Abstract—Hyperspectral unmixing (HU) is a very useful and increasingly popular preprocessing step for a wide range of hyperspectral applications. However, the HU research has been constrained a lot by three factors: (a) the number of hyperspectral images (especially the ones with ground truths) are very limited; (b) the ground truths of most hyperspectral images are not shared on the web, which may cause lots of unnecessary troubles for researchers to evaluate their algorithms; (c) the codes of most state-of-the-art methods are not shared, which may also delay the testing of new methods.

Accordingly, this paper deals with the above issues from the following three perspectives: (1) as a profound contribution, we provide a general labeling method for the HU. With it, we labeled up to 15 hyperspectral images, providing 18 versions of ground truths. To the best of our knowledge, this is the first paper to summarize and share up to 15 hyperspectral images and their 18 versions of ground truths for the HU. Observing that the hyperspectral classification (HyC) has much more standard datasets (whose ground truths are generally publicly shared) than the HU, we propose an interesting method to transform the HyC datasets for the HU research. (2) To further facilitate the evaluation of HU methods under different conditions, we reviewed and implemented the algorithm to generate a complex synthetic hyperspectral image. By tuning the hyper-parameters in the code, we may verify the HU methods from four perspectives. The code would also be shared on the web. (3) To provide a standard comparison, we reviewed up to 10 state-of-the-art HU algorithms, then selected the 5 most benchmark HU algorithms, and compared them on the 15 real hyperspectral datasets. The experiment results are surely reproducible; the implemented codes would be shared on the web.

Index Terms—Hyperspectral Unmixing (HU), Datasets, labeling, Ground Truth, Hyperspectral Classification (HyC).

I. INTRODUCTION

Hyperspectral remote sensing has been widely used in lots of applications1 since it can capture a 3D image cube at hundreds of contiguous bands across the electromagnetic spectrum, providing substantial information of the scene [1]. However, due to the microscopic material mixing, multiple scattering and low spatial resolution of hyperspectral sensors, the pixel spectra are inevitably mixed with various substances [1, 2], resulting in lots of mixed pixels. Accordingly, hyperspectral unmixing (HU) is an essential processing step for various hyperspectral image applications, such as high-resolution hyperspectral imaging [3–8], hyperspectral enhancement [9], sub-pixel mapping [10], hyperspectral compression and reconstruction [11], detection and identification substances in the scene [9, 12], hyperspectral visualization [13, 14], etc.

The HU aims to decompose each (mixed) pixel into a set of “pure” pixels (called endmembers such as the spectra of grass, water etc.), weighted by the corresponding proportions, called abundances [14–16]. Formally, given a hyperspectral image with $L$ channels and $N$ pixels, each pixel $y \in \mathbb{R}^L_N$ is assumed to be a composite of $K$ endmembers $\{m_k\}_{k=1}^K \in \mathbb{R}^L_K$. The linear combinatorial model is the most commonly used one

$$x = \sum_{k=1}^K m_k a_k, \quad \text{s.t. } a_k \geq 0 \text{ and } \sum_{k=1}^K a_k = 1,$$

where $a_k$ is the composite abundance of the $k^{th}$ endmember. In the unsupervised setting, both endmembers $\{m_k\}_{k=1}^K$ and abundances $\{a_k\}_{k=1}^K$ are unknown, which makes the HU a challenging problem [17–26].

The HU is a very hot research topic—tens or even hundreds HU papers have been published each year. However, the HU research has been constrained a lot since the commonly used datasets (especially their ground truths), are generally not shared on the web. Such case will surely hinder the development of new methods; the researchers, who are interested in the HU, will have to make great efforts to do the preparation works—they have to find the hyperspectral datasets and their ground truths, which generally ends up with failures. Instead, they have to label the hyperspectral images, which is very challenging and needs lots of techniques.

In order to promote the HU research, we write this paper from the following five perspectives:

1) This is the first paper to introduce a general method to label hyperspectral images for the HU research. We provide the method to label the endmembers & abundances (in Sections IV-A, IV-B), as well as the method to evaluate the labeling results (in Section IV-C). The hyperspectral classification (HyC) is similar to the HU task; it has much more standard datasets with publicly available ground truths. We propose an interesting method to transform the benchmark HyC datasets for the HU research. Please refer to Section IV.

2) Moreover, we are the first to summarize the information of 15 most commonly used hyperspectral images as well as their 18 versions of ground truths for the HU. All of them will be shared on the web as a standard dataset for the evaluation of new HU methods (cf. Section V, Table I and Fig. 1).

3) We reviewed and implemented the method to generate a complex synthetic hyperspectral image for the HU research in Section VI. This synthetic image is widely used in the papers [2, 12, 27–32]. The code will be also shared on the web.

4) We reviewed 10 state-of-the-art methods, summarized their main ideas and algorithms as well as pointed out their

1such as mining & oil industries, agriculture, food safety, pharmaceutical process monitoring and quality control, biomedical & biometric, surveillance, military, environment monitoring and forensic applications etc.
Fig. 1. The illustration of 7 popular hyperspectral images where we select 15 subimages (i.e., ROIs) and provide 18 versions of ground truths for the HU study. They are [a] Samson (3 ROIs), [b] Jasper Ridge (3 ROIs), [c] Urban (the full image and 2 ROIs), [d] Cuprite (1 ROI), [e] Moffett Field (2 ROIs), [f] San Diego Airport (1 ROI), and [g] Washington DC Mall (2 ROIs). In Section V and Table I, we provide very detailed information about the 15 ROIs.

TABLE I
The 15 real hyperspectral (sub-) images and their information, including the image size, the numbers of all spectral bands and the selected bands, the endmembers numbers, the illustration of scenes and their 18 versions of ground truths etc.

| Real Datasets | Image sizes | Number of bands | # Endmember | Ground truth shown in | The top-left pixel in original image | Subimages shown in |
|---------------|-------------|-----------------|-------------|----------------------|-------------------------------------|-------------------|
| 1. Samson#1   | 95 95       | 156 156         | 3           | Fig. 2a              | (252, 332)                          | Fig. 1a           |
| 2. Samson#2   | 95 95       | 156 156         | 3           | Fig. 2b              | (252, 352)                          |                   |
| 3. Samson#3   | 95 95       | 156 156         | 3           | Fig. 2c              | (252, 455)                          |                   |
| 4. Jasper Ridge#1 | 115 115   | 224 198         | 5           | Fig. 3a              | (1, 72)                            | Fig. 1b           |
| 5. Jasper Ridge#2 | 100 100   | 224 198         | 4           | Fig. 3b              | (105, 269)                          |                   |
| 6. Jasper Ridge#3 | 122 104   | 224 198         | 4           | Fig. 3c              | (1, 246)                           |                   |
| 7. Urban#4    | 307 307     | 224 198         | 6           | Figs. 4a, 4b, 4c     | (1, 1)                             | Fig. 1c           |
| 8. Urban#5    | 160 168     | 224 198         | 6           | Fig. 5a              | (1, 1)                             |                   |
| 9. Urban#6    | 160 168     | 224 198         | 6           | Fig. 5b              | (1, 1)                             |                   |
| 10. Urban#1   | 250 190     | 224 188         | 12          | Fig. 6               | (86, 425)                          | Fig. 1d           |
| 11. Urban#2   | 160 168     | 224 198         | 6           | Fig. 7a              | (54, 172)                          |                   |
| 12. Cuprite#1 | 60 60       | 224 196         | 3           | Fig. 7b              | (62, 146)                          |                   |
| 13. Moffett Field#1 | 60 60      | 224 196         | 3           | Fig. 7c              | (62, 146)                          |                   |
| 14. Moffett Field#2 | 86 86      | 224 196         | 3           | Fig. 7d              | (62, 146)                          |                   |
| 15. San Diego Airport | 160 140 | 224 189         | 3           | Figs. 8a, 8b         | (211, 261)                         | Fig. 1e           |
| 16. Washington DC Mall#1 | 150 150 | 224? 191        | 6           | Fig. 9a              | (549, 160)                         | Fig. 1f           |
| 17. Washington DC Mall#2 | 180 160 | 224? 191        | 6           | Fig. 9b              | (945, 90)                          |                   |

Accordingly, the weakly supervised methods [33, 34] were proposed. A large library of material spectra had been collected by a field spectrometer beforehand. Then, the HU task becomes the problem of finding an optimal subset of material spectra in the library that can best represent all the pixels in the hyperspectral image [33] as well as their abundance maps. Unfortunately, the library is far from optimal because the spectra in it are not standardly unified. First, for different hyperspectral sensors, the spectral signatures of the same material can be very inconsistent. Second, for the hyperspectral images recorded by different sensors, both the number of spectral bands and the electromagnetic range of recorded spectra can be largely different as well—for example, some images (like Samson) have 156 channels covering the spectra from 401 nm to 889 nm, while other images (like Cuprite and Jasper Ridge) have 224 channels covering the spectra from 370 nm to 2,480 nm. Finally, the recording

2Here supervise, weakly supervise and supervise are totally different from the those terms in the general machine learning.

3e.g., the United States Geological Survey (USGS) digital spectral library.
conditions are very different—some hyperspectral images are captured far from the outer space, while some hyperspectral images are obtained from the airplane or even in the lab. Due to the atmospheric effects etc., the different recording conditions would result in different spectral appearances. In short, the weakness of the spectral library brings side effects into this kind of methods.

More commonly, the endmembers are learned from the hyperspectral image itself to ensure the spectral coherence [9]—the unsupervised HU methods are preferred, where both the endmember and abundances are learned from the hyperspectral image. Specifically, unsupervised HUs can be categorized into two types: geometric methods [38–43] and statistical ones [44–48]. The geometric methods usually exploit the simplex to model the distribution of pixel spectra. The N-FINDR [49] and Vertex Component Analysis (VCA) [39] are the most benchmark geometric methods. For the N-FINDR, the endmembers are extracted by inflating a simplex inside the hyperspectral pixel space and treating the vertices of a simplex with the largest volume as endmembers [49]. The VCA [39] projects all the residual pixel onto a direction orthogonal to the simplex spanned by the chosen endmembers; the new endmember is identified as the extreme of the projection. Although these methods are simple and fast, they suffer from the requirement of pure pixels, which is usually unreliable in practice [12, 28, 50].

Accordingly, many statistical methods have been proposed for or applied to the HU problem, among which the Nonnegative Matrix Factorization (NMF) [51] and its extensions are the most popular. In (the following) Section III, we will review the NMF and its extensions in detail.

III. REVIEW OF THE STATE-OF-THE-ART HU METHODS

Notations. For clear modeling, the boldface uppercase letter (e.g. \( X \)) and lowercase case letter (x) are used to represent matrices and vectors respectively. Given a matrix \( X \triangleq \{X_{ln}\} \in \mathbb{R}^{L \times N} \), \( X_l \in \mathbb{R}^{1 \times N} \) is the \( l \)-th row vector and \( x_n \in \mathbb{R}^L \) denotes the \( n \)-th column vector. \( X_{ln} \) is the \((l, n)\)-th element in the matrix. \( X \geq 0 \) or \( X \in \mathbb{R}_+^{L \times N} \) represent a nonnegative matrix. The \( \ell_{2,1} \)-norm of matrices is defined as \( \|X\|_{2,1} = \sum_{l=1}^{L} \left( \sum_{n=1}^{N} X_{ln}^2 \right)^{1/2} \).

Formalization. In the HU modeling, the hyperspectral image with \( L \) channels (or bands) and \( N \) pixels, is represented by a nonnegative matrix \( X \triangleq [x_1, x_2, \cdots, x_N] \in \mathbb{R}_+^{L \times N} \). From the perspective of the linear mixture assumption, the goal of HU is to find two nonnegative matrices to well approximate \( X \) with their product:

\[
\min_{M, A} \text{loss} \{X, \tilde{X}\} + \lambda \Psi (A) + \alpha \Phi (M), \quad (2)
\]

s.t. \( \tilde{X} = MA, M \geq 0, A \geq 0 \),

where \( \tilde{X} \) is the approximation of the original hyperspectral image; \( M \triangleq [m_1, \cdots, m_K] \in \mathbb{R}_+^{L \times K} \) is the endmember matrix that consists of \( K \) pure pixel spectra and \( K \leq \min \{L, N\} \); \( A \triangleq [a_1, \cdots, a_N] \in \mathbb{R}_+^{K \times N} \) is the abundance matrix—the \( n \)-th column vector \( a_n \) contains all the \( K \) abundances at pixel \( x_n \); loss \( \{\cdot, \cdot\} \) is a loss function measuring the difference between two terms; \( \Psi (A) \) is some kind of constraints on the abundance; \( \Phi (M) \) is the constraint on the endmembers.

A. Nonnegative Matrix Factorization (i.e., NMF) [51, 52]

When the loss \( \{\cdot, \cdot\} \) is the Euclidean loss, the objective (2) becomes the standard NMF problem [14, 51, 52], which is commonly used in a wide range of applications including the HU [12, 14, 27, 28, 37, 53]. Since (2) is non-convex w.r.t. the two variables (i.e., \( M \) and \( A \)) together [51, 52], it is unrealistic to find global minima. Alternatively, Lee and Seung [51, 52] have proposed the multiplicative update rules as follows:

\[
M_{lk} \leftarrow \frac{M_{lk} (X A_l^\top)_{lk}}{(MA_l^\top)_{lk}}, \quad A_{kn} \leftarrow \frac{A_{kn} (M^\top X)_{kn}}{(M^\top MA_k)_{kn}}, \quad (3)
\]

which have been proved to be non-increasing. Apart from (3), there are other optimization algorithms to solve the problem (2), such as the active-set [54], the alternation nonnegative least squares [55] and the projected gradient descent [56].

Although NMF is well adapted to face analyses [57, 58] and documents clustering [59, 60], the objective function (2) is non-convex, naturally resulting in a large solution space [52]. Many extension methods have been proposed by employing suitable priors to restrict the solution space. For the HU task, the priors are imposed either on abundances [12, 53, 61] (cf. Section III-B) or on endmembers [27, 37] (cf. Section III-C).

B. The NMF extensions with constraints on the abundances

This section reviews the NMF extensions that impose constraints only on the abundance. That is, \( \Psi (A) \) is effective and \( \Phi (M) = 0 \); the updating rule for the endmember is given in (3) by default. Specifically, the sparse constraints [12, 14–16, 62, 63] and the spatial (like, manifold, graph) constraint [15, 61, 63, 64] are the most popular ones.

1) W-NMF [61] (or G-NMF [64]): The local neighborhood weight regularized NMF (W-NMF) [61] assumes that hyperspectral pixels are distributed on a manifold; the authors exploit appropriate weights in the local neighborhood to enhance the spectral unmixing. Specifically, W-NMF employs both the spectral and spatial information to construct the weight matrix \( W \in \mathbb{R}_+^{N \times N} \) [65], where \( W_{ij} \) is the similarity between the pixel \( i \) and \( j \). While, G-NMF [64] is a general NMF extension, which only employs the pixel feature to form the graph.

In W-NMF (or G-NMF), the constraint on the abundance is \( \Psi (A) = \text{Tr} (A \Lambda A^\top) \), where \( \Lambda = D - W \) is the graph Laplacian learned from the hyperspectral image, and \( D \) is the degree matrix whose diagonal elements are column sums of \( W \). The updating rule of abundances for the W-NMF is

\[
A_{kn} \leftarrow \frac{A_{kn} (M^\top Y + \lambda AD)_{kn}}{(M^\top MA_k + \lambda AD)^{1/2} k_n}, \quad (4)
\]

where \( \lambda \) is a balancing parameter for the Graph constraint.

2) Typical Sparsity constrained NMF (e.g., \( \ell_1 \)-NMF) [62] and \( \ell_{1/2} \)-NMF [12]): The sparsity constrained NMFs [2, 12, 14–16, 62] are the most successful methods for the HU task. Those methods assume that most hyperspectral pixels are mixed with parts of (not all) endmembers, and exploits all kinds of sparse constraints on the abundance. The \( \ell_1 \)-norm is
one of the most benchmark sparsity constraint. Considering it gives the \( \ell_1 \)-NMF \([62]\), where the constraint is \( \Psi (A) = \|A\|_1 = \sum_{k,n} |A_{kn}|. \) The updating rule for the abundance is

\[
A_{kn} \leftarrow \frac{A_{kn} (M^T Y)_{kn}}{\left(M^T MA + \lambda |A + \xi|^{-1/2}\right)_{kn}}, \quad (5)
\]

Although the \( \ell_1 \)-NMF makes sense, the lasso constraint \([66, 67]\) could not enforce further sparse when the full additivity constraint is used, limiting the effectiveness \([12]\). Accordingly, Dr. Qian proposed the state-of-the-art \( \ell_{1/2} \)-norm constrained NMF, where \( \Psi (A) = \|A\|_{1/2} \approx \sum_{k,n} |A_{kn} + \xi|^{1/2}, \) \( \xi \) is a small positive value to ensure the numerical condition. The updating rule for the abundance is given as

\[
A_{kn} \leftarrow \frac{A_{kn} (M^T Y)_{kn}}{\left(M^T MA + 0.5 \lambda (A + \xi)^{-1/2}\right)_{kn}}, \quad (6)
\]

where \( \lambda \) in Eqs. (5) and (6) are the balancing parameter for the sparsity constraints \([12, 62]\).

3) DgS-NMF \([14]\) and RRLbS \([16]\): To reduce the solution space, the state-of-the-art HU methods exploit various constraints on the abundances and on the endmembers. However, they generally employ an identical strength of constraints on all the factors, which may not meet the practical situation. Instead, Dr. Zhu \([14]\) observed that the mixed level of each pixel varies over image grids. Based on this prior, he proposed a novel method to learn a data-guided map (DgMap, i.e., \( h \in \mathbb{R}^N_x \)), which aims to describe the mixed level of each pixel. Through this DgMap, the \( \ell_p, (0 < p < 1) \) constraint is applied in an adaptive manner. For each pixel, the choice of \( p \) is tightly related to the corresponding value in the DgMap.

In DgS-NMF, the sparsity constraint on the abundance is

\[
\Psi (A) = \sum_{k,n} (A_{kn} + \xi)^{-1} h_{kn}, \quad H = 1_K h^T \in \mathbb{R}^{K \times N}, \quad h \in \mathbb{R}^N_x \text{ is the DgMap, and } \xi \text{ is a small positive value to ensure numerical conditions.}
\]

The updating rule for the abundance is

\[
A_{kn} \leftarrow \frac{A_{kn} (M^T Y)_{kn}}{\left(M^T MA + \lambda (1 - H) \circ (A + \xi)^{-1} H\right)_{kn}}, \quad (7)
\]

where \( \circ \) is the Hadamard product between matrices. Note that the \( \ell_1 \)-NMF and \( \ell_{1/2} \)-NMF are special cases of the DgS-NMF. Given a constant DgMap with all elements equal to zero, the DgS-NMF turns into the \( \ell_1 \)-NMF. Whereas if each element in the DgMap is equal to 1/2, the DgMap guided sparsity constraint turns into the \( \ell_{1/2} \) norm based sparse regularization.

DgS-NMF \([14]\) is an interesting method. However, a heuristic algorithm is proposed in \([14]\) to learn the DgMap, which is ineffective for the vast smooth areas in the image. It is expected that the more accurate DgMap constraint would bias the solution to the more satisfactory local minima. Besides, the state-of-the-art method generally ignores the badly degraded (i.e., outlier) channels in the hyperspectral image. To address the above two problems, a robust representation and learning-based sparsity (RRLbS) method is proposed in \([16]\) by emphasizing both robust representation and learning-based sparsity. Specifically, the loss \( \{\cdot, \cdot\} \) in (2) is set as the \( \ell_{2,1} \)-norm based loss. The new loss is better at preventing the outlier channels from dominating the objective. The constraint in RRLbS is \( \Psi (A) = \sum_{k,n} (A_{kn} + \xi)^{1-H_{kn}}, \) which is similar to the DgS-NMF. However, due to the robust loss function and the simultaneous learning process for DgMaps, the updating rule in RRLbS is quite different from that of DgS-NMF:

\[
M_{lk} \leftarrow \frac{M_{lk} (UXA^T)_{lk}}{(UMA A^T)_{lk}}, \quad (8)
\]

\[
A_{kn} \leftarrow \frac{A_{kn} (M^T UX)_{kn}}{\left(M^T UMA + \lambda (1 - H) \circ (A + \xi)^{-1} H\right)_{kn}}, \quad (9)
\]

where \( U \in \mathbb{R}^{\ell \times L} \) is a nonnegative diagonal matrix\(^4\). The most direct clue to estimate DgMap \( h \in \mathbb{R}^N_x \) is its crucial dependence upon abundances. Once getting the stable abundance \( A, h = \{h_n\}_{n=1}^N \) could be efficiently estimated as follows:

\[
H = 1_K h^T, \quad h_n = S(a_n), \quad \forall n \in \{1, 2, \cdots, N\} \quad (10)
\]

where \( S(a_n) \) is the Gini index \([16, 68]\) that measures the sparsity of column vectors in \( A \). The above updating process for \( \{M, A\} \) and \( h \) is iterated, which is expected to generate a sequence of ever improved estimates until convergences.

4) Structure (or Graph) and Sparsity constrained NMF (SS-NMF \([15]\) and GL-NMF \([63]\)): There are two methods \([15, 63]\) considering both the spatial (like graph) constraint and the sparsity constraint. In the Stuctured sparsed NMF (SS-NMF), the constraint is \( \Psi (A) = \text{Tr} (\text{ALA}^T) + \frac{\alpha}{2} \|A\|_1, \) where the graph Laplacian \( L \) is learned via a novel method that considers both the spectral and spatial information in the hyperspectral image. The updating rule for the abundance is as follows:

\[
A_{kn} \leftarrow \frac{A_{kn} (M^T X + \lambda AW)_{kn}}{\left(M^T MA + \lambda AD + \alpha\right)_{kn}}, \quad (11)
\]

In GL-NMF \([63]\), the constraint is \( \Psi (A) = \text{Tr} (\text{ALA}^T) + \frac{\alpha}{2} \|A\|_1, \) where the authors ignore the spatial information inherent in the hyperspectral image and only employ the spectral information to construct the graph Laplacian \( L \). The updating rule for the abundance is:

\[
A_{kn} \leftarrow \frac{A_{kn} (M^T X + \lambda AW)_{kn}}{\left(M^T MA + \lambda AD + 1/2\alpha A^{-1/2}\right)_{kn}}, \quad (12)
\]

5) Correntropy based NMF (CE-NMF) \([2]\): It is well known that the Euclidean loss is prone to outliers \([69, 70]\). Accordingly, Dr. Wang employed the correntropy metric to measure the reconstruction error; the \( \ell_1 \) norm based sparse constraint is considered, resulting in the new robust objective

\[
\min_{M \geq 0, A \geq 0} \sum_{l=1}^L \left[- \exp \left(-\frac{||X - (MA)_{l,1}||_2^2}{\sigma^2}\right)\right] + 2\lambda \|A\|_1. \quad (13)
\]

A half-quadratic optimization technique is proposed to convert the complex optimization problem (13) into an iteratively reweighted NMF problem. As a result, the optimization can

\(^4\)The \( l^\text{th} \) diagonal entry \( U_{ll} \) is set as \( U_{ll} = 1 / \left(2 \sqrt{\|MA - X\|_2^2 + \epsilon}\right) \), where \( \epsilon \) is typically set to \( 10^{-8} \) to avoid singular failures.
adaptively assign small weights to noisy channels. The updating rule is
\[
\tilde{M}_{lk} \leftarrow \frac{\tilde{M}_{lk} (X \Lambda^T)}{(M \Lambda A^T)_{lk}}, \quad A_{kn} \leftarrow \frac{A_{kn} (M^T \tilde{X})_{kn}}{(M^T M + \lambda)_{kn}}, \quad (14)
\]
where \(\tilde{X} = U^{1/2} X\), \(\tilde{M} = U^{1/2} M\) and \(U \in \mathbb{R}^{L \times L}\) is a diagonal matrix with the \(l^{th}\) element as \(U_{ll} = \exp \left( -\frac{\| \kappa - (M \Lambda A)^T \|_2^2}{\sigma^2} \right)\).

C. The NMF extensions with constraints on the endmembers

We review the NMF extensions that impose constraints only on the endmember. That is, \(\Phi(M)\) is effective and \(\Psi(A) = 0\).

1) Minimum Volume Constraint NMF (MVC-NMF) [27]:
The MVC-NMF combines the property of both the geometric methods and statistical methods. It aims to find the endmembers, which compose the minimum volume simplex that circumscribes the hyperspectral data scatters. Given the endmembers, the simplex volume is defined as
\[
\Phi(M) = \frac{1}{(K - 1)!} |\det((m_2 - m_1, \cdots, m_K - m_1))|. \quad (15)
\]
The standard Euclidean is used to measure the representation error. For the sake of easy optimization, the objective function is finalized as \(\min_{M, A} f(M, A)\) where
\[
f(M, A) = \frac{1}{2} \| X - MA \|_F^2 + \frac{\lambda}{2 (K - 1)!} \det^2 (1 + M M^T)^T
\]
where \(\tilde{M}\) is a low dimensional transform of \(M\), i.e., \(\tilde{M} = P^T (M - \mu 1_K)\); \(P \in \mathbb{R}^{L \times (K - 1)}\) is formed by the \(K - 1\) most significant principle components of hyperspectral data \(X\) and \(\mu\) is the mean of data \(X\). The updating rule is based on the projected gradient algorithm, which is given as follows:
\[
M^{(t+1)} = \max \left\{ 0, M^{(t)} - \alpha^{(t)} \tilde{\nabla}_M f(M^{(t)}, A^{(t)}) \right\} \quad (16)
\]
\[
A^{(t+1)} = \max \left\{ 0, A^{(t)} - \beta^{(t)} \tilde{\nabla}_A f(M^{(t+1)}, A^{(t)}) \right\} \quad (17)
\]
where \(\alpha^{(t)}\) and \(\beta^{(t)}\) are the learning rates. They can be fixed at small values or determined by the Armijo rule [27, 71].

2) Endmember Dissimilarity Constrained NMF (EDC-NMF) [37]: Inspired by the MVC-NMF [27], Dr. Wan proposed the Endmember Dissimilarity Constrained NMF (EDC-NMF). The core assumption is that due to the high spectral resolution of hyperspectral sensors, the endmember spectra should be smooth itself and different as much as possible from each other. The EDC constraint on the endmember is
\[
\Phi(M) = \sum_{i=1}^{K-1} \sum_{j=i+1}^{K} \| \nabla m_i - \nabla m_j \|_2^2, \quad (18)
\]
where \(\nabla m_i = [m_{2,i} - m_{1,i}, \cdots, m_{L,i} - m_{L-1,i}]\) is the gradient along the spectral dimension.

The multiplicative update rule is used to optimize the objective function, resulting in the algorithms for the abundances in (3) and for the endmembers as follows:
\[
M_{lk} \leftarrow \frac{M_{lk} (X A^T - 0.5 \lambda \nabla M \Phi(M))_{lk}}{(M A A^T)_{lk}}, \quad (19)
\]
where \(\nabla M \Phi(M) = 2HH^T M T\); \(T\) is a specially designed matrix, i.e., \(T = K I_K - 1_{K \times K}\). \(I_K\) is a \(K \times K\) identical matrix and \(1_{K \times K}\) is a matrix whose all elements are 1; \(H = [-p^T; I_{L-1}] + [-I_{L-1}; q^T]\), \(I_{L-1}\) is a \((L - 1) \times (L - 1)\) identical matrix, \(p = [1, 0, \cdots, 0]^T \in \mathbb{R}^{L - 1}\) and \(q = [0, \cdots, 0, 1]^T \in \mathbb{R}^{L - 1}\).

The derivative over endmembers, i.e., \(\nabla M \Phi(M)\), introduces negative values to the updating rule (19). To make up this problem, the negative elements in the endmember matrix are required to project to a given nonnegative value after each iteration. Consequently, the regularization parameter \(\lambda\) couldn’t be chosen freely, limiting the efficacy of EDC-NMF.

IV. HOW TO LABEL GROUND TRUTHS IN THE HU?

In the HU, the ground truth labeling consists of two parts: (1) the endmember labeling, and (2) the abundance labeling. Both parts are challenging. There is only one paper [36] briefly introducing the ground truth labeling on the HYDICE urban image. As a result, this is the first paper to provide a general method to label the hyperspectral images for the HU task.

A. The Endmember labeling

The endmember is the “pure” material in the image [1]. What is the “pure” material? From the chemistry perspective, the “pure” material (i.e. endmember) can be the pure substance. However, this ideal assumption makes the HU impossible to solve—there can be thousands of “pure” substance in the image scene. Such case leads to much more unknown variables than the observation itself in the hyperspectral image, which is obviously a severely underconstrained problem [72].

Alternatively, it is widely accepted that the notion of “pure material” is subjective and problem dependent [1], which is similar with the determination of classes in the hyperspectral classification task. Such setting requires lots of professional analyses and understandings of the image scene [73]. Taking the Cuprite image for example, it may take Earth scientists several weeks to analyze all the endmembers in the scene. The good news is that three factors make it easy to determine endmembers: (1) there are many HU papers publishing the endmember information of the common hyperspectral images, such as the HYDICE urban [12, 28, 36], Cuprite [12, 39, 63], Washington DC Mall [37], Moffett Field [45, 74, 75] etc; (2) for the images whose scenes are simple, it is easy to determine endmembers, e.g., the Samson [15], Japser Ridge [14, 16] and San Deigo Airport [76] etc. (3) the existing methods, like virtual dimensionality (VD) [77], is also helpful to determine the number of endmembers in the hyperspectral image.

After the determination of endmembers, we need their spectral signatures, which could be obtained as below: (1) They can be manually chosen from the hyperspectral image. In some images, the “pure” pixels have been explicitly marked in the published papers. For example, the endmembers for the Washington DC Mall image are explicitly marked in Fig. 13 in [37]. In the HYDICE urban, the location of the endmember pixel for “asphalt” is provided in the paper [36], which is then compared with the asphalt spectrum in the standard spectral library. In other hyperspectral images, like...
the Jasper Ridge, Samson, etc., we may have to select the endmembers from the image based on our understandings.

2). The endmember signature can be found in the USGS mineral spectral library. The Cuprite hyperspectral image is one of the typical hyperspectral image [37] in this type.

### B. The Abundance labeling

The abundance provides the composite percentage of all endmembers at each pixel (cf. Section III). Since the abundance is continuous between 0 and 1, it is impossible to manually label the abundances—Human can not tell the difference between 30% and 40% of an endmember at a pixel; besides, it is too tedious and labor intensive to label all pixels. Instead, we propose a learning method to label the abundance.

Given the determined $K$ endmembers, the abundance labeling becomes a least square problem with the nonnegative and full additivity constraints, which is indeed the supervised unmixing problem as reviewed in Section II. Such algorithm is far more easier than the unsupervised unmixing methods. The general quadratic programming algorithm with constraints (like the quadprog in Matlab) is helpful to estimate the abundances. Moreover, we may make use of the state-of-the-art unmixing methods reviewed in Section III-A and III-B by treating the endmembers as a fixed input rather than unknown variables and by only updating the abundances. The state-of-the-art unmixing methods are promising to achieve very good ground truths because of the different priors on the abundance.

### C. How to verify our labeling ground truth for the HU study?

What is a good ground truth for the given hyperspectral image? To verify it, we come up with two evaluation methods:

1). compare our labeling result with the ground truth posted on the published papers. It is necessary to ensure the high similarity between them. Note that, only the illustration of ground truths (not the ground truth itself) is posted in the paper. Accordingly, we can only compare them visually, not quantitatively. For example, we may compare the shape of the endmember signatures and the illustration of abundances.

2). check the labeling result based on our understanding of the hyperspectral image—the labeling result should be consistent with our understanding of the image.

3). it is reasonable to assume that similar pixels have similar abundances [15, 63]. Based on this assumption, we will click many pixels to evaluate their signature’s similarities, and then to check their abundance’s similarity. Those two kinds of similarities should be consistent as well.

4). for the hyperspectral images, like Cuprite, the ground truth endmembers are selected from the standard spectral library [12, 14, 37, 39, 63]. We don’t have to verify it.

We may have to try the different settings on the candidate endmembers, on the supervised unmixing algorithms for the abundance labeling, and on the hyper parameters etc (as shown in Algorithm 1), to make a better ground truth.

### D. How to transform the benchmark hyperspectral classification (HyC) datasets for the HU research?

Compared with the HU datasets, the datasets for the hyperspectral classification (HyC) task have three obvious advantages. (1) It is well known that the HyC task has much more benchmark datasets than the HU task. (2) Different from the HU datasets whose ground truths are mostly unavailable on the web, the HyC datasets and their ground truths are mostly available on the web. (3) the labeling process as well as the label itself for the HyC image are much easier to understand. Thus, it would be promising to transform the benchmark HyC datasets for the HU study. However, there is a big gap between the HyC task and the HU task, which makes the transformation process challenging. In this paper, we are the first to propose a promising method to accomplish this transformation task.

Our method consists of three process steps listed as below:

1). The number of classes should be consistent with the number of endmembers. Due to the supervised information involved in the training process, the HyC algorithms are generally able to handle more classes than the number of endmembers in the HU. Accordingly, we have to combine the similar classes to reduce the number of endmembers.

2). Candidate endmember labeling. We need the endmember signatures. Given the training labels for the HyC image, we may select a set of high quality pixels from the pixel set of each class. The endmember signature can be set as the mean or median signature of all the selected pixels in that class.

3). Abundance labeling. Suppose we are given a hyperspectral image $X \in \mathbb{R}^{L \times N}$, with $L$ channels and $N$ samples. The label of all pixels, i.e., $Y \in \mathbb{R}^{K \times N}$, could be predicted via the state-of-the-art classification algorithms [20, 78], where each column $y_n \in \mathbb{R}^K$, $\forall n \in \{1, \cdots, N\}$ is the 0-1 binary label vector. That is, if pixel $n$ is in the $k_{th}$ class, we define the label vector $y_n = [0, \cdots, 0, 1, 0, \cdots, 0]^{\top} \in \mathbb{R}^K$ with only the $k_{th}$ element equal to 1, all the others equal to 0 [79]. There are two ways to treat the candidate endmember obtained in the 2nd step. We may treat them as the final endmembers. Accordingly, the abundance labeling becomes the supervised unmixing problem with a special anchoring constraint as follows:

$$\min_{\mathbf{A}} \left\{ \mathbf{X}, \mathbf{X} \right\}_A + \lambda \mathbb{Y} \left( \mathbf{A} \right) \text{ s.t. } \mathbf{X} = \mathbf{M} \mathbf{A}, \mathbf{M} \geq 0, \mathbf{A} \geq 0,$$

where $\mathbb{Y} \left( \mathbf{A} \right) = \mathbb{Y} \left( \mathbf{A} \right) + \frac{\alpha}{2} \left\| \mathbf{A} - \mathbf{Y} \right\|_F^2$ is the constraint on the abundance; $\mathbb{Y} \left( \mathbf{A} \right)$ is the state-of-the-art sparse/spatial (or both) constraints reviewed in Section III-B; $\left\| \mathbf{A} - \mathbf{Y} \right\|_F$ is the anchoring constraint; its aim is to keeps the abundance around the high accuracy classification results.

We may also treat the candidate endmember, obtained in the 2nd step, as a good initialization of $\mathbf{M}$, and update both

---

**Algorithm 1 to label the ground truth for the HU study**

**Input** a hyperspectral image $\mathbf{X} \in \mathbb{R}^{L \times N}$, with $N$ pixels.

1. set the endmembers and the endmember number $K$.
2. repeat
3. endmember labeling via the method in Section IV-A.
4. abundance labeling via the method in Section IV-B.
5. verify the ground truth labeling.
6. until meeting the criteria in Section IV-C

**Output** the ground truth $\mathbf{M} \in \mathbb{R}^{K \times N}$ and $\mathbf{A} \in \mathbb{R}^{+ \times N}$.
Fig. 2. The three subimages selected from the Samson hyperspectral image. They are Samson#1, Samson#2 and Samson#3 respectively, as shown in Figs 2a, 2b and 2c. Besides, we provide the illustration of the ground truth endmembers and abundance maps for each subimage.

V. Summarization of the 15 Real Hyperspectral Images and their 18 Ground Truths

A. Samson’s three subimages

The Samson image is illustrated in Fig. 1a, where there are 952 × 952 pixels [14–16, 80]. Each pixel is observed at 156 bands covering the electromagnetic spectra from 401 nm to 889 nm. As a result, the spectral resolution is highly up to 3.13 nm. This data is in good condition and not degraded by the blank channels or badly noised channels. The original image is very large, which could be computationally expensive for the HU study. Accordingly, three subimages are selected due to that they contain large transitional areas which consist of lots of mixed pixels. We name the three ROIs as Samson#1, Samson#2 and Samson#3 respectively for convenience. Please refer to the Fig. 1a and Table I for their information, e.g., their locations, sizes, scenes, number of bands etc.

B. Jasper Ridge’s three subimages

Jasper Ridge is a popular hyperspectral image [81, 82], as shown in Fig. 1b. It was captured via the AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) sensor by the Jet Propulsion Laboratory (JPL). The ground sample distance (GSD) is 20 m [82]. There are 512 × 614 pixels in it. Each pixel is recorded at 224 electromagnetic bands ranging from 380 nm to 2,500 nm. The spectral resolution is highly up to 9.46 nm. Due to the dense water vapor and atmospheric effects, we remove the spectral bands 1–3, 108–112, 154–166 and 220–224, remaining 198 bands. This is a common pre-processing step if the HU methods do not focus on the robust learning [2, 16, 20, 78, 83, 84].

Since the full hyperspectral image, as shown in Fig. 1b, is too complex to get the ground truth, we consider three
HYDICE Urban and its two subimages

HYDICE Urban is one of the most widely used hyperspectral image for the HU research [2, 12, 15, 16, 28, 31, 36, 92–97]. It was recorded by the HYDICE (Hyperspectral Digital Image Collection Experiment) sensor in October 1995, whose scene is an urban area at Copperas Cove, TX, U.S. There are 307 × 307 pixels in this image; each pixel corresponds to a 2 × 2 m² region in the scene. The image has 210 spectral bands ranging from 400 nm to 2,500 nm, resulting in a very high spectral resolution of 10 nm. After removing the badly degraded bands 1–4, 76, 87, 101–111, 136–153 and 198–210 (due to dense water vapor and atmospheric effects), we remain 162 bands. There are three versions of ground truths for the HYDICE Urban hyperspectral image:

1. In the early HU papers [2, 12, 15, 16, 28, 31, 36, 94], it is widely accepted that there are four endmembers in the HYDICE Urban. They are “#1 Asphalt Road”, “#2 Grass”, “#3 Tree” and “#4 Roof”, as shown in Fig. 4a.

2. Recently, the analyse on Urban image become more and more precise. The “#1 Asphalt Road” is divided into “#1 Asphalt” and “#6 Soil” whereas the “#4 Roof” is divided into “#4 Roof₁” and “#5 Roof₂/shadow”. In other words, the number of endmembers rises to six [12, 97]; they are “#1 Asphalt”, “#2 Grass”, “#3 Tree”, “#4 Roof₁”, “#5 Roof₂/shado” and “#6 Soil”, as shown in Fig. 4c.

3. Apart from the above two versions of ground truths, we introduce a new ground truth that consists of 5 endmembers: “#1 Asphalt”, “#2 Grass”, “#3 Tree”, “#4 Roof” and “#5 Soil”, as shown in Fig. 4b. Compared with the first version, the “#1 Asphalt Road” is divided into “#1 Asphalt” and “#5 Soil”. Compared with the second version, we merge “#4 Roof₁” and “#5 Roof₂/shadow” into “#4 Roof”.

To make it more challenging, we select two subimages from the Urban image (cf. Fig. 1c). In them, the small objects, like small houses, vehicles and small grasslands etc., are the main scene, causing lots of transitional areas. Such case surely leads to lots of mixed pixels due to the low spatial resolution of hyperspectral sensors. The first subimage (i.e., Urban#1) has 160 × 168 pixels, whose first pixel is the (1, 1)-th pixel in the original image. The second subimage (i.e., Urban#2) starts from the (40, 1)-th pixel in the original image. It also has 160 × 168 pixels. There are six endmembers: “#1 Asphalt”, “#2 Grass”, “#3 Tree”, “#4 Roof₁”, “#5 Roof₂/shadow” and “#6 Soil” respectively. A new ground truth that consists of 5 endmembers: “#1 Asphalt”, “#2 Grass”, “#3 Tree”, “#4 Roof₁”, “#5 Roof₂/shadow” and “#6 Soil” respectively.
"#2 Grass", "#3 Tree", "#4 Roof", "#5 Roof2/shadow", and "6 Soil" in both Urban#1 (cf., Fig. 5a) and Urban#2 (cf., Fig. 5b).

D. Cuprite

Cuprite is the most benchmark and challenging hyperspectral image for the HU research [12, 14, 37, 39, 50, 53, 98–103], which is captured by the AVIRIS sensor. It covers a Cuprite area in Las Vegas, NV, U.S. There are 224 spectral bands in the Cuprite image, ranging from 370 nm to 2,480 nm. After removing the noisy bands (i.e., 1–2 and 221–224) and the water absorption bands (i.e., 104–113 and 148–167), it remains 188 bands. In this paper, a subimage of 250 × 190 pixels is considered, which is widely used in the state-of-the-art HU papers [12, 14, 39, 50]. Please refer to Fig. 1d for the illustration of its position, image size and scene etc.

In this subimage, there are 14 types of minerals (or endmembers), whose spectra can be obtained from the ENVI software. There are very minor differences between the variants of the same type of mineral, for example Kaolinite$_1$ and Kaolinite$_3$ are very similar in signature. The researchers have their own thoughts, resulting in different versions of ground truths. In [39], there are 14 endmembers; while there are 10 endmembers in [12]; then Dr. Lu hold that there are 12 endmembers in the Cuprite. Here we agree with Dr. Lu’s setting. Please refer to Fig. 6 for the illustration of the 12 endmembers. Because there are small differences in the setting of endmembers among the papers [12, 37, 39, 50, 53], the results of the state-of-the-art methods in those papers might be different from each other.

E. Moffett Field

The Moffett Field hyperspectral image is illustrated in Fig. 1e. It was captured via the AVIRIS sensor by the JPL in 1997 and was used in a number of state-of-the-art papers [104–109]. There are 512 × 614 pixels and 224 spectral bands, ranging the electromagnetic spectra from 400 nm to 2500 nm. The spectral resolution is up to 9.38 nm. The full image can be found in [110]. It consists of a large area of water (that appears in dark pixel at the top of the image) and a coastal area composed of trees and soil. Due to the water vapor and atmospheric effects, we remove the noisy spectral bands. After this process, there remains 196 bands.

The original image is very large in size and too complex to analyze because the spatial resolution is low and the object is very small in the scene. In this paper, we select two subimages for the HU study. Please refer to Fig. 1e for their locations, image sizes and scenes etc. The first subimage (named Moffett Field#1) consists of 60 × 60 pixels, whose first (top-left) pixel is the (54, 172)-th pixel in the original image. The second subimage (named Moffett Field#2) also has 60 × 60 pixels. Its first pixel is the (62, 146)-th pixel in the full image. A similar subimage with Moffett Field#1 was used in [109, 111], where the authors hold that there are three endmembers. In this paper, we agree with their setting. The three endmembers for both Moffett Field#1 and Moffett Field#2 are “#1 Soil”, “#2 Tree” and “#3 Water”, respectively, as shown in Figs. 7a and 7b.

F. San Diego Airport

San Diego Airport is a popular airborne hyperspectral image collected by the AVIRIS sensor [76, 112–114]. There are 400 × 400 pixels and 224 spectral bands, as illustrated in Fig. 1f. After removing the noisy bands with the water vapor and atmospheric effects, including the bands 1–6, 33–35, 97, 107–113, 153–166, and 221–224, there remains 189 spectral bands. In this paper, we select a subimage for the HU study. It has 160 × 140 pixels. The top-left pixel is the (241, 261)-th pixel in the original image, as shown in Fig. 1f. There are two endmember settings: (1) the first setting has 4 endmembers, i.e., “#1 Grass”, “#2 Tree”, “#3 Roof” and “#4 Ground & Road”, as shown in Fig. 8a; (2) the second setting has 5 endmembers, i.e., “#1 Grass”, “#2 Tree”, “#3 Roof”, “#4 Ground & Road” and “#5 Other”, as shown in Fig. 8b.
The Washington DC Mall (WDM)

The Washington DC Mall (i.e., WDM), an airborne hyperspectral image flightline over the Washington DC Mall (cf. Fig. 9g), is used to verify a lot of research works [28, 31, 37, 102, 115, 116]. It was collected by the HYDICE sensor. The full image can be downloaded from the webpage. There are 1208 × 307 pixels and 210 spectral channels, covering the electromagnetic spectra from 400 nm to 2400 nm. The spectral resolution is up to 9.52 nm. After removing the bands with water vapor and atmospheric effects, including 103–106, 138–148, and 207–210, there remains 191 bands [102].

In this paper, we select two subimages for the HU study, i.e., WDM#1 and WDM#2 respectively. Please refer to Fig. 9g and Table I for their information like locations and scenes etc. The first subimage (i.e. WDM#1) has 150 × 150 pixels, whose first pixel starts from the (549, 160)-th pixel in the original image. WDM#1 has been used in [37], where they provide the rough locations of the six endmembers; however, they did not share the ground truth on the Internet. The six endmembers are “#1 Roof”, “#2 Grass”, “#3 Water”, “#4 Tree”, “#5 Trail” and “#6 Road” respectively, as shown in Fig. 9a. The second subimage (i.e. WDM#2) has 180 × 160 pixels, whose left-top pixel is the (945, 90)-th pixel in the original image. The six endmembers are “#1 Road”, “#2 Grass”, “#3 Trail”, “#4 Water”, “#5 Tree” and “#6 Roof” respectively, as shown in Fig. 9b.

VI. THE SIMULATED HYPERSPECTRAL IMAGE

This section provides the method to generate a complex synthetic image [2, 12, 27–32] for the HU research. It consists of two steps. First, 15 true spectra are chosen from the United States Geological Survey (USGS) digital spectral library as the candidate endmembers. Their signature are shown in Fig. 10. The first K spectra are selected as the true endmember, where 2 ≤ K ≤ 15 is the number of endmembers.

Second, we use the method in [2, 12, 27, 28, 30, 32] to generate the abundance maps, which has the following steps:

1). We divide an image of z 2 × z 2 (z ∈ Z+) pixels into z × z regions. The default setting is z = 8.
2). Each region is filled up with a type of ground cover, which means the abundance of that endmember is 1 in the whole region. Note that, this process is randomly implemented.
3). To generate the mixed pixels, we use the (z + 1) × (z + 1) spatial low-pass filter. Such process causes large transitional areas, where there are lots of mixed pixels.
4). To remove pure pixels and generate sparse abundances, we replace the pixels whose largest abundance is larger than 0.8 with a mixture of two endmembers with equal abundances.

Now both endmembers M and abundances A have been generated. Based on them, a perfect hyperspectral image is created as Y = MA. To simulate a more real remote sensing image that contains complex noises, we add the zero-mean Gaussian (or some other) noise to the generate the image data. The Signal-to-Noise Ratio (SNR) is defined as follows:

\[
\text{SNR} = 10 \log_{10} \left\{ \frac{\mathbb{E} [y^\top y]}{\mathbb{E} [n^\top n]} \right\}.
\]

With the above synthetic hyperspectral image, we can verify the HU algorithms from the following four perspectives:

- Robustness analysis to the noise corruptions, e.g., SNR∈ [∞, 35, ⋯, 10 dB].
- Generalization to the pixel number, which is set by z.
- Robustness to the endmember number \{K: 2 ≤ K ≤ 15\}.
- Generalization to the number of bands in the image.

We will provide the full codes for this simulated dataset.

VII. EXPERIMENTS RESULTS OF 5 BENCHMARK HU METHODS ON THE 18 HYPERSPECTRAL IMAGES

In this section, we compare 5 benchmarks HU methods on the 15 hyperspectral images summarized in Section V. The selected 5 methods are widely compared against the state-of-the-art HU methods. The aim is to provide benchmark HU results, saving other’s effort to evaluate their new HU methods.

A. Five Benchmark Methods Compared in this Section

1). Vertex Component Analysis [39] (VCA) is a classic geometric method. The code is available on http://www.lx.it.pt/bioucas/code.htm.

2). Nonnegative Matrix Factorization [51] (NMF) is a benchmark statistical method. The code is obtained from http://www.cs.helsinki.fi/u/ophoyer/software.html.

3). Projection Pursuit Regression (PPR) [103, 104] is a robust technique. The code is available on http://www.icsi.berkeley.edu/˜bergamo/PPR/

4). Box Least Mean Squares (BLMS) [21, 22] is a hypothesis seeking model. The code is available on http://www.icsi.berkeley.edu/˜bergamo/BLMS/

5). Spectral Angle Mapping (SAM) [20] is a spectral feature based method. The code is available on http://www.icsi.berkeley.edu/˜bergamo/SAM/

6). This website is https://engineering.purdue.edu/~biehl/MultiSpec/hyperspectral.html

7). The Signal-to-Noise Ratio (SNR) is defined as follows:

\[
\text{SNR} = 10 \log_{10} \left\{ \frac{\mathbb{E} [y^\top y]}{\mathbb{E} [n^\top n]} \right\}.
\]
3). Nonnegative sparse coding [62] ($\ell_1$-NMF) is a classic sparse regularized NMF method. The code is downloadable from http://www.cs.helsinki.fi/u/phoyer/software.html.

4). $\ell_{1/2}$ sparsity-constrained NMF [12] ($\ell_{1/2}$-NMF) is a state-of-the-art method that could get sparser results than $\ell_1$-NMF. Since the code is unavailable, we implement it.

5). Graph regularized NMF [61,64] (G-NMF) is an interesting algorithm that transfers graph information inherent in the hyperspectral image into the abundances space. We will share the code for all the above five algorithms.

### B. Experiment Settings and Experiment Results

Generally, there are one or two hyper-parameters in the 5 benchmark methods. We use the coarse to fine grid search method to find the best hyper-parameters [14–16]. The initialization of the endmembers $M$ and abundances $A$ is very important to the final HU results [37]. We employ the same benchmark algorithm, i.e., VCA [39], to initialize $M$ and $A$ for the five methods on all the 15 datasets. There are two advantages: (1) it is helpful to generate the reproducible results by the VCA; (2) VCA is really simple, effective and solid.

Two evaluation metrics are used to obtain the quantitative comparison results: (1) the Spectral Angle Distance (SAD) [14–16,53,63] is used to verify the performance of the estimated endmembers; (2) the Root Mean Square Error (RMSE) [12,14–16,117] is for the evaluation of the estimated abundances. To get a valid results, each experiment is repeated 50 times and the mean result is provided in the Tables II, III, IV, V, VI, VII, VIII, IX. In short, we summarize the tables that display the results on all the 15 datasets:

1). Table II summarizes the experiment results of all the five benchmark algorithms (reviewed in Section VII-A). There are three tables, displaying the HU results on the three subimages, i.e., Samson#1, Samson#2 and Samson#3 respectively.

2). In Table III, the HU results on the Cuprite image are summarized. Note that the ground truth abundance is not available; only the endmember results are compared.

3). Table IV illustrates the HU results of 5 benchmark algorithms on the full HYDICE Urban image. There are 3 versions of ground truths for the Urban images. Accordingly, we summarize 3 versions of HU results in the 3 sub-tables.

4). In Table V, the HU results on the two subimages, Urban#1 and Urban#2, are summarized. There are six endmembers as shown in the Table V.

5). In Table VI, the experiment results on the 3 subimages from the JasperRidge is displayed. The 3 subimages are JasperRidge#1, JasperRidge#2 and JasperRidge#3.

6). Table VII illustrates the HU results on the 2 subimages selected from Moffett Field. They are Moffett Field#1 and Moffett Field#2 respectively as shown in the Table.

7). Table VIII illustrates the HU results on the 2 subimages selected from the Washington DC Mall. They are WDM#1 and WDM#2 respectively.

8). Table IX displays the HU results on the two versions of ground truths on the San Diego Airport dataset.

### VIII. Conclusion

Hyperspectral unmixing (HU) is an important preprocessing step for a great number of hyperspectral applications. To promote the HU research, we accomplished this paper by emphasizing on the generation of standard HU datasets (with ground truths) and the summarization of benchmark results of the state-of-the-art HU algorithms. Specifically, this is the first paper to propose a general labeling method for the HU. Via it, we summarized and labeled up to 15 hyperspectral images. To further enrich the HU datasets, an interesting method has been proposed to transform the hyperspectral classification...
datasets for the HU research. Besides, we reviewed and implemented a widely accepted algorithm to generate a very complex simulated hyperspectral image for the HU study. Such synthetic dataset is helpful to verify the HU algorithms under different conditions.

To summarize the benchmark HU results, we reviewed up to 10 state-of-the-art HU algorithms, and selected the five most benchmark HU algorithms for the comparison. Those HU algorithms are compared on the 15 hyperspectral images. To the best of our knowledge, this is the first paper to compare the benchmark algorithms on so many real hyperspectral images. The experient results on this paper may provide a valid baseline for the evaluation of new HU algorithms.

![Table IV](image)

**Table IV**

The performances of five benchmark HU algorithms on the HYDICE Urban image, which has three versions of ground truths.

![Table V](image)

**Table V**

The performances of five benchmark HU algorithms on the Urban1 and Urban2 hyperspectral sub-images.

---

R. Kawakami, J. Wright, Y.-W. Tai, Y. Matsuhashi, M. Ben-Ezra, and K. Ikeuchi, “High-resolution hyperspectral imaging via matrix factorization,” IEEE Conference on Computer Vision and Pattern Recognition, vol. 0, pp. 2329–2336, 2011.

C. Lanaras, E. Baltsavias, and K. Schindler, “Hyperspectral super-resolution by coupled spectral unmixing,” in IEEE International Conference on Computer Vision, 2015, pp. 3586–3594.

W. Dong, F. Pu, G. Shi, X. Cao, J. Wu, G. Li, and X. Li, “Hyperspectral image super-resolution via non-negative structured sparse representation,” IEEE Transactions on Image Processing, vol. 24, no. 2, pp. 4241–4247, 2015.

C. Ma, X. Cao, X. Tong, Q. Dai, and S. Lin, “Acquisition of high spatial and spectral resolution video with a hybrid camera system,” International journal of computer vision, vol. 110, no. 2, pp. 141–155, 2014.

N. Akhtar, F. Shaiaft, and A. Mian, “Sparse spatio-spectral representation for hyperspectral image super-resolution,” in European Conference on Computer Vision. Springer, 2014, pp. 63–78.

[1] J. Boucauats-Dias and A. Plaza, “Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 5, no. 2, pp. 534 –579, april 2012.

[2] Y. Wang, C. Pan, S. Xiang, and F. Zhu, “Robust hyperspectral unmixing with currientropy-based metric,” IEEE Transactions on Image Processing, vol. 24, no. 11, pp. 4027–4040, 2015.

[3] R. Kawakami, J. Wright, Y.-W. Tai, Y. Matsuhashi, M. Ben-Ezra, and K. Ikeuchi, “High-resolution hyperspectral imaging via matrix factorization,” IEEE Conference on Computer Vision and Pattern Recognition, vol. 0, pp. 2329–2336, 2011.

[4] C. Lanaras, E. Baltsavias, and K. Schindler, “Hyperspectral super-resolution by coupled spectral unmixing,” in IEEE International Conference on Computer Vision, 2015, pp. 3586–3594.

[5] W. Dong, F. Pu, G. Shi, X. Cao, J. Wu, G. Li, and X. Li, “Hyperspectral image super-resolution via non-negative structured sparse representation,” IEEE Transactions on Image Processing, vol. 24, no. 2, pp. 4241–4247, 2015.

[6] C. Ma, X. Cao, X. Tong, Q. Dai, and S. Lin, “Acquisition of high spatial and spectral resolution video with a hybrid camera system,” International journal of computer vision, vol. 110, no. 2, pp. 141–155, 2014.

[7] N. Akhtar, F. Shaiaft, and A. Mian, “Sparse spatio-spectral representation for hyperspectral image super-resolution,” in European Conference on Computer Vision. Springer, 2014, pp. 63–78.

---

REFERENCES

---

[1] J. Boucauats-Dias and A. Plaza, “Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 5, no. 2, pp. 534 –579, april 2012.

[2] Y. Wang, C. Pan, S. Xiang, and F. Zhu, “Robust hyperspectral unmixing with currientropy-based metric,” IEEE Transactions on Image Processing, vol. 24, no. 11, pp. 4027–4040, 2015.

[3] R. Kawakami, J. Wright, Y.-W. Tai, Y. Matsuhashi, M. Ben-Ezra, and K. Ikeuchi, “High-resolution hyperspectral imaging via matrix factorization,” IEEE Conference on Computer Vision and Pattern Recognition, vol. 0, pp. 2329–2336, 2011.

[4] C. Lanaras, E. Baltsavias, and K. Schindler, “Hyperspectral super-resolution by coupled spectral unmixing,” in IEEE International Conference on Computer Vision, 2015, pp. 3586–3594.

[5] W. Dong, F. Pu, G. Shi, X. Cao, J. Wu, G. Li, and X. Li, “Hyperspectral image super-resolution via non-negative structured sparse representation,” IEEE Transactions on Image Processing, vol. 24, no. 2, pp. 4241–4247, 2015.

[6] C. Ma, X. Cao, X. Tong, Q. Dai, and S. Lin, “Acquisition of high spatial and spectral resolution video with a hybrid camera system,” International journal of computer vision, vol. 110, no. 2, pp. 141–155, 2014.

[7] N. Akhtar, F. Shaiaft, and A. Mian, “Sparse spatio-spectral representation for hyperspectral image super-resolution,” in European Conference on Computer Vision. Springer, 2014, pp. 63–78.

---

REFERENCES

---

[1] J. Boucauats-Dias and A. Plaza, “Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 5, no. 2, pp. 534 –579, april 2012.

[2] Y. Wang, C. Pan, S. Xiang, and F. Zhu, “Robust hyperspectral unmixing with currientropy-based metric,” IEEE Transactions on Image Processing, vol. 24, no. 11, pp. 4027–4040, 2015.
### TABLE VI
THE HU RESULTS OF FIVE BENCHMARK HU ALGORITHMS ON THE JASPER RIDGE#1, JASPER RIDGE#2 AND JASPER RIDGE#3 hyperspectral images.

| Endmembers | Spectral Angle Distance (SAD) | Root Mean Square Error (RMSE) |
|-------------|-----------------------------|-----------------------------|
|             | VCA                         | NMF | \(\ell_1\)-NMF | \(\ell_2\)-NMF | G-NMF | \(\ell_1\)-NMF | \(\ell_2\)-NMF | G-NMF |
| Jasper Ridge #1 |                             |     |               |               |      |               |               |      |
| #1 Tree     | 0.3381                      | 0.1325 | 0.0553 | 0.0458 | 0.0743 | 0.2651 | 0.1327 | 0.0928 | 0.1032 | 0.0752 |
| #2 Water    | 0.2573                      | 0.2823 | 0.0300 | 0.0303 | 0.2923 | 0.2711 | 0.1384 | 0.0714 | 0.0911 | 0.1366 |
| #3 Soil     | 0.3062                      | 0.2125 | 0.1318 | 0.0534 | 0.2906 | 0.2567 | 0.2221 | 0.2256 | 0.2631 | 0.1837 |
| #4 Road     | 0.6053                      | 0.4843 | 0.6522 | 0.5525 | 0.5411 | 0.2361 | 0.1876 | 0.2064 | 0.2389 | 0.1908 |
| #5 Other    | 0.7156                      | 0.6463 | 1.0503 | 0.8187 | 0.6742 | 0.2842 | 0.1601 | 0.1357 | 0.1451 | 0.1669 |
| Avg.        | 0.4445                      | 0.3516 | 0.3805 | 0.3001 | 0.3565 | 0.2626 | 0.1682 | 0.1520 | 0.1663 | 0.1506 |
| Jasper Ridge #2 |                             |     |               |               |      |               |               |      |
| #1 Tree     | 0.2565                      | 0.2130 | 0.0602 | 0.0601 | 0.2781 | 0.3628 | 0.1402 | 0.0836 | 0.0786 | 0.1740 |
| #2 Water    | 0.2474                      | 0.2001 | 0.3815 | 0.1682 | 0.2530 | 0.3151 | 0.1106 | 0.0660 | 0.1031 | 0.1375 |
| #3 Soil     | 0.3584                      | 0.1569 | 0.0898 | 0.0506 | 0.3246 | 0.2936 | 0.2557 | 0.2463 | 0.2679 | 0.2840 |
| #4 Road     | 0.5489                      | 0.3522 | 0.4118 | 0.3670 | 0.2776 | 0.2829 | 0.2450 | 0.2344 | 0.2737 | 0.2279 |
| Avg.        | 0.3528                      | 0.2305 | 0.2378 | 0.1567 | 0.2833 | 0.3046 | 0.1879 | 0.1526 | 0.1789 | 0.2058 |
| Jasper Ridge #3 |                             |     |               |               |      |               |               |      |
| #1 Tree     | 0.1168                      | 0.0992 | 0.1069 | 0.1081 | 0.1117 | 0.1522 | 0.1080 | 0.1182 | 0.1017 | 0.1104 |
| #2 Water    | 0.2763                      | 0.3124 | 0.0513 | 0.0474 | 0.3369 | 0.1467 | 0.1248 | 0.0418 | 0.0651 | 0.1298 |
| #3 Soil     | 0.5364                      | 0.3748 | 0.3525 | 0.2419 | 0.3414 | 0.2475 | 0.2663 | 0.2536 | 0.2289 | 0.2247 |
| #4 Road     | 0.4235                      | 0.2997 | 0.3906 | 0.4098 | 0.2200 | 0.2105 | 0.2163 | 0.1707 | 0.1834 | 0.1688 |
| Avg.        | 0.3833                      | 0.2715 | 0.2228 | 0.2018 | 0.2525 | 0.1892 | 0.1769 | 0.1461 | 0.1448 | 0.1599 |

### TABLE VII
THE PERFORMANCES OF FIVE BENCHMARK HU ALGORITHMS ON THE MOFFETT FIELD#1 AND MOFFETT FIELD#2 hyperspectral images.

| Endmembers | Spectral Angle Distance (SAD) | Root Mean Square Error (RMSE) |
|-------------|-----------------------------|-----------------------------|
|             | VCA                         | NMF | \(\ell_1\)-NMF | \(\ell_2\)-NMF | G-NMF | \(\ell_1\)-NMF | \(\ell_2\)-NMF | G-NMF |
| Moffett Field #1 |                             |     |               |               |      |               |               |      |
| #1 Soil     | 0.0760                      | 0.0875 | 0.0884 | 0.0717 | 0.1006 | 0.2492 | 0.0923 | 0.0912 | 0.0943 | 0.1085 |
| #2 Tree     | 0.0418                      | 0.0432 | 0.0332 | 0.0357 | 0.0477 | 0.0992 | 0.0540 | 0.0595 | 0.0590 | 0.0842 |
| #3 Water    | 0.3946                      | 0.3528 | 0.3471 | 0.3432 | 0.3697 | 0.2751 | 0.1006 | 0.0790 | 0.0839 | 0.1219 |
| Avg.        | 0.1708                      | 0.1786 | 0.1563 | 0.1502 | 0.1727 | 0.2078 | 0.0763 | 0.0823 | 0.0791 | 0.1049 |
| Moffett Field #2 |                             |     |               |               |      |               |               |      |
| #1 Soil     | 0.0509                      | 0.0598 | 0.0560 | 0.0528 | 0.0402 | 0.1120 | 0.1685 | 0.0856 | 0.0703 | 0.1701 |
| #2 Tree     | 0.0406                      | 0.0419 | 0.0684 | 0.0592 | 0.0422 | 0.0403 | 0.0600 | 0.0286 | 0.0417 | 0.0595 |
| #3 Water    | 0.2793                      | 0.3224 | 0.0555 | 0.0726 | 0.3247 | 0.1072 | 0.1771 | 0.0348 | 0.0501 | 0.1779 |
| Avg.        | 0.1236                      | 0.1347 | 0.0553 | 0.0525 | 0.1357 | 0.0865 | 0.1352 | 0.0364 | 0.0540 | 0.1359 |
TABLE VIII
THE HU RESULTS OF FIVE BENCHMARK ALGORITHMS ON THE WDM#1 AND WDM#2 SUBIMAGES. WDM DENOTES WASHINGTON DC MALL.

| Endmembers | Spectral Angle Distance (SAD) | Root Mean Square Error (RMSE) |
|-------------|------------------------------|------------------------------|
|             | VCA NMF | ℓ₁-NMF | ℓ₂-NMF | G-NMF | VCA NMF | ℓ₁-NMF | ℓ₂-NMF | G-NMF |
| #1 Roof     | 0.1817  | 0.2343  | 0.2481  | 0.2210 | 0.21093 | 0.12021 | 0.16432 | 0.16577 | 0.16744 | 0.17293 |
| #2 Grass    | 0.2477  | 0.16813 | 0.17883 | 0.17013 | 0.17285 | 0.18837  | 0.19286 | 0.19112 | 0.19147 | 0.19202 |
| #3 Water    | 0.18327 | 0.17505 | 0.16717 | 0.17077 | 0.17099 | 0.33159  | 0.20351 | 0.20051 | 0.20236 | 0.19769 |
| #4 Tree     | 0.13241 | 0.15310 | 0.15585 | 0.14407 | 0.15898 | 0.15137  | 0.14405 | 0.14417 | 0.14449 | 0.14406 |
| #5 Tran.    | 0.17445 | 0.26442 | 0.26161 | 0.26343 | 0.27072 | 0.30302  | 0.26752 | 0.27690 | 0.28636 | 0.28636 |
| #6 Shd      | 0.42313 | 0.35212 | 0.34244 | 0.37354 | 0.36841 | 0.36034  | 0.29089 | 0.29570 | 0.30967 | 0.29885 |
| Avg.        | 0.22380 | 0.22437 | 0.22605 | 0.22384 | 0.22481 | 0.24248  | 0.21209 | 0.21212 | 0.21528 | 0.21532 |

| Endmembers | Spectral Angle Distance (SAD) | Root Mean Square Error (RMSE) |
|-------------|------------------------------|------------------------------|
|             | VCA NMF | ℓ₁-NMF | ℓ₂-NMF | G-NMF | VCA NMF | ℓ₁-NMF | ℓ₂-NMF | G-NMF |
| #1 Asphalt  | 0.17554 | 0.14735 | 0.07486 | 0.11782 | 1.44942 | 0.39304  | 0.39164 | 0.39048 | 0.39034 | 0.39048 |
| #2 Grass    | 0.98158 | 0.69179 | 0.11641 | 0.14638 | 0.63077 | 0.45647  | 0.46033 | 0.39328 | 0.44229 | 0.46026 |
| #3 Roof     | 0.15295 | 0.04880 | 0.03986 | 0.03278 | 0.04388 | 0.21752  | 0.18320 | 0.23490 | 0.21929 | 0.18256 |
| #4 Roof2/   | 0.18483 | 0.09550 | 0.11804 | 0.31077 | 0.8978  | 0.24413  | 0.25255 | 0.25370 | 0.26166 | 0.25228 |
| #5 Shadow   | 0.22996 | 0.73903 | 0.80175 | 0.93171 | 0.73604 | 0.26711  | 0.18963 | 0.16798 | 0.16017 | 0.18300 |
| #6 Soil     | 0.57558 | 0.63391 | 0.46868 | 0.55895 | 0.63210 | 0.31408  | 0.30186 | 0.26696 | 0.29722 | 0.30035 |

TABLE IX
THE PERFORMANCES OF FIVE BENCHMARK HU ALGORITHMS ON THE SAN DIEGO AIRPORT IMAGE, WHICH HAS TWO VERSIONS OF GROUND TRUTHS.

| V1.4 endmembers | Spectral Angle Distance (SAD) | Root Mean Square Error (RMSE) |
|-----------------|------------------------------|------------------------------|
|                 | VCA NMF | ℓ₁-NMF | ℓ₂-NMF | G-NMF | VCA NMF | ℓ₁-NMF | ℓ₂-NMF | G-NMF |
| #1 Grass        | 0.01767 | 0.07338 | 0.1811 | 0.0847 | 0.7763 | 0.3164  | 0.3403 | 0.2675 | 0.2692 | 0.3255 |
| #2 Tree         | 0.18914 | 0.1904  | 0.7613 | 0.8001 | 0.1709 | 0.2257  | 0.2338 | 0.2672 | 0.2851 | 0.2336 |
| #3 Roof         | 0.38448 | 0.5240  | 0.4924 | 0.4608 | 0.5224 | 0.4508  | 0.3361 | 0.3101 | 0.3491 | 0.3348 |
| #4 Ground       | 0.2187 | 0.2543  | 0.1313 | 0.0609 | 0.2487 | 0.3719  | 0.2593 | 0.2283 | 0.2964 | 0.2562 |
| #5 Other        | 0.46000 | 0.4356  | 0.3915 | 0.3423 | 0.4296 | 0.3412  | 0.2924 | 0.2683 | 0.2999 | 0.2875 |
| Avg.            | 0.18604 | 0.3528  | 0.2560 | 0.2597 | 0.2544 | 0.3249  | 0.2916 | 0.2804 | 0.2754 | 0.2906 |

| V2.5 endmembers | Spectral Angle Distance (SAD) | Root Mean Square Error (RMSE) |
|-----------------|------------------------------|------------------------------|
|                 | VCA NMF | ℓ₁-NMF | ℓ₂-NMF | G-NMF | VCA NMF | ℓ₁-NMF | ℓ₂-NMF | G-NMF |
| #1 Grass        | 0.48060 | 0.3582  | 0.3560 | 0.2597 | 0.2544 | 0.3249  | 0.2916 | 0.2804 | 0.2754 | 0.2906 |
| #2 Tree         | 0.15922 | 0.1361  | 0.1439 | 0.1624 | 0.1430 | 0.2209  | 0.2267 | 0.2212 | 0.2176 | 0.2207 |
| #3 Roof         | 0.6239  | 0.5609  | 0.6293 | 0.6205 | 0.6409 | 0.3299  | 0.2538 | 0.2626 | 0.2715 | 0.2563 |
| #4 Ground       | 0.2509  | 0.2276  | 0.2200 | 0.2047 | 0.2501 | 0.3443  | 0.2544 | 0.2622 | 0.2566 | 0.2581 |
| #5 Other        | 0.4856  | 0.5421  | 0.5586 | 0.5681 | 0.6157 | 0.3382  | 0.2665 | 0.2735 | 0.2781 | 0.2696 |
| Avg.            | 0.40000 | 0.3790  | 0.3815 | 0.3739 | 0.3808 | 0.3117  | 0.2586 | 0.2632 | 0.2599 | 0.2591 |
