interpolatory projector of $\mathbf{p}_1$ onto $\mathbf{w}_1$ is given by

$$\mathcal{P}_1 \equiv \mathbf{w}_1\left(\mathbf{P}_1^T\mathbf{w}_1\right)^{-1}\mathbf{P}_1^T \tag{13}$$

The second index p_2 is selected as the index of the largest element in $\mathbf{w}_2$ but after removing its interpolatory projection in the $\mathbf{w}_1$ direction.

$$\mathbf{r}_2 = \mathbf{w}_2 - \mathcal{P}_1\mathbf{w}_2, \quad |\mathbf{r}_2(p_2)| = \|\mathbf{r}_2\|_\infty. \tag{14}$$

Because the $\mathcal{P}_1\mathbf{w}_2$ and $\mathbf{w}_2$ are matched in the p_1 position, the $\mathbf{r}_2(p_1) = 0$. So there are non-repeated indices. Generally, for selecting the index p_i, the $i-1$ indices are available and the interpolatory projection matrix is given by

$$\mathcal{P}_{i-1} \equiv \mathbf{W}_{i-1}\left(\mathbf{P}_{i-1}^T\mathbf{W}_{i-1}\right)^{-1}\mathbf{P}_{i-1}^T \tag{15}$$

where $\mathbf{W}_{i-1}$ is the first $i-1$ left singular vectors, $\mathbf{W}_{i-1} = \begin{bmatrix} \mathbf{w}_1 & \cdots & \mathbf{w}_{i-1} \end{bmatrix}$. Then, the index p_i is computed as

$$\mathbf{r}_i \equiv \mathbf{w}_i - \mathcal{P}_{i-1}\mathbf{w}_i, \quad |\mathbf{r}_i(p_i)| = \|\mathbf{r}_i\|_\infty \tag{16}$$

The process of selecting the column indices $\mathbf{q}$ is the same above steps to the right singular vectors $\mathbf{V}$. Algorithm 1 shows the procedures to compute the selected row indices $\mathbf{p}$ by the DEIM algorithm.

Algorithm 1: Discrete empirical interpolation method algorithm

> **Input** : $\mathbf{W} \in \mathbb{R}^{m \times r}$ left singular vectors matrix $(m \geq r)$
> **Output:** $\mathbf{p}$, an integer vector with r distinct entries in $\{1, \ldots, m\}$
> 1 　$\mathbf{w} = \mathbf{W}(:, 1)$;
> 2 　$[\sim, p_1] = \max(|\mathbf{w}|)$;
> 3 　$\mathbf{p} = [p_1]$;
> 4 　**for** $i = 2, 3, \ldots, r$ **do**
> 5 　　$\mathbf{w} = \mathbf{W}(:, i)$;
> 6 　　$\mathbf{c} = \mathbf{W}(\mathbf{p}, 1 : i - 1)^{-1}\mathbf{w}(\mathbf{p})$;
> 7 　　$\mathbf{r} = \mathbf{w} - \mathbf{W}(:, 1 : i - 1)\mathbf{c}$;
> 8 　　$[\sim, p_i] = \max(|\mathbf{r}|)$;
> 9 　　$\mathbf{p} = [\mathbf{p}; p_i]$
> 10 **end**

The interpolation indices $\mathbf{p}$ used for constructing the $\mathbf{R}$ matrix of CUR decomposition are selected inductively from the basis vectors of $\mathbf{U}$. After processing all singular vectors and obtaining the row indices vector $\mathbf{p}$, the $\mathbf{R}$ matrix is constructed as $\mathbf{R} = \mathbf{Y}(\mathbf{p}, :)$. Similarly, obtain $\mathbf{q}$ and construct the $\mathbf{C}$ matrix as $\mathbf{C} = \mathbf{Y}(:, \mathbf{q})$.

4.1. Approximating the Singular Vectors

The main limitation of the selection method is requiring the top r right and left singular vectors. For moderate-size HSI, software libraries can be used to compute SVD. Unfortunately, this is inefficient for large-scale HSI. Therefore, iterative approaches can be used to compute the leading r singular vectors of HSI [53,57,58]. We utilize the QR factorization based on an incremental approach (IQR) to approximate the computation of the singular vectors and estimate the number of endmembers in the HSI. The main advantages of IQR are that (1) it is only required to pass once through the HSI matrix $\mathbf{Y}$ for approximating the SVD, leading to efficient storage of massive data sets in memory; and (2) it satisfies a deterministic error bound in approximating the SVD of low-rank data. IQR approximates a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ as $\mathbf{A} \approx \mathbf{QR}$, where $\mathbf{Q} \in \mathbb{R}^{m \times k}$ that contains k orthonormal column vectors and $\mathbf{R} \in \mathbb{R}^{k \times n}$ is the upper triangular matrix.

Based on eigenanalysis algorithms of estimating the number of endmember signatures of HSI and considering the HSI can be well approximated using a low rank, we can utilize the QR factorization method to estimate the rank of the HSI data matrix, which represents the number of main components that approximate the HSI [5,17,59]. We use the QR factorization based on incremental methods for two tasks: (1) approximating the singular vectors required by the sampling algorithm and (2) estimating the number of endmember signatures of HSI data. By using the IQR, Algorithm A1 in Appendix A, we approximate the HSI matrix $\mathbf{Y}$ as $\mathbf{Y} \approx \mathbf{QR}$, where $\mathbf{Q} \in \mathbb{R}^{L \times p}$ that contains the orthonormal column vectors, $\mathbf{R} \in \mathbb{R}^{p \times N}$ is the upper triangular matrix, and p is the estimated number of endmembers. IQR processes one pixel of $\mathbf{Y}$ at each step to make it orthogonal with the previously orthogonalized pixels. After the orthogonality step, the algorithm verifies if any row of $\mathbf{R}$ has a tiny relative norm. If there is such a row, then this row of $\mathbf{R}$ and the corresponding column of $\mathbf{Q}$ can be deleted because this column of $\mathbf{Q}$ has a small contribution to the cumulatively computed Frobenius norm of $\mathbf{R}$.

At the start, we select the first k pixels from the HSI matrix $\mathbf{Y}$ which denoted as $\mathbf{Y}_k = [\mathbf{y}_1, \ldots, \mathbf{y}_k] \in \mathbb{R}^{L \times k}$, and let $k = 2$, as an initial step. Then, $\mathbf{Y}_k = \mathbf{Q}_k \mathbf{R}_k$, with $\mathbf{Q}_k \in \mathbb{R}^{L \times k}$ and $\mathbf{R}_k \in \mathbb{R}^{k \times k}$. In step 2 of the algorithm, the norm of rows in the $\mathbf{R}$ is computed. According to its value, we can determine which row should be deleted by incrementing the value of k to obtain the next pixel (column) of $\mathbf{Y}$. To make the pixel $\mathbf{y}_{k+1}$ orthogonal with previously k orthogonalized pixels, the classical Gram–Schmidt operations can be applied as

$$\mathbf{r} = \mathbf{Q}_k^T \mathbf{y}_{k+1}, \quad \mathbf{f} = \mathbf{y}_{k+1} - \mathbf{Q}_k \mathbf{r} \tag{17}$$

For a robust and stable re-orthogonalization step, the classical Gram–Schmidt is replaced by the following steps (7–8) of Algorithm A1, as proposed in [60]. This modification provides a $\mathbf{Q}$ matrix which is numerically orthogonal to working precision, and it can be implemented in an efficient parallel way as it utilizes the Gram–Schmidt process. see [61] for the full analysis.

The output of the above steps is the orthonormal vector $\mathbf{q}_{k+1}$ and the corresponding $\mathbf{r}_{k+1}$. Then, the $\mathbf{Q}$ and $\mathbf{R}$ matrices are updated, as formulated in the algorithm as

$$\mathbf{Q} = [\mathbf{Q}_k, \mathbf{q}_{k+1}]; \quad \mathbf{R} = \begin{bmatrix} \mathbf{R} & \mathbf{r}_{k+1} \\ \mathbf{0} & \rho \end{bmatrix} \tag{18}$$

Next is checking whether there is a truncation step by checking the row of the minimum relative norm in $\mathbf{R}$. Suppose this row has a relative norm less than the tolerance parameter (i.e., *tol*, a positive value controlling the accuracy of decomposition). In that case, the corresponding column in $\mathbf{Q}$ has little contribution to the factorization. The truncation step is required here by deleting that row from $\mathbf{R}$ and replacing it with the last row corresponding to the index $k + 1$ of $\mathbf{R}$ the matrix. Similar to how the corresponding column of $\mathbf{Q}$ will delete and put it in its position in the last column of $\mathbf{Q}$ matrix. If the minimum relative norm of rows in $\mathbf{R}$ does not achieve the truncation condition, then the last column in $\mathbf{Q}$, which is $\mathbf{q}_{k+1}$, is accepted as an orthonormal vector, and the process of the algorithm is continuing to obtain the next pixel from $\mathbf{Y}$. According to the linear model (LMM), each pixel in HSI data is a linear combination of endmember signatures in the HSI scene. So the number of column vectors in $\mathbf{Q}$ is the minimum number of the subspace dimensions that can approximate the HSI data, representing the number of endmembers in the HSI scene (i.e., $p = k$). In general, the estimated number of endmembers of HSI $\mathbf{Y}$ using incremental QR factorization within a tolerance error *tol* is $p = N - d_n$, where d_n is the number of times there is satisfaction of the truncation condition (number of deletions that occurred up to the last iteration) of Algorithm A1. The tolerance parameter *tol* plays a vital role in estimating the number of endmembers. A large tolerance value will lead to a smaller endmembers number than the actual number but will require less storage and computing time. Conversely, for a smaller *tol*, the estimated number of endmembers is equal to or larger than the actual number, and the complexity becomes significant. For more details about the complexity of approximating the singular vectors using the QR factorization, please refer to Appendix A.1.

4.2. CUR-HU

In this section, we present the proposed HU framework (i.e., CUR-HU) for performing the entire unmixing process (i.e., estimating the number of endmembers, estimating the endmembers, and extracting the corresponding abundance maps). As shown in Figure 2, CUR-HU consists of three main stages, preprocessing, estimating the number of endmembers, and CUR decomposition. Algorithm 2 shows the pseudocode of the CUR-HU framework.

In the preprocessing stage, the input HSI cube is reshaped to a matrix (i.e., HSI observed data matrix $\mathbf{Y}$), and then, the noise of HSI is estimated to obtain the estimated HSI matrix $\mathbf{X}$. Various methods for noise estimation have been proposed in the signal processing area and utilized in hyperspectral imagery, such as the nearest neighbor difference

(NND) [62] and multiple regression theory [63]-based methods. The most straightforward technique is the NND method, which assumes that the signal component of the adjacent pixels is almost equal. It is also known as the shift difference method because for a given two adjacent observed vectors $\mathbf{y_1} = \mathbf{x_1} + \mathbf{n_1}$ and $\mathbf{y_2} = \mathbf{x_2} + \mathbf{n_2}$ in a homogenous area of the HSI, the covariance matrix of noise can be estimated using the difference between $\mathbf{y_1}$ and $\mathbf{y_2}$, which yields $\mathbf{y_1} - \mathbf{y_2} = (\mathbf{x_1} + \mathbf{n_1}) - (\mathbf{x_2} + \mathbf{n_2}) \approx \mathbf{n_1} - \mathbf{n_2}$, where $\mathbf{x_1}, \mathbf{x_2}$ and $\mathbf{n_1}, \mathbf{n_2}$ are signal and noise vectors, respectively. However, noise samples are independent and almost have the same statistics. In this paper, we utilize the multiple regression theory-based method [5,63]. It outperforms the shift difference technique because of the high correlation of the neighboring spectral bands in HSIs. The second stage is performing the QR factorization based on an incremental approach, as presented in Algorithm A1 in Appendix A. This stage is responsible for two tasks: (1) approximating the singular vectors for the sampling algorithm and (2) estimating the number of endmembers as presented in Section 4.1, as shown in lines 4-7 of Algorithm 2. Suppose the number of endmembers is available by any other existing method. In that case, this stage only approximates the top *r* singular vectors of the HSI data matrix. Then, the CUR-HU can be used to estimate the endmembers and the corresponding abundance maps. Therefore, CUR-HU has much flexibility to cooperate with the existing algorithms at each stage.

In the third stage, the endmembers and their abundance maps are estimated using CUR decomposition. The inputs to this stage are the estimated number of endmembers and singular vectors of HSI data by the second stage and the HSI matrix $\mathbf{X}$. At the start, the sampling algorithm (DEIM) processes the left and right singular vectors to generate the rows and columns indices for constructing the $\mathbf{C}$ and $\mathbf{R}$ matrices of CUR decomposition, where $\mathbf{C}$ contains the most informative columns that can represent all pixels of HSI and $\mathbf{R}$ contains the most informative rows (i.e., selected bands) over the whole pixels, as shown in lines 8–11 of Algorithm 2. The $\mathbf{C}$ and $\mathbf{R}$ matrices are unique, as the deterministic sampling algorithm provides unique indices of selected rows and columns of CUR. After getting the $\mathbf{C}$ and $\mathbf{R}$ matrices, the $\mathbf{U}$ matrix is constructed to make the approximation error as small as possible [64].

$$\arg\min_{\mathbf{U}} \|\mathbf{Y} - \mathbf{CUR}\|_F^2 \tag{19}$$

Different methods for constructing the $\mathbf{U}$ matrix have been proposed. The simplest method is using the pseudoinverse of the intersection matrix between $\mathbf{C}$ and $\mathbf{R}$ matrices. Given $\mathbf{C} \in \mathbb{R}^{L \times p}$ and $\mathbf{R} \in \mathbb{R}^{p \times N}$, then $\mathbf{U}$ is computed as, $\mathbf{U} = \mathbf{Y}^{\dagger}(I, J)$. Although this is the simplest solution of $\mathbf{U}$, it gives a high approximation error in approximating high-rank or noisy data. The close form for the best middle matrix of CUR decomposition regarding the Frobenius norm error is given by solving the optimization problem in (19), so the optimal $\mathbf{U}$ matrix is given by

$$\mathbf{U} = \mathbf{C}^{\dagger}\mathbf{Y}\mathbf{R}^{\dagger} \tag{20}$$

So, $\mathbf{Y} \approx \mathbf{CUR} = \mathbf{CC}^{\dagger}\mathbf{Y}\mathbf{R}^{\dagger}\mathbf{R}$, where $\mathbf{CC}^{\dagger}$ and $\mathbf{R}^{\dagger}\mathbf{R}$ are orthogonal projections onto the span of the columns of $\mathbf{C}$ and rows of $\mathbf{R}$, respectively. Therefore, using $\mathbf{U} = \mathbf{C}^{\dagger}\mathbf{Y}\mathbf{R}^{\dagger}$ gives the best CUR approximations for given $\mathbf{C}$ and $\mathbf{R}$ matrices. In other words, the $\mathbf{U}$ matrix is formed by weighting the whole elements of $\mathbf{Y}$, which leads to a small approximation error [65].

Analog to the NMF-based methods for HU, the matrix $\mathbf{C}$ represents the source matrix of NMF or $\mathbf{M}$ matrix of HU LMM, while the $\mathbf{UR}$ matrix represents the mixing matrix of NMF or $\mathbf{S}$ matrix in Equation (6). Finally, the nonnegative abundance (ANC) and the abundance sum-to-one (ASC) constraints are imposed on the $\mathbf{UR}$ matrix. For ANC, we set any negative value of $\mathbf{UR}$ to zero, while for ASC, we normalized the output from ANC constraint to $\mathbf{UR}$.

Algorithm 2: The pseudocode of CUR-HU

Input : HSI cube $\mathcal{Y} \in \mathbb{R}^{m \times n \times L}$, *tol* a positive number for controlling the factorization accuracy.

Output: Number of endmembers p, Endmember matrix $\mathbf{M} \in \mathbb{R}^{L \times p}$, and Abundance maps $\mathbf{S} \in \mathbb{R}^{p \times N}$

Preprocessing:

1 Reshape $\mathcal{Y}$ into $\mathbf{Y} \in \mathbb{R}^{L \times N}$;   `// Where` $N = m \times n$.

2 Estimate the noise matrix $\mathbf{N} \in \mathbb{R}^{L \times N}$.

3 Compute the estimated HSI, $\mathbf{X} = \mathbf{Y} - \mathbf{N}$.

Second Stage:

4 $[\mathbf{Q}, \mathbf{R}, p] \longleftarrow IQR(\mathbf{X}, tol)$.   `// Where` $\mathbf{Q} \in \mathbb{R}^{L \times k}$.

/* Compute the singular vectors   */

5 $[\hat{\mathbf{W}}, \boldsymbol{\Sigma}, \mathbf{V}] \longleftarrow SVD(\mathbf{R})$;   `// Economy-sized SVD.`

6 Set $\mathbf{W} \longleftarrow \mathbf{Q}\hat{\mathbf{W}}$;

Thrid Stage:

/* Get Row and column indices   */

7 $I \longleftarrow DEIM(U)$; Using Algorithm 1.

8 $J \longleftarrow DEIM(V)$; Using Algorithm 1.

/* Construct C, U, and R matrices   */

9 $\mathbf{C} \longleftarrow \mathbf{X}(:, J)$;   `// Where` $\mathbf{C} \in \mathbb{R}^{L \times p}$.

10 $\mathbf{R} \longleftarrow \mathbf{X}(I, :)$;   `// Where` $\mathbf{R} \in \mathbb{R}^{p \times N}$.

11 Set $\mathbf{U} = \mathbf{C}^{\dagger}\mathbf{Y}\mathbf{R}^{\dagger}$;   `// Where` $\mathbf{U} \in \mathbb{R}^{p \times p}$.

/* Construct M and S matrices   */

12 Set $\mathbf{M} \longleftarrow \mathbf{C}$;   and $\mathbf{S} \longleftarrow \mathbf{U}\mathbf{R}$;

13 Apply the abundance nonnegativity constraint to $\mathbf{S}$;

14 Apply the abundance sum to one constraint to $\mathbf{S}$;

---

5. Experiments

For a comprehensive and quantitative comparison between the different algorithms, the proposed method is evaluated on synthetic and real HSIs datasets. To evaluate the performance of each stage of the CUR-HU framework, different datasets with different experimental settings were conducted. All experiments were implemented with MATLAB R2022a on a desktop computer with an Intel(R) Core(TM) i7-10700 processor @ 2.90 GHz.

5.1. Estimating the Endmember Number

CUR-HU is evaluated for the first task HU process, which is estimating the number of endmembers in the HSI data. In order to evaluate the QR factorization based on an incremental approach for estimating the number of endmember signatures, we compare the performance with state-of-the-art methods, including HySime [18], HFC, NWHFC [5], SPICE [66], FUN [59], and GENE-AH [17] for synthetic and real HSI datasets.

5.1.1. Synthetic HSI Dataset

For HSI synthetic data, we follow the setup of experiments of the HySime method. The endmember signatures are selected randomly from the USGS spectral library (https://www.usgs.gov/labs/spectroscopy-lab/science/spectral-library, accessed on 28 December 2023), and the abundance maps are created according to a Dirichlet distribution [18]. To evaluate our proposed algorithm in a noisy environment, additive Gaussian noise is added to the generated HSI datasets. Let the diagonal elements of the noise correlation matrix follow a Gaussian shape centered at $L/2$, i.e.,

$$\sigma_i^2 = \sigma^2 \frac{e^{-\frac{(i-L/2)^2}{(2\eta^2)}}}{\sum_{j=1}^{L} e^{-\frac{(j-L/2)^2}{(2\eta^2)}}} \tag{21}$$

for $i = 1, \ldots, L$. Here, the parameter η controls the variance of the Gaussian distribution when $\eta \to \infty$ corresponds to white noise, $\eta \to 0$ corresponds to one-band noise, and the parameter σ_i^2 controls the total power of noise. The amount of additive noise to HSI is set by the signal-to-noise ratio (SNR), which is defined as

$$\text{SNR} = 10 \log_{10} \frac{E[\mathbf{x}^T \mathbf{x}]}{E[\varepsilon^T \varepsilon]} \tag{22}$$

where $\mathbf{x}$ and ε are the signal and noise, respectively. We evaluated the proposed method under changing $\text{SNR} \in \{50, 35, 25, 15\}$ dB and the number of endmembers $p \in \{3, 5, 10, 15\}$. The NWHFC is an extension method of the HFC by adding noise whitening as a preprocessing step. We report the results of both estimated and actual noise correlation matrices with the NWHFC method. The result based on the noise correlation matrix is shown in brackets. The false-alarm probability p_f parameter is set as $p_f \in \{10^{-3}, 10^{-4}, 10^{-5}\}$ for both HFC and NWHFC methods. We report the tolerance parameter (*tol*) that gives the best estimation of the number of the endmembers.

Table 1 shows the estimated dimension of the signal subspace by different methods as a function of endmembers number p and SNR for one-band noise $\eta = 0$. Although the HySime method outperforms the other methods, our proposed method has a comparable performance of estimating the number of endmember signatures with the HySime method, especially for high SNR cases. For lower subspace dimensions, such as $p = 3$ and $p = 5$, the proposed algorithm can estimate the exact number of endmembers under different types of noise and different SNR values. It outperforms NWHFC, SPICE, HFC, GENE-AH, and FUN. The proposed methods perform well for low SNR values (i.e., 25 dB and 15 dB) when setting the tolerance parameter *tol* to large values (e.g., 1×10^{-2} or 1×10^{-3}). Overall, the QR factorization based on an incremental approach can estimate the number of endmembers with a comparable performance to the HySime method.

Table A1 shows the estimated dimension of the signal subspace by different algorithms as a function of endmember number p and SNR for $\eta = 1/18$. For the noisy case ($\eta = 1/18$), the HySime method outperforms all other methods. Our proposed method outperforms the NWHFC, SPICE, HFC, FUN, and GENE-AH methods. Furthermore, it has a performance comparable to the HySime method for estimating the number of endmembers of a given HSI. From Tables 1 and A1, we can observe the effect of the *tol* parameter of IQR factorization. For a large *tol* value, the estimated number of endmember signatures is smaller than the actual number, especially for low SNR cases, while setting *tol* to a very small value, the estimated number of endmembers exceeds the actual number.

5.1.2. Real HSI Dataset

The proposed algorithm is evaluated on two popular real HSI datasets, the Indian Pines site and Cuprite datasets. The Indian Pines site was collected by AVIRIS [67]. It consists of 145×145 pixels with 185 spectral bands. The dataset contains 16 pure materials [68]. The Cuprite dataset originally contained 224 bands. A subset of the dataset was formed after removing the low SNR bands. This subset consists of 188 spectral bands and contains 250×190 pixels. The ground truth of the number of endmembers of the Cuprite dataset is not known. The previous studies frequently fixed the number of endmembers between 9 and 14 [38,69]. The tolerance parameter of the proposed method is set to 1×10^{-3}. Table 2 shows the estimated endmembers number on the real HSI datasets by different methods. The proposed method is comparable to the HySime method, but it outperforms the others for estimating the closest number of endmembers to actual numbers.

Table 1. Estimation of the number of endmembers of synthetic HSI under changing SNR and *p* for $(\eta = 0)$.

SNR	Methods	No. of Endmembers p			
		3	**5**	**10**	**15**
50 dB	**HySime**	3	5	10	15
	NWHFC ($P_f = 1 \times 10^{-3}$)	3(3)	5(5)	7(7)	10(11)
	NWHFC ($P_f = 1 \times 10^{-4}$)	3(3)	5(5)	7(7)	8(8)
	NWHFC ($P_f = 1 \times 10^{-5}$)	3(3)	4(4)	7(6)	8(8)
	HFC ($P_f = 1 \times 10^{-3}$)	3	5	7	11
	HFC ($P_f = 1 \times 10^{-4}$)	3	5	7	8
	HFC ($P_f = 1 \times 10^{-5}$)	3	4	6	8
	GENE-AH ($P_f = 1 \times 10^{-3}$)	6	6	12	15
	GENE-AH ($P_f = 1 \times 10^{-4}$)	4	6	10	15
	GENE-AH ($P_f = 1 \times 10^{-5}$)	4	5	10	15
	SPICE	3	6	10	15
	FUN ($P_f = 1$)	3	5	10	15
	Ours ($tol = 2 \times 10^{-3}$)	3	5	10	15
35 dB	**HySime**	3	5	10	14
	NWHFC ($P_f = 1 \times 10^{-3}$)	3(3)	4(4)	7(7)	9(9)
	NWHFC ($P_f = 1 \times 10^{-4}$)	3(3)	4(4)	7(6)	8(8)
	NWHFC ($P_f = 1 \times 10^{-5}$)	3(3)	4(4)	6(6)	8(8)
	HFC ($P_f = 1 \times 10^{-3}$)	3	4	7	9
	HFC ($P_f = 1 \times 10^{-4}$)	3	4	6	8
	HFC ($P_f = 1 \times 10^{-5}$)	3	4	6	8
	GENE-AH ($P_f = 1 \times 10^{-3}$)	6	8	11	15
	GENE-AH ($P_f = 1 \times 10^{-4}$)	4	7	10	15
	GENE-AH ($P_f = 1 \times 10^{-5}$)	4	5	10	14
	SPICE	5	7	11	13
	FUN ($P_f = 1$)	3	5	9	13
	Ours ($tol = 1 \times 10^{-3}$)	3	5	9	13
25 dB	**HySime**	3	5	9	14
	NWHFC ($P_f = 1 \times 10^{-3}$)	3(3)	4(5)	6(6)	9(8)
	NWHFC ($P_f = 1 \times 10^{-4}$)	3(3)	5(5)	6(6)	7(7)
	NWHFC ($P_f = 1 \times 10^{-5}$)	3(3)	4(4)	5(5)	7(7)
	HFC ($P_f = 1 \times 10^{-3}$)	3	5	6	8
	HFC ($P_f = 1 \times 10^{-4}$)	3	5	6	8
	HFC ($P_f = 1 \times 10^{-5}$)	3	4	3	2
	GENE-AH ($P_f = 1 \times 10^{-3}$)	5	7	11	12
	GENE-AH ($P_f = 1 \times 10^{-4}$)	4	5	12	13
	GENE-AH ($P_f = 1 \times 10^{-5}$)	4	5	10	13
	SPICE	8	7	15	12
	FUN ($P_f = 1$)	3	5	9	12
	Ours ($tol = 5 \times 10^{-3}$)	4	6	11	12
15 dB	**HySime**	3	5	8	13
	NWHFC ($P_f = 1 \times 10^{-3}$)	4(3)	5(5)	11(10)	9(11)
	NWHFC ($P_f = 1 \times 10^{-4}$)	3(3)	4(4)	3(3)	3(2)
	NWHFC ($P_f = 1 \times 10^{-5}$)	3(3)	4(4)	3(3)	2(2)
	HFC ($P_f = 1 \times 10^{-3}$)	3	5	4	5
	HFC ($P_f = 1 \times 10^{-4}$)	3	4	3	2
	GENE-AH ($P_f = 1 \times 10^{-3}$)	8	10	8	10
	GENE-AH ($P_f = 1 \times 10^{-4}$)	7	7	10	9
	SPICE	8	37	40	40
	FUN ($P_f = 1$)	38	25	19	23
	Ours ($tol = 1 \times 10^{-2}$)	3	6	7	11

Table 2. Estimation of the number of endmembers on real HSI datasets.

Methods	P_f/tol	Dataset	
		Indian Pines	Cuprite
HySime	-	15	16
HFC	10^{-3}	24	20
	10^{-4}	22	17
	10^{-5}	20	16
NWHFC	10^{-3}	18	11
	10^{-4}	19	10
	10^{-5}	18	9
FUN	1	11	19
SPICE	-	20	8
Ours	5×10^{-3}	9	7
	1×10^{-3}	17	12
	0.5×10^{-3}	21	14

5.2. Estimating the Endmembers and Abundance Maps

In this section, CUR-HU is evaluated for estimating the endmembers and their abundance by comparing its performance with state-of-the-art NMF-based HU algorithms for both synthetic and real HSI datasets. Two widely utilized evaluation metrics for HU algorithms are used to evaluate the performance of the CUR-HU, i.e., Root-Mean-Square Error (RMSE) and Spectral Angle Distance (SAD), which are computed as

$$RMSE_i = \sqrt{\frac{1}{N} \sum_{j=1}^{N} \left(\mathbf{S}_{ij} - \widehat{\mathbf{S}}_{ij} \right)^2} \tag{23}$$

$$SAD_i = \cos^{-1} \left(\frac{\widehat{\mathbf{m}}_i^T \mathbf{m}_i}{\|\widehat{\mathbf{m}}_i\|_2 \|\mathbf{m}_i\|_2} \right). \tag{24}$$

where $RMSE_i$ measures the error between the ith ground truth abundance map $\mathbf{S}_i$ and the corresponding extracted abundance map $\widehat{\mathbf{S}}_i$, and SAD_i measures the spectral angle between the ith extracted endmember $\widehat{\mathbf{m}}_i$ and the corresponding ground truth $\mathbf{m}_i$. Furthermore, the average values of RMSE and SAD are reported. For these experiments on synthetic and real HSI datasets, we follow a procedure similar to [70,71]. We evaluated the performance of CUR-HU by comparison with the state-of-the-art blind NMF-based HU baselines, $l_{1/2}$-NMF [72], R-CoNMF [39], Min-vol NMF [35], KbSNMF-div [70], SS-RNMF [41], MVNTF [73], and SGSNMF [38].

5.2.1. Synthetic HSI Datasets

The synthetic HSI dataset contains three endmembers, seawater, clintonite, and sodium bicarbonate, with 480 spectral bands, which were selected from the USGS spectral library. The synthetic image was generated using the hyperspectral imagery synthesis toolbox (HSIST) (https://www.ehu.eus/ccwintco/index.php/Hyperspectral_Imagery_Synthesis_tools_for_MATLAB, accessed on 28 December 2023). HSIST was used to generate the abundance maps with a size of 64×64 pixels according to the Spherical Gaussian distribution. Figure 4 shows the ground truth of the normalized endmembers and their abundance maps.
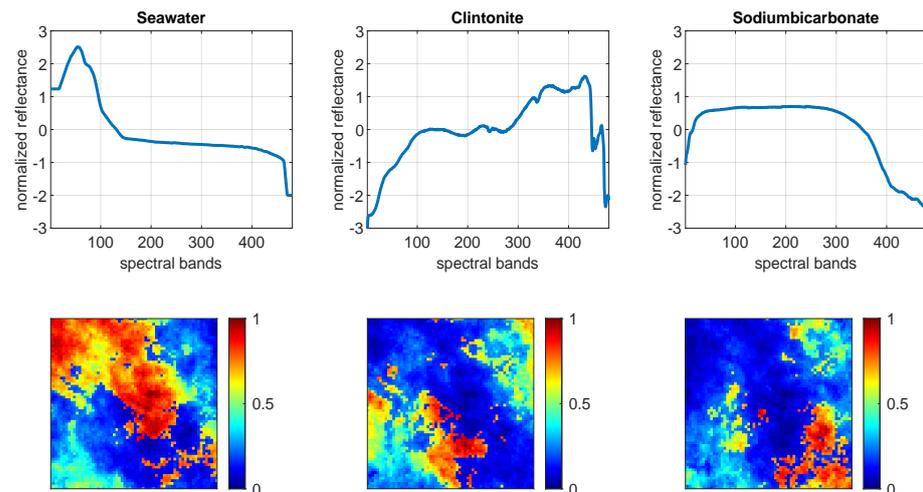
Figure 4. Ground truth of the normalized endmembers and their abundance maps of the synthetic HSI dataset.

Table 3 shows the performance of CUR-HU in terms of SAD and RMSE for a synthetic HSI dataset. In order to see the effect of ANC and ASC constraints on the abundance maps from the proposed method, we report the RMSE for considering **UR** (i.e., **S**) with ANC and **UR** with both ASC and ANC constraints as the abundance maps (i.e., **S**), which is the final output from Algorithm 2.

Table 3. The unmixing performance comparison in terms of SAD and RMSE for a simulated dataset. The best performances are in bold typeface; the second-best performances are underlined; and the third-best performances are italicized.

| Methods	Materials						Average	
	Seawater		Clintonite		Sodium Bicarbonate			
	SAD	RMSE	SAD	RMSE	SAD	RMSE	SAD	RMSE
MVNTF	0.3905	0.4088	*0.1326*	0.1506	0.0567	0.1324	0.1933	0.2306
$l_{1/2}$-NMF	0.6855	0.2488	0.2990	<u>0.0573</u>	0.0168	0.2394	0.3337	0.1818
Min-vol NMF	0.8891	*0.1052*	0.2803	0.0854	0.0188	<u>0.0745</u>	0.3961	*0.0884*
R-CoNMF	1.7408	0.1753	0.7109	0.1473	0.0614	0.1475	0.8377	0.1569
SSRNMF	*0.3017*	<u>0.0817</u>	0.1410	0.0774	*0.0160*	0.0768	<u>0.1529</u>	<u>0.0786</u>
KbSNMF-div	<u>0.2746</u>	0.3190	<u>0.1119</u>	0.1200	0.1311	0.2985	*0.1725*	0.2458
SGSNMF	0.5416	0.3579	0.2288	0.2906	<u>0.0090</u>	0.1123	0.2598	0.2538
Ours-ANC	**0**	**4.33** $\times 10^{-15}$	**0**	**9.15** $\times 10^{-17}$	**2.10** $\times 10^{-8}$	**6.98** $\times 10^{-17}$	**7.02** $\times 10^{-9}$	**1.50** $\times 10^{-15}$
Ours-ANC-ASC	**0**	**3.48** $\times 10^{-15}$	**0**	**1.55** $\times 10^{-15}$	**2.10** $\times 10^{-8}$	**2.41** $\times 10^{-15}$	**7.02** $\times 10^{-9}$	**2.48** $\times 10^{-15}$

As can be seen, using the ANC constraint with abundance maps has no effect because the CUR decomposition keeps the nonnegative property of the HSI data. The abundance matrices of **UR** with ANC and **UR** with both ASC and ANC constraints are donated as **UR1** and **UR2**, respectively. Figure 5 shows the estimated abundance maps by the HU-CUR algorithm. The proposed method can extract the exact abundance maps of the simulated dataset. The reason for the high performance of the proposed methods is that the synthetic HSI data consists of pure endmember signatures, which are distinguished from each other. Therefore, the CUR-HU can extract the exact endmembers as the selection method procedures select the most distinguished pixels in the HSI data. Considering the **R** matrix as the abundance maps without weighting it by the **U** matrix leads to a high estimation error, as shown in the second row of Figure 5.

In order to see the effect of selecting more rows and columns of CUR of the HSI dataset, Figure 6 shows the ground truth and the estimated endmembers by CUR-HU of the simulated dataset. The first six endmembers extracted by CUR-HU capture the three endmembers, e.g., the extracted endmembers *C3*, *C2*, and *C1* are almost the same as the

ground truth endmembers *GT*1, *GT*2 and *GT*3, respectively. If the number of endmembers selected by the proposed method increases, highly similar endmembers will exist (e.g., *C*5 and *C*6). Selecting only three columns from the HSI data using CUR-HU is enough to capture exactly the three endmembers of the HSI simulated data.
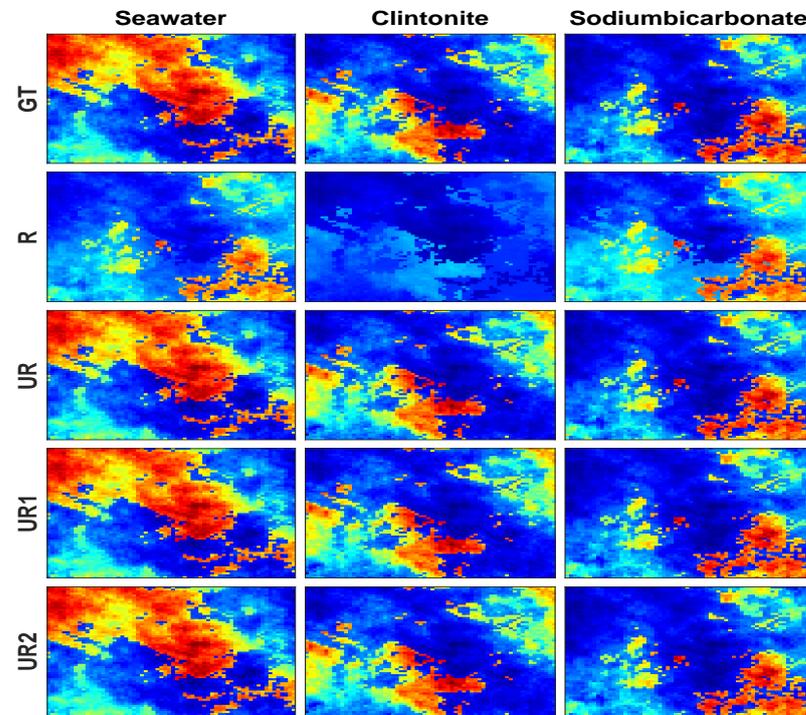
Figure 5. Ground truth and extracted abundance maps by CUR-HU of simulated HSI: 1st row, ground truth (GT); the 2nd, 3rd, 4th, and 5th rows contain the extracted abundance maps by **R**, **UR**, **UR1**, and **UR2**, respectively.
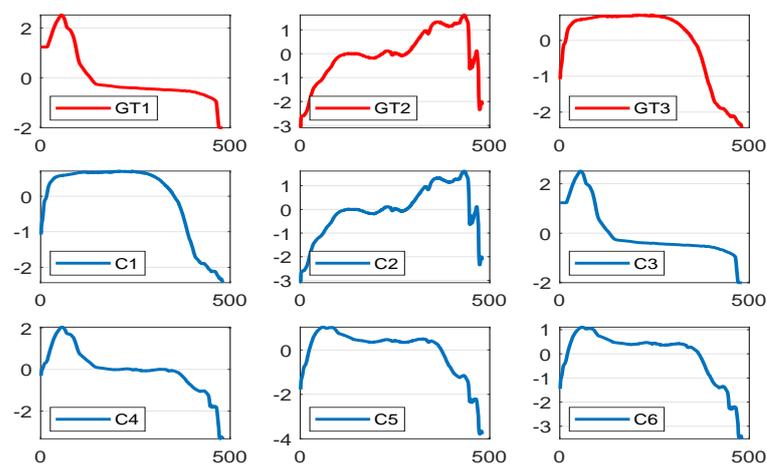
Figure 6. Normalized ground truth and extracted endmembers by CUR-HU of simulated HSI: 1st row, ground truth endmember signatures; the 2nd and 3rd rows contain the first 6 endmembers extracted by CUR-HU.

As can be seen, the proposed method can extract the exact abundance maps and endmember signatures. CUR-HU outperforms other methods for extracting the endmembers and their corresponding abundances of the synthetic HSI dataset.

5.2.2. Real HSI Datasets

In order to evaluate the performance of the CUR-HU method in the real scenario, a set of real HSIs, including the Samson, Cuprite, and Urban datasets from [14], were used.

- Samson HSI dataset: It consists of 95×95 pixels with 156 spectral bands and has three reference endmembers. The ground truth of normalized endmembers and their abundance maps of Samson HSI are shown in Figure 1b.
- Urban HSI dataset: After pre-processing, it consists of 307×307 pixels with 162 spectral bands ranging from 400 to 2500 nm and has a 10 nm spectral resolution. It has different ground truth versions. We use the one that has five endmembers.
- The Cuprite dataset: After pre-processing, it consists of 188 spectral bands covering wavelengths ranging from 370 to 2480 nm with 250×190 pixels, and 12 endmembers are selected as a reference.

Figure 7 shows the extracted endmembers using CUR-HU for the Samson HSI dataset. Extracted endmembers *C3*, *C1* or *C8*, and *C2* or *C6* are almost similar to the ground truth of endmember signatures of water, tree, and soil, respectively.
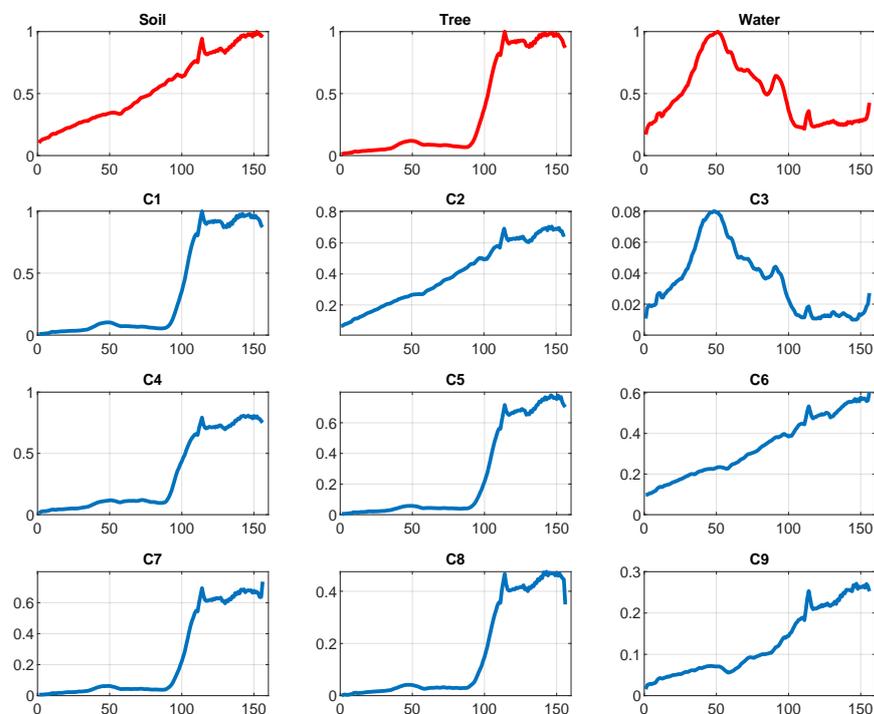
Figure 7. Normalized ground truth and extracted endmembers by CUR-HU of Samson HSI: 1st row, ground truth endmember signatures; the 2nd, 3rd, and 4th rows contain the first 9 endmembers extracted by CUR-HU.

Figure 8 shows the estimated abundance map of **R**, **UR**, **UR1**, and **UR2** matrices for the Samson dataset. As can be seen, the estimated abundance maps after applying both ANC and ASC constraints (i.e., **UR2**, final **S** from Algorithm 2) are very close to their ground truth.

Table 4 reports the quantitative comparisons regarding SAD and RMSE metrics for the Samson dataset. As can be seen, CUR-HU outperforms the method in terms of SAD for estimating the endmembers. It has the best SAD value for estimating the first endmember (i.e., soil) and the second-best value for estimating the tree and water endmembers. However, it has the best overall average performance. CUR-HU outperforms the MVNTF, $l_{1/2}$-NMF, R-CoNMF, and SSRNMF methods for extracting the abundance maps in terms of the RMSE metric.
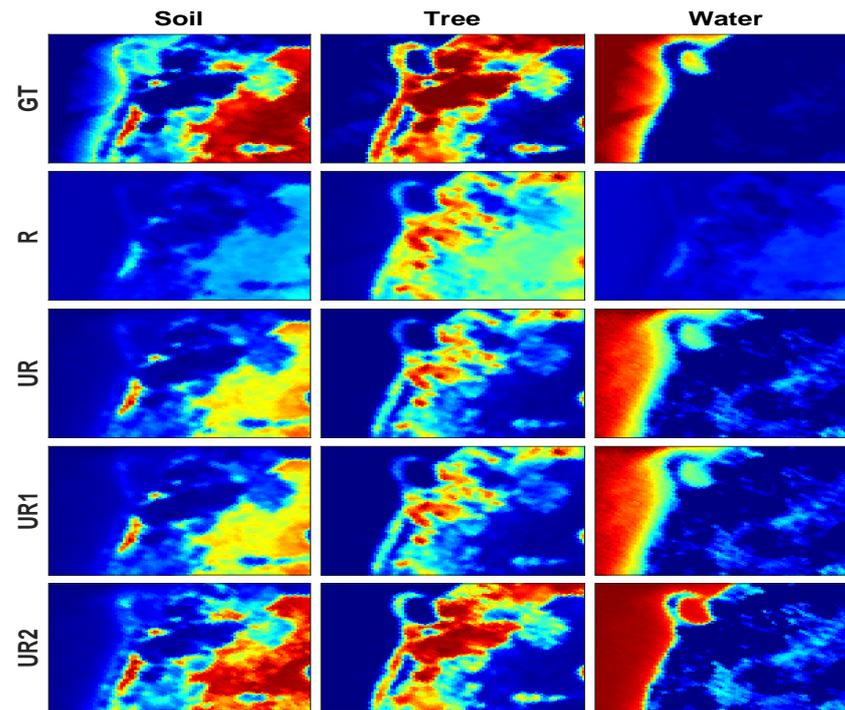
Figure 8. Ground truth and extracted abundance maps by CUR-HU of Samson HSI: 1st row, ground truth (GT); the 2nd, 3rd, 4th, and 5th rows contain the extracted abundance maps by **R**, **UR**, **UR1**, and **UR2** matrices, respectively.
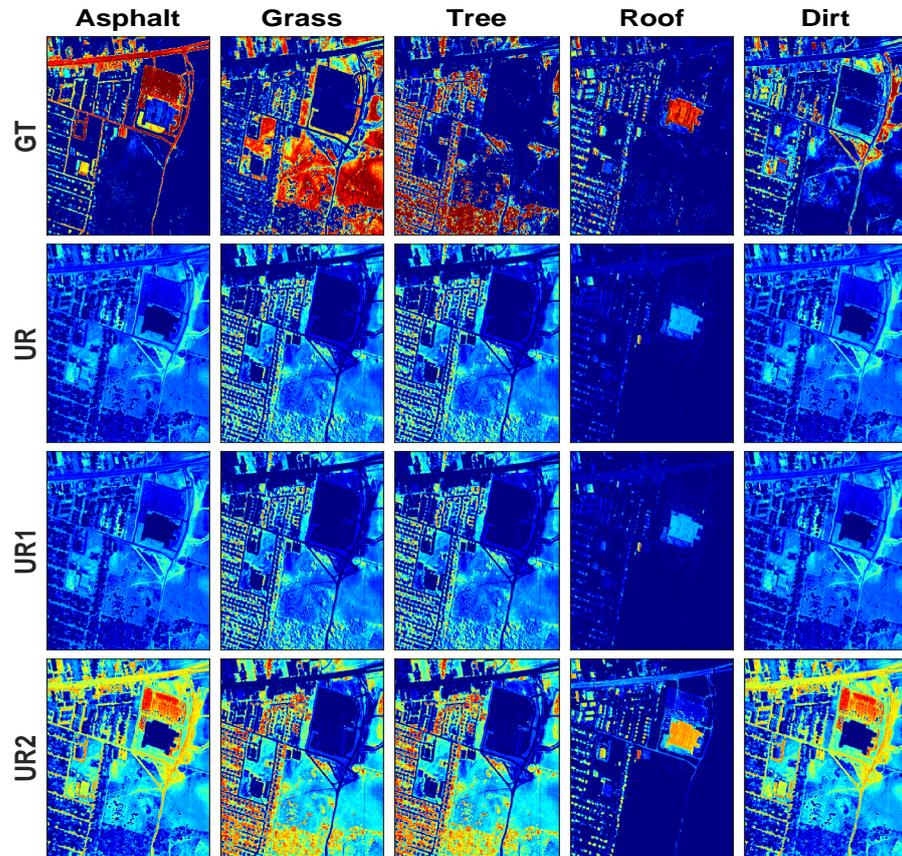
Table 4. The unmixing performance comparison in terms of SAD and RMSE for a Samson HSI dataset. The best performances are in bold typeface; the second-best performances are underlined; and the third-best performances are italicized.

Methods	Materials						Average	
	Soil		Tree		Water			
	SAD	*RMSE*	*SAD*	*RMSE*	*SAD*	*RMSE*	*SAD*	*RMSE*
MVNTF	*0.1488*	0.3517	0.0944	0.2454	**0.0887**	0.4162	0.1106	0.3378
$l_{1/2}$-NMF	0.3455	0.4217	0.1433	*0.0432*	0.3513	0.2359	0.2800	0.2336
Min-vol NMF	0.3463	*0.0967*	0.2315	0.1245	0.2429	**0.0432**	0.2736	**0.0881**
R-CoNMF	0.1253	**0.0431**	**0.0105**	**0.0118**	0.3219	0.5321	*0.1526*	0.1957
SSRNMF	0.6091	0.3832	0.0750	0.0325	*0.1624*	0.1654	0.2822	0.1937
KbSNMF-div	0.2078	0.1574	*0.0647*	0.0911	0.2014	*0.0967*	0.1580	*0.1151*
SGSNMF	0.3743	0.0532	0.1721	0.0882	0.2941	0.1432	0.2802	0.0949
Ours-UR	**0.0404**	0.2189	0.0219	0.2765	0.1189	0.1267	**0.0604**	0.2073
Ours-ANC	**0.0404**	0.2189	0.0219	0.2764	0.1189	0.0950	**0.0604**	0.1968
Ours-ANC-ASC	**0.0404**	0.1217	0.0219	0.1040	0.1189	0.1676	**0.0604**	0.1311

As mentioned above, the Urban HSI version that we used has five endmembers as the ground truth (i.e., Asphalt, Grass, Tree, Roof, and Dirt). Figure A1 shows the extracted endmembers using CUR-HU for the Urban dataset. As can be seen, CUR-HU can estimate accurate endmembers with almost the same shape and close to the ground truth. Figure 9 shows the estimated abundance map of R, UR, UR1, and UR2 matrices for the Urban dataset.

Table 5 reports the quantitative comparisons regarding SAD and RMSE metrics for the Urban dataset. We report the RMSE value of the estimated abundance maps (i.e., UR2), which is the final output after applying both ANC and ASC constraints. CUR-HU outperforms all other methods for extracting the endmembers of the Uraban HSI dataset, and for estimating the abundance maps, it has the best average RMSE value compared to other methods. Although the quantitative results of the RMSE values do not necessarily reflect a high estimation accuracy of abundance maps, the proposed framework can be

utilized with different HU algorithms as an initialization step to provide them the initial values of the number of endmembers, endmember signatures, and abundance maps.
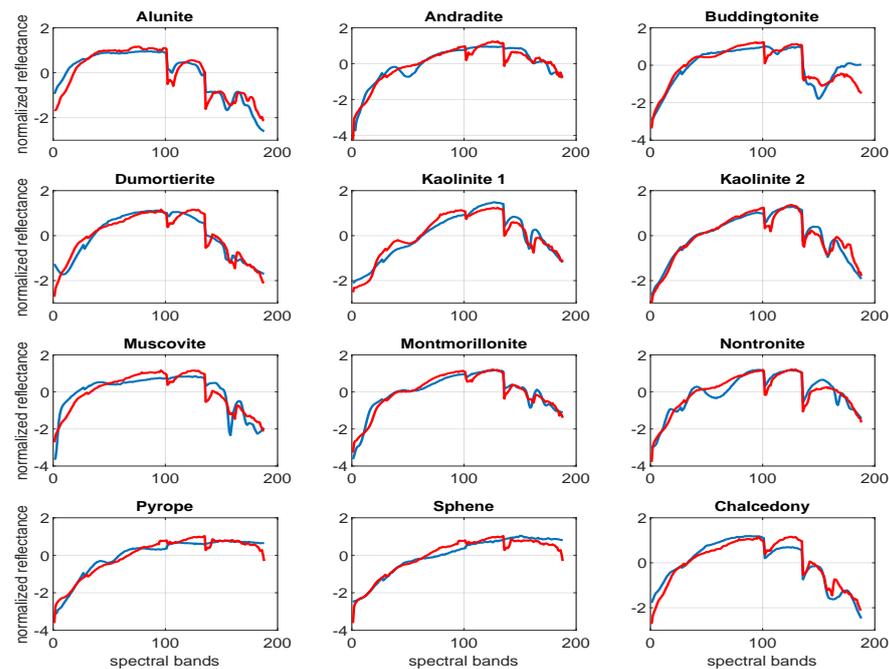

Figure 9. Ground truth and extracted abundance maps by CUR-HU of the Urban HSI: 1st row, ground truth (GT); the 2nd, 3rd, 4th, and 5th rows contain the extracted abundance maps by **R**, **UR**, **UR1**, and **UR2**, respectively.

Table 5. The unmixing performance comparison in terms of SAD and RMSE for a real data set, "Urban HSI". The best performances are in bold typeface; the second-best performances are underlined; and the third-best performances are italicized.

Methods	Materials										Average	
	Asphalt		Grass		Tree		Roof		Dirt			
	SAD	*RMSE*	*SAD*	*RMSE*	*SAD*	*RMSE*	*SAD*	*RMSE*	*SAD*	*RMSE*	*SAD*	*RMSE*
MVNTF	0.1721	0.3354	0.2080	0.3164	0.2310	0.3183	0.3941	<u>0.1858</u>	0.3689	**0.2004**	0.2748	0.2713
$l_{1/2}$-NMF	0.2966	0.3197	0.4993	*0.2410*	0.1603	0.4607	0.2518	0.5350	0.3379	0.4293	0.3092	0.3917
Min-vol NMF	0.1849	**0.1376**	**0.1045**	**0.1490**	0.1798	**0.1064**	*0.1930*	0.2130	<u>0.1521</u>	0.5409	<u>0.1628</u>	<u>0.2294</u>
R-CoNMF	0.2017	*0.1980*	0.2786	0.3010	0.2125	<u>0.1258</u>	0.2478	*0.1986*	0.2435	0.5210	0.2368	*0.2689*
SSRNMF	**0.0599**	0.3237	*0.1780*	0.2437	0.1795	0.2660	0.4217	0.2851	*0.1534*	*0.3589*	0.1985	0.2955
KbSNMF-div	*0.1252*	<u>0.1654</u>	0.2821	<u>0.2354</u>	*0.1498*	0.4240	<u>0.1621</u>	0.4584	0.1742	0.5789	*0.1787*	0.3724
SGSNMF	0.4173	0.5578	0.3434	0.2545	*0.1499*	0.2692	0.3822	0.2256	0.3359	0.5610	0.3257	0.3736
Ours	<u>0.1197</u>	0.2886	<u>0.1485</u>	0.4110	**0.0736**	*0.2073*	**0.0981**	**0.1276**	**0.0611**	<u>0.2794</u>	**0.1003**	**0.2628**

We only report the unmixing performance regarding the SAD values for the Cuprite dataset because the ground truth of Cuprite HSI contains only endmember signatures [14]. Figure 10 shows the extracted endmembers using CUR-HU for the Cuprite dataset. As can be seen, CUR-HU almost estimates the same shape of the ground truth endmembers.

Figure 10. Ground truth endmember spectra and extracted endmember spectra of the "Cuprite dataset": the blue line is the ground truth, and the red line represents the estimated endmembers by CUR-HU.

To quantify the estimated endmembers by CUR-HU with the other blind NMF-based HU methods, Table 6 reports the SAD values for each estimated endmember and the average SAD value over the endmembers. The EM_1 to EM_{12} represent the Alunite, Andradite, Buddingtonite, Dumortierite, Kaolinite 1, Kaolinite 2, Muscovite, Montmorillonite, Nontronite, Pyrope, Sphene, and Chalcedony endmember signatures of the Cuprite HSI dataset, respectively. As can be seen, the proposed method outperforms the MVNTF, $l_{1/2}$-NMF, R-CoNMF, and SGSNMF methods for estimating the endmembers of the Cuprite dataset in terms of SAD value. It has comparable performance to KbSNMF-div and Min-vol NMF methods. Overall, CUR-HU can estimate the endmember and abundance map matrices with comparable accuracy to the state-of-the-art HU method based on the NMF.

Table 6. The unmixing performance comparison in terms of SAD for a real data set, "Cuprite HSI". The best performances are in bold typeface; the second-best performances are underlined; and the third-best performances are italicized.

Methods	Materials												Average
	EM_1	EM_2	EM_3	EM_4	EM_5	EM_6	EM_7	EM_8	EM_9	EM_{10}	EM_{11}	EM_{12}	
MVNTF	0.3938	0.3118	0.5634	0.3161	0.4766	0.4595	*0.2555*	0.3130	0.2899	0.3846	0.4048	0.3648	0.3778
$l_{1/2}$-NMF	0.9145	0.5638	0.8468	0.7002	<u>0.2366</u>	0.5869	0.4558	**0.1127**	1.0758	0.9745	1.0424	**0.2291**	0.6449
Min-vol NMF	0.4747	**0.1161**	*0.1908*	**0.1316**	0.4535	0.4703	0.3358	*0.1651*	<u>0.1845</u>	0.4885	<u>0.2821</u>	0.3511	*0.3037*
R-CoNMF	0.5745	0.2054	0.1936	*0.2515*	0.5683	0.4079	0.3534	0.2266	0.4106	0.4181	*0.3770*	0.5228	0.3758
SSRNMF	**0.2521**	<u>0.1760</u>	<u>0.1885</u>	1.0213	**0.1375**	<u>0.1625</u>	<u>0.2240</u>	<u>0.1311</u>	**0.1249**	**0.0595**	0.6160	<u>0.2394</u>	**0.2777**
KbSNMF-div	<u>0.3162</u>	*0.1977*	0.2731	<u>0.1864</u>	0.2842	**0.1459**	**0.2086**	0.4222	0.5450	<u>0.2630</u>	0.3174	0.4818	<u>0.3035</u>
SGSNMF	0.7456	0.5926	**0.1182**	0.7599	0.5824	0.7393	0.5179	0.5160	*0.2102*	0.4753	0.4875	0.6865	0.5360
Ours	*0.3504*	0.3774	0.4062	0.2939	*0.2701*	*0.2256*	0.4268	0.3089	0.3060	*0.3096*	**0.2803**	0.3071	0.3222

5.3. Spectral Variability Point of View

In this section, we evaluate the unmixing performance of CUR-HU under the spectral variability HSI datasets. We compare the performance with the state-of-the-art methods, such as the fully constrained least squares (FCLS), the Perturbed LMM model (PLMM), the Extended Linear Mixing Model (ELMM), the Generalized LMM (GLMM) [74], and DeepGUn [75]. In all experiments, the VCA algorithm was utilized to extract the end-

members of HSIs data and initialize the methods requiring the endmembers reference matrix. The abundance maps from the FCLS algorithm were utilized to initialize all other methods. CUR-HU does not require initialization for the endmembers or abundance maps. To evaluate the performance of unmixing algorithms, we consider the Normalized Root-Mean-Squared Error (NRMSE) for the estimated abundance maps (NRMSE$_A$) and reconstructed images (NRMSE$_Y$). The NRMSE between a true, generic matrix $\mathbf{X}$ and its estimate $\widehat{X}$ is computed as

$$\text{NRMSE}_X = \sqrt{\frac{\|\mathbf{X} - \hat{\mathbf{X}}\|_F^2}{\|\mathbf{X}\|_F^2}} \tag{25}$$
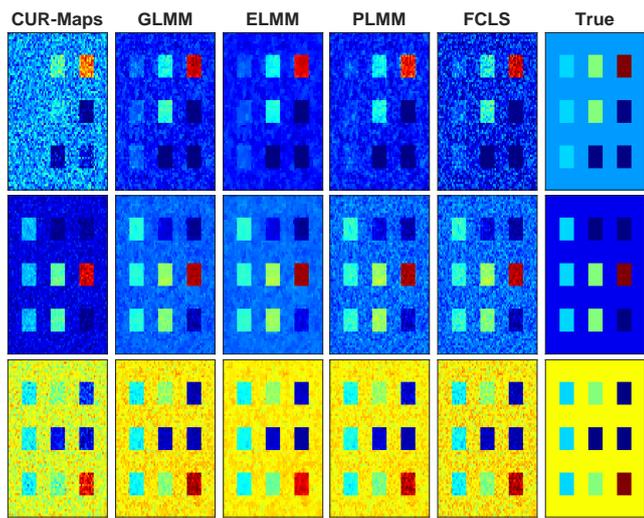
5.3.1. Synthetic HSI Datasets

For a comprehensive and quantitative comparison between the different algorithms, we generated four synthetic datasets, namely (SD0, SD2, SD3, and SD4), with 50×50 pixels for (SD1, SD2, and SD3) and 70×70 pixels for (SD0). These synthetic datasets were created using three and five 224-band endmember signatures for (SD0, SD1, SD2) and (SD3), respectively. The endmember signatures extracted from the USGS library and their abundance maps are represented in the last column of Figure 11. We adopted the generation of those datasets as used in [75,76]. The spectral variability of endmembers was imposed using different strategies for each dataset. For the SD0 dataset, we utilized the variability model based on pixel-wise multiplicative factors generated by random piecewise-linear functions as used in [77], while for SD1, we adopted the variability model used in [74], forming from band-dependent scaling factors that were smoothly varying in the spectral and spatial dimensions. For SD2, we considered a simple strategy used in [75,78] to emulate the real scenario of errors in atmospheric compensation as a function of viewing the geometry given the diffuse and direct light on the scene and the solar path transmittance. For SD3, the endmembers pure pixels of five materials (tree, asphalt, metal, dirt, and roof) were extracted manually from HSI to emulate the realistic spectral variability. Finally, white Gaussian noise was added to all synthetic datasets to set the SNR to 30 dB. The optimal parameters of each method were selected by using the grid search for each dataset. The grid search ranges of each parameter were chosen according to the ranges discussed and tested by authors in the original publications. For both GLMM and ELMM, the parameters were chosen among the following ranges: $\lambda_M, \lambda_S \in [0.01, 0.1, 1, 5, 10, 50]$, $\lambda_A \in [0.001, 0.01, 0.05]$, and $\lambda_\Psi, \lambda_\psi \in [10^{-6}, 10^{-3}, 1, 10^3]$. For the PLMM, the grid search ranges of α, γ, and β were selected as $[0.01, 0.1, 0.35, 0.7, 1.4, 25]$, $[10^{-2}, 0.1, 1, 10, 10^2]$, and $[10^{-9}, 10^{-5}, 10^{-4}, 10^{-3}]$, respectively.

The execution time of each algorithm is reported as the average running time over ten iterations. For a fair comparison, the running time performance of the FCLS method is not included because the FCLS method is used as an initialization step for other methods. We also used the VCA method to initialize and provide the number of endmember signatures for all methods.

Figure 11 shows the qualitative results of estimating the abundance maps for the synthetic datasets. For SD3 and SD4, the estimated abundance maps by CUR-HU (CUR-Maps) are much closer to the ground truth than the other methods, as shown in Figure 11c,d. Table 7 shows the quantitative results of the estimated abundance map, reconstruction error of the HSI, and algorithm running time of different methods for synthetic datasets. CUR-HU outperforms the other methods for estimating the abundance maps for SD3 and SD4 datasets and has comparable performance with other methods for estimating abundance maps of SD0 and SD1. Furthermore, CUR-HU has comparable accuracy for the reconstruction of the HSI data. According to the running time, our method outperforms all other methods.

Table 7. Performance comparison of different HU methods for synthetic HSIs. The best performances are in bold typeface; the second-best performances are underlined; and the third-best performances are italicized.
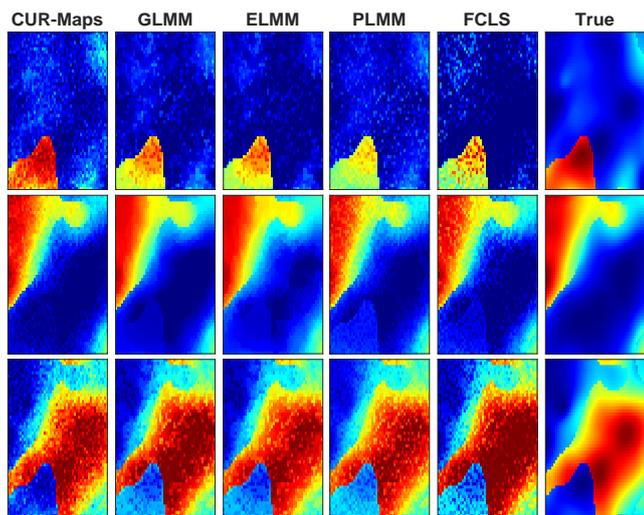
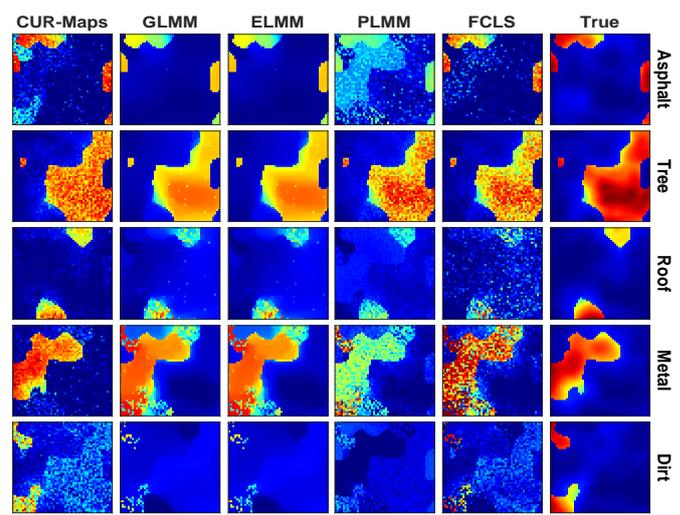|	Data Sets											
Methods	**SD0**			**SD1**			**SD2**			**SD3**		
	$NRMSE_A$	$NRMSE_Y$	*Time (s)*	$NRMSE_A$	$NRMSE_Y$	*Time (s)*	$NRMSE_A$	$NRMSE_Y$	*Time (s)*	$NRMSE_A$	$NRMSE_Y$	*Time (s)*
PLMM [77]	0.2604	**0.0007**	100.6	0.1197	0.0336	*18.87*	0.2028	*0.0302*	37.91	0.5066	0.0320	303.5
ELMM [79]	0.2554	0.0321	49.06	0.1110	0.0231	22.64	0.1997	0.0238	*25.81*	0.4385	**0.0106**	*21.97*
GLMM [74]	*0.2480*	0.0235	*41.50*	0.1146	**0.0226**	25.60	0.1841	**0.0226**	36.77	*0.4371*	*0.0108*	26.45
DeepGUn [75]	**0.0566**	0.0448	75.20	0.0969	0.0384	36.40	0.1613	0.0457	48.96	0.2550	0.1403	96.23
Ours	0.1896	*0.0320*	**0.27**	0.3017	0.0378	**0.17**	**0.1173**	0.0450	**0.11**	**0.2532**	0.0587	**0.10**

(**a**) Abundance maps of the synthetic HSI dataset (SD0)
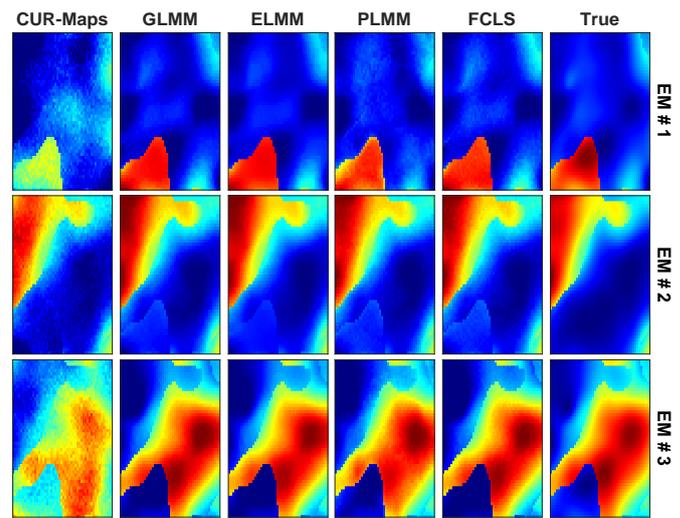
(**b**) Abundance maps of the synthetic HSI dataset (SD1)

(**c**) Abundance maps of the synthetic HSI dataset (SD2)

(**d**) Abundance maps of the synthetic HSI dataset (SD3)

Figure 11. Abundance maps of synthetic data sets using different HU methods (FCLS, PLMM, ELMM, GLMM, and CUR-HU). The colors represent the abundance values from blue ($s_k = 0$) to red ($s_k = 1$).

5.3.2. Real HSI Data

To evaluate the performance of CUR-HU by comparing it with other methods, we considered three real HSI datasets (i.e., Samson, Houston, and Jasper Ridge datasets). The

Samson dataset is the same HSI dataset that was utilized in the previous section. Houston and the Jasper Ridge HSIs were captured by the AVIRIS instrument with 224 spectral bands. After preprocessing, the Houston and the Jasper Ridge HSIs contain 188 and 198 spectral bands, respectively, and each of them has four endmember signatures as a reference [79]. The number of pixels of the Houston and the Jasper Ridge HSIs is 105×128 pixels and 100×100 pixels, respectively. Figure 12 shows the estimated abundance map of the Samson HSI dataset using different HU methods (i.e., FCLS, PLMM, ELMM, GLMM, and CUR-HU). As can be seen, the estimated abundance map from the proposed method has a clear performance improvement over the other methods, especially for the soil and water abundance maps.
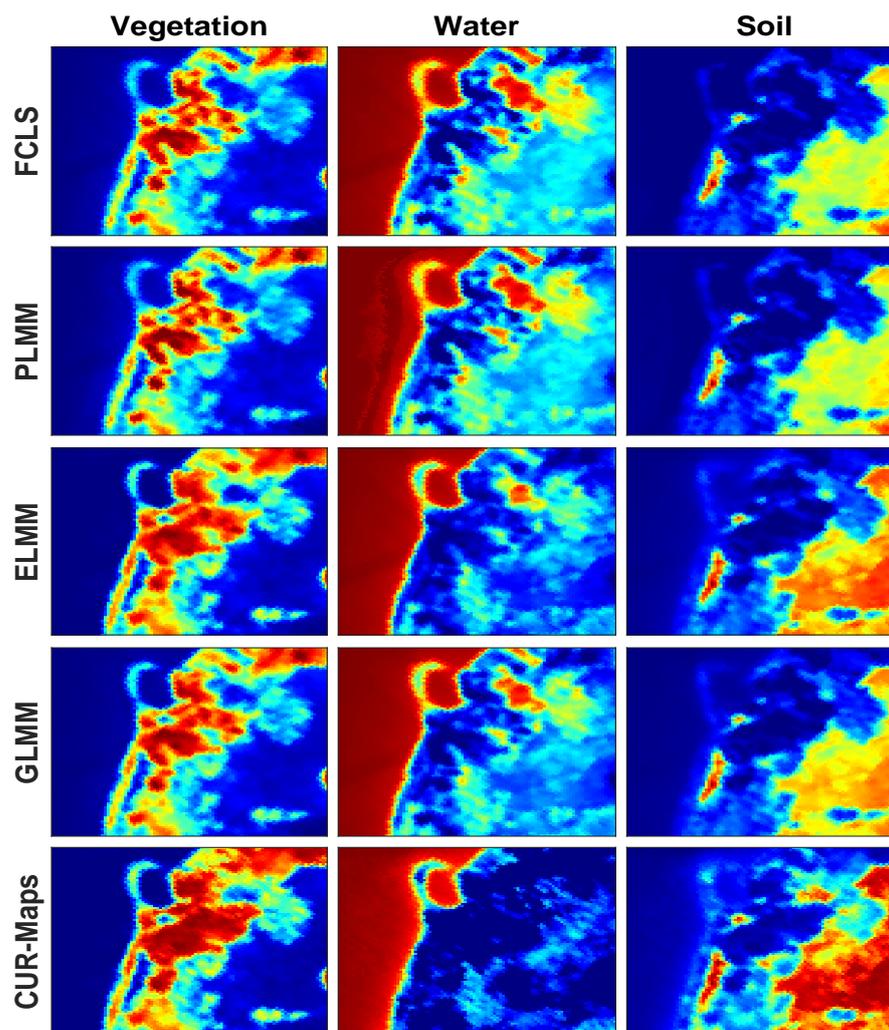
Figure 12. Abundance maps of Samson HSI for all tested methods. The colors represent the abundance values from blue ($s_k = 0$) to red ($s_k = 1$).

Figure A2a shows the estimated abundance map of the Houston HSI dataset. Although the performance of the proposed method for estimating the oncrete abundance map is weak, it leads to a considerably stronger Met, Roofs and Vegetation abundance maps but a comparable performance for Asphalt abundance map. The estimated abundance maps for the Jasper Ridge HSI dataset are shown in Figure A2b. The performance of the proposed method is comparable with other methods for estimating abundance maps (i.e., Water, Vegetation, and Soil). Although there is confusion between Water and Road abundance maps, the estimated road region is best compared to other maps.

Table 8 shows the quantitative results for different methods on real HSIs. Because the ground truth of the abundance maps is not available, we use the reconstruction error $NRMSE_Y$ metric to evaluate the quality. CUR-HU has a comparable performance with other methods. Although the ELMM, GLMM, and PLMM methods can achieve arbitrarily small $NRMSE_Y$, it does not necessarily reflect a high estimation accuracy of abundance maps, as shown in Figures A2a,b and 12. For the execution times, the proposed method outperforms all other algorithms.The speedup of CUR-HU over PLMM, ELMM, GLMM, and DeepGUn methods ranged between 9.4 to 249.5 times estimating the endmembers and their abundance maps of Samson, Houseton, and Jasper Ridge HSI datasets.

Table 8. Performance comparison of different HU methods for real HSIs. The best performances are in bold typeface; the second-best performances are underlined; and the third-best performances are italicized.

| Methods | Data Sets | | | | | |
| | Samson | | Houseton | | Jasper Ridge | |
	$NRMSE_Y$	*Time (s)*	$NRMSE_Y$	*Time (s)*	$NRMSE_Y$	*Time (s)*
PLMM [77]	*0.0239*	81.68	0.0713	412.65	*0.0553*	157.23
ELMM [79]	<u>0.0119</u>	*33.19*	<u>0.0171</u>	*72.50*	<u>0.0278</u>	90.87
GLMM [74]	**0.0011**	45.86	**0.0016**	75.88	**0.0032**	*74.23*
DeepGUn [75]	0.0862	121.88	0.2355	259.60	0.1094	209.64
CUR-HU	0.0425	**3.54**	*0.0588*	**5.13**	0.0773	**0.84**

Overall, the proposed method yielded a comparable unmixing performance with no parameter tuning and at high computational efficiency. Furthermore, the proposed framework can be utilized in the HU algorithms as an initialization step to provide the number of endmembers, endmember signatures, and abundance maps.

6. Conclusions

In this paper, we propose a new hyperspectral unmixing (HU) framework based on CUR matrix decomposition called CUR-HU that performs the entire HU process, which consists of three stages (i.e., estimating the optimal number of endmembers, estimating the endmembers and their corresponding abundance maps). However, most of the existing HU algorithms mainly focus on only one or two tasks of the HU process. CUR-HU is a fully unsupervised method, consists of three main stages, and incorporates several techniques to perform the HU process with high computational efficiency. For estimating the number of endmembers, we utilized a QR factorization based on an incremental method (IQR), used for approximating the singular vectors of the hyperspectral image (HSI) that are required for the deterministic selecting method, which significantly improves the scalability of the proposed method. Then, for the blind HU that estimates the endmembers and their abundance maps simultaneously, we proposed a method based on CUR matrix decomposition with a deterministic sampling method (i.e., DEIM). A unique advantage of our method is that the CUR decomposition based on IQR decomposition provides the matrix rank information for estimating the number of endmembers, and at the same time, we can compute the low-rank approximation of the original data with meaningful interpretation. Extensive experiments were conducted on various synthetic and real HSIs datasets to demonstrate the effectiveness of each stage of the proposed framework. The experiment results show that CUR-HU performs comparable to state-of-the-art methods for estimating the number of endmembers and abundance maps. However, it outperforms other methods for estimating the endmembers and the computational efficiency. Furthermore, CUR-HU shows robustness to spectral variability and is simple to implement and applicable to parallelizing. The proposed framework can be utilized with the existing HU algorithms to initialize the number of endmembers, signatures, and abundance maps. It can also be used to provide supervised information for HU methods based on deep learning techniques. CUR-HU has a 9.4 to 249.5 times speedup over different state-of-the-art methods for estimating the endmembers and their abundance maps of different real HSI datasets.

Appendix A. QR Factorization Based on an Incremental Method

In general, at step t of the algorithm, the approximated *QR* factorization is given by

$$\mathbf{Y}_t \approx \mathbf{Q}_t \mathbf{R}_t$$

where $\mathbf{Q}_t$ contains the orthonormal columns, $\mathbf{R}_t$ is the triangular factor at step t, and $\mathbf{Y}_t$ contains the first t columns of HSI data matrix $\mathbf{Y}$ [55]. So,

$$\|\mathbf{Y}_t - \mathbf{Q}_t \mathbf{R}_t\|_F \le tol \cdot d_t \cdot \|\mathbf{R}_t\|_F$$

where $\|.\|_F$ is the Frobenius norm, and d_t is the number of row/column deletions (number of truncation steps) that have been happened up to and including step t. At this point, the dimension of $\mathbf{Q}_t$ is $L \times (t - d_t)$, $\mathbf{R}_t$ is $(t - d_t) \times t$, and $d_t = \min\{L, N\} - k$, where $k = \text{rank}(\mathbf{QR})$.

Assume that the difference error matrix at the step t is $\mathbf{E}_t = \mathbf{A}_t - \mathbf{Q}_t \mathbf{R}_t$, and let

$$\|\mathbf{E}_t\|_F \le tol \cdot d_t \cdot \|\mathbf{R}_t\|_F$$

By using Gram–Schmidt to orthogonalize the next column $t + 1$ of $\mathbf{Y}$, the *QR* factorization will update as

$$\mathbf{Y}_{t+1} = \mathbf{Q}_{t+1} \mathbf{R}_{t+1} + [\mathbf{E}_t, \mathbf{0}]$$

In the case of no truncation occurring, then $d_{t+1} = d_t$ and $\|\mathbf{R}_t\|_F \le \|\mathbf{R}_{t+1}\|_F$. The norm of the error matrix is

$$\|\mathbf{E}_{t+1}\|_F = \|[\mathbf{E}_t, \mathbf{0}]\|_F \le tol \cdot d_t \cdot \|\mathbf{R}_t\|_F \le tol \cdot d_{t+1} \cdot \|\mathbf{R}_{t+1}\|_F.$$

In other words, each step has a row and column deleted from or added to $\mathbf{R}$ and $\mathbf{Q}$, respectively, according to the truncation condition.

Assume that after the number deletions d_t, the matrix $\mathbf{R}_t$ has dimension $m \times j$ (i.e., $m = j - d_j$) and the index of the row of the minimum norm is i. Furthermore, assume that $\widehat{\mathbf{R}}_{t+1}$ and $\widehat{\mathbf{Q}}_{t+1}$ are obtained from $\mathbf{R}_{t+1}$ and $\mathbf{Q}_{t+1}$ by removing the ith row and column, respectively. The truncation case occurs if $\mathbf{r}_i^T = \mathbf{e}_i^T \mathbf{R}_{t+1}$ satisfies $\|\mathbf{r}_i^T\| \le tol \cdot \|\widehat{\mathbf{R}}_{t+1}\|_F$, then deleting row i of $\mathbf{R}_{t+1}$ and column i of $\mathbf{Q}_{t+1}$ column i of $\mathbf{Q}_{j+1}$ and row i of $\mathbf{R}_{j+1}$ and replacing $\mathbf{Q}_{t+1}$ and $\mathbf{R}_{t+1}$ with $\widehat{\mathbf{Q}}_{t+1}$ and $\widehat{\mathbf{R}}_{t+1}$. So,

$$\widehat{\mathbf{Q}}_{t+1} \widehat{\mathbf{R}}_{j+1} = \mathbf{Q}_{t+1} \left(\mathbf{R}_{t+1} - \mathbf{e}_i \mathbf{r}_i^T \right)$$

Hence,

$$\mathbf{Y}_{t+1} = \mathbf{Q}_{t+1} \left(\mathbf{R}_{t+1} - \mathbf{e}_i \mathbf{r}_i^T \right) + [\mathbf{E}_t, \mathbf{0}] + \mathbf{Q}_{t+1} \mathbf{e}_i \mathbf{r}_i^T$$

Furthermore, $\widehat{\mathbf{R}}_{t+1}$ contains row $m+1$ of $\mathbf{R}_{t+1}$, which has a norm larger than the row determined for deletion. At the end of the truncation step, the $\mathbf{Q}_{t+1}$ and $\mathbf{R}_{t+1}$ replaced with $\widehat{\mathbf{Q}}_{t+1}$ and $\widehat{\mathbf{R}}_{t+1}$ to obtain the approximation within

$$\|\mathbf{Y}_{t+1}\| \leq tol \cdot d_{t+1} \cdot \|\mathbf{R}_{t+1}\|_F$$

where $d_{t+1} = d_t + 1$.

Algorithm A1: Estimating the endmembers number

> **Input** : HSI data matrix $\mathbf{Y} \in \mathbb{R}^{L \times N}$ and *tol*.
> **Output:** $\mathbf{Q} \in \mathbb{R}^{L \times p}$, $\mathbf{R} \in \mathbb{R}^{p \times N}$, and the number of endmembers p.
> **1** Select $k \ll \min(L, N)$;
> **2** Compute **QR** decomposition for the first k pixels of $\mathbf{Y}$, and $N(i) = \|\mathbf{R}(i, :)\|^2$ for $i = 1, \ldots, k$;
> **3** Set $t = k + 1$;
> **4** **while** *$(t \leq n)$* **do**
> **5** $\mathbf{y} = \mathbf{Y}(: t)$;
>   /* Re-orthogonalization step   */
> **6** $\mathbf{r} = \mathbf{Q}^T \mathbf{y}$;
> **7** $\mathbf{f} = \mathbf{y} - \mathbf{Q}\mathbf{r}$; $\mathbf{c} = \mathbf{Q}^T \mathbf{f}$; $\mathbf{f} = \mathbf{f} - \mathbf{Q}\mathbf{c}$;
> **8** $\mathbf{r} = \mathbf{r} + \mathbf{c}$; $\rho = \|\mathbf{f}\|$; $\mathbf{q} = \mathbf{f}/\rho$;
>   /* Updating step   */
> **9** $\mathbf{Q} = [\mathbf{Q}, \mathbf{q}]$; $\mathbf{R} = \begin{bmatrix} \mathbf{R} & \mathbf{r} \\ \mathbf{0} & \rho \end{bmatrix}$;
> **10** $N(i) = N(i) + \mathbf{r}(i)^2$ for $i = 1, \ldots, k$;
> **11** $N(k+1) = \rho^2$; $F_{normR} = sum(N)$;
> **12** $[\sigma, i_{\min}] = \min(N(1 : k+1))$;
> **13**   **if** $\sigma < (tol^2) * (F_{normR} - N(i_{\min}))$ **then**
>   /* truncation required   */
> **14**   **if** $i_{\min} < k + 1$ **then**
> **15** $\mathbf{R}(i_{\min}, :) = \mathbf{R}(k + 1, :)$;
> **16** $\mathbf{Q}(:, i_{\min}) = \mathbf{Q}(:, k + 1)$;
> **17** $N(i_{\min}) = N(k + 1)$
> **18**   **end**
> **19** $\mathbf{Q} = \mathbf{Q}(:, 1 : k)$;
> **20** $\mathbf{R} = \mathbf{R}(1 : k, :)$;
> **21**   **else**
>   /* no truncation   */
> **22** $k = k + 1$;
> **23**   **end**
> **24** $t = t + 1$;
> **25** **end**
> **26** Set $p \equiv k$;

Appendix A.1. The Complexity of Computing the Singular Vectors of HSI

The optimal computing of SVD of a large HSI dataset $\mathbf{Y}$ of size L by N requires $\mathcal{O}(LN^2 + L^2N + N^3)$. By approximating the HSI $\mathbf{Y}$ using the QR factorization as follows, $\mathbf{Y} \approx \mathbf{QR}$ where $\mathbf{Q} \in \mathbb{R}^{L \times p}$ that contains the orthonormal column vectors, $\mathbf{R} \in \mathbb{R}^{p \times N}$ is the upper triangular matrix, and p is the estimated number of endmembers. Therefore, computing the SVD based on the QR factorization requires $\mathcal{O}(LNp^2)$, which is more efficient [58].

Table A1. Estimation of the number of endmembers of synthetic HSI under changing SNR and p for ($\eta = 1/18$).

SNR	Methods	No. of Endmembers p			
		3	**5**	**10**	**15**
50 dB	**HySime**	3	5	10	15
	NWHFC ($P_f = 1 \times 10^{-3}$)	9(3)	10(5)	12(7)	10(11)
	NWHFC ($P_f = 1 \times 10^{-4}$)	48(3)	33(5)	54(10)	34(10)
	NWHFC ($P_f = 1 \times 10^{-5}$)	43(3)	28(5)	41(9)	27(10)
	GENE-AH ($P_f = 1 \times 10^{-3}$)	4	7	11	16
	GENE-AH ($P_f = 1 \times 10^{-4}$)	5	5	11	15
	GENE-AH ($P_f = 1 \times 10^{-5}$)	3	5	10	20
	SPICE	3	5	11	14
	FUN ($P_f = 1$)	3	5	10	14
	Ours ($tol = 2 \times 10^{-3}$)	3	6	9	14
35 dB	**HySime**	3	5	10	14
	NWHFC ($P_f = 1 \times 10^{-3}$)	3(3)	4(4)	7(7)	9(9)
	NWHFC ($P_f = 1 \times 10^{-4}$)	9(3)	9(5)	11(7)	8(10)
	NWHFC ($P_f = 1 \times 10^{-5}$)	3(3)	4(4)	6(6)	8(8)
	GENE-AH ($P_f = 1 \times 10^{-3}$)	4	8	18	23
	GENE-AH ($P_f = 1 \times 10^{-4}$)	3	6	15	21
	GENE-AH ($P_f = 1 \times 10^{-5}$)	5	5	19	18
	SPICE	4	5	9	13
	FUN ($P_f = 1$)	3	5	9	13
	Ours ($tol = 1 \times 10^{-3}$)	3	6	9	13
25 dB	**HySime**	3	5	9	13
	NWHFC ($P_f = 1 \times 10^{-3}$)	4(3)	5(5)	11(10)	9(11)
	NWHFC ($P_f = 1 \times 10^{-4}$)	4(3)	5(5)	11(9)	9(10)
	NWHFC ($P_f = 1 \times 10^{-5}$)	4(3)	5(5)	11(8)	8(10)
	GENE-AH ($P_f = 1 \times 10^{-3}$)	5	13	19	26
	GENE-AH ($P_f = 1 \times 10^{-4}$)	8	7	19	17
	GENE-AH ($P_f = 1 \times 10^{-5}$)	4	5	17	26
	SPICE	18	35	30	17
	FUN ($P_f = 1$)	4	8	12	14
	Ours ($tol = 5 \times 10^{-3}$)	3	7	11	16
	Ours ($tol = 1 \times 10^{-2}$)	3	5	7	11
15 dB	**HySime**	3	5	9	13
	NWHFC ($P_f = 1 \times 10^{-3}$)	4(3)	5(5)	11(10)	9(11)
	NWHFC ($P_f = 1 \times 10^{-4}$)	4(3)	5(5)	11(9)	9(10)
	NWHFC ($P_f = 1 \times 10^{-5}$)	4(3)	5(5)	11(8)	8(10)
	GENE-AH ($P_f = 1 \times 10^{-3}$)	7	8	17	29
	GENE-AH ($P_f = 1 \times 10^{-4}$)	9	10	19	12
	GENE-AH ($P_f = 1 \times 10^{-5}$)	5	9	21	24
	SPICE	36	40	40	40
	FUN ($P_f = 1$)	14	23	17	30
	Ours ($tol = 2 \times 10^{-2}$)	5	8	11	17

Appendix A.2. DEIM for Approximating the Nonlinear Functions

The discrete empirical interpolation method (DEIM) was proposed in [56] to efficiently approximate the nonlinear function $f(\tau)$ by projecting it onto a subspace that approximates space of $f(\tau)$ and is spanned by the standard basis $\mathbf{W}$ (left singular vectors). In other words, the DEIM method approximates $f(\tau)$ by combining projection with interpolation, as it constructs specially selected interpolation indices to specify the interpolation-based projection that provides a nearly optimal subspace approximation to $f(\tau)$.

The approximation from projecting $f(\tau)$ onto the subspace spanned by the basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_r\}$ is given as $\tilde{f}(\tau) \approx \mathbf{W}\mathbf{z}(\tau)$, where $\mathbf{z}(\tau)$ is the corresponding coefficients. To determine $\mathbf{z}(\tau)$, DEIM constructs r row indices based on the standard basis $\mathbf{W}$, at which the function $f(\tau)$ is evaluated. Suppose the selecting matrix $\mathbf{P}$ defined as $\mathbf{P} = [\mathbf{e}_{t_1}, \ldots, \mathbf{e}_{t_r}] \in \mathbb{R}^{n \times r}$, where $\mathbf{e}_{t_j}$ is the t_j-th column of the identity matrix $\mathbf{I}_n \in \mathbb{R}^{n \times n}$. If $\mathbf{P}^T\mathbf{W} \in \mathbb{R}^{r \times r}$ is nonsingular, then the coefficient vector $\mathbf{z}(\tau)$ can be computed uniquely from

$$\mathbf{P}^T f(\tau) = (\mathbf{P}^T \mathbf{W})\mathbf{z}(\tau)$$

So,

$$\mathbf{z}(\tau) = \left(\mathbf{P}^T\mathbf{W}\right)^{-1}\mathbf{P}^T f(\tau)$$

Therefore, the final approximation of $f(\tau)$ is

$$\tilde{f}(\tau) \approx \mathbf{W}\mathbf{z}(\tau) = \mathbf{W}\left(\mathbf{P}^T\mathbf{W}\right)^{-1}\mathbf{P}^T f(\tau)$$

where $\mathcal{P} \equiv \mathbf{W}\left(\mathbf{P}^T\mathbf{W}\right)^{-1}\mathbf{P}^T$ is called as the DEIM projector. So, given the DEIM projector $\mathcal{P}$, DEIM approximates the function $f(\tau)$ as $\tilde{f}(\tau) = \mathcal{P}f(\tau)$. To obtain the approximation of the nonlinear function using the DEIM method, we need to find the standard projection basis $\{\mathbf{W}_1, \ldots, \mathbf{W}_r\}$ and the interpolation indices $\{t_1, \ldots, t_r\}$ that used to form the matrix $\mathbf{P}$. Those interpolation indices used to determine the DEIM projector are selected inductively on the basis $\{\mathbf{W}_1, \ldots, \mathbf{W}_r\}$ by the DEIM procedures as shown in Algorithm 1. It treats the continuous domain as a finite set of discrete points in this continuous domain. The selection procedure of the DEIM method essentially involves minimizing the approximation error through the selected interpolation index in each iteration.

Appendix B

Figure A2a shows the estimated abundance map of the Houston HSI dataset. Although the performance of the proposed method for estimating the Concrete abundance map is weak, it leads to a considerably stronger Met. Roofs and Vegetation abundance maps but a comparable performance for Asphalt abundance map.

The estimated abundance maps for the Jasper Ridge HSI dataset are shown in Figure A2b. The performance of the proposed method is comparable with other methods for estimating abundance maps (i.e., Water, Vegetation, and Soil). Although there is confusion between Water and Road abundance maps, the estimated road region is best compared to other maps.
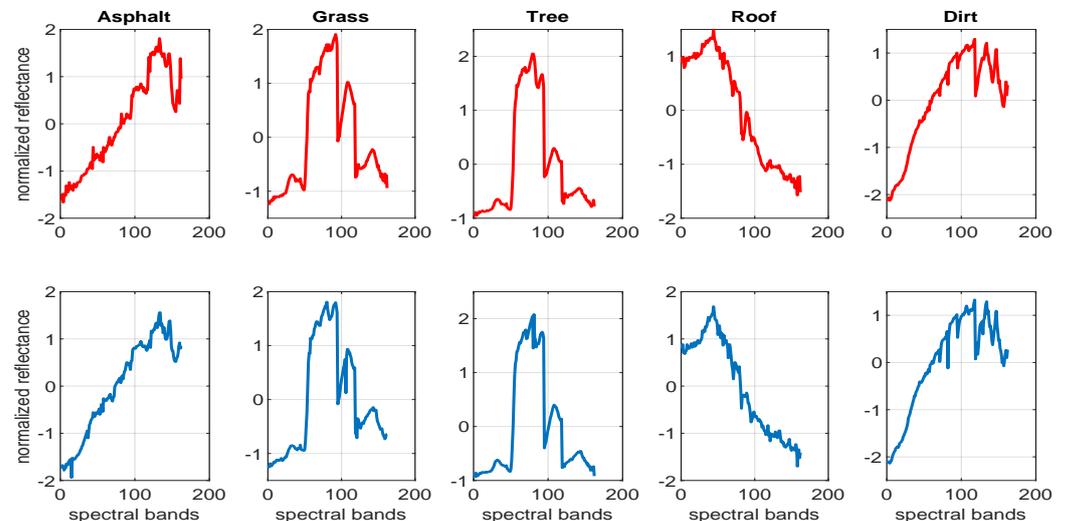
Figure A1. Ground truth and extracted endmembers of the Urban HSI: the 1st row, ground truth (GT); and the 2nd row contain the estimated endmembers.
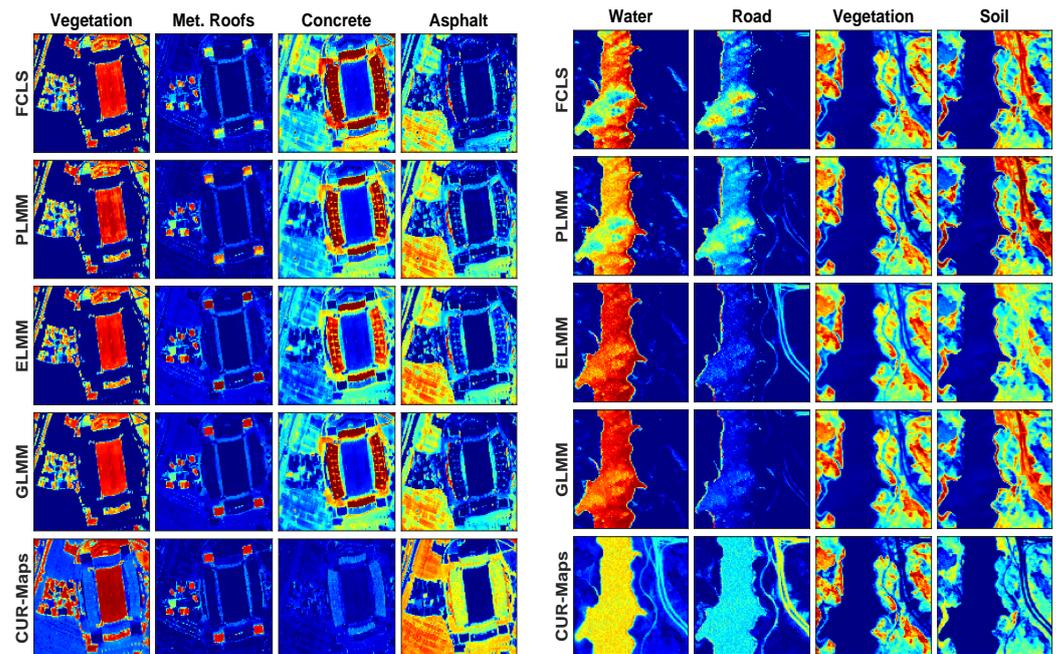
(**a**) Abundance maps of Houston HSI for all tested methods. The colors represent the abundance values from blue ($s_k = 0$) to red ($s_k = 1$).

(**b**) Abundance maps of Jaster Ridge HSI for all tested methods. The colors represent the abundance values from blue ($s_k = 0$) to red ($s_k = 1$).

Figure A2. Abundance maps of real HSI datasets for all tested methods.

References

1. Goetz, A.F. Three decades of hyperspectral remote sensing of the Earth: A personal view. *Remote Sens. Environ.* **2009**, *113*, S5–S16. [CrossRef]
2. Borsoi, R.A.; Imbiriba, T.; Bermudez, J.C.M.; Richard, C.; Chanussot, J.; Drumetz, L.; Tourneret, J.Y.; Zare, A.; Jutten, C. Spectral variability in hyperspectral data unmixing: A comprehensive review. *IEEE Geosci. Remote Sens. Mag.* **2021**, *9*, 223–270. [CrossRef]
3. Keshava, N.; Mustard, J.F. Spectral unmixing. *IEEE Signal Process. Mag.* **2002**, *19*, 44–57. [CrossRef]
4. Bioucas-Dias, J.M.; Plaza, A.; Dobigeon, N.; Parente, M.; Du, Q.; Gader, P.; Chanussot, J. Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5*, 354–379. [CrossRef]
5. Chang, C.I.; Du, Q. Estimation of number of spectrally distinct signal sources in hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 608–619. [CrossRef]
6. Nascimento, J.M.P.; Bioucas-Dias, J.M. Vertex component analysis: A fast algorithm to unmix hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 898–910. [CrossRef]
7. Heinz, D.C.; Change, C.-I. Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 529–545. [CrossRef]
8. Ahmed, A.M.; Duran, O.; Zweiri, Y.; Smith, M. Hybrid spectral unmixing: Using artificial neural networks for linear/non-linear switching. *Remote Sens.* **2017**, *9*, 775. [CrossRef]
9. Shen, X.; Bao, W. Hyperspectral endmember extraction using spatially weighted simplex strategy. *Remote Sens.* **2019**, *11*, 2147. [CrossRef]
10. Heinz, D.; Chang, C.I.; Althouse, M.L. Fully constrained least-squares based linear unmixing [hyperspectral image classification]. In Proceedings of the IEEE 1999 International Geoscience and Remote Sensing Symposium. IGARSS'99 (Cat. No. 99CH36293), Hamburg, Germany, 28 June–2 July 1999; Volume 2, pp. 1401–1403.
11. Zhong, Y.; Wang, X.; Zhao, L.; Feng, R.; Zhang, L.; Xu, Y. Blind spectral unmixing based on sparse component analysis for hyperspectral remote sensing imagery. *ISPRS J. Photogramm. Remote Sens.* **2016**, *119*, 49–63. [CrossRef]
12. Lee, D.; Seung, H.S. Algorithms for non-negative matrix factorization. *Adv. Neural Inf. Process. Syst.* **2000**, *13*. Available online: https://papers.nips.cc/paper_files/paper/2000/hash/f9d1152547c0bde01830b7e8bd60024c-Abstract.html (accessed on 28 December 2023).
13. Kishore Kumar, N.; Schneider, J. Literature survey on low rank approximation of matrices. *Linear Multilinear Algebra* **2017**, *65*, 2212–2244. [CrossRef]
14. Zhu, F. Hyperspectral unmixing: Ground truth labeling, datasets, benchmark performances and survey. *arXiv* **2017**, arXiv:1708.05125.

15. Mahoney, M.W.; Drineas, P. CUR matrix decompositions for improved data analysis. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 697–702. [CrossRef]

16. Chiu, J.; Demanet, L. Sublinear randomized algorithms for skeleton decompositions. *SIAM J. Matrix Anal. Appl.* **2013**, *34*, 1361–1383. [CrossRef]

17. Ambikapathi, A.; Chan, T.H.; Chi, C.Y.; Keizer, K. Hyperspectral data geometry-based estimation of number of endmembers using p-norm-based pure pixel identification algorithm. *IEEE Trans. Geosci. Remote Sens.* **2012**, *51*, 2753–2769. [CrossRef]

18. Bioucas-Dias, J.M.; Nascimento, J.M.P. Hyperspectral subspace identification. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 2435–2445. [CrossRef]

19. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **1978**, *6*, 461–464. [CrossRef]

20. Winter, M.E. N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data. In Proceedings of the Imaging Spectrometry V. SPIE, Denver, CO, USA, 19–21 July 1999; Volume 3753, pp. 266–275.

21. Boardman, J.W. Automating spectral unmixing of AVIRIS data using convex geometry concepts. In Proceedings of the JPL, Summaries of the 4th Annual JPL Airborne Geoscience Workshop, Washington, DC, USA, 25–29 October 1993; Volume 1. AVIRIS Workshop.

22. Chang, C.I.; Wu, C.C.; Liu, W.; Ouyang, Y.C. A new growing method for simplex-based endmember extraction algorithm. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2804–2819. [CrossRef]

23. Bioucas-Dias, J.M. A variable splitting augmented Lagrangian approach to linear spectral unmixing. In Proceedings of the 2009 First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, Lisbon, Portugal, 26–28 August 2009; pp. 1–4.

24. Iordache, M.D.; Bioucas-Dias, J.M.; Plaza, A. Sparse unmixing of hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 2014–2039. [CrossRef]

25. Xu, X.; Shi, Z.; Pan, B. ℓ_0-based sparse hyperspectral unmixing using spectral information and a multi-objectives formulation. *ISPRS J. Photogramm. Remote Sens.* **2018**, *141*, 46–58. [CrossRef]

26. Tang, W.; Shi, Z.; Wu, Y.; Zhang, C. Sparse unmixing of hyperspectral data using spectral a priori information. *IEEE Trans. Geosci. Remote Sens.* **2014**, *53*, 770–783. [CrossRef]

27. Comon, P.; Jutten, C.; Herault, J. Blind separation of sources, Part II: Problems statement. *Signal Process.* **1991**, *24*, 11–20. [CrossRef]

28. Chen, C.H.; Zhang, X. Independent component analysis for remote sensing study. In Proceedings of the Image and Signal Processing for Remote Sensing V. SPIE, Florence, Italy, 22–24 September 1999; Volume 3871, pp. 150–158.

29. Liu, L.; Wang, B.; Zhang, L.; Zhang, J.Q. Decomposition of mixed pixels using Bayesian Self-Organizing Map (BSOM) neural networks. In Proceedings of the 2007 IEEE International Geoscience and Remote Sensing Symposium, IEEE, Barcelona, Spain, 23–28 July 2007; pp. 2014–2017.

30. Nascimento, J.M.P.; Bioucas-Dias, J.M. Does independent component analysis play a role in unmixing hyperspectral data? *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 175–187. [CrossRef]

31. Nascimento, J.M.P.; Bioucas-Dias, J.M. Dependent component analysis: A hyperspectral unmixing algorithm. In Proceedings of the Iberian Conference on Pattern Recognition and Image Analysis, Barcelona, Spain, 23–28 July 2007; Springer: Berlin/Heidelberg, Germany, 2007; pp. 612–619.

32. Barros, A.K. The independence assumption: Dependent component analysis. In *Advances in Independent Component Analysis*; Springer: Berlin/Heidelberg, Germany, 2000; pp. 63–71.

33. Caiafa, C.F.; Salerno, E.; Proto, A.N.; Fiumi, L. Blind spectral unmixing by local maximization of non-Gaussianity. *Signal Process.* **2008**, *88*, 50–68. [CrossRef]

34. Zhu, F.; Wang, Y.; Xiang, S.; Fan, B.; Pan, C. Structured sparse method for hyperspectral unmixing. *ISPRS J. Photogramm. Remote Sens.* **2014**, *88*, 101–118. [CrossRef]

35. Leplat, V.; Ang, A.M.; Gillis, N. Minimum-volume rank-deficient nonnegative matrix factorizations. In Proceedings of the ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 12–17 May 2019; pp. 3402–3406.

36. Qian, Y.; Jia, S.; Zhou, J.; Robles-Kelly, A. L1/2 sparsity constrained nonnegative matrix factorization for hyperspectral unmixing. In Proceedings of the 2010 International Conference on Digital Image Computing: Techniques and Applications, Sydney, Australia, 1–3 December 2010; pp. 447–453.

37. Lu, X.; Wu, H.; Yuan, Y.; Yan, P.; Li, X. Manifold regularized sparse NMF for hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2012**, *51*, 2815–2826. [CrossRef]

38. Wang, X.; Zhong, Y.; Zhang, L.; Xu, Y. Spatial group sparsity regularized nonnegative matrix factorization for hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 6287–6304. [CrossRef]

39. Li, J.; Bioucas-Dias, J.M.; Plaza, A.; Liu, L. Robust collaborative nonnegative matrix factorization for hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6076–6090. [CrossRef]

40. Cai, D.; He, X.; Han, J.; Huang, T.S. Graph regularized nonnegative matrix factorization for data representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *33*, 1548–1560.

41. Zhou, L.; Zhang, X.; Wang, J.; Bai, X.; Tong, L.; Zhang, L.; Zhou, J.; Hancock, E. Subspace structure regularized nonnegative matrix factorization for hyperspectral unmixing. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 4257–4270. [CrossRef]

42. Lu, X.; Wu, H.; Yuan, Y. Double constrained NMF for hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2013**, *52*, 2746–2758. [CrossRef]

43. Yuan, Y.; Feng, Y.; Lu, X. Projection-based NMF for hyperspectral unmixing. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 2632–2643. [CrossRef]

44. Yu, Y.; Ma, Y.; Mei, X.; Fan, F.; Huang, J.; Li, H. Multi-stage convolutional autoencoder network for hyperspectral unmixing. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *113*, 102981. [CrossRef]

45. Su, Y.; Li, J.; Plaza, A.; Marinoni, A.; Gamba, P.; Chakravortty, S. DAEN: Deep autoencoder networks for hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 4309–4321. [CrossRef]

46. Pattathal V., A.; Sahoo, M.M.; Porwal, A.; Karnieli, A. Deep-learning-based latent space encoding for spectral unmixing of geological materials. *ISPRS J. Photogramm. Remote Sens.* **2022**, *183*, 307–320.

47. Licciardi, G.A.; Del Frate, F. Pixel unmixing in hyperspectral data by means of neural networks. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 4163–4172. [CrossRef]

48. Wang, M.; Zhao, M.; Chen, J.; Rahardja, S. Nonlinear unmixing of hyperspectral data via deep autoencoder networks. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 1467–1471. [CrossRef]

49. Manolakis, D.; Marden, D.; Shaw, G.A. Hyperspectral image processing for automatic target detection applications. *Linc. Lab. J.* **2003**, *14*, 79–116.

50. Luo, B.; Chanussot, J.; Douté, S.; Zhang, L. Empirical automatic estimation of the number of endmembers in hyperspectral images. *IEEE Geosci. Remote Sens. Lett.* **2012**, *10*, 24–28.

51. Févotte, C.; Dobigeon, N. Nonlinear hyperspectral unmixing with robust nonnegative matrix factorization. *IEEE Trans. Image Process.* **2015**, *24*, 4810–4819. [CrossRef]

52. Goreinov, S.A.; Tyrtyshnikov, E.E.; Zamarashkin, N.L. A theory of pseudoskeleton approximations. *Linear Algebra Its Appl.* **1997**, *261*, 1–21. [CrossRef]

53. Stewart, G.W. Four algorithms for the the efficient computation of truncated pivoted QR approximations to a sparse matrix. *Numer. Math.* **1999**, *83*, 313–323. [CrossRef]

54. Drineas, P.; Kannan, R.; Mahoney, M.W. Fast Monte Carlo algorithms for matrices I: Approximating matrix multiplication. *SIAM J. Comput.* **2006**, *36*, 132–157. [CrossRef]

55. Sorensen, D.C.; Embree, M. A DEIM induced CUR factorization. *SIAM J. Sci. Comput.* **2016**, *38*, A1454–A1482. [CrossRef]

56. Chaturantabut, S.; Sorensen, D.C. Nonlinear model reduction via discrete empirical interpolation. *SIAM J. Sci. Comput.* **2010**, *32*, 2737–2764. [CrossRef]

57. Lehoucq, R.B.; Sorensen, D.C.; Yang, C. *ARPACK Users' Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods*; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 1998.

58. Baker, C.G.; Gallivan, K.A.; Van Dooren, P. Low-rank incremental methods for computing dominant singular subspaces. *Linear Algebra Its Appl.* **2012**, *436*, 2866–2888. [CrossRef]

59. Guerra, R.; Santos, L.; López, S.; Sarmiento, R. A new fast algorithm for linearly unmixing hyperspectral images. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 6752–6765. [CrossRef]

60. Daniel, J.W.; Gragg, W.B.; Kaufman, L.; Stewart, G.W. Reorthogonalization and stable algorithms for updating the Gram–Schmidt QR factorization. *Math. Comput.* **1976**, *30*, 772–795. [CrossRef]

61. Giraud, L.; Langou, J.; Rozložník, M.; Eshof, J.v.d. Rounding error analysis of the classical Gram–Schmidt orthogonalization process. *Numer. Math.* **2005**, *101*, 87–100. [CrossRef]

62. Green, A.A.; Berman, M.; Switzer, P.; Craig, M.D. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Trans. Geosci. Remote Sens.* **1988**, *26*, 65–74. [CrossRef]

63. Roger, R.; Arnold, J. Reliably estimating the noise in AVIRIS hyperspectral images. *Int. J. Remote Sens.* **1996**, *17*, 1951–1962. [CrossRef]

64. Boutsidis, C.; Woodruff, D.P. Optimal CUR matrix decompositions. In Proceedings of the the Forty-Sixth Annual ACM Symposium on Theory of Computing, New York, NY, USA, 31 May–3 June 2014; pp. 353–362.

65. Hamm, K.; Huang, L. Perturbations of CUR decompositions. *SIAM J. Matrix Anal. Appl.* **2021**, *42*, 351–375. [CrossRef]

66. Zare, A.; Gader, P. Sparsity promoting iterated constrained endmember detection in hyperspectral imagery. *IEEE Geosci. Remote Sens. Lett.* **2007**, *4*, 446–450. [CrossRef]

67. Vane, G.; Green, R.O.; Chrien, T.G.; Enmark, H.T.; Hansen, E.G.; Porter, W.M. The airborne visible/infrared imaging spectrometer (AVIRIS). *Remote Sens. Environ.* **1993**, *44*, 127–143. [CrossRef]

68. Landgrebe, D. *Multispectral Data Analysis: A Signal Theory Perspective*; Purdue Univiersity: West Lafayette, IN, USA, 1998.

69. Miao, L.; Qi, H. Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 765–777. [CrossRef]

70. Ekanayake, M.; Herath, H. Constrained nonnegative matrix factorization for blind hyperspectral unmixing incorporating endmember independence. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 11853–11869. [CrossRef]

71. Jia, S.; Qian, Y. Spectral and spatial complexity-based hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 3867–3879.

72. Qian, Y.; Jia, S.; Zhou, J.; Robles-Kelly, A. Hyperspectral unmixing via $l_{1/2}$ sparsity-constrained nonnegative matrix factorization. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 4282–4297. [CrossRef]

73. Qian, Y.; Xiong, F.; Zeng, S.; Zhou, J.; Tang, Y.Y. Matrix-vector nonnegative tensor factorization for blind unmixing of hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2016**, *55*, 1776–1792. [CrossRef]

74. Imbiriba, T.; Borsoi, R.A.; Bermudez, J.C.M. Generalized linear mixing model accounting for endmember variability. In Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, BC, Canada, 15–20 April 2018; pp. 1862–1866.

75. Borsoi, R.A.; Imbiriba, T.; Bermudez, J.C.M. Deep generative endmember modeling: An application to unsupervised spectral unmixing. *IEEE Trans. Comput. Imaging* **2019**, *6*, 374–384. [CrossRef]

76. Chang, C.I. *Hyperspectral Data Processing: Algorithm Design and Analysis*; John Wiley & Sons: Hoboken, NJ, USA, 2013.

77. Thouvenin, P.A.; Dobigeon, N.; Tourneret, J.Y. Hyperspectral unmixing with spectral variability using a perturbed linear mixing model. *IEEE Trans. Signal Process.* **2015**, *64*, 525–538. [CrossRef]

78. Uezato, T.; Murphy, R.J.; Melkumyan, A.; Chlingaryan, A. A novel endmember bundle extraction and clustering approach for capturing spectral variability within endmember classes. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6712–6731. [CrossRef]

79. Drumetz, L.; Veganzones, M.A.; Henrot, S.; Phlypo, R.; Chanussot, J.; Jutten, C. Blind hyperspectral unmixing using an extended linear mixing model to address spectral variability. *IEEE Trans. Image Process.* **2016**, *25*, 3890–3905. [CrossRef]