UNIVERSIDAD **DE EXTREMADURA**

TESIS DOCTORAL

Procesamiento Eficiente y Profundo de Imágenes Hiperespectrales y Aplicación en Tareas de Desmezclado Espectral

Xuanwen Tao

Programa de doctorado en Tecnologías Informáticas

Con la conformidad:

Dr. Antonio Plaza Miguel   *Dr. Juan Mario Haut Hurtado*

2023

Acknowledgements

I am grateful to the China Scholarship Council (CSC) for providing me with a scholarship and the opportunity to study in Spain. The support of the CSC is the biggest guarantee in my doctoral study life. I am also really grateful to Spain, my second home. During these four years, Spain has brought me not only countless surprises and unparalleled beauty, but also the warmth, help, and selfless kindness of the Spanish people.

I would like to take this opportunity to express my deep gratitude to my supervisor Prof. Antonio Plaza for his invaluable guidance and profound knowledge. Even though he was very busy, he patiently helped me revise each paper. It is my great honor and pleasure to study under his guidance and supervision. He helped me solve various problems when I was living abroad for the first time. I have benefited immensely from his abilities and character, which I will treasure in my whole life. My gratitude to him knows no bounds. I hope to have the opportunity to continue studying under his guidance in the future.

I wish to express my sincere gratitude to my supervisor Juan Mario Haut and my best Spanish friend Mercedes. In the first two years, I was the only Chinese in the lab. They helped me revise my papers and took me to play, relax and eat delicious food. They not only guided and inspired me in my research but also helped me with various trivial matters. Their selfless help and kindness warmed me like sunshine. The time I spent with them will be my most valuable and fondest memories.

I am really grateful to Zhaoyue, you are like your name, the most beautiful moon. I will never forget the days we opened our hearts, shared our lives, and discussed every scientific idea passionately. I also express my thanks to all my friends, Leijian, Lirong, Chuanling, Bo Fang, Zhong Zhang, Lianhui, and Dongdong. They enrich my life every day. Because of their company, every day for me is positive and happy.

Last but not least, I am extremely grateful to my parents and Wensen for their constant support and for giving me unlimited confidence in scientific research and life.

Abstract

With the development of spectral imaging technology, hyperspectral remote sensing data have been widely used in ground monitoring tasks. Notwithstanding its interesting contributions to the scientific field, hyperspectral remote sensing data generally have the problem of high spectral resolution and low spatial resolution due to imaging instrument technical limitations. Meanwhile, influenced by atmospheric scattering and other conditions, mixed pixels are common in hyperspectral remote sensing images, which greatly hinders spectral data analysis and is not conducive to subsequent tasks such as hyperspectral image (HSI) classification and tiny target detection. How to extract and separate the spectral of various substances from mixed pixels and determine their corresponding proportions, i.e., the hyperspectral unmixing technique, become the key to analysis and quantification.

Hyperspectral unmixing generally aims to determine the number of endmembers, capture endmember signatures, and estimate their corresponding abundance maps. Among them, the total number of endmembers is frequently employed as background information for subsequent endmember extraction and abundance estimation. In recent years, hyperspectral unmixing algorithms have been greatly advanced and developed along with the increasingly mature hyperspectral imaging technology. Endmember extraction and abundance estimate are often handled as independent jobs by traditional hyperspectral unmixing algorithms, which unavoidably results in the performance of the following abundance estimation being heavily dependent on the endmember extraction results. Moreover, deep learning (DL) has achieved impressive results in various fields due to its powerful fitting and feature extraction capabilities, and it has been adopted with great enthusiasm within the remote sensing community. Among the different neural models, the autoencoder as a classic algorithm of DL can complete the two tasks of hyperspectral unmixing, namely endmember extraction and abundance estimation. Especially, they usually extract the weight and output of a particular activation layer as endmembers and abundances, respectively. Nonetheless, most of them only utilize spectral information and ignore spatial information.

In this thesis, inspired by traditional machine learning algorithms, several novel and excellent unmixing algorithms have been developed for determining the number of endmembers, extracting endmember signatures, and estimating their corresponding abundances at

each pixel. Moreover, some new networks based on DL have also been designed for effective and efficient hyperspectral unmixing. Overall, this thesis first gives the related concepts of hyperspectral unmixing, then describes the current research status of traditional unmixing methods and DL-based unmixing methods, and finally introduces the methods proposed in this thesis. The primary works and achievements of this thesis are as follows:

- The majority of hyperspectral unmixing algorithms typically employ two separate procedures for endmember extraction and abundance estimation. As a result, the performance of abundance estimation will be impacted by the endmember extraction results. To complete the well-combination and fast realization of endmember extraction and abundance estimation, this thesis designs orthogonal projection endmember (OPE) and orthogonal projection abundance (OPA) for extracting endmembers and estimating their abundance maps, respectively. In particular, the pixel with the largest orthogonal projection on any pixel is considered an endmember. OPE extracts the first endmembers by selecting a pixel from the hyperspectral data to calculate the orthogonal projection of all pixels. To avoid extracting the same endmember, OPE computes the orthogonal projection of all pixels on the previously extracted endmember, and the pixel with the largest projection is then taken into consideration as the next endmember. Moreover, OPA also makes use of the idea of orthogonal projection to find a matrix whose multiplication with the endmember matrix is both a square matrix and a diagonal matrix. Then, some specific matrix operations are used to estimate corresponding abundance maps. The related results show that OPE and OPA are capable of extracting endmembers and rapidly estimating abundance maps.

- Endmember extraction and abundance estimate are often treated as two different phases in traditional unmixing algorithms, with the results of other endmember extraction procedures being required for the completion of abundance estimation. Several DL-based techniques have been proposed in the literature for HSI processing considering the powerful learning and data fitting capabilities of convolutional neural networks (CNNs). Models that exclusively using 2D kernels are unable to acquire strong discriminative feature maps from spectral information, while 3D CNN can capture spectral and spatial information. However, models that solely rely on 3D CNN will be a lot more intricate and expensive in terms of computational overloads. To address these issues, this thesis develops a novel cross convolution unmixing network (CrossCUN) for precise spectral unmixing that combines two-dimensional (2D) and two-dimensional (3D) convolutions to concurrently learn spectral and spatial information and conduct spectral unmixing effectively and efficiently. The associated results demonstrate that the proposed CrossCUN does not only outperform previous DL-based unmixing methods

in terms of results but also does not need any prior information (e.g., the results of other endmember extraction algorithms) to estimate abundance maps.

- Most existing endmember extraction methods require the number of endmembers to be known in advance, and the number of endmembers is usually estimated by a separate procedure. In this sense, the accuracy of determining the number of endmembers has a significant impact on the performance of endmember extraction. To address this issue, this thesis interprets HSI scenes as a hyperspectral pixel set and defines the subset of pixels that differ the most from each other as the divergent subset (DS) of the hyperspectral pixel set. The characteristic of DS is that any extra pixels will increase the similarity within DS and thus reduce its divergence. In this thesis, DS is regarded as an endmember set, and the number of endmembers is the cardinality of the subset. To provide a practical computational scheme for identifying DS, it is reformulated in terms of a quadratic optimization problem with a numerical solution. Furthermore, the proposed DS method can also be used together with other endmember extraction techniques, transforming them into new and more efficient schemes. Compared with other endmember estimation methods, the experimental results verify the effectiveness of the proposed DS method in simultaneously counting and extracting endmembers.

- Furthermore, in this thesis, a maximum distance analysis (MDA) method is proposed to address the shortcoming of traditional endmember extraction, i.e., to determine the number of endmembers and extract endmembers at the same time. In particular, the proposed MDA extracts endmembers by finding the furthest pixel point from a specific pixel point, line, plane, or affine hull formed by previously extracted endmembers. The farthest pixel point from the coordinate origin is the first endmember used to create the above point. In addition to providing an endmember estimation algorithm by itself, MDA can also be generalized as a more efficient scheme in cooperation with existing endmember extraction techniques that do not have pure pixel assumptions. The related results show that both MDA and its cooperative methods exhibit remarkable performance in hyperspectral unmixing.

- Finally, on the basis of the previous research, this thesis develops a new geometric distance-based approach to improve the traditional MDA for estimating endmembers from hyperspectral remote sensing images, which does not need to know the number of endmembers in advance. This strategy, called New MDA (NMDA), optimizes the widely used MDA from two aspects. Firstly, different from the endmember extraction process of traditional MDA, the new strategy NMDA only needs to calculate the maximum distance between any pixel point and a specific affine hull. Hence,

NMDA provides a simpler way to estimate endmembers. Secondly, NMDA utilizes a new distance calculation between any pixel point and the affine hull, requiring only normal vectors (compared to the traditional MDA method using normal vectors and offsets). The new distance calculation in the proposed NMDA is more efficient than the traditional MDA method. Experimental results show that NMDA gets better results in terms of accuracy and efficiency compared with the original MDA method.

To conclude, this thesis compiles these contributions in five articles published in the highest impact journals in the field of remote sensing, which are characterised by their exhaustive peer-review procedure by anonymous experts in the field.

Resumen

Con el desarrollo de la tecnología de imagen espectral, los datos de teledetección hiperespectrales se han utilizado ampliamente en tareas de monitoreo terrestre. A pesar de sus interesantes contribuciones en el campo científico, los datos de teledetección hiperespectrales generalmente tienen el problema de alta resolución espectral y baja resolución espacial debido a limitaciones técnicas del instrumento de imagen. Mientras tanto, influenciados por la dispersión atmosférica y otras condiciones, los píxeles mixtos son comunes en las imágenes de teledetección hiperespectrales, lo que dificulta enormemente el análisis de los datos espectrales y no es propicio para tareas posteriores como la clasificación de imágenes hiperespectrales (HSI) y la detección de objetivos pequeños. Cómo extraer y separar el espectro de varias sustancias de píxeles mixtos y determinar sus proporciones correspondientes, es decir, la técnica de desmezclado hiperespectral, se convierte en la clave para el análisis y cuantificación.

El proceso de desmezclado hiperespectral tiene como objetivo determinar el número de *endmembers*, capturar las firmas espectrales de los mismos y estimar los mapas de abundancia correspondientes. El número total de *endmembers* se emplea con frecuencia como información para la extracción de los mismos y la estimación de abundancias. En los últimos años, los algoritmos de desmezclado para escenas hiperespectrales han avanzado considerablemente y se han desarrollado conjuntamente con una tecnología de imágenes hiperespectrales cada vez más madura. La extracción de *endmembers* y la estimación de la abundancia a menudo se manejan como trabajos independientes por los algoritmos tradicionales de desmezclado hiperespectral, lo que inevitablemente resulta en que el rendimiento de la estimación de la abundancia posterior dependa en gran medida de los resultados de la extracción de *endmembers*. Además, el aprendizaje profundo (DL, por sus siglas en inglés) ha logrado resultados impresionantes en varios campos debido a sus potentes capacidades de ajuste y extracción de características, y ha sido adoptado con gran entusiasmo dentro de la comunidad de teledetección. Entre los diferentes modelos neuronales, el autoencoder como algoritmo clásico de DL puede completar las dos tareas de desmezclado hiperespectral, es decir, la extracción de *endmembers* y la estimación de la abundancia. Especialmente, suelen extraer el peso y la salida de una capa de activación particular como *endmembers* y

abundancias, respectivamente. No obstante, la mayoría de ellos solo utilizan información espectral y ignoran la información espacial.

En esta tesis, inspirada en los algoritmos tradicionales de aprendizaje automático, se han desarrollado varios algoritmos de desmezclado novedosos y excelentes para determinar el número de endmembers, extraer las firmas espectrales de los *endmembers* y estimar sus abundancias correspondientes en cada píxel. Además, se han diseñado algunas redes nuevas basadas en DL para un desmezclado hiperespectral efectivo y eficiente. En general, esta tesis primero presenta los conceptos relacionados con el desmezclado hiperespectral, luego describe el estado actual de la investigación de los métodos de desmezclado tradicionales y los métodos de desmezclado basados en DL, y finalmente presenta los métodos propuestos en esta tesis. Los principales trabajos y logros de esta tesis son los siguientes:

- La mayoría de los algoritmos de desmezclado hiperespectral suelen emplear dos procedimientos separados para la extracción de endmembers y la estimación de abundancias. Como resultado, el rendimiento de la estimación de abundancia se verá afectado por los resultados de la extracción de endmembers. Para lograr la combinación efectiva y rápida de la extracción de endmembers y la estimación de abundancias, esta tesis diseña el endmember de proyección ortogonal (OPE, por sus siglas en inglés) y la abundancia de proyección ortogonal (OPA, por sus siglas en inglés) para extraer endmembers y estimar sus mapas de abundancia, respectivamente. En particular, el píxel con la mayor proyección ortogonal sobre cualquier otro píxel se considera un endmember. OPE extrae los primeros endmembers seleccionando un píxel de los datos hiperespectrales para calcular la proyección ortogonal de todos los píxeles. Para evitar la extracción del mismo endmember, OPE calcula la proyección ortogonal de todos los píxeles en el endmember previamente extraído, y luego se considera el píxel con la proyección más grande como el siguiente endmember. Además, OPA también utiliza la idea de proyección ortogonal para encontrar una matriz cuya multiplicación con la matriz de endmembers sea tanto una matriz cuadrada como una matriz diagonal. Luego, se utilizan algunas operaciones específicas de matrices para estimar los mapas de abundancia correspondientes. Los resultados relacionados muestran que OPE y OPA son capaces de extraer endmembers y estimar mapas de abundancia rápidamente.

- La extracción de endmembers y la estimación de abundancia a menudo se tratan como dos fases diferentes en los algoritmos tradicionales de desmezclado, siendo necesarios los resultados de otros procedimientos de extracción de endmembers para completar la estimación de abundancia. Varios técnicas basadas en DL han sido propuestas en la literatura para el procesamiento de HSI considerando las potentes capacidades de aprendizaje y ajuste de datos de las redes neuronales convolucionales (CNN). Los

modelos que utilizan exclusivamente *kernels* 2D son incapaces de adquirir mapas de características discriminativas sólidas a partir de información espectral, mientras que las CNN 3D pueden capturar información espectral y espacial. Sin embargo, los modelos que dependen únicamente de CNN 3D serán mucho más complejos y costosos en términos de sobrecarga computacional. Para abordar estos problemas, esta tesis desarrolla una nueva red de desmezclado de convolución cruzada (CrossCUN) para un desmezclado espectral preciso que combina convoluciones bidimensionales (2D) y tridimensionales (3D) para aprender simultáneamente información espectral y espacial y llevar a cabo el desmezclado espectral de manera efectiva y eficiente. Los resultados asociados demuestran que el CrossCUN propuesto no solo supera a los métodos de desmezclado basados en DL anteriores en términos de resultados, sino que también no necesita ninguna información previa (por ejemplo, los resultados de otros algoritmos de extracción de endmembers) para estimar mapas de abundancia.

- La mayoría de los métodos de extracción de endmembers existentes requieren que se conozca su número de antemano. Este número de endmembers generalmente se estima mediante procedimientos separados. Por lo tanto, la precisión de determinar el número de endmembers tiene un impacto significativo en el rendimiento de la extracción de endmembers. Para abordar este problema, esta tesis interpreta las escenas HSI como un conjunto de píxeles hiperespectrales y define el subconjunto de píxeles que difieren más entre sí como el subconjunto divergente (DS) del conjunto de píxeles hiperespectrales. La característica del DS es que cualquier píxel adicional aumentará la similitud dentro de DS y, por lo tanto, reducirá su divergencia. Además, DS se considera como un conjunto de endmembers donde el número de endmembers es la cardinalidad del subconjunto. Para proporcionar un esquema computacional práctico para identificar DS, se realiza una reformulación para un problema de optimización cuadrática con una solución numérica. Además, el método propuesto también se puede usar junto con otras técnicas de extracción de endmembers, transformándolas en esquemas nuevos y más eficientes. En comparación con otros métodos de estimación de endmembers, los resultados experimentales verifican la efectividad del método DS propuesto para contar y extraer endmembers simultáneamente.

- Con el fin de abordar la limitación en la extracción de endmembers tradicional, es decir, determinar el número de endmembers y extraer endmembers al mismo tiempo, se propone un método de análisis de distancia máxima (MDA, por sus siglas en inglés) . En particular, el MDA propuesto extrae endmembers encontrando el punto de píxel más lejano a partir de un punto de píxel específico, una línea, un plano o un conjunto

afín formado por endmembers previamente extraídos. El punto de píxel más lejano desde el origen de coordenadas es el primer endmember utilizado para crear el punto mencionado anteriormente. Además de proporcionar un algoritmo de estimación de endmembers por sí solo, el MDA también se puede generalizar como un esquema más eficiente en cooperación con técnicas de extracción de endmembers existentes que no tienen asumen píxeles puros. Los resultados relacionados muestran que tanto el MDA como sus métodos cooperativos exhiben un rendimiento notable en el proceso de desmezclado hiperespectral.

- Sobre la base de la investigación expuesta anteriormente, se desarrolla un nuevo enfoque basado en el cálculo de la distancia geométrica con el fin de mejorar el método tradicional utilizado para análisis de mezclas espectrales (MDA, por sus siglas en inglés). Este método se utiliza para estimar endmembers en imágenes de teledetección hiperespectrales sin necesidad de conocer el número de endmembers de antemano. La propuesta, llamada Nuevo MDA (NMDA), optimiza el MDA en dos aspectos. En primer lugar, a diferencia del proceso de extracción de endmembers del MDA tradicional, la nueva estrategia NMDA solo necesita calcular la distancia máxima entre cualquier punto de píxel y un conjunto afín específico. Por lo tanto, NMDA proporciona una forma más sencilla de estimar los endmembers. En segundo lugar, NMDA utiliza un nuevo cálculo de distancia entre cualquier punto de píxel y el conjunto afín. Este cálculo únicamente requiere vectores normales (en comparación con el método MDA tradicional que utiliza vectores normales y desplazamientos). Como resultado, el nuevo cálculo de distancia en el NMDA propuesto es más eficiente que el método MDA tradicional. Esto queda expuesto en los resultados experimentales, donde se demuestra que NMDA obtiene mejores resultados en términos de precisión y eficiencia en comparación con el método MDA original.

En resumen, esta tesis compila estas contribuciones en cinco artículos publicados en las revistas de mayor impacto en el campo de la teledetección, caracterizadas por su exhaustivo proceso de revisión por expertos en el campo de investigación.

Table of contents

List of figures

Nomenclature

Acronyms / Abbreviations

1D   One-dimensional

2D   Two-dimensional

3D   Three-dimensional

3DCUN 3D Convolution Unmixing Network

CrossCUN Cross Convolution Unmixing Network

AEC   Autoencoder Cascade

ANC   Abundance Nonnegativity Constraint

ASC   Abundance Sum-to-one Constraint

CNNs Convolutional Neural Networks

CyCU-Net Cycle Consistency Unmixing Network

DAEN Deep Autoencoder Network

Deep HSNet Deep Half-siamese Network

DL   Deep Learning

DMBU Deep Multitask Bilinear Unmixing

DNAE Deep Nonlinear Autoencoder

DNN   Deep Neural Network

DS   Divergent Subset

GBM   Generalized Bilinear Model

HSIs   Hyperspectral Images

IEA   Iterative Error Analysiss

LMM   Linear Mixing Model

LSTM   Long Short-term Memory

MDA   Maximum Distance Analysis

MDPI   Multidisciplinary Digital Publishing Institute

MSIs   Multispectral Images

MTAEU   Multitask Autoencoder Unmixing

MVSA   Minimum Volume Simplex Analysis

NAN   Nonsymmetric Autoencoder Network

NLMM   Nonlinear Mixing Model

NMDA   New Maximum Distance Analysis

NMF   Nonnegative Matrix Factorization

NNSAE   Stacked Nonnegative Sparse Autoencoder

OPA   Orthogonal Projection Abundance

OPE   Orthogonal Projection Endmember

PCA   Principal Component Analysis

PNMM   Postnonlinear Mixing Model

PPI   Pixel Purity Index

PPNMM   Polynomial PNMM

RE   Reconstruction Error

RMSE   Root Mean Square Error

RNN   Recurrent Neural Network

SAD   Spectral Angle Distance

SCNL Spectral Convolution Network Incorporating Spectral Library

SGA   Simplex Growing Algorithm

SNSA Stacked NNSAE

TAU   Traditional Autoencoder Unmixing

VCA   Vertex Component Analysis

Compendium of Publications

The current thesis is a compendium of the following publications:

1. **X. Tao**, M. E. Paoletti, L. Han, J. M. Haut, P. Ren, J. Plaza, and A. Plaza. Fast Orthogonal Projection for Hyperspectral Unmixing. IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1–13, February 2022. DOI: 10.1109/TGRS.2022.3150263. [IF(2022)=8.125]. Google Scholar Citations: 2.

2. **X. Tao**, M. E. Paoletti, L. Han, Z. Wu, P. Ren, J. Plaza, A. Plaza, and J. M. Haut. A New Deep Convolutional Network for Effective Hyperspectral Unmixing. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15, pp. 6999–7012, August 2022. DOI: 10.1109/JSTARS.2022.3200733. [IF(2022)=4.715]. Google Scholar Citations: 2.

3. **X. Tao**, T. Cui, A. Plaza, and P. Ren. Simultaneously Counting and Extracting Endmembers in a Hyperspectral Image Based on Divergent Subsets. IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 12, pp. 8952-8966, December 2020, DOI: 10.1109/TGRS.2020.2992542. [IF(2020)=5.6]. Google Scholar Citations: 13.

4. **X. Tao**, M. E. Paoletti, J. M. Haut, P. Ren, J. Plaza, and A. Plaza. Endmember Estimation with Maximum Distance Analysis. Remote Sensing, vol. 13, no. 4, p. 713, February 2021. DOI: 10.3390/rs13040713. [IF(2021)=5.349]. Google Scholar Citations: 9.

5. **X. Tao**, M. E. Paoletti, J. M. Haut, L. Han, P. Ren, J. Plaza, and A. Plaza. Endmember Estimation From Hyperspectral Images Using Geometric Distances. IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1–5, August 2021. DOI: 10.1109/LGRS.2021.3102076. [IF(2021)=5.343]. Google Scholar Citations: 4.

Chapter 1

Introduction

1.1 Research Background and Significance

The development of remote sensing technology gained momentum in the 1960s and has continued to evolve to this day. It refers to the long-distance, non-contact measurement of the electromagnetic wave reflection and radiation characteristics of objects, and the analysis of objects based on information such as reflectance. Remote sensing instruments perform air-to-ground detection by aircrafts and satellites platforms. The collected remote sensing images contain rich ground object information, which can provide an important amount of valuable and useful information for many related research fields, even in a relatively complex target environment, such as flood detection, land survey, material analysis, agricultural monitoring, food safety, and natural resource survey.
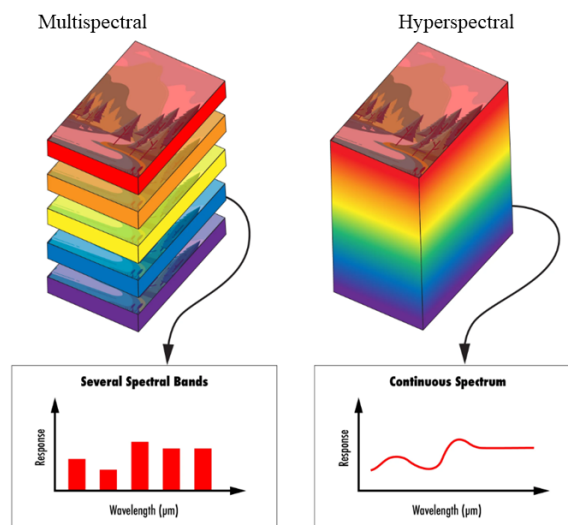
Fig. 1.1 The comparison of MSIs and HSIs (from [36] ©2018 IOP Publishing).

From the 1980s, spectral resolution has increased due to the fast development of imaging spectrometers, and multispectral images (MSIs) [127, 61, 141] and hyperspectral images (HSIs) [5, 38, 31, 108, 106, 36] have been successfully collected. Fig. 1.1 graphically describes the difference between HSIs and MSIs. It can be seen that, usually, the MSI has tens of bands and the spectral curve is discrete, whilst the HSI has hundreds of bands and, consequently, the spectral curve is continuous. Generally, a color image only contains spectral information of red, green, and blue wavelengths, thus it can only characterize the color characteristics of the image, whilst an HSI captures the detailed spectral information of each pixel along the range between the visible and shortwave infrared regions of the spectrum. As shown in Fig. 1.2, HSIs obtained by the remote sensor are recorded in units of pixels, and each pixel corresponds to only one spectral signal. Each pixel in an HSI represents radiation on a series of narrow and continuous wavelengths and bands. Indeed, the wavelength range that HSIs can cover includes visible light, near-infrared and short-wave infrared, and other spectral bands (wavelength range from 0.3 μm to 2.5 μm [39]). Compared to MSI, HSI is a huge data cube, which includes 3D information, i.e., 2D information on spatial data and 1D information on spectral data. Therefore, HSIs not only have a certain ability to describe the spatial characteristics of ground objects, but also have a good ability to describe the spectral characteristics. As a result, the spatial information and spectral information can be well integrated. This thesis focuses on HSIs.

Fig. 1.2 Detail of the hyperspectral remote sensing image (from [12] ©2012 IEEE).
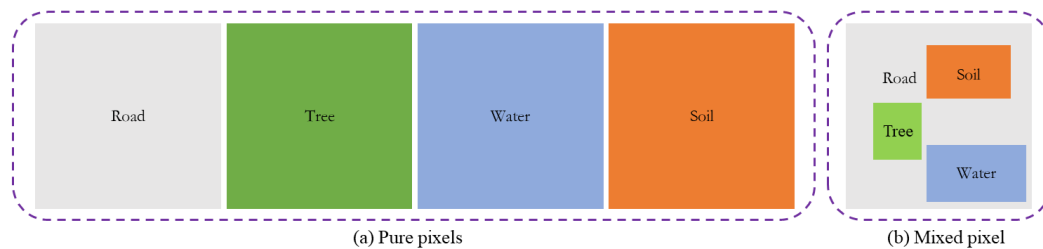
Fig. 1.3 Schematic diagram of pure pixels (a) and mixed pixels (b).

1.1.1 The Generation of Mixed Pixels
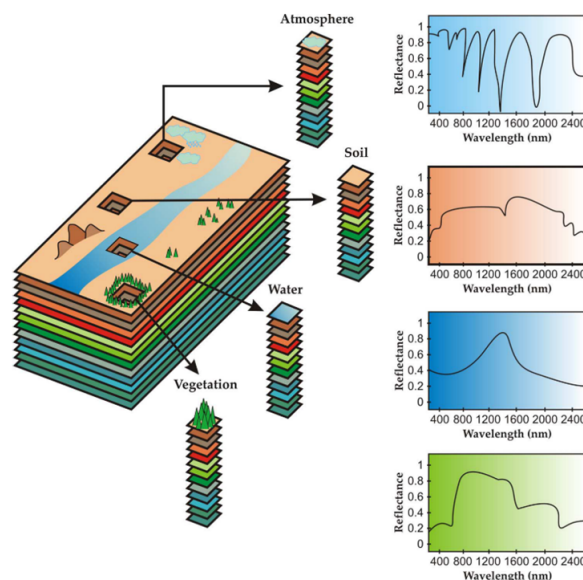
Due to the relatively low spatial resolution of the sensor produced by technical limitations, some pixels in the image may be composed of different substances, thus forming a mixed pixel [52, 65, 99, 16, 84, 125, 60, 69, 64]. In this context, a mixed pixel contains multiple types of substances, whilst a pure pixel only contains one substance. Fig. 1.3(a) shows pure pixels, and each pure pixel only has one substance represented by a single color. On the contrary, Fig. 1.3(b) represents a mixed pixel, which contains four substances, i.e., road, tree, water, and soil. Generally, a mixed pixel contains two or more substances in varying quantities. There are three main reasons for generating mixed pixels:

- The spatial resolution of the sensor is low, and multiple substances may occupy the same pixel so that the spectral measurement result is a mixture of the spectral reflectance of these substances. This situation commonly exists on high-altitude remote sensing platforms.

- When different substances are uniformly mixed to form a mixture, it will also cause mixed pixels, such as sand, dead leaves, or water mixed to form soil.

- Mixing effects that occur in atmospheric transport.

The challenge of mixed pixels commonly exists in HSIs, which is detrimental to applications such as the classification of pixel and sub-pixel level objects, and target recognition. For example, the traditional HSI classification methods are to classify a pixel into a land-cover category. When dealing with mixed pixels, it is often impossible to accurately classify the pixel as a certain type of ground object due to its high spectral variability, which affects their performance in classification. In addition, the mixed pixel is also one of the main obstacles to the quantitative development of remote sensing. In the scenario, some hyperspectral unmixing techniques have been developed to address the issue.
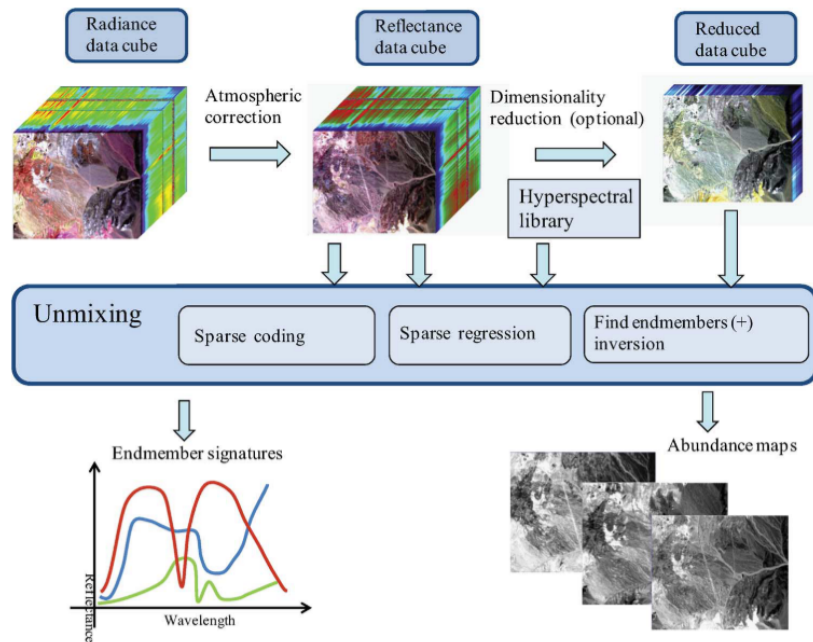
Fig. 1.4 Schematic diagram of hyperspectral unmixing process (from [12] ©2012 IEEE).

1.1.2 Hyperspectral Unmixing

Hyperspectral unmixing [15, 130, 62, 55, 28] is a hot topic and a difficult problem in the field of HSI processing. The schematic diagram of the hyperspectral unmixing process is shown in Fig 1.4. In general, hyperspectral unmixing needs to complete four tasks:

- Reduction of HSIs dimension. HSIs are 3D data, which contain a large number of highly correlated pixels and have obvious information redundancy. The high dimension and complexity of HSIs lead to high computational complexity when processing them, which is not conducive to analysis and processing. In this context, dimension reduction techniques are often seen as a pre-processing technique to improve the accuracy of subsequent data analysis. Generally, dimension reduction techniques can not only reduce the amount of data but also express the spectral features within the data with almost no loss of information.

- Determination of the number of endmembers contained in HSIs. The number of endmembers is the number of pure substances present in HSIs. Correctly estimating the number of endmembers is very important for subsequent endmember extraction and abundance estimation. The number of endmembers can also be estimated by field inspection of how many pure substances there are. For instance, assuming that an HSI

has pure substances such as water, tree, river, and grass, it can be considered that the number of endmembers in the HSI is 4.

- Extraction of the spectral curves of the corresponding pure pixels, i.e., endmembers, of each type of surface object. The quality of the endmembers directly determines the accuracy of unmixing and the effect of subsequent applications. Endmembers can be defined as ideal independent signals of a class of spectra. This is a very general definition, and it is almost impossible to find a pure pixel containing only one endmember in a real image. Therefore, only the signal closest to the endmember spectral can be found. There are two ways to obtain endmembers: (1) The spectral library; (2) Endmember extraction methods. Due to the influence of the imaging spectrometer's own noise, time-space transformation, and other conditions, the spectral obtained through the spectral library is quite different from the real spectral in the image, which cannot meet the requirements of actual use. Therefore, currently, endmember extraction techniques are mainly used to obtain endmember spectra from remote sensing images.

- Estimation of the proportion of each endmember in the mixed pixel spectral (abundance). The vector composed of the abundance coefficients corresponding to different endmembers is called the abundance vector. Given HSI observation data and actual extracted endmember spectral signals, abundance estimation is usually solved as a constrained optimization problem. In earlier studies, endmember extraction and abundance estimation were usually regarded as two independent steps. Nonetheless, many current hyperspectral unmixing methods can accomplish these two tasks simultaneously.

In the unmixing process, whether the dimension of data is reduced depends on the algorithm itself, which is not necessary for the unmixing process. Nevertheless, the number of endmembers in HSI must be determined before unmixing, and then the extraction and abundance estimation of endmembers can be performed.

1.2 Research Objectives and Contents

Most traditional unmixing methods can only accomplish endmember extraction or abundance estimation, and usually researchers need to choose two excellent methods for extracting endmembers and estimating their corresponding abundance maps at each pixel to complete the whole unmixing process. In this sense, it is very time-consuming to find two excellent algorithms and combine them to accomplish accurate unmixing. Furthermore, the accuracy of the final abundance estimation heavily depends on the extracted endmembers. In addition,

most unmixing methods based on DL usually consider spectral information but ignore spatial information, which leads to unsatisfactory unmixing results. In this scenario, this thesis is dedicated to designing effective and efficient methods for endmember extraction and abundance estimation. Moreover, this thesis aims to design new frameworks to accomplish these two tasks simultaneously. More specifically, this thesis has the following scientific objectives:

1. Traditional hyperspectral unmixing algorithms generally consider endmember extraction and abundance estimation as two independent processes to obtain good performance in their respective fields, and then they search for a good combination of the two processes of algorithms to yield the best performance. Another strategy is the nonnegative matrix factorization (NMF)-based method, which completes the two tasks simultaneously. Whether seeking the best combination of the two tasks or completing unmixing directly, the two types of hyperspectral unmixing techniques both bring render considerable computational burden, making them inefficient for tackling large-size remote sensing images. To address the problem, this thesis aims to establish new frameworks for completing endmember extraction and abundance estimation simultaneously.

2. Most existing DL-based unmixing approaches usually use autoencoder networks or 2D CNNs to learn the features contained in HSIs, but only using 2D CNN techniques cannot capture a good discriminative feature map from a spectral perspective. Moreover, although 3D CNN has the ability to learn spectral and spatial information simultaneously, models only using 3D CNN are significantly more complex and expensive in terms of computation. In this sense, this thesis aims to combine 2D and 3D convolution to design a new network for hyperspectral unmixing.

3. Most existing endmember extraction algorithms require prior knowledge about the number of endmembers, yet the algorithms used to count endmembers are frequently separate techniques that may not always be advantageous to endmember extraction algorithms. Thus, this thesis provides a unified endmember estimate framework to perform both tasks simultaneously in order to bridge these two seemingly separate (but highly connected, indeed) processes.

Following these objectives, this thesis presents results obtained throughout the PhD period in five different publications, which have been submitted and published in international journals with high impact indices within the research area:

- The first article, entitled *Fast Orthogonal Projection for Hyperspectral Unmixing* [121], designs a novel orthogonal projection framework to conduct fast hyperspectral

unmixing. It addresses both endmember extraction and abundance estimation with OPE and OPA, respectively. The related results show that OPE and OPA are not only very efficient, but also very effective in endmember extraction and abundance estimation, providing outstanding results at the level of current state of the art (SoTA).

- The second article, entitled *A New Deep Convolutional Network for Effective Hyperspectral Unmixing* [122], combines 3D and 2D convolutions to propose `CrossCUN` for hyperspectral unmixing. The related results show that the proposed `CrossCUN` can effectively and efficiently estimate endmembers and their corresponding abundances at each pixel of the HSI scene without any prior knowledge about endmember extraction compared with traditional unmixing methods.

- The third article, entitled *Simultaneously Counting and Extracting Endmembers in a Hyperspectral Image Based on Divergent Subsets* [119], proposes a DS-based method for counting and extracting endmembers at the same time. Without any prior information, the proposed DS approach can not only extract endmembers, but it can also be combined with other endmember extraction methods to create more powerful schemes.

- The fourth article, entitled *Endmember Estimation with Maximum Distance Analysis* [124], proposes an MDA-based method for estimating the number of endmembers and the corresponding endmember signatures. The proposed MDA method captures endmembers without any prior knowledge. Moreover, the proposed MDA may be included in existing unmixing techniques like minimal volume simplex analysis (MVSA) to avoid the restriction that the number of endmembers needs to be known in advance.

- The fifth article, entitled *Endmember Estimation From Hyperspectral Images Using Geometric Distances* [123], develops a new geometric distance-based algorithm NMDA to optimize the original MDA method. The main difference between NMDA and traditional MDA can be summarized as follows: (1) traditional MDA algorithm obtains endmembers by searching for the farthest pixel from one specific pixel, line, plane, or affine hull composed by the endmembers that have been formerly extracted. The proposed NMDA method optimizes the process of endmember estimation and yields endmembers by finding the farthest pixel from one specific affine hull, and (2) the distance computation of the traditional MDA method needs the normal vector and offset, and the proposed NMDA algorithm exploits a new distance computation that only requires the normal vector for endmember estimation.

1.3 Dissertation Outline

This thesis focuses on these three tasks, i.e., determining the number of endmembers, extracting endmember signatures, and estimating abundance maps of every endmember at each pixel, and the corresponding research content mainly includes five aspects:

- Chapter 1 introduces the research background and significance of hyperspectral unmixing, explains the research objective, and clarifies the innovations of this thesis.

- Chapter 2 describes two commonly used unmixing models and reviews of some classical traditional unmixing methods to further explain the motivation of the proposed methods in this thesis.

- Chapter 3 introduces the layers commonly used in neural networks, describes in detail some classical autoencoder networks, and reviews many typical networks for hyperspectral unmixing to exploit how to design effective and efficient networks for endmember extraction and abundance estimation.

- Chapter 4 briefly presents the five hyperspectral unmixing methods proposed in this thesis and simultaneously introduces their advantages and defects in hyperspectral unmixing.

- Chapter 5 summarizes the paper and gives some future directions for hyperspectral unmixing.

Chapter 2

Review of Traditional Unmixing Methods

2.1 Mixing Models

In the field of remote sensing, hyperspectral unmixing has been a very challenging inverse problem without a closed-form solution. The *spectral mixing model* is the premise of the hyperspectral unmixing technique. According to the consideration of spectral mixing factors, the spectral mixing model can be divided into linear mixing model (LMM) and nonlinear mixing model (NLMM) [66, 89, 109, 87, 56]. Next, these two models will be described in detail.

2.1.1 Linear Mixing Model

The condition for the establishment of LMM is that the incident light from the sun only interacts with one kind of material, and the light reflected by the material is received directly by the sensor. At this point, the limited spatial resolution of the HSI itself is the primary cause of the creation of mixed pixels. Spatial resolution refers to the size of the actual area represented by each pixel in the image. The actual area represented by each pixel in an HSI is very large and often contains several different substances. As a result, the light reflected by them will be received by the sensor and reflected on the spectrum of this pixel.

A graphical example of LMM is shown in Fig. 2.1. According to the LMM, a pixel in an HSI can be represented as a linear combination of a number of endmembers and their associated abundance fractions. In this regaLet $\mathbf{Y} \in \mathbb{R}^{L \times N}$ be an HSI, where L denotes the number of bands, and N represents the number of pixels. Suppose the endmember matrix is $\mathbf{M} = [\mathbf{m}_1, \cdots, \mathbf{m}_c] \in \mathbb{R}^{L \times c}$, the abundance vector is $\mathbf{a}_i = [a_{i1}, \cdots, a_{ic}]^{\mathrm{T}}$, and the i-th pixel $\mathbf{y}_i$
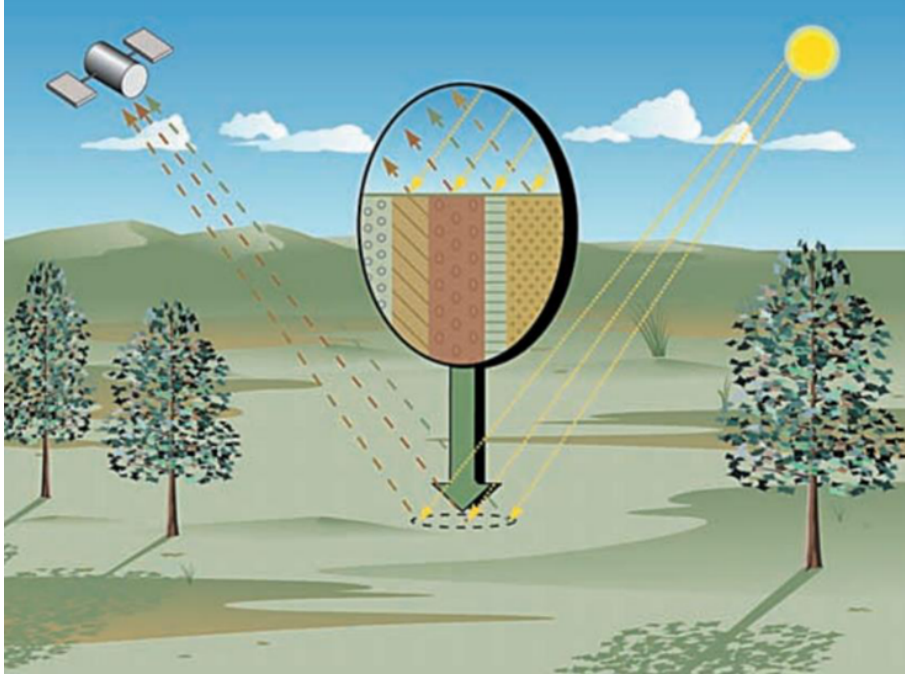
Fig. 2.1 Illustration of LMM (from [66] ©2002 IEEE).

can be expressed as follows:

$$\mathbf{y}_i = \sum_{j=1}^{c} \mathbf{m}_j a_{ij} + \mathbf{n}_i = \mathbf{M}\mathbf{a}_i + \mathbf{n}_i \qquad (2.1)$$

where c is the number of endmembers, $\mathbf{n}_i$ is the noise, and $\mathbf{m}_j \geq 0$ denotes the measured reflectance of endmembers at the i-th spectral signature. In addition, considering the practical situation, the LMM needs to guarantee abundance nonnegativity constraint (ANC) and abundance sum-to-one constraint (ASC) given by:

$$\text{ANC}: a_{ij} \geq 0; \forall i, j \qquad (2.2)$$

and

$$\text{ASC}: \sum_{j=1}^{c} a_{ij} = 1 \qquad (2.3)$$

Assuming that there are N pixels in the image, the mathematical expression of the LMM corresponding to these N pixels is:

$$\mathbf{Y} = \mathbf{M}\mathbf{A} + \mathbf{N} \qquad (2.4)$$

where $\mathbf{A}$ is the abundance matrix and $\mathbf{N}$ is the noise matrix.
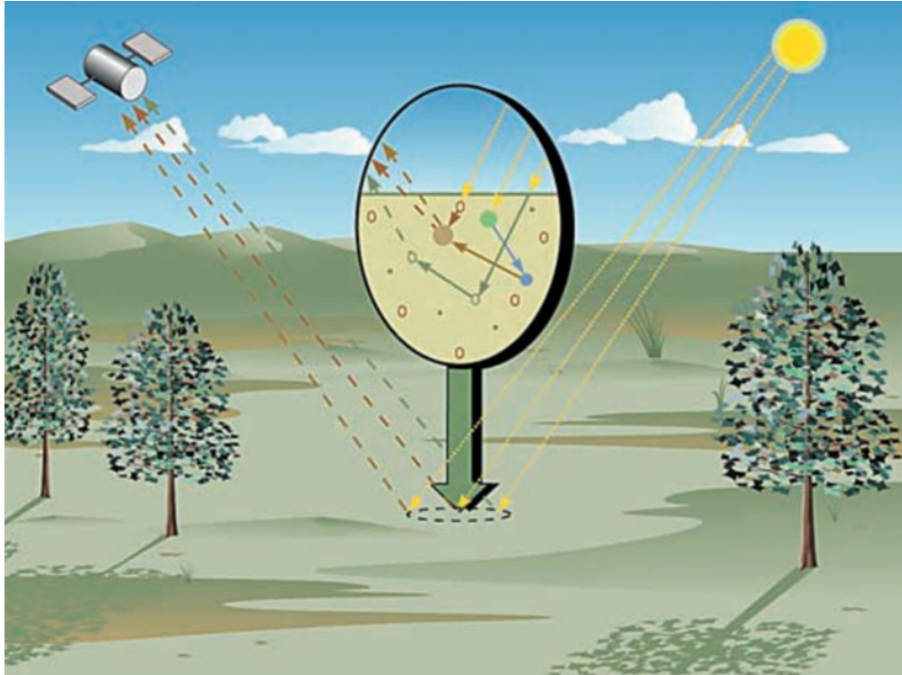
Fig. 2.2 Illustration of NLMM (from [66] ©2002 IEEE).

2.1.2 Nonlinear Mixing Model

Nonlinear mixing is usually caused by the physical interactions of light scattered by various substances in the scene. Light interaction in the traditional sense means that the incident light from the sun hits on one object and then the reflected light from the object hits on another object, which is finally received by the sensor. Since the NLMM is closer to the actual hyperspectral remote sensing image, the results of the nonlinear unmixing method are often better for the actual HSI. Fig. 2.2 depicts an illustration of NLMM. The typical NLMMs includes bilinear mixing model [13, 149, 43, 2, 29], postnonlinear mixing model (PNMM) [4, 118, 3, 21, 163], generalized bilinear model (GBM) [44, 53, 75, 78, 159], and Hapke model [46, 47, 148, 105], among others.

In NLMM, the Fan Model [33] uses the Hadamard product of the characteristic spectrum to represent the scattering effect between photons. The GBM model has a clear physical meaning. The PNMM model performs a nonlinear transformation on top of a linear mixture. Polynomial PNMM (PPNMM) is established using the polynomial progression technique. In a mathematical sense, they all add nonlinear items on the basis of the LMM to complete the establishment of the model. Therefore, the basic concepts in LMM must still be used by these models, although new constraints need to be added to ensure its physical meaning; on the other hand, new coefficients for adjusting nonlinear mixing are included to the corresponding model.

In summary, the NLMM has to consider the radiation between various ground objects, and the model structure is relatively complex, which brings difficulties to practical application. Nonetheless, the LMM has a clear physical meaning and has achieved a good unmixing result in practical applications. As a result, LMM has been widely used in the application of HSIs.

2.2 Current Status of Traditional Unmixing Methods

Hyperspectral unmixing usually faces two challenges, i.e., extracting endmembers and estimating their corresponding abundances at each pixel. Most traditional unmixing methods treat these two challenges as separate problems. In this section, some of the traditional unmixing methods developed for these two tasks will be reviewed.

(1) Pixel Purity Index

The core idea of the pixel purity index (PPI) algorithm [16, 131, 144] has the following two points: firstly, all the pixels in the image are gathered inside a simplex; secondly, the endmembers are just the vertices of the simplex formed by the image. PPI makes a random straight line projection of all pixels toward a feature space and then calculates the position of all data in the projection direction. The point falling on the outermost point of the projection direction is most likely to be an endmember. During the calculation of the algorithm, the number of times each pixel falls on both ends of the line segment in all random projection directions is recorded. The more times the point is projected to both ends, the more likely the point is to be an endmember. Therefore, the PPI of each pixel is used to judge whether the pixel is an endmember or not. The projection from a point to a straight line in the feature space can be completed by the inner product operation of vectors. Especially, the flow of the PPI algorithm is as follows:

1. Suppose $\mathbf{Y}$ is an HSI, c is the number of endmembers, k is the maximum number of cycles, and $\mathbf{M}$ is the endmember matrix.

2. Suppose the set of pure pixel indices is $\{p_i\}_{i=1}^n$, where $p_i = 0, \forall i$.

3. For $k(k = 1, \cdots, k_{max})$, sequentially obtain the corresponding random unit vector β_k of k, calculate the projection $R_k^{'} = \beta_k^{\mathrm{T}} R$ of pixels in the data set in the β_k direction, then use the following formula to obtain the positions of the maximum and minimum values of $R_k^{'}$: $i_{max} = \arg\max(R_k^{'}), i_{min} = \arg\min(R_k^{'})$, and finally record pure pixel indices by $p_{i_{max}} = p_{i_{max}} + 1, p_{i_{min}} = p_{i_{min}} + 1$.

4. If the non-zero value in the PPI set is less than c, all the pixels with non-zero PPI are endmembers ($\mathbf{M} = \{\mathbf{y}_i | p_i > 0, i = 1, 2, \cdots, n\}$). Otherwise, the c pixels with the largest PPI are selected as endmembers.

From the above process, it is not difficult to find that the implementation of the PPI algorithm is simple and flexible. Especially, PPI has been used as a standardized algorithm in ENVI software. However, PPI also faces some drawbacks:

- It is necessary to pre-set the number of random unit vectors and the threshold required for selecting endmembers (i.e., the number of endmembers needs to be known in advance).

- The randomness of the vector makes the results of each solution inconsistent.

- The number of test vectors used each time is relatively large, so the speed of the PPI algorithm is not very fast.

(2) N-FINDR

Unlike PPI, N-FINDR obtains endmembers through the largest simplex in the feature space. Due to its simple design principle, it is used as the basis for many endmember extraction algorithms. According to the definition of the simplex volume, N-FINDR believes that the simplex formed with the endmembers as vertices in the feature space of HSI has the largest volume [139, 120, 20, 37, 146]. Based on this assumption, the algorithm calculates the volume of the simplex in an iterative way and finally finds the maximum volume to obtain the endmembers in the HSI. The specific flow of N-FINDR is as follows:

1. Use principal component analysis (PCA) or other methods [140, 143, 48] to reduce the dimension of HSI.

2. Randomly select c pixel points from the original hyperspectral data as endmembers, and these c pixel points are used as the initial endmember set. Also, the augmented matrix is constructed as follows:

$$\mathbf{E} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_c \end{bmatrix} \tag{2.5}$$

Then, the volume of the corresponding simplex can be expressed as:

$$V(\mathbf{E}) = \frac{1}{(c-1)!} |\det(\mathbf{E})| \tag{2.6}$$

where $|\cdot|$ denotes the absolute value.

3. Any point in the set of endmembers is replaced one by one with the remaining pixels in the image to form a simplex, and the volume of the new simplex is calculated. If the volume increases, the pixel replaces the existing endmember and becomes a new endmember. This procedure is continued until no more endmembers are substituted, and these pixels with the maximum volume are endmembers.

Although N-FINDR is simple in principle and easy to operate, there are still some challenges to be addressed:

- It is difficult to converge to the global optimal solution due to the limitation of the search strategy, and the time complexity is high when the number of pixels is large.

- The number of endmembers needs to be known in advance.

- The final obtained endmember extraction result is related to the pixels initially selected as endmembers.

(3) Vertex Component Analysis

The fundamental principle of vertex component analysis (VCA) [88, 24, 14, 111, 6] is to repeatedly project the data set to the direction orthogonal to the subspace, which is formed by the determined endmembers. The new endmember corresponds to the extreme value of the projection direction, and the algorithm continues to iterate until all endmembers are found. VCA is a method based on projection, which is mainly based on two facts: (1) the endmembers are located at the vertices of the simplex; (2) a simplex is still a simplex after an affine transformation. The detailed process of VCA is as follows:

1. Initialize the number of endmembers c, and set a random projection vector $\mathbf{w}$, e.g., $\mathbf{w} = \{1, 0, \cdots, 0\}^{\mathrm{T}}$. Also, initialize the empty endmember matrix $\mathbf{M}$, and set iteration counter $j = 0$.

2. The iteration counter increments by 1, if $j <= m$, project the image onto $\mathbf{w}$ and calculate the projection length of each pixel. Suppose the image matrix is $\mathbf{Y}$, and the projection length vector is $\mathbf{L} = |\mathbf{w}^{\mathrm{T}}\mathbf{Y}|$.

3. Find the position $i_{max} = \arg\max(\mathbf{L})$ corresponding to the maximum value in the vector $\mathbf{L}$, and add $\mathbf{Y}_{i_{max}}$ to the endmember matrix: $\mathbf{M} = [\mathbf{M}, \mathbf{Y}_{i_{max}}]$.

4. Update the vector $\mathbf{w}$ using Schmidt orthogonalization. $\mathbf{w}$ is orthogonal to each endmember vector in $\mathbf{M}$. Go to step 1 and continue to loop this operation, and the algorithm ends when all endmembers are found.

Compared with PPI and N-FINDR, VCA performs better and requires less processing, and it has clear advantages when the data set contains a lot of pixels. In addition, VCA has strong anti-noise ability in endmember extraction, but the selection of the initial endmember will also affect the result of the final endmember extraction.

(4) Simplex Growing Algorithm

The principle of the simplex growth algorithm (SGA) [18, 17, 19] is the same as that of N-FINDR and can be regarded as an improvement of N-FINDR. N-FINDR needs to first select c pixels as initial endmembers and then optimize each initial endmember sequentially based on the volume of the simplex until the volume no longer increases. Conversely, SGA first extracts an endmember and then extracts each endmember sequentially based on the volume of the simplex. In general, N-FINDR requires initial endmembers, but the initial endmembers do not necessarily appear in the final endmember set, and it is possible that the initial endmember sets are replaced by new endmembers that meet the requirements more. However, SGA can do the enumeration search, and the final extraction result is unique, and the extraction efficiency of endmembers is higher than that of N-FINDR. The specific process of extracting endmembers by SAG is as follows:

1. Suppose the hyperspectral data is $\mathbf{Y}$, the first c components of $\mathbf{Y}$ are expressed as $\mathbf{Y}_i^{(c)} = [\mathbf{Y}_{1i}, \mathbf{Y}_{2i}, \cdots, \mathbf{Y}_{ci}]$, and c is the number of endmembers.

2. Randomly select a pixel from the hyperspectral remote sensing image, denoted as $\mathbf{t}$.

3. Calculate the first endmember position by:

$$p = \arg\max_{i=1,\cdots,N} \begin{vmatrix} 1 & 1 \\ \mathbf{t} & \mathbf{Y}_i^{(1)} \end{vmatrix} \tag{2.7}$$

where N is the number of pixels.

4. Get the spectral vector of the first endmember: $\mathbf{m}_1 = \mathbf{Y}_p$.

5. Repeat the above steps to extract the remaining endmembers by

$$p = \arg\max_{i=1,\cdots,N} \begin{vmatrix} 1 & \cdots & 1 & 1 \\ \mathbf{m}_1^{(c-1)} & \cdots & \mathbf{m}_{c-1}^{(c-1)} & \mathbf{Y}_i^{(c-1)} \end{vmatrix} \tag{2.8}$$

(5) Iterative Error Analysis

The main idea of the iterative error analysis (IEA) algorithm [90, 135, 107, 110] is to iteratively calculate the error between the reconstructed data and the original data to obtain endmembers. ICA first selects the average value of all pixels as the endmembers to perform constraint unmixing, calculates the reconstructed data after obtaining the endmembers and abundance matrix, obtains the error between the reconstructed data and the original data, and finally the pixel with the largest error is regarded as the new endmember. In order to reduce the interference of outliers and noise, multiple pixels with small spectral angles and large errors can be selected as new endmembers. This process is repeated until the error falls below a certain threshold or reaches a predetermined number of endmembers. The specific process of IEA is as follows:

1. Select the mean vector of the original image as the initial endmember: $\mathbf{m}_1 = \frac{1}{N}\sum_{i=1}^{N}\mathbf{y}_i$, where $\mathbf{M} = [\mathbf{m}_1]$, N is the number of pixels, and $\mathbf{y}_i$ represent the pixel vector in the hyperspectral data $\mathbf{Y}$. Initialize the counter j and the number of endmembers c.

2. When $j <= m$, find the fully constrained least square (FCLS) [50, 51, 134, 25, 23] solution of the original image about the endmembers, i.e., the abundance matrix $\mathbf{A}$. Obtain the reconstructed image according to $\hat{\mathbf{Y}} = \mathbf{MA}$, and calculate the error between the reconstructed matrix and the original image.

3. Find the pixel $\mathbf{y}_{max}$ corresponding to the maximum error vector, and update endmember matrix $\mathbf{M} = [\mathbf{M}, \mathbf{y}_{max}]$. Go to step 2 and continue the calculation until c endmembers are found.

In IEA, candidate endmembers are selected by unmixing errors only, making the algorithm risky in some cases. Moreover, the above methods, i.e., PPI, N-FINDR, VCA, SGA, and ICA, all require a combination of abundance estimation methods for the final unmixing, and FCLS is the most commonly used abundance inversion model combined with the endmember extraction algorithm. The problem solved by FCLS is

$$
\mathbf{X} = \arg\max \frac{1}{2}\|\mathbf{Y} - \mathbf{MA}\|
$$
$$
s.t. : \mathbf{1}^{\mathsf{T}}\mathbf{a}_i = \mathbf{1}; a_{ij} \geq 0
$$

(2.9)

where $\mathbf{a}_i$ is the abundance vector of the i-th pixel, and a_{ij} is the abundance of the j-th endmember at the i-th pixel pixel. After the endmember matrix $\mathbf{M}$ is known, the abundance estimation problem becomes a problem of solving a linear equation system, the expression is

as follows:

$$A = (M^TM)^{-1}M^TY \tag{2.10}$$

Add the ASC and ANC constraints, and construct the matrix as follows:

$$M' = \begin{bmatrix} M \\ \delta 1^T \end{bmatrix}, Y' = \begin{bmatrix} Y \\ \delta 1^T \end{bmatrix} \tag{2.11}$$

where δ is a constant used to control the error, and 1^T is a column vector whose elements are all 1. The original M and Y are replaced with M' and Y', respectively. The operation ensures that the obtained abundance matrix A satisfies ASC and ANC constraints. Moreover, some excellent methods using these two constraints [153, 63, 76, 34, 150, 165] have also developed for simultaneously completing endmember extraction and abundance estimation, and most of them are based on NMF. Typical algorithms include minimum volume constrained nonnegative matrix factorization [85], spatial group sparsity regularized NMF [137], structure constrained sparse NMF [82], graph regularized $L_{1/2}$-NMF [83], etc.

(6) Fast Unmixing

Compared with the above methods, the fast unmixing (FUN) method [41] can simultaneously complete these three tasks, i.e., determining the number of endmembers, extracting endmember signatures, and estimating abundance maps. FUN regards the pixel with the largest orthogonal projection relative to the centroid pixel as the first endmember, and then uses the modified Gram–Schmidt orthogonal method to perform orthogonal projection on the HSI, and the pixel with the largest projection is regarded as an endmember. At the same time, the number of endmembers is calculated by judging the proportion of endmember information in the new hyperspectral dataset as the stopping condition. While obtaining the endmember set, the matrix Q is obtained by applying c times of Gram–Schmidt, and the abundance of the corresponding endmember is obtained by $A = Q^TM$. The specific steps of this algorithm are as follows:

1. Input: HSI data $Y = [y_1, y_2, \cdots, y_N]$, $\alpha = 1$.

2. Endmember Matrix: $M = []$, endmembers' Gram–Schmidt orthogonalization matrix: $Q = []$, endmembers' Gram–Schmidt orthogonal normalization matrix: $U = []$.

3. Replication matrix for hyperspectral data $X = [x_1, x_2, \cdots, x_N] = Y$.

4. $m_1 = x_j$ (x_j is the pixel selected as the first endmember).

5. $\mathbf{q}_1 = \mathbf{m}_1$, $\mathbf{u}_1 = \mathbf{q}_1/||\mathbf{q}_1||$.

6. $\mathbf{M} = [\mathbf{m}_1]$, $\mathbf{Q} = [\mathbf{q}_1]$, $\mathbf{U} = [\mathbf{u}_1]$.

7. $c = 1$, exit =0.

8. While exit = 0 do

    For $j = 1$ to N do

        $\mathbf{x}_j = \mathbf{x}_j - (\mathbf{x}_j \bullet \mathbf{u}_c) \bullet \mathbf{u}_c$

        $s_j = 100 \cdot ||\mathbf{x}_j||/||\mathbf{y}_j||$

    End

    if $\max(s) \leq \alpha$ then

        Exit = 1

    Else

        $j_{max} = \arg\max(s_j)$

        c = c+1

        $\mathbf{q}_c = \mathbf{x}_{jmax}$

        $\mathbf{u}_c = \mathbf{x}_{jmax}/||\mathbf{x}_{jmax}||$

        $\mathbf{m}_c = \mathbf{x}_{jmax}$

    End

  End

9. Output: the number of endmembers: c

   endmember matrix: $\mathbf{M}$

   endmembers' Gram–Schmidt orthogonalization matrix: $\mathbf{Q}$

   endmembers' Gram–Schmidt orthogonal normalization matrix: $\mathbf{U}$

From the steps of FUN, it can be observed that FUN only needs some matrix operations to complete the whole unmixing, and it is very effective for hyperspectral unmixing. However, it is not difficult to find that FUN requires the pure pixel assumption.

2.3 Summary and Thinking of Traditional Unmixing Methods

After reviewing these typical traditional unmixing methods, it is easy to find that they still face many challenges. The detailed summary and thinking are as follows.

- They both require the number of endmembers to extract enmembers, but the number of endmembers usually needs to combine the related algorithms to determine. The defect

leads to the performance of endmember extraction algorithms being largely dependent on other algorithms of counting the number of endmembers. In this sense, exploiting a comprehensive framework for directly determining the number of endmembers and simultaneously extracting their spectral signatures is very necessary.

- Most of them need to enumerate all the pixels to find the simplex with the largest volume and then extract the vertices as endmembers. This operation is extremely time-consuming. In this sense, it is very significant to develop effective and efficient methods to avoid enumerating all pixels to find the simplex with the largest volume to extract endmember.

- Except for FUN, they both can only extract endmembers but not estimate the abundance maps of each endmember at every pixel. Generally, they need to combine other excellent methods (e.g., FCLS) to complete the final abundance estimation. In this sense, firstly, they cannot complete the overall spectral unmixing task; secondly, how to find an abundance estimation method to cooperate with the corresponding endmember extraction algorithms is very vital and time-consuming. To this end, developing an effective and efficient framework for extracting endmembers and estimating their corresponding abundance maps at each pixel simultaneously is very essential for hyperspectral unmixing.

2.4 Evaluation Metrics for Unmixing

Regarding the comparison of the performance of the unmixing algorithm, the accuracy of the final unmixing is generally judged visually by drawing the extracted endmember spectral curves and abundance maps. However, when the improvement of the unmixing algorithm does not vary much or the difference observed on the image is not intuitive, specific metrics are needed to quantitatively measure the advantage of the unmixing effect. The performance evaluation metrics mainly include three aspects: endmember information, abundance information, and image quality.

Spectral angle distance (SAD) is frequently used to evaluate the effectiveness of endmember extraction. In fact, it represents the angle between the extracted endmember and the actual endmember in radians. The specific solution expression is:

$$J_{SAD} = \arccos \frac{\hat{\mathbf{m}}_i^{\mathrm{T}} \mathbf{m}_i}{\|\hat{\mathbf{m}}_i\| \, \|\mathbf{m}_i\|} \tag{2.12}$$

where $\hat{\mathbf{m}}_i$ is estimated endmembers by unmixing methods, and $\mathbf{m}_i$ denote real endmembers. Endmember extraction gets more accurate results when SAD is smaller. Because SAD is measured by the vector angle index, the results of endmember extraction pay more attention to the similarity between them, which is intuitively the shape between them.

Also, root mean square error (RMSE) is often used to evaluate the performance of abundance estimation. RMSE calculates the difference between estimated and real abundances using their respective standard deviations. Results from abundance estimation are more accurate when RMSE is lower. RMSE is calculated using:

$$J_{RMSE} = \sqrt{\frac{1}{Nc} \sum_{i=1}^{N} ||\hat{\mathbf{a}}_i - \mathbf{a}_i||^2} \qquad (2.13)$$

where c is the nunmber of endmembers, N is the number of pixels, $\hat{\mathbf{a}}_i$ is estimated abundances by unmixing methods, and $\mathbf{a}_i$ denotes real abundances.

Finally, reconstruction error (RE) is employed to measure the difference between two images, and the extracted endmembers and the estimated abundance by the unmixing algorithm are used to reconstruct pixels. RE reflects the difference between the reconstructed image and the original image. RE is computed by

$$J_{RE} = \sqrt{\frac{1}{NL} \sum_{i=1}^{N} ||\hat{\mathbf{y}}_i - \mathbf{y}_i||^2} \qquad (2.14)$$

where L is the nunmber of bands, $\hat{\mathbf{y}}_i$ is reconstructed pixels by unmixing methods, and $\mathbf{y}_i$ denote real pixels. The reconstructed image will be closer to the original image when RE is smaller.

2.5 Hyperspectral Datasets

At this point, there are four commonly used datasets, i.e., Samson, Jasper, Urban, and Cuprite, for hyperspectral unmixing. Fig. 2.3 displays the four hyperspectral datasets, and the details are described as follows:

- Samson dataset: it has three endmembers and contains 952×952 pixels, where each one has 156 bands covering the wavelengths from 401nm to 889nm. Considering that the original image is too large and computationally expensive, a region of 95×95 pixels is used in this thesis for hyperspectral unmixing.

(a) Samson (b) Jasper Ridge (c) Urban (d) Cuprite
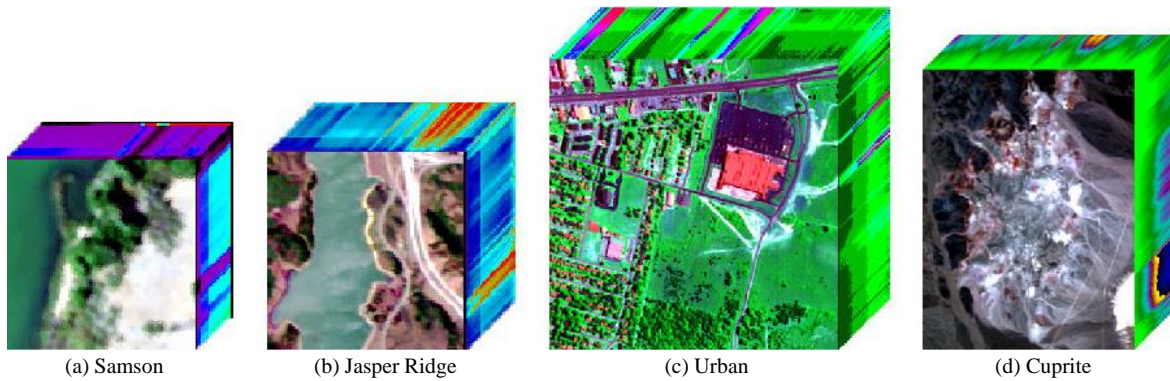
Fig. 2.3 Four real datasets for hyperspectral unmixing.

- Jasper dataset: it contains four endmembers, 210 bands, and 512×614 pixels with wavelengths ranging from 380nm to 2500nm. The original data is cropped into a sub-image of 100 × 100 pixels. The position of the first pixel is (105, 269). Bands 1-3, 108-112, 154-166, and 220-224 are removed due to dense water vapor and atmospheric effects, and 188 bands are retained for hyperspectral unmixing.

- Urban dataset: it has four endmembers, and 307×307 pixels with wavelengths ranging from 400nm to 2500nm. Due to dense water vapor and atmospheric effects, the bands 1-4, 76, 87, 101-111, 136-153, and 198-210 are removed. In this sense, the modified data has 162 bands for hyperspectral unmixing.

- Cuprite dataset: it contains 224 bands with 250×190 pixels. Regarding the spectral information, its wavelengths are ranging from 370nm to 2480nm. The noisy bands (1-2 and 221-224) and water absorption bands (104, 113, 148, and 167) are removed. In this sense, the modified data has 188 bands for hyperspectral unmixing.

Chapter 3

Review of DL-based Unmixing Methods

3.1 Layers of DL used in unmixing

DL-based spectral unmixing methods design their corresponding network architectures by combining different layers, i.e., fully connected layers, convolutional layers, and activation layers. In this scenario, to better understand the DL-based spectral unmixing networks, the different layers will be introduced briefly in this chapter.

3.1.1 Fully Connected Layers

Fully connected layers [7, 26, 157, 80, 147] are used to connect every neuron in the l-th layer to every neuron in the $l+1$-th layer. The corresponding diagram can be observed in Fig. 3.1. Suppose that $\mathbf{Y}^{l-1}$, $\mathbf{W}^l$ and $\mathbf{b}^l$ are the input data, weight, and bias of the l-th layer, respectively. The fully connected layers exploit a linear transformation in the l-th layer to
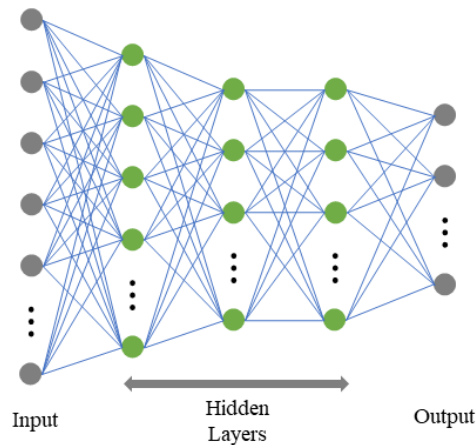
Fig. 3.1 Diagram of fully connected layers.

obtain the input data of the l+1-th layer by:

$$\mathbf{Y}^l = \mathbf{W}^l \cdot \mathbf{Y}^{l-1} + \mathbf{b}^l \tag{3.1}$$

Generally, fully connected layers contain a large number of neurons, which means that many parameters need to be adjusted. More parameters will inevitably lead to higher computational overloads. The number of parameters is computed by:

$$n_{para} = \sum_{i=0}^{L-1} (n_{neurs}^l \cdot n_{neurs}^{l+1} + 1) \tag{3.2}$$

where L is the number of layers, n_{neurs}^l and n_{neurs}^{l+1} are the number of the neurons of the l-th layer and the l+1-th layer, respectively. Especially, the input data and extracted features will be limited to the vector representation of the input data, which will lose the potential of the spatial-contextual information to some extent.

3.1.2 Convolutional Layers

As shown in Fig. 3.2, the convolutional layer [10] defines a block of neurons that acts as a linear kernel, as these neurons are connected and applied to a smaller predefined area of the input data. The convolutional layer is a basic building block of CNN [40, 77, 142, 1, 98]. It is designed to automatically learn the spatial hierarchy of features from the original input data (e.g., images or audio). The key idea of the convolutional layer is to apply a set of learnable filters (also called kernels or weights) to a small region of the input data, called the perceptual field, and then combine the results to produce the feature map. The convolution operation is essentially a sliding window that moves the filter over the input data, computing the dot product between the filter and the receptive field at each position. The resulting values are then summed and passed through a nonlinear activation function, such as rectified linear unit (ReLU). This process produces a new feature map that represents a higher-level
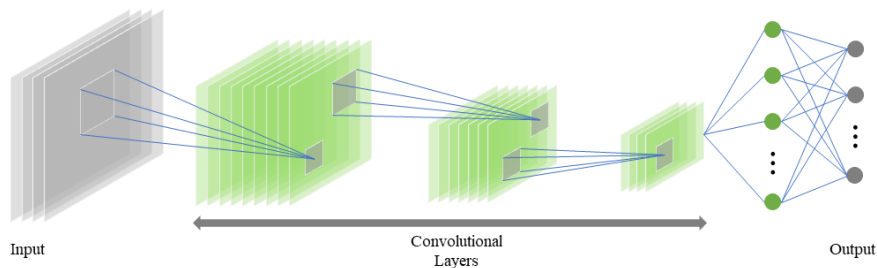
Fig. 3.2 Diagram of the convolutional architecture of a DNN model.

representation of the input data, with each filter capturing a specific pattern or feature.

Compared with the traditional fully connected layers, the convolution operation mainly includes three superiorities: (i) the local connectivity learns the correlation between surrounding pixels; (ii) the weight sharing mechanism and the sparse connectivity of the same feature map decrease the number of parameters and (iii) the location of the feature does not vary. Considering the advantages of the convolution operation, some researchers exploit it, replacing the fully connected layers to accelerate the learning process [92, 117].

3.1.3 Activation Layers

The data transformation of the fully connected and convolutional layers is considered as the feature extractor stage of the DNN [86, 116, 72, 79, 8]. If DNN does not use the activation function, the input to each of its layers will be the linear output of the previous layer. In this case, no matter how many layers the DNN contains, its final output is a linear combination of inputs, which is equivalent to the effect of no hidden layer. This is the case with the traditional multilayer perceptron [91, 100, 126, 96]. Therefore, the activation function is an important part of the DNN. Basically, four classic activation functions are commonly used for spectral unmixing in DNN. The output of the l-th layer is $\mathbf{Y}^l = f(\mathbf{Y}^{l-1})$, and $f(\cdot)$ is the corresponding activation function. The graphical visualization of these four activation functions is shown in Fig. 3.3. In the following, how to compute the outputs of the different activation functions, i.e., Sigmoid, TanH, ReLU, and Leaky ReLU [152, 32, 151], and their advantages and disadvantages are briefly described.
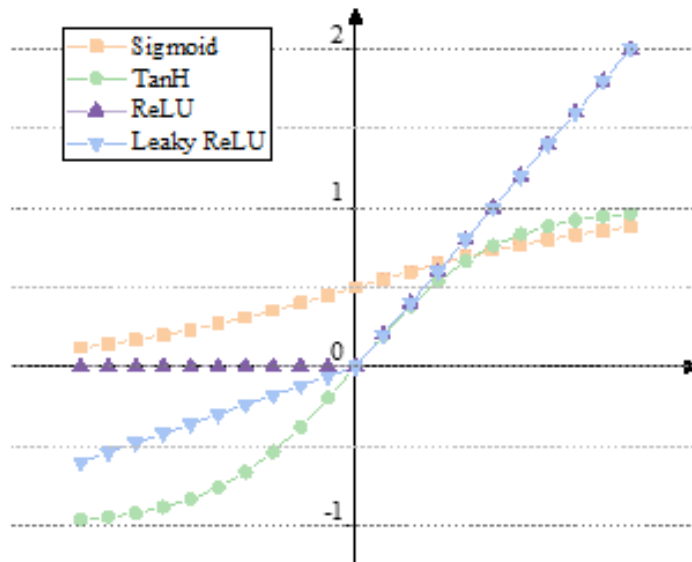
Fig. 3.3 Graphical visualization of different activation functions.

1. Sigmoid

    Output: $f(y) = \frac{1}{1+e^{-y}}$.

    Advantages: The output range is [0,1].

    Disadvantages: (a) The gradient vanishes easily. (b) The output is not zero-centered, which reduces the efficiency of updating weights. (c) The computation of the exponential function is slow and complex for computers.

2. TanH

    Output: $f(y) = \frac{e^y - e^{-y}}{e^y + e^{-y}}$.

    Advantages: It addresses the problem that the output faced by Sigmoid is not zero-centered.

    Disadvantages: It retains two problems of Sigmoid, i.e., gradient vanishing and the computation of the exponential function.

3. ReLU

    Output: $f(y) = max\{0, y\} = \begin{cases} 0, y \leq 0 \\ y, y > 0 \end{cases}$.

    Advantages: (a) It addresses the gradient vanishing problem (in the positive range). (b) The computation speed is very fast, and it only needs to judge whether the input is greater than zero. (c) The convergence speed is much faster than that of Sigmoid and TanH.

    Disadvantages: (a) The output is not zero-centered. (b) Some neurons may never be activated, causing the corresponding parameters to never be updated, and consequently their only output values is 0 (dying ReLU), the reasons for this problem are generally a bad parameter initialization and the high learning rate.

4. Leaky ReLU

    Output: $f(y) = max\{\alpha y, y\} = \begin{cases} y, y \geq 0 \\ \alpha y, y < 0 \end{cases}$.

    Advantages: (a) Neurons do not die. (b) Neurons do not saturate. (c) It shares the two advantages of ReLU, i.e., the computation and convergence are fast.

    Disadvantages: The parameter α needs to be manually set through prior knowledge.

Recently, researchers exploit Softmax as the activation function to estimate abundance maps. The output of Softmax is computed by:

$$f(\mathbf{y}_i) = \frac{e^{y_i}}{\sum_{i=1}^{C} e^{y_i}}. \tag{3.3}$$

One problem associated with the use of Softmax for spectral unmixing is that it does not produce sparse outputs, but researchers can add a sparse regularization to obtain an approximate sparse solution. Generally, ReLU and Softmax are the most commonly used activation functions in spectral unmixing.

3.2 Introducing Different Autoencoders

Autoencoder models [103, 27, 162, 67, 132, 74, 58] are popular in many fields due to their unsupervised learning characteristics. They learn data features by encoding the input data, then decode the feature data to obtain reconstructed data, and optimize the reconstructed data through the backpropagation algorithm to keep it close to the input data. From input to output, the autoencoder does not need to use any other information, and feature extraction can be done through the original data only, so the autoencoder is an unsupervised feature learning model.
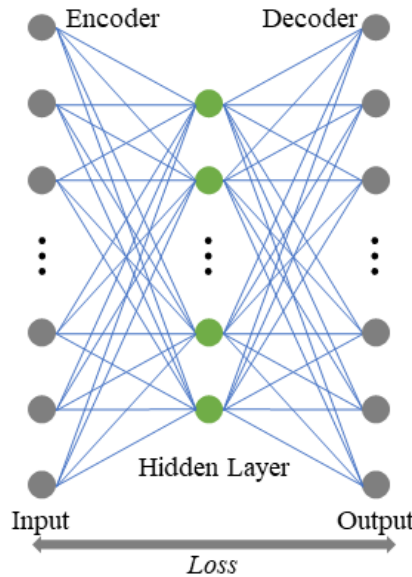
Fig. 3.4 Diagram of the traditional autoencoder.

3.2.1 Traditional Autoencoder

Traditional autoencoder model [129, 156] contains an encoder and a decoder, and its structure diagram is shown in Fig. 3.4. Details of each part are provided below.

Encoder

Assuming that one observed pixel can be expressed as $\mathbf{y}_i = \mathbf{M}\mathbf{a}_i + \mathbf{n}_i$, as indicated by Eq. (2.1), the encoder part compresses the input data $\mathbf{y}_i$ into a low dimensional representation $\mathbf{h}_i$ by:

$$\mathbf{h}_i = f_E(\mathbf{y}_i) = f_E(\mathbf{W}^{(e)} \cdot \mathbf{y}_i + \mathbf{b}^{(e)}) \tag{3.4}$$

where f_E is an activation function, e.g., the Sigmoid, TanH, and $\mathbf{W}^{(e)}$ and $\mathbf{b}^{(e)}$ are the weight and bias in the encoder part, respectively.

Decoder

The decoder part employs the low dimensional representation $\mathbf{h}_i$ as the input and applies a hidden layer to obtain an approximate reconstruction of the original input vector $\mathbf{x}_i$ by:

$$\hat{\mathbf{y}}_i = f_D(\mathbf{h}_i) = f_D(\mathbf{W}^{(d)} \cdot \mathbf{h}_i + \mathbf{b}^{(d)}) \tag{3.5}$$

where $\hat{\mathbf{y}}_i$ denotes the reconstructed pixel, and $\mathbf{W}^{(d)}$ and $\mathbf{b}^{(d)}$ represent the weight and bias of the decoder part, respectively. The goal of autoencoder training is to minimize the error between the i-th input and its corresponding output, thus, the loss function for a single sample is:

$$Loss = \frac{1}{2}||\mathbf{y}_i - \hat{\mathbf{y}}_i|| \tag{3.6}$$

When inputting N sample data, its loss function can be written as:

$$Loss = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{2}||\mathbf{y}_i - \hat{\mathbf{y}}_i|| \tag{3.7}$$

After the training process is over, the unmixing results are extracted by:

$$\begin{aligned} \text{Abundance estimation}: \mathbf{h}_i &\Rightarrow \hat{\mathbf{a}}_i \\ \text{Endmember extraction}: \mathbf{W}^{(d)} &\Rightarrow \hat{\mathbf{M}} \end{aligned} \tag{3.8}$$

where $\hat{\mathbf{a}}_i$ is the estimtated abundance vector of the i-th pixel $\mathbf{y}_i$, and $\hat{\mathbf{M}}$ is the estimated endmember matirx.
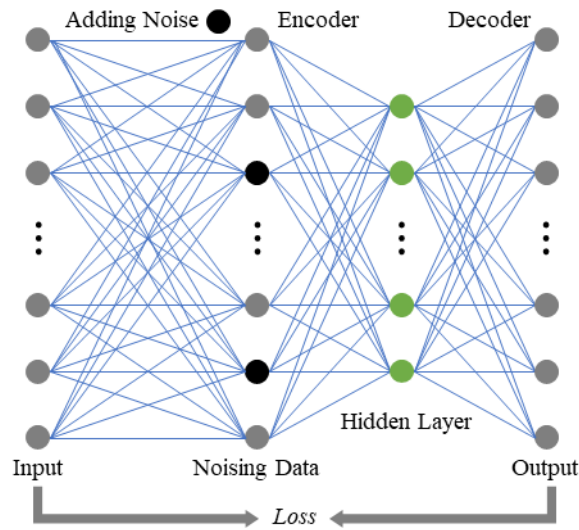
Fig. 3.5 Diagram of denoising autoencoder.

3.2.2 Denoising Autoencoder

In a large number of practical application scenarios, original data usually contain a certain degree of noise. If the noise data is used in the follow-up research analysis, it will bring great errors. In the process of image analysis, in order to restore and repair images affected by noise and ensure more accurate recognition of images, many researchers have improved traditional autoencoders and developed denoising autoencoders [145, 71, 104, 103, 73]. A simple denoising autoencoder is shown in Fig. 3.5.

The denoising autoencoders start from the robustness of the data. During the model training initialization stage, random noise is added to the original data to cause certain damage to the original data, and then the noise data is used for feature extraction and reconstruction. If the model is robust to noise, the extracted hidden layer features will be closer to those extracted from the original pure data. Thus, the reconstructed data using the hidden layer features will be closer to the original pure data. However, the denoising autoencoders require a random amount of noise to be added before each training, which leads to a significant increase in the running time of the model. Therefore, denoising autoencoders are frequently used on data that is heavily affected by noise.

3.2.3 Stacked Autoencoder

The data obtained in the real scene usually has some outliers. In this context the denoising autoencoder has little influence on the outliers and does not perform well on very few data with many outliers. In order to solve this problem, researchers propose stacked autoencoders
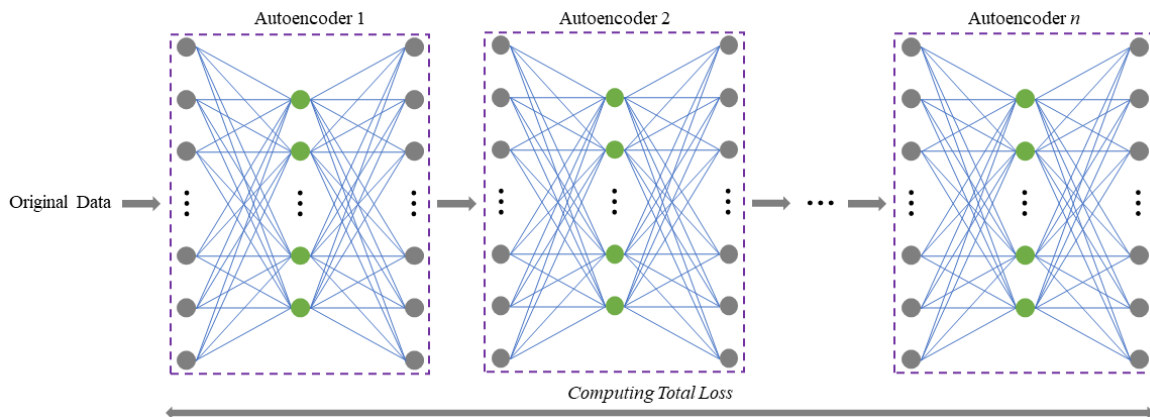
Fig. 3.6 Diagram of the stacked autoencoder.

[155, 133, 164, 102] to extract more accurate data features to obtain more realistic original data.

Fig. 3.6 shows a simple stacked autoencoder model, which is an autoencoder network model composed of multiple layers of traditional autoencoders, where each layer is implemented by a DNN and the number of nodes in each layer is usually different. Among them, the output of the hidden layer of the previous layer of autoencoder will be used as the input of the next layer of autoencoder, and the hidden layer of the last layer of autoencoder will be the result of the entire stacked autoencoder. Stacking training with multiple autoencoders can prevent certain randomness phenomena, so it has been widely used in object classification and feature extraction. In addition, since the number of stacked autoencoder layers may be large, the training time of each layer of the autoencoder needs to be considered, so its running cost is a big challenge. In practical applications, the stacked autoencoder is rarely used alone, and it is frequently used to initialize the original data and then provide it to other applications to improve the accuracy of the entire model.

3.2.4 Variational Autoencoder

Different from traditional autoencoders, variational autoencoder [158, 154, 22, 59, 68] is a generative network model that inherits the basic architecture of traditional autoencoders. The encoder and decoder of the traditional autoencoder correspond to the discriminative network and the generation network in the variational autoencoder, respectively. The hidden layer corresponds to the hidden variables in the variational autoencoder. The encoder is used to identify the model to obtain the variational probability distribution of low-dimensional features, whilst the decoder is employed to generate the model to obtain the probability distribution of the output.
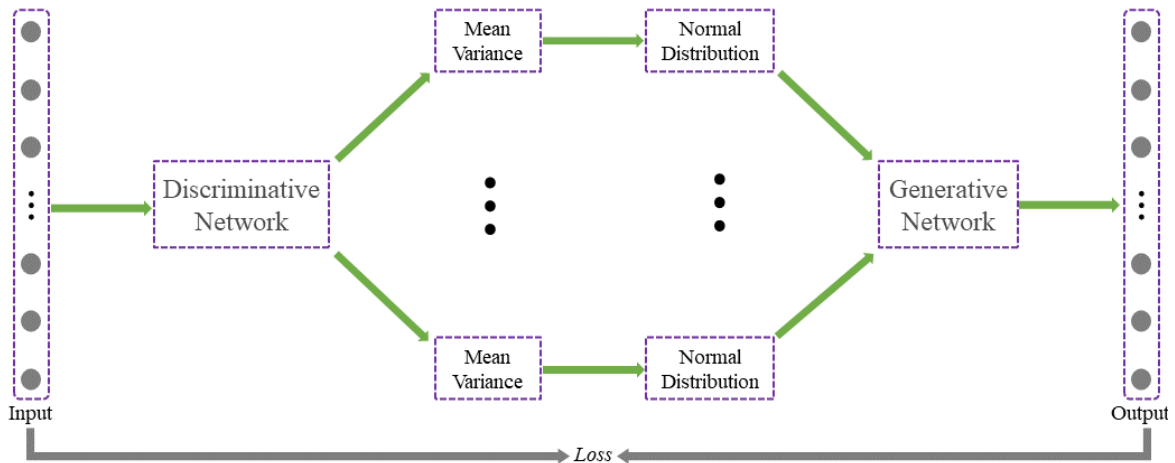
Fig. 3.7 Diagram of the variational autoencoder.

Fig. 3.7 depicts a typical variational autoencoder model, which first estimates the mean and variance of the original data by the discriminative network. Compared with the traditional autoencoder that uses an encoder to represent the hidden layer, the variational autoencoder uses a normal distribution to describe the probability distribution of hidden variables. The operation can make the hidden layer learn continuous and smooth features so that the features between adjacent elements in the hidden space will be more similar and more in line with a large number of actual situations.

3.2.5 Convolutional Autoencoder

CNN has achieved excellent results in a large number of fields. In particular, the combination of CNN and autoencoder derived convolutional autoencoder [30, 70, 81, 161, 11]. Same as the mechanism of traditional autoencoder, the convolutional autoencoder also obtains more accurate features of deep hidden layer data by reducing the reconstruction error. The difference is that the network structure of the traditional autoencoder usually uses a fully connected layer, which is not conducive to processing some multi-dimensional data, and the number of parameters of the fully connected layer in the autoencoder is more than that of convolutional layers. This makes training autoencoders difficult and time-consuming. Furthermore, the fully connected layer is a simple affine transformation after removing the bias term, projecting the data into a more separable space to simplify the estimation process. However, when the dimension of the data increases, the irregularity of the data will make it difficult to implement unsupervised methods. The convolutional autoencoder uses the convolutional layer and the pooling layer instead of the fully connected layer, which not only makes the learning times of the trainable parameters less but also retains the multi-
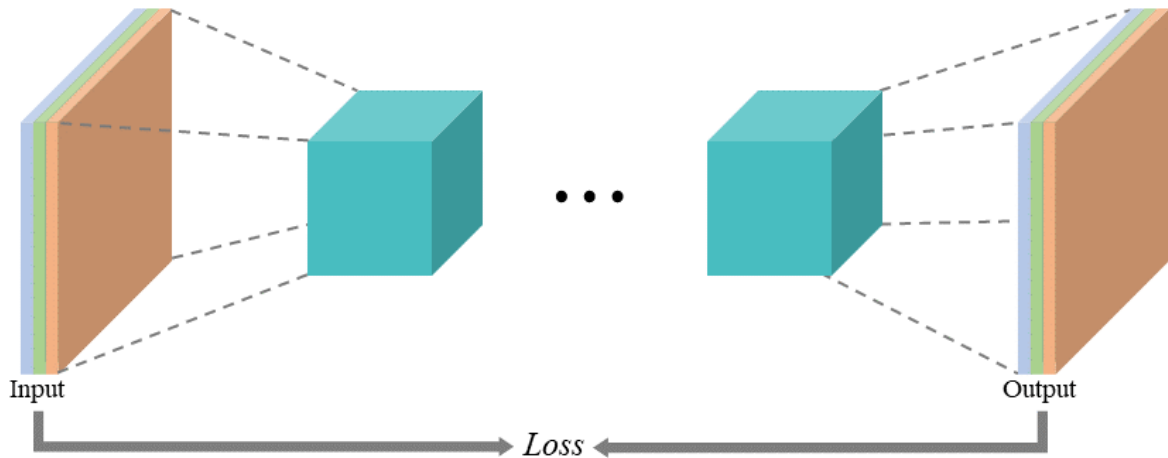
Fig. 3.8 Diagram of the convolutional autoencoder.

dimensional spatial information of the original data to a certain extent, thereby extracting more representative data features.

The structure of the convolutional autoencoder is shown in Fig. 3.8, and its encoding stage is the same as that of the CNN. They both use a convolution kernel of a certain size to scan the entire image, and the size of the feature data can be jointly determined by the size and step size of the convolution kernel. Usually, the convolution operation keeps the size of the data constant and extracts features by reducing the number of channels. The pooling layer is used for feature selection to compress the size of the data, thereby reducing the parameters of the network. In addition, a weight sharing strategy is introduced into the network to reduce parameters and increase its generalization ability.

3.3 DL-Based Unmixing Methods

After introducing some typical DL frameworks, e.g., autoencoder frameworks, in this sub-chapter, some typical unmixing methods based on DL will be reviewed to learn how DL can play a role in unmixing.

(1) Autoencoder Cascade Network

In [42], Guo *et al.* propose an autoencoder cascade (AEC) framework to address the spectral unmixing problem by concatenating a marginalized denoising autoencoder and a non-negative sparse autoencoder to consider the sparsity and non-negative properties. The proposed AEC is a neural network composed of multiple layers of autoencoders, where the output of a denoising layer is fed to the input of successive unmixing layers. The proposed AEC is

the first attempt that applies DL-based related methods to spectral unmixing applications, and it considers the observation noise and the sparsity prior to enhance the performance of unmixing for hyperspectral images with high noise. However, the framework also has some shortcomings, which are as follows.

- It employs the same encoder and decoder to perform unmixing, thus the accuracy of the estimated endmembers will be constrained by the usage of the tied-weights.

- The reconstruction error is high due to the errors introduced in the cascade layer and its element-wised way of optimizing the network.

- It only considers spectral information and ignores spatial information, and it cannot adaptively adjust according to different scenarios concerning various degrees of corruption as well as spatial correlation.

- It aims to address the linear unmixing problems, it thus fails to use the superior potential of neural networks to address nonlinear unmixing problems. Especially, the linear unmixing problems can be easily performed by some traditional unmixing methods.

- The strength of the proposed AEC is antinoise. However, it will have strong limitations in the case of outliers. The reason is that outliers may cause initialization failure, and the existence of outliers will strongly interfere with the unmixing results.

To improve AEC, in [103], Qu *et al.* propose a part-based denoising autoencoder with an unique structure to address the unmixing problem for hyperspectral images with high noise. The framework applies the $l_{2,1}$ norm and denoising constraints on the network to better handle noise and reduce the redundant endmembers and the reconstruction error simultaneously. The main advantages of improved AEC are: (a) the framework optimizes the network by incorporating denoising capability into it in the form of a denoising constraint, instead of cascading another denoising preprocessor to avoid introducing additional reconstruction errors, (b) a part-based autoencoder integrates a denoising part to denoise the input data and estimates the number of endmembers in the hyperspectral image by automatically removing redundant hidden nodes, and (c) the framework proposes a back-propagation method to effectively force the endmembers to be non-negative and the abundance vector to be sum-to-one by using simple matrix operations. However, the improved AEC also only considers spectral information and is suitable for addressing linear unmixing problems. Moreover, it initializes the weight of the decoder with endmembers estimated by VCA. In this sense, the network is not stable enough due to the different number of endmembers and random initialization of VCA.

(2) Stacked Nonnegative Sparse Autoencoder and Deep Autoencoder Networks

In [114], Su *et al.* present a stacked nonnegative sparse autoencoder (NNSAE) framework, namely stacked NNSAE (SNSA), for spectral unmixing to handle the outliers and low noise-signal ratio contained in the data. The framework first uses a set of NNSAE to detect the outliers contained in the data and then employs a final autoencoder to extract endmembers and estimate their corresponding fractional abundances at each pixel. The framework considers the sparsity and non-negative properties in the abundance vector to improve the performance of unmixing.

Given that SNSA performs well in unmixing, in [113] and [112], Su *et al.* propose a new unsupervised unmixing technique based on a deep autoencoder network (DAEN). Firstly, DAEN employs stacked autoencoders to learn the spectral signatures to generate a good initialization for their designed network. Secondly, it uses a variational autoencoder to extract endmembers and estimate abundance maps. DAEN mainly has two advantages: (a) the framework takes the advantage of non-negative sparse self-encoding to address the problems of the outliers and low noise-signal ratio contained in the data to obtain good unmixing results, and (b) the framework employs a variational autoencoder to consider non-negative and sum-to-one constraints to obtain good results of unmixing. However, SNSA and DAEN are designed for linear spectral unmixing problems, and they both only consider spectral information but ignore spatial information and may not have good results for the linear spectral unmixing problems. Moreover, they also employ VCA as the initialization method, and their performance will be affected by the number of endmembers and random initialization of VCA.

(3) Traditional and Multitask Autoencoder Unmixing Networks

In [93], Palsson *et al.* propose a DL-based framework, namely traditional autoencoder unmixing (TAU), for blind hyperspectral unmixing[1] in the form of an asymmetrical autoencoder. Unlike other unmixing methods based on autoencoders, the encoder of TAU is tested using different activation functions and it is evaluated using different objective functions. The proposed TAU contains four fully connected layers as the hidden layers to complete unmixing. The strong points of the TAU framework are summarized below:

- The proposed TAU uses a custom layer and weight constraints to force the abundance to be non-negative and sum-to-one.

---

[1]https://github.com/dv-fenix/HyperspecAE

- The proposed TAU considers the sparsity in the abundance vector by adopting an adaptive thresholding form optimized for the objective function of the network.

- Compared with the traditional methods, the framework does not have parameters to be tuned except for estimating the number of endmembers.

- The proposed TAU shows that a more deep and complex encoder does not give significantly better results than a single-layer encoder. This is very useful during the design process of unmixing network architectures.

- The proposed TAU is tested using different objective functions, which can help researchers choose a good objective function to train their framework in the future.

Despite TAU having many advantages, it is not difficult to find that it still only exploits spectral information and fails to utilize spatial information. Moreover, it is designed to address the linear spectral unmixing problems. In [94], Palsson *et al.* improve TAU and present a novel multitask autoencoder unmixing (MTAEU) framework. The proposed method is inspired by multitask learning and employs many autoencoders to decompose a whole neighborhood of pixels simultaneously to learn spatial information. Especially, multitask learning has these advantages, e.g., incorporating spatial information, improving stability, reducing the risk of overfitting, and accelerating learning. Compared with [93], the advantages and improvements of MTAEU are:

- MTAEU exploits the spatial information by operating on image patches instead of a single pixel. Importantly, it is the first method of its kind to directly use the spatial structure of the data to estimate the signatures of endmembers.

- MTAEU uses Softmax as the activation function to enforce the abundance to satisfy the constraint of sum-to-one.

- MTAEU makes better use of batch normalization and dropout than TAU.

However, MTAEU pays more attention to the pixel-level reconstruction loss, and some detailed physical information about certain materials will be lost in the reconstruction process.

(4) Deep Nonlinear Autoencoder Network

In [136], Wang *et al.* propose a deep nonlinear autoencoder (DNAE) for hyperspectral unmixing. The proposed DNAE completes endmember extraction and abundance estimation in a tractable manner. Moreover, DNAE employs a pretraining technique to improve the

performance of unmixing. Particularly, the decoder of the framework contains a linear part and a nonlinear part. The linear part denotes the linear model of spectral unmixing, and its corresponding weights are endmembers that need to be extracted. The nonlinear part denotes nonlinear iterations between endmembers. Moreover, the proposed DNAE framework proposes a decoder network as the combination of endmember matrix with nonlinear neural network layer to learn postnonlinear mixtures, and it only depends on the current data without conducting additional training.

Nonetheless, in addition to the two shortcomings of only considering the spectral information and the performance will be affected by VCA, the proposed DNEA also faces the following two defects:

- It models the post-nonlinearity by the decoder part of the autoencoder. The effectiveness of the decoder is guaranteed by pre-training and learning rate adjustment techniques. Moreover, the nonlinear model denoted by the decoder cannot cover many nonlinear situations due to it is not sufficiently ordinary.

- It takes into account the nonlinear effects and uses a multilayer decoder to reconstruct the data. However, due to the constraints on the code and the reconstructed data being insufficient, the decoder easily reconstructs the data and is insensitive to the codes, which will inevitably lead to the instability and uncertainty of unmixing.

(5) Weakly-supervised and Deep Half-siamese Networks

In [54], Hong *et al.* propose a weakly-supervised unmixing network (WU-Net) based on DL. The proposed WU-Net employs an endmember extraction algorithm (e.g., VCA) and an abundance estimation algorithm (e.g., FCLS) to address the problem that most unmixing networks based on DL tend to obtain trivial unmixing results due to the lack of the ability to handle spectral variability and extract physically meaningful endmembers. WU-Net is a two-stream deep network that contains an endmember network and an unmixing network. The proposed WU-Net learns the pure endmember network or nearly-pure endmembers and uses the parameters of the endmember network to correct the weights of the unmixing framework to get accurate and interpretable unmixing performance. Simultaneously, the proposed WU-Net also contains a regularization term to avoid overfitting. In this scenario, WU-Net gets more physically meaningful unmixing results compared to other unmixing networks based on DL. However, WU-Net introduces an endmember network to correct the parameters of an unmixing network, and its performance will be affected by the results of endmember extraction and abundance estimation used in the endmember network.

In [45], Han *et al.* present a novel deep half-siamese network (Deep HSNet) for spectral unmixing. The motivation of the Deep HSNet is to address the issue that most existing networks based on DL lack the ability of considering physically meaningful endmember information. Unlike [54] which utilizes the result of single endmember extraction algorithm as candidate endmembers, the Deep HSNet considers the results of many endmember extraction algorithms (e.g., VCA, N-FINDR, and PPI) to improve the performance of unmixing. The Deep HSNet contains an endmember-guided network and a reconstruction network. The endmember-guided network learns the endmember representations to map the extracted endmembers to their corresponding abundances. The abundances are acquired by the LMM-based method [57]. In [45], the proposed Deep HSNet uses the scaled constrained least squares (SCLS) method to get the abundance maps. Especially, the reconstruction network is a simple autoencoder architecture that is designed to reconstruct the HSI data. To embed the physical meaning of endmember information into the reconstruction network, the endmember-guided network and the encoder part of the reconstruction network share the same weights. In this sense, the framework gets more physically meaningful unmixing results, and it is the first one that considers the results of different endmember extraction algorithms to embed the endmember information into the deep unmixing network. However, the framework keeps all disadvantages of [54]. Moreover, although it considers the results of many endmember extraction algorithms, its performance is still limited by the results of the endmember extraction and abundance estimation.

(6) Deep Multitask Bilinear Unmixing Network

In [115], considering that nonlinear models need detailed knowledge about the physical interactions between sunlight scattered by multiple materials and the bilinear mixed model can achieve good results via a relatively simple scattering model, Su *et al.* present an unsupervised deep multitask bilinear unmixing (DMBU) framework. Especially, the proposed DMBU framework mainly contains two tasks: (i) using a deep autoencoder to estimate the endmember signatures and their corresponding abundances at every pixel, and (ii) adopting another deep autoencoder to simultaneously update bilinear components. It is worth noting that the two deep autoencoders are trained in a mutually interdependent manner under multitask learning framework, and the stopping criterion is based on the relative reconstruction error. Moreover, to learn the hidden representations from the original HSI data and further optimize the performance of unmixing, the framework also integrates some aspects of deep semi-NMF [128] into the aforementioned autoencoders. However, the proposed DMBU framework does not consider spatial information and spectral variability, and it is designed to address the linear unmixing problems and cannot guarantee accurate results for nonlinear

unmixing problems. Moreover, the proposed DMBU framework obtains endmember initialization by VCA, and the stability of DMBU will be affected by the number of endmembers and the random initialization of VCA.

(7) Cycle Consistency Unmixing Network

In [35], since the traditional autoencoder-based architectures usually pay attention to the problem of the pixel-level reconstruction loss, which will easily lead to the loss of some significant detailed information of certain materials, Gao *et al.* present a cycle consistency unmixing network (CyCU-Net)[2] to reduce the loss by relaxing the original pixel-level reconstruction assumptions to the cyclic consistency dominated by two cascaded autoencoders. The cyclic consistency is obtained by the newly proposed self-perception loss, which contains two spectral reconstruction items and one abundance reconstruction item. The high-level semantic information is preserved and the effectiveness of the CyCU-Net unmixing framework is improved by learning the two cascaded autoencoders in an end-to-end manner. In this scenario, the proposed CyCU-Net framework has many good points:

- The proposed CyCU-Net estimates the abundance maps more efficiently and effectively in comparison with other methods and obtains excellent performance in unmixing.

- The proposed CyCU-Net is the first framework that embeds high-level semantic information into a DL-based unmixing network.

- The performance of the proposed CyCU-Net is evaluated by extensive ablation studies, i.e., different convolution filter sizes, abundance constraints, objective functions, and more complex unmixing scenes.

Although CyCU-Net has so many advantages, it also faces the problem of neglecting the features of the spectral dimensions.

(8) Pixel-based CNN and Cube-based CNN

In [160], Zhang *et al.* present an end-to-end spectral unmixing framework[3], which contains a pixel-based CNN and cube-based CNN. The proposed framework uses a CNN architecture with two stages, which are used to extract features and map them from the extracted features to obtain the abundance maps. Moreover, pixel-based CNN and cube-based CNN both adopt dropout to avoid overfitting. It is worth noting that the framework is the first strategy

---

[2]https://github.com/danfenghong/IEEE_TGRS_CyCU-Net
[3]https://web.xidian.edu.cn/xrzhang/paper.html

that addresses spectral unmixing using a deep CNN, and it can automatically learn and interpret discriminative feature representations. Furthermore, the framework is a simple but fast end-to-end model, which can avoid manual preprocessing and subsequent processing. Simultaneously, the framework uses patches to consider both the spectral and the spatial information and further enhances the performance in unmixing.

However, the framework still exists many problems:

- The framework generates labels by theoretical models, which may not match the real scene and has low generalizability.

- The framework needs a large amount of data with known ground truth to complete the process of training and these data are usually generated by other unmixing methods.

- The framework only pays attention to the local area and does not extract the non-local features of the original data.

(9) Spectral Convolution Network With Spectral Library

In [101], Qi *et al.* present a deep spectral convolution network with spectral library called SCNL to perform the unmixing task. The proposed SCNL framework uses the convolutional layers to extract features and then estimates the abundance maps via fully connected layers, and finally employs another fully connected layer to obtain reconstructed pixels. The proposed SCNL framework allows the network after training to be used for multiple HSIs instead of a single one, and the framework drops the waste of resources and enhances the utilization of the network. The endmember signatures from the spectral library added to the framework also improve the representation capabilities of the network. Moreover, the framework uses spectral convolution to learn spectral features to make full use of the spectral information of the original data and employs a deeper network to merge the spectral library because it requires capturing more features, and the deeper network acquires better feature representations and better estimation results. Meanwhile, SCNL develops a new loss that contains pixel reconstruction error, abundance sparsity, and abundance cross-entropy to train the network.

(10) Nonsymmetric Autoencoder Network

In [162], Zhao *et al.* present a novel nonsymmetric autoencoder network (NAN) based on the long short-term memory network (LSTM) and DNN for spectral unmixing, and the framework uses the general modeling capabilities of DNNs to learn the nonlinear relationships from the original HSI data. The proposed NAN framework employs LSTM structure to learn spectral

correlation information and uses spatial regularization to enhance the spatial continuity of results. As an RNN architecture, LSTM allows the memory of the previous step to remain in the current state, thereby affecting the output of the current step. The proposed NAN framework also uses an attention mechanism to further improve the performance of unmixing. In this case, the framework has many advantages and improvements:

- The proposed NAN framework employs an RNN-based network to process HSI as sequential data. This is the first time that RNN is used to address the blind unmixing problem.

- The proposed NAN framework uses LSTM-DNN to perform nonlinear tasks. The output and weight of a particular layer of the framework are used to extract endmembers and estimate abundance maps, respectively.

- The proposed NAN framework uses an attention mechanism to selectively focus the input part during the learning process and map the input sequence to the required dimensions.

- The proposed NAN framework imposes the total variation regularization to integrate the spatial information in the original data. By using the total variation regularization, the framework can learn the piece-wise smooth structure of the abundance maps and further improve the performance of unmixing.

3.4 Summary and Thinking of DL-Based Unmixing Methods

In this chapter, this thesis has reviewed many typical frameworks based on DL for hyperspectral unmixing. After analyzing their corresponding frameworks and principles, the main problems are summarized as follows.

- They usually use fully connected layers to construct autoencoder networks for hyperspectral unmixing. This way of building the network is commonly used to reconstruct the spectrum to complete the reconstruction of the entire image. It is not difficult to find that the operation can only capture spectral information but ignore spatial information. Nevertheless, HSIs usually contain rich and detailed spectral and spatial information, which is very significant for processing HSIs. In this sense, exploiting a new network for capturing spectral and spatial information simultaneously is very vital for enhancing the performance of hyperspectral unmixing.

- Most of them employ a liner activation function for reconstructing the original image, and endmember signatures and their abundance maps are extracted from the output and weight of a particular layer, respectively. However, most HSIs are generally nonlinear mixed. In this sense, how to add nonlinearity to the corresponding networks is crucial for further improving their corresponding networks.

- Most existing networks employ VCA to initialize the weight of a particular layer to obtain accurate results in hyperspectral unmixing. Notwithstanding its results, VCA also faces the problem of random initialization. In this sense, the performance of unmixing will be affected by VCA. To this end, how to find the optimal way to initialize the weights plays a vital role in hyperspectral unmixing.

Chapter 4

Contributions of This Thesis to The Scientific Field

4.1 Motivations

In Chapter 2, this thesis has reviewed some of the most popular traditional hyperspectral unmixing methods, most of them complete endmember extraction and abundance estimation separately, and then they search for a good combination of the two processes of algorithms to yield the best performance in hyperspectral unmixing. Another strategy is to use NMF to extract endmember signatures and estimate their corresponding abundance maps simultaneously. Whether seeking the optimal combination of the two tasks or directly accomplishing the unmixing, these two hyperspectral unmixing techniques both impose considerable computational burdens, making them inefficient when dealing with large-scale remote sensing images. In this sense, to aim the problem, the work [121] has been proposed for completing endmember extraction and abundance estimation at the same time.

In Chapter 3, some typical unmixing methods based on DL have been introduced. By reviewing these methods, it is easy to obtain that most of these methods use autoencoder-based networks to reconstruct the original image. In this sense, they only learn spectral information and neglect spatial information. But the spatial information is very significant for HSI processing. Recently, CNN has been developed for HSI processing due to its powerful fitting and feature extraction capabilities. Nonetheless, most hyperspectral unmixing methods are based on 2D CNN architectures, thus they cannot learn good discriminative feature map from a spectral perspective. In addition, although 3D CNN has the ability to extract spectral and spatial features, modes only using 3D CNN bring high computational overloads. To learn spectral and spatial information at the same time and perform hyperspectral unmixing

effectively and efficiently, the work [122] has been developed to combine 2D CNN and 3D CNN for hyperspectral unmixing.

Furthermore, most traditional endmember extraction techniques need to know the number of endmembers in advance, and the algorithms for counting endmembers tend to be characterized by independent methods that may not seamlessly benefit the endmember extraction algorithms. In this sense, how to develop a new framework to bridge these two seemingly independent but actually highly related procedures is the goal of this thesis. To this end, three works [119, 124, 123] have been developed to determine the number of endmembers and extract their corresponding spectral signatures simultaneously.

In general, this thesis is dedicated to presenting effective and efficient frameworks for addressing these three challenges associated with hyperspectral unmixing, i.e., determining the number of endmembers, extracting endmember signatures, and estimating the abundances of each endmember at every pixel, and five novel frameworks have been developed. A brief description of how they accomplish unmixing is given below.

4.2 Fast Orthogonal Projection

Most existing hyperspectral unmixing techniques treat endmember extraction and abundance estimation as two independent tasks, and the result of endmember extraction will greatly affect the performance of abundance estimation. Another strategy is to exploit NMF to simultaneously complete the two independent tasks. However, the two strategies both will introduce high computational overloads, making them inefficient when dealing with large-scale remote sensing images. To address this issue, this work entitled *Fast Orthogonal Projection for Hyperspectral Unmixing* [121] develops a novel orthogonal projection framework to conduct fast hyperspectral unmixing. It addresses both endmember extraction and abundance
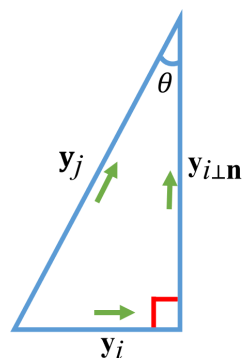

Fig. 4.1 Schematic diagram of orthogonal projection calculation (from [121] ©2022 IEEE).

estimation with OPE and OPA, respectively. Especially, OPE needs to use some techniques to filter out the noise contained in HSIs in advance.

4.2.1 Endmember Extraction

As shown in Fig. 4.1, $\mathbf{y}_i$ and $\mathbf{y}_j$ represent two pixels of HSI, $\mathbf{y}_{i\perp\mathbf{n}}$ denote the orthogonal projection vector of $\mathbf{y}_j$ to $\mathbf{y}_i$, and the green arrow indicates the direction of pixels. $\mathbf{y}_{i\perp\mathbf{n}}$ can be obtained by:

$$
\begin{aligned}
||\mathbf{y}_{i\perp\mathbf{n}}||_2 &= ||\mathbf{y}_j||_2 |\cos\theta| \\
&= ||\mathbf{y}_j||_2 |\frac{\mathbf{y}_{i\perp\mathbf{n}} \cdot \mathbf{y}_j}{||\mathbf{y}_{i\perp\mathbf{n}}||_2 ||\mathbf{y}_j||_2}| \\
&= |\frac{\mathbf{y}_{i\perp\mathbf{n}} \cdot \mathbf{y}_j}{||\mathbf{y}_{i\perp\mathbf{n}}||_2}|
\end{aligned}
\tag{4.1}
$$

where $||\cdot||_2$ denotes the L_2 norm, assuming $\mathbf{n}$ is the normal vector of $\mathbf{y}_i$ and $\mathbf{n}_{\text{unit}}$ is the unit vector of $\mathbf{n}$, the Eq. (4.1) is transformed into:

$$
||\mathbf{y}_{i\perp\mathbf{n}}||_2 = |\mathbf{n}_{\text{unit}} \cdot \mathbf{y}_j|
\tag{4.2}
$$

OPE first uses Eq. (4.2) to find the index of the maximum value $\mathbf{y}_{i\perp\mathbf{n}}$ to extract the first endmembers. Assuming that the endmember matrix is $\mathbf{M}_{k-1} = [\mathbf{m}_1, \cdots, \mathbf{m}_{k-1}] \in \mathbb{R}^{L\times(k-1)}$ and $\mathbf{M}_{k-1}$ denotes the extracted $(k-1)$-th endmembers, to extract the k-th endmember, the normal vector $\mathbf{n}_k$ of $\mathbf{M}_{k-1}$ first is obtained by:

$$
\mathbf{M}_{k-1}^{\text{T}} \mathbf{n}^k = 0
\tag{4.3}
$$

and the unit vector $\mathbf{n}_{\text{unit}}^k$ of $\mathbf{n}^k$ is computed by:

$$
\mathbf{n}_{\text{unit}}^k = \frac{\mathbf{n}^k}{||\mathbf{n}^k||_2}
\tag{4.4}
$$

The index $\hat{\mathbf{I}}_k$ of the k-th endmember is obtained by:

$$
\hat{\mathbf{I}}_k = \underset{\mathbf{I}_k \in \{1,2,\cdots,N\}}{\arg\max} |\mathbf{n}_{\text{unit}}^k \cdot \mathbf{y}_{\mathbf{I}_k}|
\tag{4.5}
$$

and endmember matrix is updated as $\mathbf{M}_k = [\mathbf{m}_1, \cdots, \mathbf{m}_{k-1}, \mathbf{m}_k] \in \mathbb{R}^{L\times k}$. OPE repeats the above process to extract all endmembers.

4.2.2 Abundance Estimation

Under the LMM, the data matrix $\mathbf{Y}$ can be represented by the multiplication of the endmember matrix $\mathbf{M}$ and the abundance matrix $\mathbf{A}$ by:

$$\mathbf{Y} = \mathbf{MA} \tag{4.6}$$

In general, the abundance matrix $\mathbf{A}$ can be acquired by:

$$\mathbf{A} = \mathbf{M}^{-1}\mathbf{Y} \tag{4.7}$$

and $\mathbf{M}^{-1}$ is the inverse matrix of $\mathbf{M}$. In general, it is unrealistic to compute abundance matrix $\mathbf{A}$ by Eq. (4.7) because the endmember matrix $\mathbf{M}$ is usually not a square matrix. In this scenario, OPA exploits a diagonal matrix $\mathbf{Q}$ ($\mathbf{P}^{\mathrm{T}}\mathbf{M} = \mathbf{Q}$) to complete the task of abundance estimation. Especially, OPA acquires $\mathbf{p}_i$ by:

$$[\mathbf{m}_1, \mathbf{m}_2, \cdots, \mathbf{m}_{i-1}, \mathbf{m}_{i+1}, \cdots, \mathbf{m}_c]^{\mathrm{T}}\mathbf{p}_i = \mathbf{0} \tag{4.8}$$

The final $\mathbf{p}_i$ is obtained by summing all the solutions obtained by Eq. (4.8). It is easy to find the following two facts:

$$\mathbf{m}_i \cdot \mathbf{p}_i \neq 0 \tag{4.9}$$

and

$$\mathbf{m}_i \cdot \mathbf{p}_j = 0, i \neq j \tag{4.10}$$

The diagonal matrix $\mathbf{Q}$ is obtained by

$$\mathbf{Q} = \mathbf{P}^{\mathrm{T}}\mathbf{M} \tag{4.11}$$

Eqs. (4.9) and (4.10) reflect $Q_{ii} = \mathbf{m}_i \cdot \mathbf{p}_i \neq 0$, and $Q_{ij} = \mathbf{m}_i \cdot \mathbf{p}_j = 0$. The fact shows that the diagonal matrix $\mathbf{Q}$ has inverse matrix $\mathbf{Q}^{-1}$. In this sense, Eq. (4.6) is transferred into:

$$\begin{aligned} \mathbf{P}^{\mathrm{T}}\mathbf{Y} &= \mathbf{P}^{\mathrm{T}}\mathbf{MA} \\ &= \mathbf{QA} \end{aligned} \tag{4.12}$$

Obviously, the abundance matrix can be easily computed by

$$\mathbf{A} = \mathbf{Q}^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{Y} \tag{4.13}$$

In general, ANC and ASC need to be imposed on the abundance matrix **A**. The related results show that all negative values appearing in **A** obtained by the above process are really close to zero. Therefore, OPA meets the ANC constraint by setting all negative abundances to zero or its opposite value, and then it divides each abundance vector by its sum to ensure that the abundance estimated by the above process meets the ASC constraint. The contributions of this work [121] are as follows:

- This work establishes a novel orthogonal projection framework for simultaneously completing endmember extraction and abundance estimation, and it exhibits effectively and efficiently in hyperspectral unmixing.

- The proposed OPE algorithm extracts endmembers by utilizing the concept of the pixel with the maximum orthogonal projection on any pixel as an endmember. Since it only involves straightforward vector-based computations, its computational complexity is very low.

- OPA estimates the abundance matrix by computing the multiplication of matrices. In this scenario, it has high efficiency in terms of estimating the abundance of endmembers at every pixel.

4.3 3D Convolution Network and Cross Convolution Network

Due to the impressive learning and data fitting capabilities of CNNs, DL-based hyperspectral unmixing techniques have rapidly developed in the literature. Existing unmixing methods usually only employ 2D convolution kernels to extract features, and most of them cannot learn good discriminative feature maps from spectral information. In this sense, 3D CNN has drawn more attention due to its ability to simultaneously extract spectral and spatial features. Especially, 2D CNN can also capture spectral and spatial information by adjusting the number of filters and the size of kernels, but 3D CNN uses 3D convolutional kernels that can learn spectral and spatial information more flexibly. However, only using 3D convolution is significantly more complex and expensive in computational overloads. To this end, this work entitled *A New Deep Convolutional Network for Effective Hyperspectral Unmixing* [122] proposes `CrossCUN` that combines 3D and 2D convolutions to simultaneously extract spectral and spatial information and perform hyperspectral unmixing more effective and efficient. Especially, `CrossCUN` builds 3D patches to further improve the performance of unmixing.
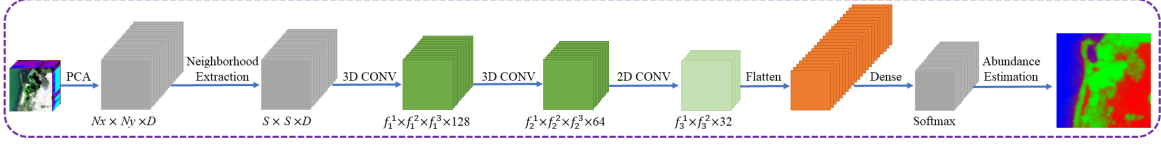
Fig. 4.2 The framework of the proposed `CrossCUN` (from [122] ©2022 IEEE).

Let $\mathbf{Y} \in \mathbb{R}^{Nx \times Ny \times Nb}$ denote HSI data, where Nx is the width, Ny is the height, and Nb is the number of spectral bands. Let $\mathbf{X} \in \mathbb{R}^{Nx \times Ny \times D}$ be the reduced data by PCA (in our work, D is empirically set to 15).

On the one hand, `CrossCUN` involves 2D convolutions, and the convolved features employ an activation function to introduce nonlinearity in the model, and its activation value $v_{i,j}^{x,y}$ at spatial position (x,y) in the j-th feature map of the i-th layer is obtained by:

$$v_{i,j}^{x,y} = \varphi \left(b_{i,j} + \sum_{\tau=1}^{d_{i-1}} \sum_{\beta=-\rho}^{\rho} \sum_{\alpha=-\gamma}^{\gamma} w_{i,j,\tau}^{\alpha,\beta} \times v_{i-1,\tau}^{x+\alpha,y+\beta} \right) \tag{4.14}$$

where φ represents the activation function, $w_{i,j}$ denotes the value of the weight parameter for the j-th feature map of the i-th layer, $b_{i,j}$ denotes the bias for the j-th feature map of the i-th layer, d_{i-1} is the number of feature maps in the $(i-1)$th layer, and $2\rho+1$ and $2\gamma+1$ are the width and height of kernel, respectively.

On the other hand, `CrossCUN` utilizes 3D convolutions, the activation function is introduced to learn the non-linear features, and the activation value $v_{i,j}^{x,y,z}$ at spatial position (x,y,z) in the j-th feature map of the i-th layer is obtained by:

$$v_{i,j}^{x,y,z} = \varphi \left(b_{i,j} + \sum_{\tau=1}^{d_{l-1}} \sum_{\lambda=-\eta}^{\eta} \sum_{\beta=-\rho}^{\rho} \sum_{\alpha=-\gamma}^{\gamma} w_{i,j,\tau}^{\alpha,\beta,\lambda} \times v_{i-1,\tau}^{x+\alpha,y+\beta,z+\lambda} \right) \tag{4.15}$$

where $2\eta+1$ is the depth of the convolution kernel, respectively.

Fig. 4.2 shows the framework of the proposed `CrossCUN`. From Fig. 4.2, it can be observed that the proposed `CrossCUN` has two 3D convolutional layers, one 2D convolutional layer, one flatten layer, and one dense layer. The kernel sizes of the two 3D convolutional layers are $3 \times 3 \times 7$ (with $f_1^1 = f_1^2 = 3$ and $f_1^3 = 7$), and $3 \times 3 \times 5$ (where $f_2^1 = f_2^2 = 3$ and $f_2^3 = 5$). The kernel size of the 2D convolutional layer is $f_3^1 = f_3^2 = 3$. To satisfy ASC and ANC, SoftMax is employed as the activation function of the last dense layer, following

$$\hat{a}_{ij} = \frac{e^{z_j}}{\sum_{j=1}^{c} e^{z_j}} \tag{4.16}$$

where $\hat{a}_{ij}$ denotes the estimated abundance of the j-th endmember on the i-th pixel, z_j represents the output value of the dense layer, and c is the number of endmembers. The cross-entropy is used as the loss function to train `CrossCUN`, and it can be computed by:

$$L = -\frac{1}{N} \sum_{i=1}^{N} [\mathbf{a}_i \log \hat{\mathbf{a}}_i] \tag{4.17}$$

where $\hat{\mathbf{a}}_i$ is the abundance estimated for the i-th pixel, $\mathbf{a}_i$ represents the real abundance vector of the i-th pixel, and N is the number of pixels. The contributions of the proposed `CrossCUN` are listed below:

- The proposed `CrossCUN` can not only complete the abundance estimation task directly without any prior knowledge about endmember extraction but also obtain more accurate results than DL-based unmixing methods and traditional unmixing methods.

- Most existing DL-based methods usually use autoencoder networks or 2D convolutional networks to accomplish abundance estimation, and the proposed `CrossCUN` is the first one to exploit cross convolution combining 2D and 3D convolutions to perform abundance estimation.

- The proposed `CrossCUN` shows the potential of cross convolution compared to using 2D or 3D convolution alone in abundance estimation and obtains better results than other traditional and DL-based methods.

4.4 Divergent Subset

The accuracy of estimating the number of endmembers will have an impact on the results of endmember extraction. The majority of hyperspectral unmixing approaches typically take into account the number of endmembers as previous information to complete endmember extraction. To this end, the work entitled *Simultaneously Counting and Extracting Endmembers in a Hyperspectral Image Based on Divergent Subsets* [119] proposes a new DS method for counting and extracting endmembers simultaneously. The proposed DS method considers the subset of pixels that are most different from each other to extract endmembers, and the cardinality of the subset is the number of endmembers.

This thesis obtains the DS by the following proposition:

Proposition 1: The DS is computed in terms of solving the following quadratic optimization problem:

$$\hat{\mathbf{y}} = \arg\max_{\mathbf{y}} \frac{1}{2}\mathbf{y}^{\mathrm{T}}\mathbf{D}\mathbf{y}$$
$$\text{subject to} \quad y_i \geq 0; \quad \sum_{i=1}^{N} y_i = 1. \tag{4.18}$$

The non-zero elements in the optimal $\hat{\mathbf{y}} \in \mathbb{R}^{N \times 1}$ represent the DS. It is important to note that the solution y_i of the optimization problem Eq. (4.18) not only extracts endmembers according to its non-zero elements but also expresses the number of endmembers by counting the non-zero elements. Therefore, the proposed DS method can simultaneously count and extract endmembers. Especially, the proof of Proposition 1 can be found in [97].

Although Proposition 1 offers a DS in a principled manner, it is also challenging to come up with an analytical solution. To address the problem, this thesis follows the strategy of replicator dynamics in evolutionary game theory [138] to numerically calculate the solution to the optimization problem (4.18) as follows:

$$y_i^{\mathrm{new}} = y_i \frac{(\mathbf{D})_i^{\mathrm{T}} \mathbf{y}}{\mathbf{y}^{\mathrm{T}}\mathbf{D}\mathbf{y}}. \tag{4.19}$$

where the i-th column of $\mathbf{D}$ is represented by $(\mathbf{D})_i$. Iteratively updating in accordance with Eq. (4.18) makes $\mathbf{y}$ converge to the ideal value. The Baum-Eagon inequality [9] can guarantee the convergence of the iterative computation.

Also, the proposed DS method can be utilized to enhance the current endmember extraction methods. The number of endmembers M in an HSI is typically taken into consideration as prior knowledge when applying the endmember extraction techniques. Based on the information, the endmembers of M are identified by the endmember extraction methods. If an estimated number of endmembers $M_l < M$ needs to be determined for endmember extraction methods, then at most M_l correct endmembers can be captured and at least $M - M_l$ endmembers will be neglected. On the other hand, if capturing endmember methods, e.g., VCA, is informed of the estimated number of endmembers $M_g > M$, it extracts M_g candidate endmembers that include the M true endmembers. This observation is relevant for most endmember extraction techniques, as they tend to extract a pixel that is most different from the others as an endmember from an iteration such that M endmembers should be selected before the iteration in which any of the $(M_g - M)$ non-endmembers are extracted. Another finding is that extracted non-endmembers $(M_g - M)$ are less discriminative than the M endmembers. A

subset of $(M_g - M)$ non-endmembers is included in the subset of M endmembers, increasing the similarity of M endmembers and decreasing their divergent degree.

These findings led this thesis to use the DS as a post-processing step to finely prune the M_g candidate endmembers obtained by the original endmember extraction method. While doing endmember extraction, the candidate endmembers are treated as a set and their DS is calculated. In summary, the enhanced endmember extraction method is completed by:

1. Considering that M_g, a quantity significantly higher than M, represents the number of endmembers.

2. Using endmember extraction methods to provide a collection of M_g candidate endmembers, such as VCA.

3. Computing the DS of the candidate endmember set. The cardinality of the DS is equal to the number of endmembers since the hyperspectral pixels in the DS are endmembers.

The main contributions of the proposed DS are as follows:

- The proposed DS method proposes a comprehensive endmember estimation framework to concurrently count and extract endmembers. Specifically, a DS is defined in this thesis to extract endmembers, and the number of endmembers may be naturally calculated based on the cardinality of the DS. The development of an automated framework for concurrently counting endmembers and extracting their related spectral signatures is the first attempt in this study.

- The proposed DS method offers a method for improving on current endmember extraction techniques such that they may be used to extract endmembers without prior knowledge of the number of endmembers. As a result, the proposed DS method may be seen as the basis for developing novel endmember extraction schemes based on existing arbitrary endmember extraction methods. The resultant endmember extraction strategies have the potential to produce more accurate endmembers than the original endmember extraction methods, but they do not need knowledge of the number of endmembers.

4.5 Maximum Distance Analysis

The majority of endmember extraction techniques that are being currently used require previous knowledge of the number of endmembers, which is a crucial step in the unmixing process. To bridge this, the work entitled *Endmember Estimation with Maximum Distance*

Analysis [124] is proposed to simultaneously estimate the number and the spectral signatures of endmembers without any prior information on experimental data containing pure pixel spectral signatures and no noise, being based on the assumption that endmembers form a simplex with the greatest volume over all pixel combinations. The simplex includes the farthest pixel point from the coordinate origin in the spectral space, which implies that:

1. An endmember is the pixel that is the farthest away from all other pixels, and the distance between any two pixels i and j can be computed by:

$$d_{ij} = \| \mathbf{y}_i - \mathbf{y}_j \|_2$$
$$= \sqrt{(y_{1i} - y_{1j})^2 + \cdots + (y_{Li} - y_{Lj})^2} \tag{4.20}$$

   where L is the number of bands. The Eq. (4.20) is used to find the farthest pixel point from the coordinate origin as the first endmember A. After that, the Eq. (4.20) is still used to determine the second endmember B by determining the farthest pixel point from the first extracted endmember.

2. The farthest pixel from any line will be an endmember. The two extracted endmembers A and B form a line, and the distance between any pixel and the line $L_{\{AB\}}$ formed by the two endmembers A and B is computed by:

$$d_{i\_L_{\{AB\}}} = \| \overrightarrow{Bi} \|_2 \sin \alpha \tag{4.21}$$

   The index of the third endmember C is obtained by:

$$\text{Index}_C = \underset{i \in \{1,2,\cdots,N\}}{\arg\max} \, d_{i\_L_{\{AB\}}} \tag{4.22}$$

   then Index_C is used to extract the third endmemeber C.

3. An endmember is the pixel that is farthest away from any plane (or affine hull). Assuming that the extracted endmembers has $\{ABC\cdots\}$, which form an affine hull $\text{Aff}\{ABC\cdots\}$. The proposed MDA defines the matrix $\mathbf{Y}_{\{ABC\cdots\}} = \{\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC}, \cdots\} \in \mathbb{R}^{N \times c}$, where c is the number of extracted endmembers and N is the total number of pixels. Let $\mathbf{w}$ be the normal vector of the affine hull $\text{Aff}\{ABC\cdots\}$, b is the offset or distance between coordinate origin and the affine hull, the normal vector $\mathbf{w}$ can be computed by:

$$\mathbf{Y}_{\{ABC\cdots\}}^{\mathrm{T}} \mathbf{w} = \mathbf{0} \tag{4.23}$$

Assuming that $\mathbf{y}_i$ is a pixel point of the affine hull $\text{Aff}\{\text{ABC}\cdots\}$, the affine hull is written as:

$$\mathbf{w}^{\mathrm{T}}\mathbf{y}_i + b = 0 \tag{4.24}$$

The distance between $\mathbf{y}_i$ and the considered affine hull is computed by

$$d = \frac{|\mathbf{w}^{\mathrm{T}}\mathbf{y}_i + b|}{||\mathbf{w}||_2} \tag{4.25}$$

Specifically, support vector machines [49] provide the distance calculation given by Eq. (4.25) between any pixel point $\mathbf{y}_i$ and the affine hull. The index of the next endmember by

$$\text{Index}_{next} = \arg\max_{i \in \{1,2,\cdots,N\}} \frac{|\mathbf{w}^{\mathrm{T}}\mathbf{y}_i + b|}{||\mathbf{w}||_2} \tag{4.26}$$

The index Index_{next} is employed to extract the next endmember.

Fig. 4.3 illustrates the process of extracting four endmembers. Especially, this thesis employs the stopping criterion to complete the process of endmember extraction, i.e.: 1) if a pixel point belongs to an affine hull, the distance between it and the affine hull is zero, and 2) if one point does not belong to one affine hull, the distance between it and the affine hull is greater than zero. In other words, the stopping criterion of endmember extraction is expressed as: the maximum distance between all pixel points and the affine hull produced by the extracted endmember is zero, and the number of endmembers is determined by counting the number of extracted pixel points.
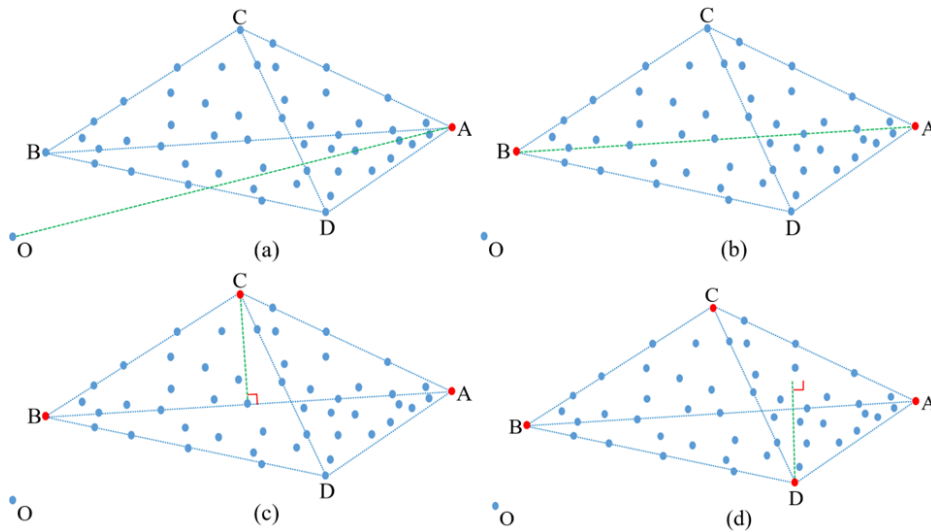

Fig. 4.3 Illustration using MDA to extract four endmembers (from [124] ©2021 MDPI).

In addition to behaving as a stand-alone endmember estimating algorithm, MDA can also work with current endmember extraction techniques without pure pixel assumptions by transforming them into more efficient schemes. In [124], MDA is combined with MVSA (called MDA-MVSA) to address the problem that the original MVSA requires prior knowledge of the number of endmembers. According to the associated results, MDA and MDA-MVSA can precisely count and extract endmembers, and they do so with extremely high efficiency when compared to other conventional endmember extraction techniques. Overall, the contribution of this work [124] consists of two main aspects:

- MDA has the ability to simultaneously count and extract endmembers. In particular, the proposed MDA has a relatively low computational cost since it only uses fundamental vector-based operations. Furthermore, dimension reduction is not necessary for MDA.

- MDA can extend the majority of existing endmember extraction techniques (MVSA is employed in this study) into more precise ones without prior knowledge of the number of endmembers. Specifically, MVSA needs to know the number of endmembers in advance and utilizes the endmembers extracted by VCA as the initialization endmembers. As a result, the proposed MDA provides a broad foundation for creating new endmember extraction schemes based on arbitrary current endmember extraction techniques without the pure pixel assumption in addition to counting and extracting endmembers simultaneously.

4.6 New Maximum Distance Analysis

This work entitled *Endmember Estimation from Hyperspectral Images Using Geometric Distances* [123] proposes a new geometric distance-based algorithm derived from MDA, which is called NMDA, to optimize the widely used MDA method.

By reviewing MDA, it can be known that the farthest pixel points in MDA are acquired by computing the maximum distances between any pixel point and one specific pixel point, line, plane, or affine hull, and the maximum distances are computed by some mathematical method. In particular, the maximum distance in MDA between any pixel point and the considered affine hull is computed by:

$$d = \frac{|\mathbf{w}^{\mathrm{T}}\mathbf{y}_i + b|}{||\mathbf{w}||_2} \tag{4.27}$$

where $\mathbf{y}_i$ is a pixel point in the pixel space, $\mathbf{w}$ is the normal vector of the affine hull, and b is the offset. MDA utilizes Eq. (4.27) to compute the maximum distances between all pixel
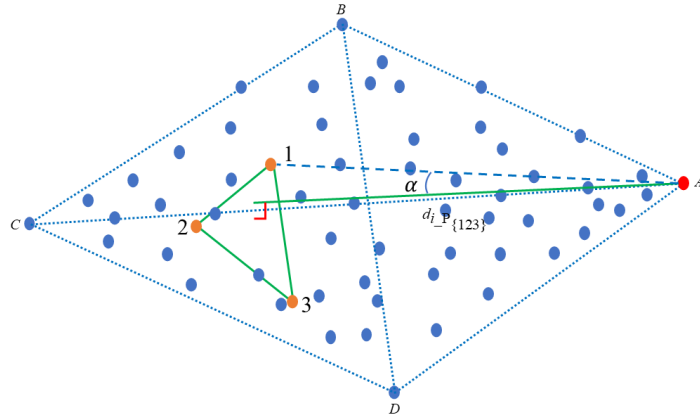
Fig. 4.4 Schematic diagram illustrating how to use NMDA to extract the first endmember (from [123] ©2021 IEEE).

points and the specific affine hull formed by all endmembers extracted by previous steps, and it considers the endmember extraction task to be accomplished when the maximum distance is zero.

In contrast to MDA, NMDA benefits from the requirement that the farthest pixel point from any plane or affine hull must be an endmember. In this sense, NMDA first forms the initial plane $P_{\{123\}}$ by randomly selecting three pixel points (1, 2, and 3) from the pixel space, as shown in Fig. 4.4. The first endmember is the pixel point that is farthest away from the plane $P_{\{123\}}$. The red dots indicate the first extracted endmember, and the blue dots are the pixel points in the pixel space. The vector $\mathbf{n}$ is the normal vector of the plane $P_{\{123\}}$, and the angle α between the two vectors $\overrightarrow{1i}$ and $\mathbf{n}$ is trigonometrically computed by:

$$\alpha = \arccos \frac{\overrightarrow{1i} \cdot \mathbf{n}}{|| \overrightarrow{1i} ||_2 ||\mathbf{n}||_2} \tag{4.28}$$

The distance $d_{i\_P_{\{123\}}}$ between any pixel point i and the plane $P_{\{123\}}$ can be computed by:

$$\begin{aligned} d_{i\_P_{\{123\}}} &= || \overrightarrow{1i} ||_2 |\cos \alpha| \\ &= \frac{| \overrightarrow{1i} \cdot \mathbf{n}|}{||\mathbf{n}||_2} \end{aligned} \tag{4.29}$$

Then the index I_1 of the first endmember is obtained by:

$$I_1 = \underset{i \in \{1,2,\cdots,N\}}{\arg\max} \ d_{i\_P_{\{123\}}} \tag{4.30}$$
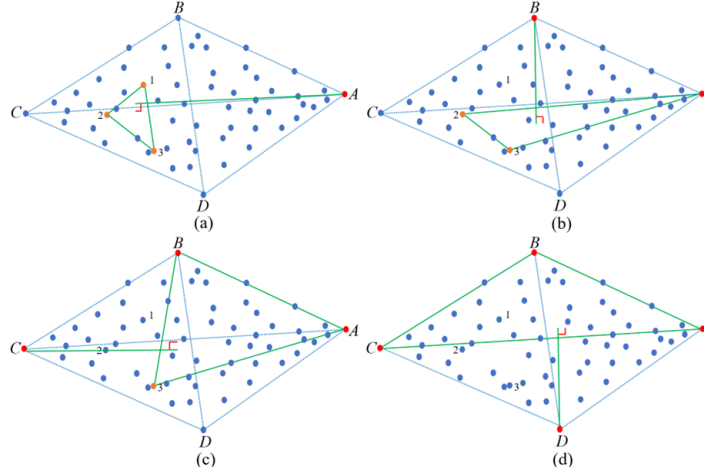
Fig. 4.5 Step-by-step process illustrating how to use NMDA to extract four endmembers (from [123] ©2021 IEEE).

The index I_1 is then used to extract the first endmember A. In summary, Eq. (4.29) provides a new distance computation method, and it can be used to compute the distance between any pixel point and one specific plane or affine hull. Since one endmember A cannot form a plane or affine hull, and NMDA uses two random pixel points (2 and 3) to help the pixel point A to form the plane $P_{\{23A\}}$. NMDA compute the distance $d_{i\_P_{\{23A\}}}$ between any pixel point i and the plane $P_{\{23A\}}$ to find the farthest pixel point from the plane $P_{\{23A\}}$ as the second endmember B. Now NMDA has the plane $P_{\{3AB\}}$, and it then computes the distance $d_{i\_P_{\{3AB\}}}$ to extract the third endmember C. These three endmembers can now combine to form a plane and Eqs. (4.29) and (4.30) are updated as follows:

$$
\begin{aligned}
d_{i\_P_{\{ABC\cdots\}}} &= || \overrightarrow{Ai} ||_2 |\cos \alpha| \\
&= \frac{| \overrightarrow{Ai} \cdot \mathbf{n}|}{||\mathbf{n}||_2}
\end{aligned}
\tag{4.31}
$$

and

$$
I = \underset{i \in \{1,2,\cdots,N\}}{\arg\max} \; d_{i\_P_{\{ABC\cdots\}}}
\tag{4.32}
$$

The proposed NMDA extracts the remaining endmembers by using Eqs. (4.31) and (4.32) to compute the distance between any pixel point and the specific plane, and finalizes the endmember extraction process when the maximum distance is zero. Fig. 4.5 graphically illustrates the process of extracting four endmembers by NMDA. Particularly, by observing Eqs. (4.27) and (4.31), it can be found that Eq. (4.31) provides a more efficient way to

compute the distance between any pixel point and one specific affine hull. Overall, the major contribution of this work [123] are as follows:

- MDA computes the maximum distances between any pixel and one specific pixel, line, plane, or affine hull composed by the endmembers that have been formerly extracted to capture endmembers, and the proposed NMDA only needs to compute the maximum distance between any pixel and one specific affine hull. This operation provides a much more simple way to estimate endmembers.

- MDA needs the normal vector and offset to obtain the maximum distance, but the proposed NMDA only requires the normal vector to complete the tasks. In this sense, NMDA provides a more simple and more efficient way for endmember estimation than MDA

4.7 Discussions

In this chapter, five unmixing methods including three traditional methods and two DL-based methods have been briefly introduced. They both present good results in hyperspectral unmixing, and corresponding results can be found in related publications. Unfortunately, there are certain unaddressed problems that might become problematic in the future. Here are the specifics:

- Firstly, these traditional methods, i.e., MDA, NMDA, OPE, and OPA, are both designed for linear mixed problems. However, in the real sense, HSIs tend to be nonlinearly mixed. Therefore, how to re-design and make them into nonlinear methods should be exploited in the future. Moreover, they both complete the two tasks, i.e., determining the number of endmembers and extracting endmember signatures, but do not exploit the function of abundance estimation. Therefore, how to generalize them to complete abundance estimation also should be developed.

- Secondly, although 3D CUN and `CrossCUN` have been developed in this thesis, they are both supervised methods and need real abundance maps to complete the final task. Moreover, 2D CNN also has the ability that learns spectral and spatial information simultaneously by adjusting the number of filters and the size of kernels, but 3D CNN captures spectral and spatial information by 3D convolutional kernels. In this sense, 3D CNN extracts spectral and spatial information more flexibly than 2D CNN, and vital research for comparing 2D CNN and 3D CNN should be developed for illustrating this situation.

- Thirdly, although 3D CUN and `CrossCUN` are unmixing methods based on DL, they also have the problem faced by traditional unmixing methods, i.e., require the number of endmembers as prior knowledge. Especially, the number of nodes in a particular layer must be equal to the number of endmembers. Therefore, how to address the problem and design a new framework for automatically determining the number of nodes in the layer should be developed.

Except for the above challenges, this thesis has some curiosity about designing a framework based on DL for hyperspectral unmixing:

- Some DL-based methods use endmember extraction results obtained by VCA to initialize the weight of a particular layer. However, DL can generally update weights automatically, so the problems should be considered: (1) a good initialization for DL-based hyperspectral unmixing methods is significant? (2) If a good initialization is very significant, why did researchers choose VCA (it is not the existing best unmixing method) as the initialization method? (3) Does a good initialization just make the network converge faster?

- Some DL-based methods extract the weight and output of a specific layer as endmember features and abundance maps, respectively. It is worth noting that some methods add additional layers after the specific layer. This operation may be done to obtain good reconstruction results. However, the idea of extracting the weight and output of the specific layer as endmember signatures and abundance maps is to build on the fact that the output of the last layer is the reconstructed image. Therefore, how to get the endmember signatures and abundance maps after adding other layers?

In the future, the author of this thesis will address the shortcomings of the proposed methods and exploit these interesting questions to propose effective and efficient frameworks for hyperspectral unmixing.

Chapter 5

Conclusion and Future Work

As hyperspectral imager technology has advanced, HSIs have found widespread usage in remote sensing and other fields. As a hotspot and difficulty in HSI processing, hyperspectral unmixing technique has been deeply studied. Traditional unmixing algorithms usually regard endmember extraction and abundance estimation as two independent processes, therefore, the performance of abundance estimation largely depends on the quality of endmember extraction. In addition, DL has made a splash in the field of HSI processing with its excellent feature extraction ability. But the existing hyperspectral unmixing frameworks based on DL rarely consider both spatial and spectral information, leading to the unideal unmixing effect. To address the above issues, this thesis aims to develop fast, effective, and stable traditional hyperspectral unmixing methods as well as explore how to efficiently use spectral and spatial information for hyperspectral unmixing. The main contents of this paper are summarized as follows:

- This thesis proposes a new and fast orthogonal projection framework to accomplish the whole hyperspectral unmixing task of extracting endmembers and estimating the abundance of each endmember at each pixel. The proposed endmember extraction and abundance estimation algorithms are called OPE and OPA, respectively. OPE extracts endmember signatures by finding the orthogonal vectors of the extracted endmembers, and selects the pixel with the largest projection to the orthogonal vector as the endmembers. OPA computes a diagonal matrix whose product with an endmember matrix is not only a square matrix but also a diagonal matrix, and it then utilizes some specific operations on the matrix to estimate the abundances of each endmember at each pixel. Since OPE and OPA require only fast matrix operations, they are highly efficient in endmember extraction and abundance estimation tasks, respectively.

- This thesis introduces a new CrossCUN for hyperspectral unmixing. The proposed CrossCUN combines 2D convolution and 3D convolution to address the issue that only using 2D CNN cannot learn a good discriminative feature map from a spectral perspective and only using 3D CNN will be significantly more complex and expensive in terms of computational overloads. In this sense, the proposed CrossCUN can capture spectral and spatial information and perform spectral unmixing effectively and efficiently. Experimental results show that the proposed CrossCUN not only obtains more accurate results than DL-based methods but also accomplishes abundance estimation directly without any prior results on endmember extraction compared to traditional unmixing methods.

- This thesis proposes a new DS method for determining the number of endmembers and extracting endmembers at the same time. The proposed DS method addresses the common problem faced by most endmember extraction techniques, i.e. extracting endmembers needs to know the number of endmembers in advance. Experimental results show the effectiveness and efficiency of the proposed DS in determining the number of endmembers and extracting their corresponding spectral signatures. Unfortunately, in the case of noisy and interfering observations, the DS technique is unable to scale to a high number of endmembers due to the limited capability of pairwise representations. To address the issue, the proposed DS method combines with an existing endmember extraction method (such as VCA) to construct a new DS-based strategy to extract an arbitrarily large number of candidate endmembers and then the proposed DS method is used to count and extract real endmembers in the candidates. The related results also show that this innovative technique achieves successful endmember estimation.

- This thesis develops an MDA method to estimate the number of endmembers and extract their corresponding endmember signatures. In contrast to other conventional unmixing techniques, MDA extracts endmember signatures without any prior knowledge about the number of endmembers. In addition, when the experimental data is noisy, MDA cannot produce the most accurate endmember estimate results since the distance calculation in Eq. (4.25) is provided in the scenario of the LMM. To address this issue, this thesis creates a new endmember estimation framework by integrating MDA into the original MVSA method. This procedure is more effective for the overall endmember estimate problem since it gets beyond the original MVSA algorithm's restriction of needing to know the number of endmembers beforehand. The new approach is known as MDA-MVSA.

- This thesis designs a new endmember estimation method called NMDA to improve the original MDA and concurrently determine the number of endmembers and extract their signatures. Especially, MDA extracts endmembers by finding the farthest pixel from one specific pixel, line, plane, or affine hull composed by the endmembers that have been formerly extracted, and it needs the normal vector and offset to complete the corresponding computation of distance. Therefore, the process of extracting endmembers is complex. To this end, NMDA finds the farthest pixel from one specific affine hull and proposes a new distance computation to optimize the process of endmember estimation. In general, the proposed NMDA improves of original MDA from two aspects: (1) it provides a simpler procedure for extracting endmembers and (2) it improves the efficiency of endmember extraction. The related results show that the NMDA has higher efficiency and better results than the original MDA.

Although an in-depth and detailed study of hyperspectral unmixing techniques has been conducted in this thesis, the current research content of this thesis is not comprehensive enough due to the limitations of existing knowledge and time. Future research directions for hyperspectral unmixing are as follows:

- Generally, hyperspectral unmixing methods get the abundance maps by the output of a specific activation layer. However, the number of nodes of the activation layer must be set to the number of endmembers. This means that the number of endmembers must be obtained by other techniques in advance. How to automatically determine the number of endmembers to improve networks and enhance the performance of unmixing will be a further research direction.

- Most methods mainly focus on abundance estimation and pixel reconstruction and rarely try to extract endmembers. Furthermore, they frequently use some endmember extraction algorithms (e.g., VCA and N-FINDR) or spectral library to initialize the parameters in unmixing to obtain abundance maps and reconstructed pixels. In this context, the unmixing performance of these methods will be limited by these endmember extraction algorithms or spectral library. Therefore, how to design unmixing networks based on DL without any assistive technique and any prior knowledge will be a further research direction.

- Some methods usually require a large amount of training data with known ground truth to perform the unmixing task, and these training data are usually generated by other unmixing methods. Therefore, how to develop some semi-supervised or unsupervised unmixing techniques will be a further research direction.

- The performance of some unmixing methods is usually evaluated on some data sets with low resolution, which may not fully reflect the advantages of these proposed methods in unmixing. Therefore, how to design some techniques (e.g., hypersharpening) for increasing the resolution of the experimental data will be a further research direction.

- Few frameworks capture spectral and spatial information, and most frameworks can only extract spectral or spatial information. However, the detailed spectral and spatial information contained in the original data is very significant to enhance the unmixing performance. Therefore, how to propose some frameworks capable of extracting spectral and spatial information will be a further research direction.

- Most existing frameworks based on DL are designed to address linear unmixing problems. However, most materials are nonlinearly mixed in real scenes. Therefore, how to design new frameworks based on DL considering the nonlinearity of the original data will be a further research direction.

- The performance of DL algorithms is closely related to the initialization of network parameters. A good initialization can make the model converge quickly and achieve better generalization performance. Therefore, exploring the initialization of DL networks will be a further research direction.

- Considering better objective functions, such as endmember simplex volume constraints, abundance smoothness constraints, and regularization constraints on network weights, will be a further research direction.

- In general, the performance of the proposed network needs to be adjusted accordingly for the number of layers and the size of the convolution kernels for different data sets. Therefore, how to choose the number of layers of the network and the size of the convolution kernels will be a further research direction.

Overall, in addition to the comprehensive analysis and future direction of hyperspectral unmixing techniques, this thesis also provides the scientific community with five works published in journals in the fields of remote sensing and computer science. Moreover, this thesis has also been very productive in terms of scientific contributions to the HSI remote sensing community, contributing 16 scientific articles. Among these works, some are known for certain merits due to their impact on the scientific community. For example, [95] has been selected as the cover of the Multidisciplinary Digital Publishing Institute Remote Sensing journal of april 2020.

Last but not less important, the author of this thesis has been awarded as one of the five best reviewers of the IEEE Geoscience and Remote Sensing Letters in 2020, and the author has reviewed many works for hyperspectral community including 51 journal papers in IEEE Transactions on Geoscience and Remote Sensing, 57 journal papers in IEEE Geoscience and Remote Sensing Letters, and 6 journal papers in MDPI Remote Sensing.

References

[1] Aghdam, H. H. and Heravi, E. J. (2017). Guide to convolutional neural networks. *New York, NY: Springer*, 10(978-973):51.

[2] Altmann, Y., Dobigeon, N., and Tourneret, J.-Y. (2011a). Bilinear models for nonlinear unmixing of hyperspectral images. In *2011 3rd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pages 1–4. IEEE.

[3] Altmann, Y., Halimi, A., Dobigeon, N., and Tourneret, J.-Y. (2011b). A post nonlinear mixing model for hyperspectral images unmixing. In *2011 IEEE International Geoscience and Remote Sensing Symposium*, pages 1882–1885. IEEE.

[4] Altmann, Y., Halimi, A., Dobigeon, N., and Tourneret, J.-Y. (2012). Supervised nonlinear spectral unmixing using a postnonlinear mixing model for hyperspectral imagery. *IEEE Transactions on Image Processing*, 21(6):3017–3025.

[5] Amigo, J. M., Babamoradi, H., and Elcoroaristizabal, S. (2015). Hyperspectral image analysis. a tutorial. *Analytica chimica acta*, 896:34–51.

[6] Barberis, A., Danese, G., Leporati, F., Plaza, A., and Torti, E. (2012). Real-time implementation of the vertex component analysis algorithm on gpus. *IEEE Geoscience and Remote Sensing Letters*, 10(2):251–255.

[7] Basha, S. S., Dubey, S. R., Pulabaigari, V., and Mukherjee, S. (2020). Impact of fully connected layers on performance of convolutional neural networks for image classification. *Neurocomputing*, 378:112–119.

[8] Bau, D., Zhu, J.-Y., Strobelt, H., Lapedriza, A., Zhou, B., and Torralba, A. (2020). Understanding the role of individual units in a deep neural network. *Proceedings of the National Academy of Sciences*, 117(48):30071–30078.

[9] Baum, L. E. and Eagon, J. A. (1967). An inequality with applications to statistical estimation for probabilistic functions of markov processes and to a model for ecology. *Bulletin of the American Mathematical Society*, 73(3):360–363.

[10] Bayar, B. and Stamm, M. C. (2016). A deep learning approach to universal image manipulation detection using a new convolutional layer. In *Proceedings of the 4th ACM workshop on information hiding and multimedia security*, pages 5–10.

[11] Bhosle, K. and Musande, V. (2022). Evaluation of cnn model by comparing with convolutional autoencoder and deep neural network for crop classification on hyperspectral imagery. *Geocarto International*, 37(3):813–827.

[12] Bioucas-Dias, J. M., Plaza, A., Dobigeon, N., Parente, M., Du, Q., Gader, P., and Chanussot, J. (2012). Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *IEEE journal of selected topics in applied earth observations and remote sensing*, 5(2):354–379.

[13] Borel, C. C. and Gerstl, S. A. (1994). Nonlinear spectral mixing models for vegetative and soil surfaces. *Remote sensing of environment*, 47(3):403–416.

[14] Callicó, G. M., Lopez, S., Aguilar, B., López, J. F., and Sarmiento, R. (2014). Parallel implementation of the modified vertex component analysis algorithm for hyperspectral unmixing using opencl. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(8):3650–3659.

[15] Chan, T.-H., Ma, W.-K., Ambikapathi, A., and Chi, C.-Y. (2011). A simplex volume maximization framework for hyperspectral endmember extraction. *IEEE Transactions on Geoscience and Remote Sensing*, 49(11):4177–4193.

[16] Chang, C.-I. (2002). Target signature-constrained mixed pixel classification for hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 40(5):1065–1081.

[17] Chang, C.-I., Li, H.-C., Wu, C.-C., and Song, M. (2016). Recursive geometric simplex growing analysis for finding endmembers in hyperspectral imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(1):296–308.

[18] Chang, C.-I., Wu, C.-C., Liu, W., and Ouyang, Y.-C. (2006). A new growing method for simplex-based endmember extraction algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, 44(10):2804–2819.

[19] Chang, C.-I., Wu, C.-C., Lo, C.-S., and Chang, M.-L. (2009). Real-time simplex growing algorithms for hyperspectral endmember extraction. *IEEE Transactions on Geoscience and Remote Sensing*, 48(4):1834–1850.

[20] Chang, C.-I., Wu, C.-C., and Tsai, C.-T. (2010). Random n-finder (n-findr) endmember extraction algorithms for hyperspectral imagery. *IEEE Transactions on Image Processing*, 20(3):641–656.

[21] Chen, J., Richard, C., and Honeine, P. (2013). Estimating abundance fractions of materials in hyperspectral images by fitting a post-nonlinear mixing model. In *2013 5th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pages 1–4. IEEE.

[22] Chen, Z., Tong, L., Qian, B., Yu, J., and Xiao, C. (2021). Self-attention-based conditional variational auto-encoder generative adversarial networks for hyperspectral classification. *Remote Sensing*, 13(16):3316.

[23] Chouzenoux, E., Legendre, M., Moussaoui, S., and Idier, J. (2013). Fast constrained least squares spectral unmixing using primal-dual interior-point optimization. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(1):59–69.

[24] Deán-Ben, X. L., Deliolanis, N. C., Ntziachristos, V., and Razansky, D. (2014). Fast unmixing of multispectral optoacoustic data with vertex component analysis. *Optics and Lasers in Engineering*, 58:119–125.

[25] Deborah, H., Ulfarsson, M. O., and Sigurdsson, J. (2021). Fully constrained least squares linear spectral unmixing of the scream (verso, 1893). In *2021 11th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pages 1–5. IEEE.

[26] Ding, X., Xia, C., Zhang, X., Chu, X., Han, J., and Ding, G. (2021). Repmlp: Re-parameterizing convolutions into fully-connected layers for image recognition. *arXiv preprint arXiv:2105.01883*.

[27] Dou, Z., Gao, K., Zhang, X., Wang, H., and Wang, J. (2020). Hyperspectral unmixing using orthogonal sparse prior-based autoencoder with hyper-laplacian loss and data-driven outlier detection. *IEEE Transactions on Geoscience and Remote Sensing*, 58(9):6550–6564.

[28] Drumetz, L., Meyer, T. R., Chanussot, J., Bertozzi, A. L., and Jutten, C. (2019). Hyperspectral image unmixing with endmember bundles and group sparsity inducing mixed norms. *IEEE Transactions on Image Processing*, 28(7):3435–3450.

[29] Eches, O. and Guillaume, M. (2013). A bilinear–bilinear nonnegative matrix factorization method for hyperspectral unmixing. *IEEE Geoscience and Remote Sensing Letters*, 11(4):778–782.

[30] Elkholy, M. M., Mostafa, M., Ebied, H. M., and Tolba, M. F. (2020). Hyperspectral unmixing using deep convolutional autoencoder. *International Journal of Remote Sensing*, 41(12):4799–4819.

[31] ElMasry, G. and Sun, D.-W. (2010). Principles of hyperspectral imaging technology. In *Hyperspectral imaging for food quality analysis and control*, pages 3–43. Elsevier.

[32] Fan, E. (2000). Extended tanh-function method and its applications to nonlinear equations. *Physics Letters A*, 277(4-5):212–218.

[33] Fan, W., Hu, B., Miller, J., and Li, M. (2009). Comparative study between a new nonlinear model and common linear model for analysing laboratory simulated-forest hyperspectral data. *International Journal of Remote Sensing*, 30(11):2951–2962.

[34] Févotte, C. and Dobigeon, N. (2015). Nonlinear hyperspectral unmixing with robust nonnegative matrix factorization. *IEEE Transactions on Image Processing*, 24(12):4810–4819.

[35] Gao, L., Han, Z., Hong, D., Zhang, B., and Chanussot, J. (2021). Cycu-net: Cycle-consistency unmixing network by learning cascaded autoencoders. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–14.

[36] Giannoni, L., Lange, F., and Tachtsidis, I. (2018). Hyperspectral imaging solutions for brain tissue metabolic and hemodynamic monitoring: past, current and future developments. *Journal of Optics*, 20(4):044009.

[37] González, C., Mozos, D., Resano, J., and Plaza, A. (2011). Fpga implementation of the n-findr algorithm for remotely sensed hyperspectral image analysis. *IEEE transactions on geoscience and remote sensing*, 50(2):374–388.

[38] Grahn, H. and Geladi, P. (2007). *Techniques and applications of hyperspectral image analysis.* John Wiley & Sons.

[39] Green, R. O., Eastwood, M. L., Sarture, C. M., Chrien, T. G., Aronsson, M., Chippendale, B. J., Faust, J. A., Pavri, B. E., Chovit, C. J., Solis, M., et al. (1998). Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (aviris). *Remote sensing of environment*, 65(3):227–248.

[40] Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., et al. (2018). Recent advances in convolutional neural networks. *Pattern recognition*, 77:354–377.

[41] Guerra, R., Santos, L., López, S., and Sarmiento, R. (2015). A new fast algorithm for linearly unmixing hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 53(12):6752–6765.

[42] Guo, R., Wang, W., and Qi, H. (2015). Hyperspectral image unmixing using autoencoder cascade. In *Proc. 7th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pages 1–4. IEEE.

[43] Halimi, A., Altmann, Y., Dobigeon, N., and Tourneret, J.-Y. (2011a). Nonlinear unmixing of hyperspectral images using a generalized bilinear model. *IEEE Transactions on Geoscience and Remote Sensing*, 49(11):4153–4162.

[44] Halimi, A., Altmann, Y., Dobigeon, N., and Tourneret, J.-Y. (2011b). Unmixing hyperspectral images using the generalized bilinear model. In *2011 IEEE International Geoscience and Remote Sensing Symposium*, pages 1886–1889. IEEE.

[45] Han, Z., Hong, D., Gao, L., Zhang, B., and Chanussot, J. (2020). Deep half-siamese networks for hyperspectral unmixing. *IEEE Geoscience and Remote Sensing Letters*.

[46] Hapke, B. (1981). Bidirectional reflectance spectroscopy: 1. theory. *Journal of Geophysical Research: Solid Earth*, 86(B4):3039–3054.

[47] Hapke, B. (2012). *Theory of reflectance and emittance spectroscopy.* Cambridge university press.

[48] He, X. and Niyogi, P. (2003). Locality preserving projections. *Advances in neural information processing systems*, 16.

[49] Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., and Scholkopf, B. (1998). Support vector machines. *IEEE Intelligent Systems and their applications*, 13(4):18–28.

[50] Heinz, D. C. et al. (2001). Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 39(3):529–545.

[51] Heylen, R., Burazerovic, D., and Scheunders, P. (2011). Fully constrained least squares spectral unmixing by simplex projection. *IEEE Transactions on Geoscience and Remote Sensing*, 49(11):4112–4122.

[52] Heylen, R., Parente, M., and Gader, P. (2014). A review of nonlinear hyperspectral unmixing methods. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6):1844–1868.

[53] Heylen, R. and Scheunders, P. (2012). Calculation of geodesic distances in nonlinear mixing models: Application to the generalized bilinear model. *IEEE Geoscience and Remote Sensing Letters*, 9(4):644–648.

[54] Hong, D., Chanussot, J., Yokoya, N., Heiden, U., Heldens, W., and Zhu, X. X. (2019). Wu-net: A weakly-supervised unmixing network for remotely sensed hyperspectral imagery. In *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*, pages 373–376. IEEE.

[55] Hong, D., Gao, L., Yao, J., Yokoya, N., Chanussot, J., Heiden, U., and Zhang, B. (2021). Endmember-guided unmixing network (egu-net): A general deep learning framework for self-supervised hyperspectral unmixing. *IEEE Transactions on Neural Networks and Learning Systems*, 33(11):6518–6531.

[56] Hong, D., Yokoya, N., Chanussot, J., and Zhu, X. X. (2018). An augmented linear mixing model to address spectral variability for hyperspectral unmixing. *IEEE Transactions on Image Processing*, 28(4):1923–1938.

[57] Hong, D. and Zhu, X. X. (2018). Sulora: Subspace unmixing with low-rank attribute embedding for hyperspectral data analysis. *IEEE Journal of Selected Topics in Signal Processing*, 12(6):1351–1363.

[58] Hosseini-Asl, E., Zurada, J. M., and Nasraoui, O. (2015). Deep learning of part-based representation of data using sparse autoencoders with nonnegativity constraints. *IEEE transactions on neural networks and learning systems*, 27(12):2486–2498.

[59] Hou, X., Shen, L., Sun, K., and Qiu, G. (2017). Deep feature consistent variational autoencoder. In *2017 IEEE winter conference on applications of computer vision (WACV)*, pages 1133–1141. IEEE.

[60] Imbiriba, T., Bermudez, J. C. M., Richard, C., and Tourneret, J.-Y. (2015). Nonparametric detection of nonlinearly mixed pixels and endmember estimation in hyperspectral images. *IEEE Transactions on Image Processing*, 25(3):1136–1151.

[61] Jameel, S. M., Gilal, A. R., Rizvi, S. S. H., Rehman, M., and Hashmani, M. A. (2020). Practical implications and challenges of multispectral image analysis. In *2020 3rd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, pages 1–5. IEEE.

[62] Jia, P., Zhang, M., and Shen, Y. (2021). Deep spectral unmixing framework via 3d denoising convolutional autoencoder. *IET Image Processing*, 15(7):1399–1409.

[63] Jia, S. and Qian, Y. (2008). Constrained nonnegative matrix factorization for hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 47(1):161–173.

[64] Jin, J., Wang, B., and Zhang, L. (2010). A novel approach based on fisher discriminant null space for decomposition of mixed pixels in hyperspectral imagery. *IEEE Geoscience and Remote Sensing Letters*, 7(4):699–703.

[65] Kaur, S., Bansal, R., Mittal, M., Goyal, L. M., Kaur, I., Verma, A., and Son, L. H. (2019). Mixed pixel decomposition based on extended fuzzy clustering for single spectral value remote sensing images. *Journal of the Indian Society of Remote Sensing*, 47(3):427–437.

[66] Keshava, N. and Mustard, J. F. (2002). Spectral unmixing. *IEEE signal processing magazine*, 19(1):44–57.

[67] Khajehrayeni, F. and Ghassemian, H. (2020). Hyperspectral unmixing using deep convolutional autoencoders in a supervised scenario. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13:567–576.

[68] Kingma, D. P., Welling, M., et al. (2019). An introduction to variational autoencoders. *Foundations and Trends® in Machine Learning*, 12(4):307–392.

[69] Kosaka, N., Uto, K., and Kosugi, Y. (2005). Ica-aided mixed-pixel analysis of hyperspectral data in agricultural land. *IEEE Geoscience and Remote Sensing Letters*, 2(2):220–224.

[70] La Grassa, R., Re, C., Cremonese, G., and Gallo, I. (2022). Hyperspectral data compression using fully convolutional autoencoder. *Remote Sensing*, 14(10):2472.

[71] Lan, R., Li, Z., Liu, Z., Gu, T., and Luo, X. (2019). Hyperspectral image classification using k-sparse denoising autoencoder and spectral–restricted spatial characteristics. *Applied Soft Computing*, 74:693–708.

[72] Larochelle, H., Bengio, Y., Louradour, J., and Lamblin, P. (2009). Exploring strategies for training deep neural networks. *Journal of machine learning research*, 10(1).

[73] Lee, W.-H., Ozger, M., Challita, U., and Sung, K. W. (2021). Noise learning-based denoising autoencoder. *IEEE Communications Letters*, 25(9):2983–2987.

[74] Lemme, A., Reinhart, R. F., and Steil, J. J. (2010). Efficient online learning of a non-negative sparse autoencoder. In *ESANN*. Citeseer.

[75] Li, C., Liu, Y., Cheng, J., Song, R., Peng, H., Chen, Q., and Chen, X. (2018). Hyperspectral unmixing with bandwise generalized bilinear model. *Remote Sensing*, 10(10):1600.

[76] Li, J., Bioucas-Dias, J. M., Plaza, A., and Liu, L. (2016). Robust collaborative nonnegative matrix factorization for hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 54(10):6076–6090.

[77] Li, Z., Liu, F., Yang, W., Peng, S., and Zhou, J. (2021). A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE transactions on neural networks and learning systems*.

[78] Li, Z. and Zhang, H. (2022). Line spectral estimation: Generalized bilinear modeling and hybrid inference method. *Signal Processing*, 195:108479.

[79] Liu, C., Arnon, T., Lazarus, C., Strong, C., Barrett, C., Kochenderfer, M. J., et al. (2021). Algorithms for verifying deep neural networks. *Foundations and Trends® in Optimization*, 4(3-4):244–404.

[80] Liu, K., Kang, G., Zhang, N., and Hou, B. (2018). Breast cancer classification based on fully-connected layer first convolutional neural networks. *IEEE Access*, 6:23722–23732.

[81] Liu, Y., Li, X., Hua, Z., Xia, C., and Zhao, L. (2022). A band selection method with masked convolutional autoencoder for hyperspectral image. *IEEE Geoscience and Remote Sensing Letters*, 19:1–5.

[82] Lu, X., Wu, H., and Yuan, Y. (2013). Double constrained nmf for hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 52(5):2746–2758.

[83] Lu, X., Wu, H., Yuan, Y., Yan, P., and Li, X. (2012). Manifold regularized sparse nmf for hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 51(5):2815–2826.

[84] Ma, K. Y. and Chang, C.-I. (2021). Kernel-based constrained energy minimization for hyperspectral mixed pixel classification. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–23.

[85] Miao, L. and Qi, H. (2007). Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization. *IEEE Transactions on Geoscience and Remote Sensing*, 45(3):765–777.

[86] Montavon, G., Samek, W., and Müller, K.-R. (2018). Methods for interpreting and understanding deep neural networks. *Digital signal processing*, 73:1–15.

[87] Nascimento, J. M. and Bioucas-Dias, J. M. (2009). Nonlinear mixture model for hyperspectral unmixing. In *Image and Signal Processing for Remote Sensing XV*, volume 7477, pages 157–164. SPIE.

[88] Nascimento, J. M. and Dias, J. M. (2005). Vertex component analysis: A fast algorithm to unmix hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 43(4):898–910.

[89] Nash, D. and Conel, J. (1974). Spectral reflectance systematics for mixtures of powdered hypersthene, labradorite, and ilmenite. *Journal of Geophysical Research*, 79(11):1615–1621.

[90] Neville, R. (1999). Automatic endmember extraction from hyperspectral data for mineral exploration. In *International Airborne Remote Sensing Conference and Exhibition, 4 th/21 st Canadian Symposium on Remote Sensing, Ottawa, Canada*.

[91] Noriega, L. (2005). Multilayer perceptron tutorial. *School of Computing. Staffordshire University*, 4:5.

[92] Oquab, M., Bottou, L., Laptev, I., and Sivic, J. (2015). Is object localization for free?-weakly-supervised learning with convolutional neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 685–694.

[93] Palsson, B., Sigurdsson, J., Sveinsson, J. R., and Ulfarsson, M. O. (2018). Hyperspectral unmixing using a neural network autoencoder. *IEEE Access*, 6:25646–25656.

[94] Palsson, B., Sveinsson, J. R., and Ulfarsson, M. O. (2019). Spectral-spatial hyperspectral unmixing using multitask learning. *IEEE Access*, 7:148861–148872.

[95] Paoletti, M. E., Haut, J. M., Tao, X., Miguel, J. P., and Plaza, A. (2020). A new gpu implementation of support vector machines for fast hyperspectral image classification. *Remote Sensing*, 12(8):1257.

[96] Parlos, A. G., Chong, K. T., and Atiya, A. F. (1994). Application of the recurrent multilayer perceptron in modeling complex process dynamics. *IEEE Transactions on Neural Networks*, 5(2):255–266.

[97] Pavan, M. (2004). A new graph-theoretic approach to clustering, with applications to computer vision. *PhD Thesis*.

[98] Pinaya, W. H. L., Vieira, S., Garcia-Dias, R., and Mechelli, A. (2020). Convolutional neural networks. In *Machine learning*, pages 173–191. Elsevier.

[99] Plaza, A., Martinez, P., Perez, R., and Plaza, J. (2004). A new approach to mixed pixel classification of hyperspectral imagery based on extended morphological profiles. *Pattern Recognition*, 37(6):1097–1116.

[100] Popescu, M.-C., Balas, V. E., Perescu-Popescu, L., and Mastorakis, N. (2009). Multilayer perceptron and neural networks. *WSEAS Transactions on Circuits and Systems*, 8(7):579–588.

[101] Qi, L., Li, J., Wang, Y., Lei, M., and Gao, X. (2020). Deep spectral convolution network for hyperspectral image unmixing with spectral library. *Signal Processing*, 176:107672.

[102] Qi, Y., Wang, Y., Zheng, X., and Wu, Z. (2014). Robust feature learning by stacked autoencoder with maximum correntropy criterion. In *2014 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pages 6716–6720. IEEE.

[103] Qu, Y., Guo, R., and Qi, H. (2017). Spectral unmixing through part-based non-negative constraint denoising autoencoder. In *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 209–212. IEEE.

[104] Qu, Y. and Qi, H. (2018). udas: An untied denoising autoencoder with sparsity for spectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 57(3):1698–1712.

[105] Rasti, B., Koirala, B., and Scheunders, P. (2022). Hapkecnn: Blind nonlinear unmixing for intimate mixtures using hapke model and convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–15.

[106] Ren, L. (2021). Unmixing techniques of hyperspectral remote sensing image based on sparse regression[d]. *Southwest Jiaotong University*.

[107] Sánchez, S., Paz, A., and Plaza, A. (2011). Real-time spectral unmixing using iterative error analysis on commodity graphics processing units. In *2011 IEEE International Geoscience and Remote Sensing Symposium*, pages 1767–1770. IEEE.

[108] Shippert, P. (2003). Introduction to hyperspectral image analysis. *Online Journal of Space Communication*, 2(3):8.

[109] Singer, R. B. and McCord, T. B. (1979). Mars-large scale mixing of bright and dark surface materials and implications for analysis of spectral reflectance. In *Lunar and Planetary Science Conference Proceedings*, volume 10, pages 1835–1848.

[110] Song, A., Choi, J., Chang, A., and Kim, Y. (2015). Change detection using spectral unmixing and iea (iterative error analysis) for hyperspectral images. *Korean Journal of Remote Sensing*, 31(5):361–370.

[111] Song, D., Sun, N., Xu, M., Wang, B., and Zhang, L. (2020). Fast unmixing of noisy hyperspectral images based on vertex component analysis and singular spectrum analysis algorithms. *Canadian Journal of Remote Sensing*, 46(1):34–48.

[112] Su, Y., Li, J., Plaza, A., Marinoni, A., Gamba, P., and Chakravortty, S. (2019). Daen: Deep autoencoder networks for hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 57(7):4309–4321.

[113] Su, Y., Li, J., Plaza, A., Marinoni, A., Gamba, P., and Huang, Y. (2018a). Deep auto-encoder network for hyperspectral image unmixing. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, pages 6400–6403. IEEE.

[114] Su, Y., Marinoni, A., Li, J., Plaza, J., and Gamba, P. (2018b). Stacked nonnegative sparse autoencoders for robust hyperspectral unmixing. *IEEE Geoscience and Remote Sensing Letters*, 15(9):1427–1431.

[115] Su, Y., Xu, X., Li, J., Qi, H., Gamba, P., and Plaza, A. (2020). Deep autoencoders with multitask learning for bilinear hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 59(10):8615–8629.

[116] Sze, V., Chen, Y.-H., Yang, T.-J., and Emer, J. S. (2017). Efficient processing of deep neural networks: A tutorial and survey. *Proceedings of the IEEE*, 105(12):2295–2329.

[117] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9.

[118] Tang, M., Zhang, B., Marinoni, A., Gao, L., and Gamba, P. (2018). Multiharmonic postnonlinear mixing model for hyperspectral nonlinear unmixing. *IEEE Geoscience and Remote Sensing Letters*, 15(11):1765–1769.

[119] Tao, X., Cui, T., Plaza, A., and Ren, P. (2020). Simultaneously counting and extracting endmembers in a hyperspectral image based on divergent subsets. *IEEE Transactions on Geoscience and Remote Sensing*, 58(12):8952–8966.

[120] Tao, X., Cui, T., and Ren, P. (2019). Cofactor-based efficient endmember extraction for green algae area estimation. *IEEE Geoscience and Remote Sensing Letters*, 16(6):849–853.

[121] Tao, X., Paoletti, M. E., Han, L., Haut, J. M., Ren, P., Plaza, J., and Plaza, A. (2022a). Fast orthogonal projection for hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–13.

[122] Tao, X., Paoletti, M. E., Han, L., Wu, Z., Ren, P., Plaza, J., Plaza, A., and Haut, J. M. (2022b). A new deep convolutional network for effective hyperspectral unmixing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15:6999–7012.

[123] Tao, X., Paoletti, M. E., Haut, J. M., Han, L., Ren, P., Plaza, J., and Plaza, A. (2021a). Endmember estimation from hyperspectral images using geometric distances. *IEEE Geoscience and Remote Sensing Letters*, 19:1–5.

[124] Tao, X., Paoletti, M. E., Haut, J. M., Ren, P., Plaza, J., and Plaza, A. (2021b). Endmember estimation with maximum distance analysis. *Remote Sensing*, 13(4):713.

[125] Tao, X., Wang, B., and Zhang, L. (2009). Orthogonal bases approach for the decomposition of mixed pixels in hyperspectral imagery. *IEEE Geoscience and Remote Sensing Letters*, 6(2):219–223.

[126] Taud, H. and Mas, J. (2018). Multilayer perceptron (mlp). *Geomatic approaches for modeling land change scenarios*, pages 451–455.

[127] Tran, T. N., Wehrens, R., and Buydens, L. M. (2005). Clustering multispectral images: a tutorial. *Chemometrics and Intelligent Laboratory Systems*, 77(1-2):3–17.

[128] Trigeorgis, G., Bousmalis, K., Zafeiriou, S., and Schuller, B. W. (2016). A deep matrix factorization method for learning attribute representations. *IEEE transactions on pattern analysis and machine intelligence*, 39(3):417–429.

[129] Tschannen, M., Bachem, O., and Lucic, M. (2018). Recent advances in autoencoder-based representation learning. *arXiv preprint arXiv:1812.05069*.

[130] Uezato, T., Murphy, R. J., Melkumyan, A., and Chlingaryan, A. (2016). A novel endmember bundle extraction and clustering approach for capturing spectral variability within endmember classes. *IEEE Transactions on Geoscience and Remote Sensing*, 54(11):6712–6731.

[131] Valencia, D. and Plaza, A. (2006). Fpga-based hyperspectral data compression using spectral unmixing and the pixel purity index algorithm. In *Computational Science–ICCS 2006: 6th International Conference, Reading, UK, May 28-31, 2006, Proceedings, Part I 6*, pages 888–891. Springer.

[132] Vincent, P., Larochelle, H., Bengio, Y., and Manzagol, P.-A. (2008). Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th international conference on Machine learning*, pages 1096–1103.

[133] Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., Manzagol, P.-A., and Bottou, L. (2010). Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *Journal of machine learning research*, 11(12).

[134] Wang, L., Liu, D., and Wang, Q. (2012). Geometric method of fully constrained least squares linear spectral mixture analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 51(6):3558–3566.

[135] Wang, L., Wei, F., Liu, D., Wang, Y., and Wang, Q. (2013). Endmember extraction based on modified iterative error analysis. In *2013 IEEE International Geoscience and Remote Sensing Symposium-IGARSS*, pages 1071–1074. IEEE.

[136] Wang, M., Zhao, M., Chen, J., and Rahardja, S. (2019). Nonlinear unmixing of hyperspectral data via deep autoencoder networks. *IEEE Geoscience and Remote Sensing Letters*, 16(9):1467–1471.

[137] Wang, X., Zhong, Y., Zhang, L., and Xu, Y. (2017). Spatial group sparsity regularized nonnegative matrix factorization for hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 55(11):6287–6304.

[138] Weibull, J. W. (1996). Evolutionary game theory. *MIT Press*, 78(3):850–853.

[139] Winter, M. E. (1999). N-findr: An algorithm for fast autonomous spectral end-member determination in hyperspectral data. In *Imaging Spectrometry V*, volume 3753, pages 266–275. SPIE.

[140] Wold, S., Esbensen, K., and Geladi, P. (1987). Principal component analysis. *Chemometrics and intelligent laboratory systems*, 2(1-3):37–52.

[141] Wu, C., Du, B., and Zhang, L. (2013a). Slow feature analysis for change detection in multispectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 52(5):2858–2874.

[142] Wu, J. (2017). Introduction to convolutional neural networks. *National Key Lab for Novel Software Technology. Nanjing University. China*, 5(23):495.

[143] Wu, J.-z., Yan, W.-d., Ni, W.-p., and Bian, H. (2013b). Feature extraction for hyperspectral data based on mnf and singular value decomposition. In *2013 IEEE International Geoscience and Remote Sensing Symposium-IGARSS*, pages 1430–1433. IEEE.

[144] Wu, X., Huang, B., Plaza, A., Li, Y., and Wu, C. (2013c). Real-time implementation of the pixel purity index algorithm for endmember identification on gpus. *IEEE Geoscience and Remote Sensing Letters*, 11(5):955–959.

[145] Xing, C., Ma, L., and Yang, X. (2016). Stacked denoise autoencoder based feature extraction and classification for hyperspectral images. *Journal of Sensors*, 2016.

[146] Xiong, W., Chang, C.-I., Wu, C.-C., Kalpakis, K., and Chen, H. M. (2011). Fast algorithms to implement n-findr for hyperspectral endmember extraction. *IEEE Journal of selected topics in applied earth observations and remote sensing*, 4(3):545–564.

[147] Xu, Q. and Pan, G. (2017). Sparseconnect: regularising cnns on fully connected layers. *Electronics Letters*, 53(18):1246–1248.

[148] Yan, B., Wang, R., Gan, F., and Wang, Z. (2010). Minerals mapping of the lunar surface with clementine uvvis/nir data based on spectra unmixing method and hapke model. *Icarus*, 208(1):11–19.

[149] Yang, B., Wang, B., and Wu, Z. (2017). Nonlinear hyperspectral unmixing based on geometric characteristics of bilinear mixture models. *IEEE Transactions on Geoscience and Remote Sensing*, 56(2):694–714.

[150] Yang, Z., Zhou, G., Xie, S., Ding, S., Yang, J.-M., and Zhang, J. (2010). Blind spectral unmixing based on sparse nonnegative matrix factorization. *IEEE Transactions on Image Processing*, 20(4):1112–1125.

[151] Yarotsky, D. (2017). Error bounds for approximations with deep relu networks. *Neural Networks*, 94:103–114.

[152] Yin, X., Goudriaan, J., Lantinga, E. A., Vos, J., and Spiertz, H. J. (2003). A flexible sigmoid function of determinate growth. *Annals of botany*, 91(3):361–371.

[153] Yokoya, N., Yairi, T., and Iwasaki, A. (2011). Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 50(2):528–537.

[154] Yu, W., Zhang, M., and Shen, Y. (2020). Spatial revising variational autoencoder-based feature extraction method for hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 59(2):1410–1423.

[155] Zabalza, J., Ren, J., Zheng, J., Zhao, H., Qing, C., Yang, Z., Du, P., and Marshall, S. (2016). Novel segmented stacked autoencoder for effective dimensionality reduction and feature extraction in hyperspectral imaging. *Neurocomputing*, 185:1–10.

[156] Zhang, C., Liu, Y., and Fu, H. (2019). Ae2-nets: Autoencoder in autoencoder networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2577–2585.

[157] Zhang, C.-L., Luo, J.-H., Wei, X.-S., and Wu, J. (2018a). In defense of fully connected layers in visual representation transfer. In *Advances in Multimedia Information Processing–PCM 2017: 18th Pacific-Rim Conference on Multimedia, Harbin, China, September 28-29, 2017, Revised Selected Papers, Part II 18*, pages 807–817. Springer.

[158] Zhang, J., Xu, Y., Zhan, T., Wu, Z., and Wei, Z. (2021). Anomaly detection in hyperspectral image using 3d-convolutional variational autoencoder. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pages 2512–2515. IEEE.

[159] Zhang, J., Zhang, X., Meng, H., Sun, C., Wang, L., and Cao, X. (2022). Nonlinear unmixing via deep autoencoder networks for generalized bilinear model. *Remote Sensing*, 14(20):5167.

[160] Zhang, X., Sun, Y., Zhang, J., Wu, P., and Jiao, L. (2018b). Hyperspectral unmixing via deep convolutional neural networks. *IEEE Geoscience and Remote Sensing Letters*, 15(11):1755–1759.

[161] Zhao, C., Cheng, H., and Feng, S. (2021a). A spectral–spatial change detection method based on simplified 3-d convolutional autoencoder for multitemporal hyperspectral images. *IEEE Geoscience and Remote Sensing Letters*, 19:1–5.

[162] Zhao, M., Yan, L., and Chen, J. (2021b). Lstm-dnn based autoencoder network for nonlinear hyperspectral image unmixing. *IEEE Journal of Selected Topics in Signal Processing*, 15(2):295–309.

[163] Zhong, L., Luo, W., and Gao, L. (2016). A particle swarm optimization algorithm for unmixing the polynomial post-nonlinear mixing model. In *2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, pages 596–600. IEEE.

[164] Zhou, P., Han, J., Cheng, G., and Zhang, B. (2019). Learning compact and discriminative stacked autoencoder for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(7):4823–4833.

[165] Zhou, X., Zhang, Y., Zhang, J., and Shi, S. (2020). Alternating direction iterative nonnegative matrix factorization unmixing for multispectral and hyperspectral data fusion. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13:5223–5232.

Thesis publications

The following publications have been achieved in the context of this thesis work (it should be noticed that publications from 1 to 6 are part of the thesis compendium while the rest of publications have not been included for practical purposes):

Journal Papers:

1. **X. Tao**, M. E. Paoletti, L. Han, J. M. Haut, P. Ren, J. Plaza, and A. Plaza. Fast Orthogonal Projection for Hyperspectral Unmixing. IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1–13, February 2022. DOI: 10.1109/TGRS.2022.3150263. [IF(2022)=8.125]. Google Scholar Citations: 1.

2. **X. Tao**, M. E. Paoletti, L. Han, Z. Wu, P. Ren, J. Plaza, A. Plaza, and J. M. Haut. A New Deep Convolutional Network for Effective Hyperspectral Unmixing. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15, pp. 6999–7012, August 2022. DOI: 10.1109/JSTARS.2022.3200733. [IF(2022)=4.715]. Google Scholar Citations: 1.

3. **X. Tao**, T. Cui, A. Plaza, and P. Ren. Simultaneously Counting and Extracting Endmembers in a Hyperspectral Image Based on Divergent Subsets. IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 12, pp. 8952-8966, December 2020, DOI: 10.1109/TGRS.2020.2992542. [IF(2020)=5.6]. Google Scholar Citations: 13.

4. **X. Tao**, M. E. Paoletti, J. M. Haut, P. Ren, J. Plaza, and A. Plaza. Endmember Estimation with Maximum Distance Analysis. Remote Sensing, vol. 13, no. 4, p. 713, February 2021. DOI: 10.3390/rs13040713. [IF(2021)=5.349]. Google Scholar Citations: 8.

5. **X. Tao**, M. E. Paoletti, J. M. Haut, L. Han, P. Ren, J. Plaza, and A. Plaza. Endmember Estimation From Hyperspectral Images Using Geometric Distances. IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1–5, August 2021. DOI: 10.1109/LGRS.2021.3102076. [IF(2021)=5.343]. Google Scholar Citations: 4.

6. M. E. Paoletti, J. M. Haut, **X. Tao**, J. Plaza, and A. Plaza, FLOP-Reduction Through Memory Allocations Within CNN for Hyperspectral Image Classification. IEEE Transactions on Geoscience and Remote Sensing, vol. 59, no. 7, pp. 5938-5952, July 2020. DOI: 10.1109/TGRS.2020.3024730. [IF(2020)=5.6]. Google Scholar Citations: 22.

7. P. Li, L. Han, **X. Tao**, X. Zhang, C. Grecos, A. Plaza, and P. Ren. Hashing Nets for Hashing: A Quantized Deep Learning to Hash Framework for Remote Sensing Image Retrieval. IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 10, pp. 7331-7345, October. 2020. DOI: 10.1109/TGRS.2020.2981997. [IF(2020)=5.6]. Google Scholar Citations: 45.

8. Z. Wu, H. Su, **X. Tao**, L. Han, M. E. Paoletti, J. M. Haut, J. Plaza, and A. Plaza. Hyperspectral Anomaly Detection With Relaxed Collaborative Representation. IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-17, 2022, Art no. 5533417. DOI: 10.1109/TGRS.2022.3190327. [IF(2022)=8.125]. Google Scholar Citations: 13.

9. M. E. Paoletti, J. M. Haut, **X. Tao**, J. Plaza and A. Plaza. A new GPU implementation of support vector machines for fast hyperspectral image classification. Remote Sensing, vol. 12, no. 8, 1257, April 2020. DOI: 10.3390/rs12081257. [IF(2018)=4.118]. Google Scholar Citations: 33. Selected as journal cover.

10. M. E. Paoletti, J. M. Haut, **X. Tao**, S. Moreno-Álvarez, and A. Plaza. Deep Mixed Precision for Hyperspectral Image Classification. The Journal of Supercomputing, vol. 77, pp. 9190-9201, February 2021. DOI: 10.1007/s11227-021-03638-2. [IF(2021)=2.557]. Google Scholar Citations: 3.

11. L. Han, M. E. Paoletti, **X. Tao**, Z. Wu, J. M. Haut, J. Plaza, and A. Plaza. Central Cohesion Gradual Hashing for Remote Sensing Image Retrieval. IEEE Geoscience and Remote Sensing Letters, vol. 20, pp. 1-5, 2023, Art no. 6002505. DOI: 10.1109/LGRS.2023.3241849. [IF(2023)=5.343]. Google Scholar Citations: 0.

12. A. Redondo, J. M. Haut, M. E. Paoletti, **X. Tao**, J. Plaza and A. Plaza. Analysis of Remotely Sensed Images Through Social Media. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 3026-3039, 2021, DOI: 10.1109/JSTARS.2021.3062116. [IF(2021)=4.715]. Google Scholar Citations: 1.

Submited Journal Papers:

1. M. E. Paoletti, **X. Tao**, L. Han, Z. Wu, S. Moreno-Alvarez, SK Roy, A. Plaza, J. M. Haut. Parameter-Free Attention Network for Spectral-Spatial Hyperspectral Image Classification. IEEE Transactions on Geoscience and Remote Sensing. [IF(2023)=8.125].

2. X. Pan, D. Meng, P. Ren, Y. Xiao, K. Kim, **X. Tao**, R. Liu, Q. Wang, J. Ryu. An Automatic Model for Macroalgae Monitoring from Satellite Optical Images. Ecological Indicators. [IF(2023)=6.263].

Peer-Reviewed International Conference Papers:

1. **X. Tao**, L. Han, M. E. Paoletti, S.K. Roy, J. Plaza, J. M. Haut, and A. Plaza. Multiple incremental kernel convolution for land cover classification of remotely sensed images. IEEE Geoscience and Remote Sensing Symposium (IGARSS'21), Brussels, Belgium, 2021. DOI: 10.1109/IGARSS47720.2021.9555028.

2. **X. Tao**, M. E. Paoletti, L. Han, Z. Wu, LI Jiménez, J. M. Haut, P. Ren, J. Plaza, and A. Plaza. A New 3D Convolution Network for Hyperspectral Unmixing. IEEE Geoscience and Remote Sensing Symposium (IGARSS'22), Kuala Lumpur, Malaysia, 2022. DOI: 10.1109/IGARSS46834.2022.9883506.

3. M.E. Paoletti, **X. Tao**, L. Han, Z. Wu, S. Moreno-Álvarez, and J.M. Haut. Deep Attention-Driven HSI Scene Classification Based on Inverted Dot-Product. IEEE Geoscience and Remote Sensing Symposium (IGARSS'22), Kuala Lumpur, Malaysia, 2022. DOI: 10.1109/IGARSS46834.2022.9883028.

4. Z. Wu, H. Su, **X. Tao**, L. Han, M.E. Paoletti, J.M. Haut, J. Plaza, and A Plaza. Adaptive Dictionary Construction for Hyperspectral Anomaly Detection Based on Collaborative Representation. IEEE Geoscience and Remote Sensing Symposium (IGARSS'22), Kuala Lumpur, Malaysia, 2022. DOI: 10.1109/IGARSS46834.2022.9884182.

Copyright notice

Escuela Politecnica
Av. de la Universidad, S/N, 10003
Caceres, Spain
Phone: 0034927257000. Ext. 51662
Email: aplaza,juanmariohaut{@unex.es}

Dr. Antonio Plaza Miguel y Dr. Juan Mario Haut Hurtado como directores de la tesis titulada "Procesamiento Eficiente y Profundo de Imágenes Hiperespectrales y Aplicación en Tareas de Desmezclado Espectral", certifican el factor de impacto y la categorización de la siguiente publicación, incluída en la tesis doctoral. Del mismo modo, se especifica la aportación del doctorado.

Antonio Plaza Miguel PhD, and Juan Mario Haut Hurtado PhD, as directors of the Phd thesis titled "Efficient and Deep Processing of Hyperspectral Images and Application in Spectral Unmixing Tasks", certify the impact factor and the categorization of the following publication, included in the doctoral thesis. In the same way, the contribution of the doctorate is specified.

Artículo / Paper

Autores/Authors: **X. Tao**, M. E. Paoletti, J. M. Haut, P. Ren, J. Plaza and A. Plaza.
Title: Fast Orthogonal Projection for Hyperspectral Unmixing.
Journal: IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-13, February 2022
DOI: 10.1109/TGRS.2022.3150263.
Impact factor 2021: 8.125. Q1

Abstract: Spectral unmixing plays a vital role in hyperspectral image analysis. It mainly consists of two procedures, i.e., endmember extraction and abundance estimation. Although most algorithms for each of the two procedures may exhibit good performance, few studies have been done considering both problems simultaneously. Therefore, hyperspectral unmixing accuracy is normally achieved by exploring all possible combinations of the two types of algorithms, which renders high computational overloads. We propose a novel orthogonal projection framework to conduct fast hyperspectral unmixing. It addresses both endmember extraction and abundance estimation with orthogonal projection endmember (OPE) and orthogonal projection abundance (OPA). Especially, the pixel with the largest orthogonal projection on any pixel is considered to be an endmember. We randomly choose one pixel from the hyperspectral data to compute the orthogonal projections of all pixels and extract the pixel with the largest projection as the first endmember. To avoid extracting the same endmembers, we compute orthogonal projections of all pixels to endmembers that have been previously extracted, and the pixel with the largest projection is considered as the next endmember. In terms of abundance estimation, we also utilize the concept of orthogonal projection and search for a diagonal matrix whose multiplication with the endmember matrix is not only a square matrix but also a diagonal matrix. Then, we exploit some specific matrix operations to estimate the abundance of each endmember at every pixel. We have evaluated the proposed OPE and OPA algorithms on synthetic and real data, and the experimental results have validated their effectiveness and efficiency in hyperspectral unmixing.

Contribución del doctorado: Planteamiento de la hipótesis, desarrollo práctico, análisis y discusión de los resultados, elaboración y escritura del manuscrito.

Firma / Signature
Marzo / Marzo, 2023

Antonio Plaza Miguel   Juan Mario Haut Hurtado

Fast Orthogonal Projection for Hyperspectral Unmixing

Xuanwen Tao, *Graduate Student Member, IEEE*, Mercedes E. Paoletti, *Senior Member, IEEE*,
Lirong Han, *Student Member, IEEE*, Juan M. Haut, *Senior Member, IEEE*,
Peng Ren, *Senior Member, IEEE*, Javier Plaza, *Senior Member, IEEE*,
and Antonio Plaza, *Fellow, IEEE*

Abstract— Spectral unmixing plays a vital role in hyperspectral image analysis. It mainly consists of two procedures, i.e., endmember extraction and abundance estimation. Although most algorithms for each of the two procedures may exhibit good performance, few studies have been done considering both problems simultaneously. Therefore, hyperspectral unmixing accuracy is normally achieved by exploring all possible combinations of the two types of algorithms, which renders high computational overloads. We propose a novel orthogonal projection framework to conduct fast hyperspectral unmixing. It addresses both endmember extraction and abundance estimation with orthogonal projection endmember (OPE) and orthogonal projection abundance (OPA). Especially, the pixel with the largest orthogonal projection on any pixel is considered to be an endmember. We randomly choose one pixel from the hyperspectral data to compute the orthogonal projections of all pixels and extract the pixel with the largest projection as the first endmember. To avoid extracting the same endmembers, we compute orthogonal projections of all pixels to endmembers that have been previously extracted, and the pixel with the largest projection is considered as the next endmember. In terms of abundance estimation, we also utilize the concept of orthogonal projection and search for a diagonal matrix whose multiplication with the endmember matrix is not only a square matrix but also a diagonal matrix. Then, we exploit some specific matrix operations to estimate the abundance of each endmember at every pixel. We have evaluated the proposed OPE and OPA algorithms on synthetic and real data, and the experimental results have validated their effectiveness and efficiency in hyperspectral unmixing.

Index Terms— Abundance estimation, endmember extraction, hyperspectral unmixing, orthogonal projection.

I. INTRODUCTION

HYPERSPECTRAL images (HSIs) contain detailed and rich spectral and spatial information, which has become the key to identifying the composition of ground scenes and has been applied to numerous fields [1]–[5]. A typical problem associated with HSIs is the low spatial resolution of sensors, which inevitably leads to the existence of mixed pixels. Especially, hyperspectral unmixing [6]–[10] is exploited as an outstanding tool to address the problem by decomposing mixed pixels into a set of pure spectra (i.e., endmembers) and their corresponding proportions (i.e., abundances). Over the past decades, many hyperspectral unmixing techniques have been developed in the literature. These techniques are divided into three main groups, i.e., determining the number of endmembers, extracting their spectral signatures, and estimating the corresponding abundances. Generally, most endmember extraction algorithms assume that the number of endmembers has been known in advance through other strategies. Therefore, in our work, we mainly focus on the two tasks of endmember extraction and abundance estimation. In Sections I-A and I-B, we review the literature about endmember extraction and abundance estimation, respectively. In Section I-C, we briefly describe the major contributions of our work in terms of presenting a new hyperspectral unmixing framework that simultaneously extracts endmembers and estimates the abundance of each endmember at every pixel.

A. Endmember Extraction

The related algorithms for endmember extraction are mainly divided into two groups according to whether there are pure pixels in the HSI or not. Some typical algorithms implemented using the pure pixel assumption include orthogonal subspace projection (OSP) [11], vertex component analysis (VCA) [12], and N-FINDR [6], [13], among others. OSP first selects the pixel vector with the largest length in the scene as the first endmember and then finds the pixel vector with the maximum absolute projection as the second endmember in the space orthogonal to the space linearly spanned by the initial pixel. The third endmember is extracted by applying an orthogonal subspace projector to the original image, where the signature

that has the maximum orthogonal projection in the space orthogonal to the space linearly spanned by the first two endmembers. OSP repeats the process until all endmembers are found. Both VCA and N-FINDR are devised to search for a simplex with the maximum volume, and endmembers are the vertices of the simplex. VCA iteratively projects data onto an orthogonal direction to the subspace spanned by the endmembers already extracted, and the pixel with the extreme of the projection is the next endmember. N-FINDR randomly chooses some pixels to form an initial endmember matrix and then replaces the pixels with other pixels to find a set of pixels with the largest volume to extract endmembers. To avoid identifying the same endmember, an automatic target generation process (ATGP) [14] is proposed, and it utilizes the notion of OSP to extract endmembers. To address the impact of noise for endmember extraction, alternating decoupled volume max-min (ADVMM) and successive decoupled volume max-min (SDVMM) [15] are developed for dealing with the worst case simplex volume maximization problem by alternating optimization and successive optimization, respectively. Moreover, considering that spectral variability plays a vital role in spectral unmixing, Drumetz *et al.* [16] proposed to reinterpret the concept of convex geometry used in spectral unmixing in the cases of illumination induced and intrinsic variability in a linear mixing model (LMM) framework and discussed the validity and extension of concepts, such as endmember extraction, intrinsic dimensionality, and pixel purity. In addition, based on the insights gained from this analysis, Drumetz *et al.* [16] proposed an integrated unmixing chain for this situation, which is still based on the concept of convex geometry but adapts it to the new geometry introduced by the presence of variable endmembers. Most unmixing methods require the number of endmembers and have high computational overloads. To address these problems, Das *et al.* [17] proposed two new frameworks for spectral unmixing under the known and unknown number of endmembers, respectively. Das *et al.* [18] presented a new and fast convex geometric framework for spectral unmixing. The new framework detects the extreme points to extract endmember signatures.

If there are no pure pixels in HSIs, endmember extraction will become a very challenging task. The difficulty lies in the fact that HSIs may not contain any endmembers or at least some of them. Some popular algorithms implemented without the pure pixel assumption include minimum-volume enclosing simplex (MVES) [19], minimum-volume simplex analysis (MVSA) [20], [21], robust MVSA (RMVSA) [22], polytope decomposition (POD) [23], and so on. Among representative algorithms in this category, MVES incorporates convex analysis and Craig's criterion to extract endmembers, and it searches for a simplex containing all the dimensionally reduced pixels by minimizing the simplex volume subject to the constraint. MVSA fits a minimum-volume simplex to HSIs by constraining the abundance fractions to belong to the probability simplex. RMVSA is a version of MVSA that is robust to noise, and it develops the relaxation of nonlinear chance constraint, which greatly reduces the computational complexity of nonlinear chance constraint problems. Especially, RMVSA can

be used in any unmixing algorithm that uses nonlinear chance constraints. POD provides an effective method to get the abundance in a nonlinear mixture, which can be approximated by a polynomial model and is independent of the order of the nonlinearity.

In the above two cases, it is a great challenge to directly identify realistic endmember signatures from the HSIs without prior knowledge. In this scenario, some methods exploit endmember signatures from a previously available spectral library, such as the U.S. Geological Survey (USGS),[1] which is potentially very large. Then, a best subset of the endmembers of the library will be chosen to model each pixel [24]. Tang *et al.* [25] developed a new unmixing method called sparse unmixing using spectral *a priori* information (SUnSPI), which uses the spectral *a priori* information in the HSI to alleviate the problem of high mutual coherence of the spectral library faced by sparse unmixing. Shi *et al.* [26] presented a novel algorithm termed collaborative sparse hyperspectral unmixing using l_0-norm (CSUnL0) for the sparse hyperspectral unmixing problem to address the NP-hard problem caused by the l_0-norm constraint. Xu and Shi [27] transformed sparse unmixing to a biobjective optimization problem to present a new multiobjective-based sparse unmixing method, which addresses the sparse unmixing problem without any relaxation. Das *et al.* [28] employed a sparsity measure for dictionary pruning and presented a new sparsity-based dictionary-aided unmixing framework, which is nonparametric and noise robust with satisfactory performance on the spectral library with high mutual coherence. However, endmember signatures from spectral libraries usually present inconsistent acquisition conditions compared to the image pixel signatures, and this fact may affect the result of abundance estimation, which is the subsequent step of endmember extraction.

B. Abundance Estimation

The main models for abundance estimation include the LMM or nonlinear mixing model (NLMM) [29]–[32]. LMM focuses on the assumption that incident light interacts with a single component, and the spectrum of each pixel is formulated as a linear combination of endmembers and corresponding abundances. NLMM works on the actual mechanism of hyperspectral sensors, although it has high computational complexity and requires prior knowledge of the scene. The most popular algorithm for abundance estimation is the fully constrained least squares (FCLS) [33]. FCLS imposes the abundance nonnegative constraint (ANC) and abundance sum-to-one constraint (ASC) and minimizes the square error in the linear approximation of HSIs. In [34], FCLS has been exploited by introducing a variable splitting and solving the optimization problem with the alternating direction method of multipliers (ADMM), denoted as SUnSAL. In [35] and [36], perturbed linear mixing model (PLMM) and augmented linear mixing model (ALMM) approaches are proposed to estimate abundances when the HSI ensures spectral variability.
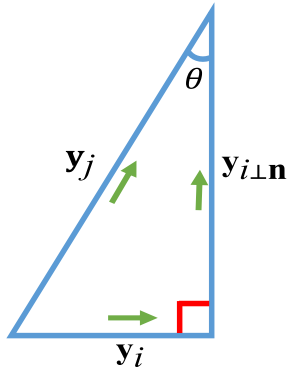
[1]http://speclab.cr.usgs.gov/spectral-lib.html

Fig. 1. Schematic representation about the computation of the orthogonal projection of $\mathbf{y}_j$ to $\mathbf{y}_i$, where $\mathbf{y}_i$ and $\mathbf{y}_j$ are two HSI pixels, $\mathbf{y}_{i\perp\mathbf{n}}$ is the orthogonal projection vector of $\mathbf{y}_j$ to $\mathbf{y}_i$, and the green arrow indicates the direction of pixels.

Recently, some researchers consider hyperspectral unmixing as a blind source separation (BSS) problem and bring the application of nonnegative matrix factorization (NMF). NMF-based unmixing not only provides fully unsupervised without the pure pixel assumption but also simultaneously extracts endmembers and estimates their abundances. However, the standard NMF model is nonconvex and shows scale indeterminacy. In this regard, it is not enough to obtain a stable solution by only considering ANC. Therefore, a series of additional constraints [37] is introduced into the standard NMF model. Taking into account the nonconvexity and computational complexity of the simplex volume and the fact that the smallest-volume simplex is equal to the compact simplex possible, MDCNMF [38] has been proposed. MDC-NMF exploits the distances between endmembers to measure compactness. Concerning the constraints on the endmembers, minimum-volume constrained nonnegative matrix factorization (MVCNMF) [39] has been introduced. MVCNMF adopts nonnegative matrix factorization and a volume-based constraint to extract endmembers and estimate the abundance of each endmember at every pixel.

The aforementioned algorithms present good performance in the specific task of the unmixing process. However, they also show some drawbacks in computational complexity. Moreover, most hyperspectral unmixing techniques need to combine an endmember extraction algorithm and an abundance estimation algorithm to complete the overall process of unmixing. Therefore, the accuracy of unmixing relies on the combination of the two types of algorithms. However, the fact is that this strategy results in high computation burden by exploiting different combinations of endmember extraction and abundance estimation algorithms to obtain the best performance in hyperspectral unmixing.

C. Contributions

As reviewed in Sections I-A and I-B, hyperspectral unmixing algorithms generally consider endmember extraction and abundance estimation as two independent processes, obtaining good performance in their respective fields, and then, they

---

Algorithm 1 Pseudocode of OPE for Endmember Extraction

Input:
  HSI data set: $\mathbf{Y} \in \mathbb{R}^{L \times N}$
  The number of endmembers: c
1: I = []
2: $\mathbf{M} = \mathbf{Y}(:, 1)$ {The randomly selected pixel}
3: **for** $i = 1$ to c **do**
4:   $\mathbf{M}^T \mathbf{n} = \mathbf{0}$
5:   $\mathbf{n}_{\text{unit}} = \frac{\mathbf{n}}{||\mathbf{n}||_2}$
6:   **for** $j = 1$ to N **do**
7:     $\hat{I} = \text{argmax}|\mathbf{n} \cdot \mathbf{y}_j|$
8:   **end for**
9:   I = [I, $\hat{I}$]
10:   $\mathbf{M} = [\mathbf{Y}(:,I)]$
11: **end for**
Output:
  Endmember Matrix: $\mathbf{M} \in \mathbb{R}^{L \times c}$

Algorithm 2 Pseudocode of OPA for Abundance Estimation

Input:
  HSI data set: $\mathbf{Y} \in \mathbb{R}^{L \times N}$
  Endmember Matrix: $\mathbf{M} \in \mathbb{R}^{L \times c}$
  The number of endmembers: c
  The number of pixels: N
1: $\mathbf{P} = []$
2: **for** $i = 1$ to c **do**
3:   $\mathbf{M}(:, i) = []$
4:   $\mathbf{M}^T \mathbf{p} = 0$
5:   $\mathbf{P} = [\mathbf{P},\mathbf{p}];$
6: **end for**
7: $\mathbf{Q} = \mathbf{P}^T \mathbf{M}$
8: $\mathbf{A} = \mathbf{Q}^{-1}\mathbf{P}^T\mathbf{Y}$ {Eq. (22)}
9: $\mathbf{A} = |\mathbf{A}|$
   {Abundance nonnegative constraint (ANC)}
10: **for** $j = 1$ to N **do**
11:   $\mathbf{A}(:,j) = \mathbf{A}(:,j)/(\sum_{k=1}^{c} \mathbf{A}(k,j))$
   {Abundance sum-to-one constraint (ASC)}
12: **end for**
Output:
  Abundance Matrix: $\mathbf{A} \in \mathbb{R}^{c \times N}$

---

search for a good combination of the two processes of algorithms to yield the best performance. Another strategy is the NMF-based, which simultaneously completes the two tasks, i.e., endmember extraction and abundance estimation. Whether seeking for the best combination of the two tasks or directly completing unmixing, the two types of hyperspectral unmixing techniques both bring render considerable computation burden, making them inefficient for tackling large-size remote sensing images. To address the problem, we establish a unified and fast orthogonal projection framework for hyperspectral unmixing. To evaluate our proposed framework, we refer to the algorithms of endmember extraction and abundance estimation as orthogonal projection endmember (OPE) and orthogonal projection abundance (OPA), respectively. The contributions of our work are as follows.

1) We establish a novel orthogonal projection framework for hyperspectral unmixing, which can simultaneously complete endmember extraction and abundance estimation, and it exhibits good effectiveness and high efficiency in hyperspectral unmixing.

2) The proposed OPE algorithm extracts endmembers by utilizing the concept of the pixel with the maximum orthogonal projection on any pixel is an endmember. In particular, to avoid repeatedly extracting the same endmembers, our proposed OPE algorithm exploits the normal vector of all endmembers that have been extracted to compute the projections of all pixels to it and then chooses the pixel with the maximum projection as an endmember. Since it only involves straightforward vector-based computations, its computational complexity is very low. Especially, the high efficiency of our proposed OPE algorithm has been validated in the experiments.

3) Endmembers and their proportions at every pixel are generally considered to be linearly mixed. However, we cannot estimate abundances by computing the multiplication of the data matrix and the inverse of endmember matrix since the endmember matrix is usually not a square matrix. To this end, our proposed OPA algorithm exploits a diagonal matrix, and its multiplication with the endmember matrix is not only a square matrix but also a diagonal matrix. We then obtain the abundance matrix by computing the multiplication of matrices. In this scenario, our proposed OPA algorithm has high efficiency in terms of solving the abundance of endmembers at every pixel. Especially, experimental results have validated that our proposed OPA algorithm has high efficiency in abundance estimation.

D. Article Outline

The rest of this article is organized as follows. In Section II, we detail the proposed OPE algorithm for endmember extraction and the proposed OPA algorithm for abundance estimation. In Section III, we empirically evaluate the performance of the proposed OPE and OPA algorithms in endmember extraction and abundance estimation on synthetic data and Cuprite data. In Section IV, we conclude this article with some remarks and hints at plausible future research lines.

II. METHODOLOGY

Let $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_i, \ldots, \mathbf{y}_N] \in \mathbb{R}^{L \times N}$ denote the HSI data matrix with L bands and N pixels, where $\mathbf{y}_i = [y_{1i}, x_{2i}, \ldots, y_{Li}]^T \in \mathbb{R}^{L \times 1}$ denotes a pixel represented by an L-dimensional vector. As shown in Fig. 1, $\mathbf{y}_i$ and $\mathbf{y}_j$ are two pixels of HSI and $\mathbf{y}_{i \perp \mathbf{n}}$ is the orthogonal projection vector of $\mathbf{y}_j$ to $\mathbf{y}_i$. We have

$$\cos\theta = \frac{\mathbf{y}_{i\perp\mathbf{n}} \cdot \mathbf{y}_j}{||\mathbf{y}_{i\perp\mathbf{n}}||_2 ||\mathbf{y}_j||_2} = \frac{||\mathbf{y}_{i\perp\mathbf{n}}||_2}{||\mathbf{y}_j||_2} \quad (1)$$

where $||\cdot||_2$ represents the L_2-norm. We obtain the orthogonal projection of $\mathbf{y}_j$ to $\mathbf{y}_i$ by

$$||\mathbf{y}_{i\perp\mathbf{n}}||_2 = ||\mathbf{y}_j||_2 |\cos\theta|$$

$$= ||\mathbf{y}_j||_2 |\frac{\mathbf{y}_{i\perp\mathbf{n}} \cdot \mathbf{y}_j}{||\mathbf{y}_{i\perp\mathbf{n}}||_2 ||\mathbf{y}_j||_2}|$$

$$= |\frac{\mathbf{y}_{i\perp\mathbf{n}} \cdot \mathbf{y}_j}{||\mathbf{y}_{i\perp\mathbf{n}}||_2}|. \quad (2)$$

Assuming that $\mathbf{n}$ is the normal vector of $\mathbf{y}_i$ and $\mathbf{n}_{\text{unit}}$ is the unit vector of $\mathbf{n}$, (2) is transformed into

$$||\mathbf{y}_{i\perp\mathbf{n}}||_2 = |\mathbf{n}_{\text{unit}} \cdot \mathbf{y}_j| \quad (3)$$

and $|\cdot|$ denotes the absolute value.

A. Endmember Extraction

We use some techniques to filter out the noise in advance. From a geometric point of view, we know that the pixel with the largest orthogonal projection is an endmember. We randomly select one pixel $\mathbf{y}_i$ from HSI dataset and obtain its normal vector $\mathbf{n}^1 \in \mathbb{R}^{L \times 1}$ by addressing

$$\mathbf{y}_i \cdot \mathbf{n}^1 = 0. \quad (4)$$

Especially, it will have many solutions if we solve the normal vector according to (4). In our work, we sum all the solutions of the normal vector as its real solution. Next, we obtain the unit vector $\mathbf{n}^1_{\text{unit}}$ of $\mathbf{n}^1$ by

$$\mathbf{n}^1_{\text{unit}} = \frac{\mathbf{n}^1}{||\mathbf{n}^1||_2}. \quad (5)$$

We utilize (3) to compute the orthogonal projections of all pixels of HSI to $\mathbf{y}_i$ and obtain the index $\hat{I}_1$ of the first endmember by

$$\hat{I}_1 = \underset{I_1 \in \{1,2,\ldots,N\}}{\arg\max} \left| \mathbf{n}^1_{\text{unit}} \cdot \mathbf{y}_{I_1} \right| \quad (6)$$

and we then use the index $\hat{I}_1$ to extract the first endmember $\mathbf{m}_1$ from the HSI dataset. At this moment, the endmember matrix is $\mathbf{M}_1 = [\mathbf{m}_1] \in \mathbb{R}^{L \times 1}$. Next, we get the normal vector $\mathbf{n}^2 \in \mathbb{R}^{L \times 1}$ of $\mathbf{M}_1$ by

$$\mathbf{M}_1^T \mathbf{n}^2 = 0 \quad (7)$$

and the symbol T denotes transpose. The unit vector $\mathbf{n}^2_{\text{unit}}$ of $\mathbf{n}^2$ is computed by

$$\mathbf{n}^2_{\text{unit}} = \frac{\mathbf{n}^2}{||\mathbf{n}^2||_2}. \quad (8)$$

We obtain the index $\hat{I}_2$ of the second endmember by

$$\hat{I}_2 = \underset{I_2 \in \{1,2,\ldots,N\}}{\arg\max} \left| \mathbf{n}^2_{\text{unit}} \cdot \mathbf{y}_{I_2} \right|. \quad (9)$$

Similarly, we use the index $\hat{I}_2$ to extract the second endmember $\mathbf{m}_2$ from the HSI dataset. Now, the endmember matrix is updated to $\mathbf{M}_2 = [\mathbf{m}_1, \mathbf{m}_2] \in \mathbb{R}^{L \times 2}$. Especially, the reason for constantly updating the endmember matrix is that the projections of all endmembers to their orthogonal vector are zero. Therefore, we compute the orthogonal vector of the new endmember matrix to avoid extracting the same endmembers and reduce the computational complexity of extracting endmembers. We extract the remaining endmember according to the method of extracting the second endmember, and the endmember matrix is updated to $\mathbf{M}_{k-1} = [\mathbf{m}_1, \ldots, \mathbf{m}_{k-1}] \in \mathbb{R}^{L \times (k-1)}$,

and $\mathbf{M}_{k-1}$ denotes that we have extracted $k-1$ endmembers. To extract the kth endmember, we get the normal vector $\mathbf{n}_k$ of $\mathbf{M}_{k-1}$ by

$$\mathbf{M}_{k-1}^T \mathbf{n}^k = 0 \tag{10}$$

and the unit vector $\mathbf{n}_{\text{unit}}^k$ of $\mathbf{n}^k$ is computed by

$$\mathbf{n}_{\text{unit}}^k = \frac{\mathbf{n}^k}{||\mathbf{n}^k||_2}. \tag{11}$$

We get the index I_k of the kth endmember by

$$\hat{I}_k = \underset{I_k \in \{1,2,\dots,N\}}{\arg\max} \left| \mathbf{n}_{\text{unit}}^k \cdot \mathbf{y}_{I_k} \right|. \tag{12}$$

We use the index $\hat{I}_k$ to extract the kth endmember and update the endmember matrix to $\mathbf{M}_k = [\mathbf{m}_1, \dots, \mathbf{m}_{k-1}, \mathbf{m}_k] \in \mathbb{R}^{L \times k}$. We repeat the above process until all endmembers are extracted. We refer to the method of extracting endmembers as OPE. The detailed process of the OPE algorithm is shown in Algorithm 1. From Algorithm 1, we see that our proposed OPE algorithm only involves matrix operations, and its computational complexity is very low. Especially, in Section III, the experimental results also validate that the proposed OPE algorithm exhibits high efficiency in the task of extracting endmembers on synthetic data and Cuprite data.

B. Abundance Estimation

Let $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_i, \dots, \mathbf{a}_N] \in \mathbb{R}^{c \times N}$ represent abundance matrix and c is the number of endmembers. Under the LMM, the data matrix is represented by the multiplication of the endmember matrix and the abundance matrix by

$$\mathbf{Y} = \mathbf{MA}. \tag{13}$$

From a traditional view, abundance matrix $\mathbf{A}$ is computed by

$$\mathbf{A} = \mathbf{M}^{-1} \mathbf{Y} \tag{14}$$

and $\mathbf{M}^{-1}$ is the inverse matrix of $\mathbf{M}$. Broadly speaking, it is unrealistic to computing abundance matrix $\mathbf{A}$ by (14) because the endmember matrix $\mathbf{M}$ is usually not a square matrix. In this scenario, we find a diagonal matrix $\mathbf{Q}$ ($\mathbf{P}^T \mathbf{M} = \mathbf{Q}$) to complete the task of abundance estimation. Next, we first introduce how to get the matrix $\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_i, \dots, \mathbf{p}_c] \in \mathbb{R}^{L \times c}$. Especially, we obtain $\mathbf{p}_1$ by

$$[\mathbf{m}_2, \mathbf{m}_3, \dots, \mathbf{m}_c]^T \mathbf{p}_1 = \mathbf{0}. \tag{15}$$

In this sense, $\mathbf{p}_1$ is not unique. We sum all the solutions obtained by (15) as the final $\mathbf{p}_1$. We get $\mathbf{p}_2$ by addressing

$$[\mathbf{m}_1, \mathbf{m}_3, \dots, \mathbf{m}_c]^T \mathbf{p}_2 = \mathbf{0}. \tag{16}$$

We also sum all the solutions obtained by (16) to obtain the final $\mathbf{p}_2$. Similarly, we acquire $\mathbf{p}_i$ by

$$[\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_{i-1}, \mathbf{m}_{i+1}, \dots, \mathbf{m}_c]^T \mathbf{p}_i = \mathbf{0}. \tag{17}$$

We get the final $\mathbf{n}_i$ by summing all the solutions obtained by (17). Obviously, we have

$$\mathbf{m}_i \cdot \mathbf{p}_i \neq 0 \tag{18}$$

and

$$\mathbf{m}_i \cdot \mathbf{p}_j = 0, \quad i \neq j. \tag{19}$$

We acquire the diagonal matrix $\mathbf{Q}$ by

$$\mathbf{Q} = \mathbf{P}^T \mathbf{M}. \tag{20}$$

From (18) and (19), we obtain $Q_{ii} = \mathbf{m}_i \cdot \mathbf{p}_i \neq 0$ and $Q_{ij} = \mathbf{m}_i \cdot \mathbf{p}_j = 0$. The fact reflects that the diagonal matrix $\mathbf{Q}$ exists inverse matrix $\mathbf{Q}^{-1}$. Therefore, we transfer (13) into

$$\mathbf{P}^T \mathbf{Y} = \mathbf{P}^T \mathbf{MA}$$
$$= \mathbf{QA}. \tag{21}$$

Obviously, the abundance matrix can be easily computed by

$$\mathbf{A} = \mathbf{Q}^{-1} \mathbf{P}^T \mathbf{Y}. \tag{22}$$

Especially, when we do not impose the ANC and the ASC on the abundance matrix $\mathbf{A}$, we observe that all negative values appearing in $\mathbf{A}$ are really close to zero from experimental results. Considering that fact, first, we make that the abundance estimated by OPA meets the ANC constraint by setting all negative abundances to zero or its opposite value. We then divide each abundance vector by their sum to ensure that the abundance estimated by OPA meets the ASC constraint. We refer to the algorithm of estimating abundance as OPA and present its pseudocode in Algorithm 2. From Algorithm 2, we observe that the proposed OPA algorithm for abundance estimation is very efficient because it only involves matrix operations. It is worth noting that OPE and OPA are applicable to the LMM model, and OPE will present excellent performance in endmember extraction under the pure pixel assumption. However, OPA needs an accurate endmember extraction algorithm to extract endmember signatures, and then, it can obtain good abundance estimation results. Especially, in Section III, experimental results validate that our proposed OPA algorithm has low computational complexity in the task of abundance estimation on synthetic data and real Cuprite data.

III. EXPERIMENTAL EVALUATION

We use both synthetic and real Cuprite data and conduct a series of experiments on them to evaluate the effectiveness and efficiency of our proposed OPE and OPA algorithms in the tasks of extracting endmembers and estimating the abundance of each endmember at every pixel. We perform the comparison tests using MATLAB 2018b on an i5-8300 CPU (2.3G). The endmember extraction capacity of our proposed OPE algorithm is assessed with five typical endmember extraction algorithms, i.e., OSP, VCA [12], ATGP [14], ADVMM, and SDVMM [15]. We use certain algorithms, e.g., hyperspectral subspace identification by minimum error (Hysime) [40], to filter the noise of HSI data in advance. In addition, we experimentally evaluate the performance of the proposed OPA algorithm and SUnSPI, CSUnL0, and typical FCLS algorithm [33] in the task of estimating abundance. Especially, considering that SUnSPI and CSUnL0 require the spectral library to complete the final abundance estimation,
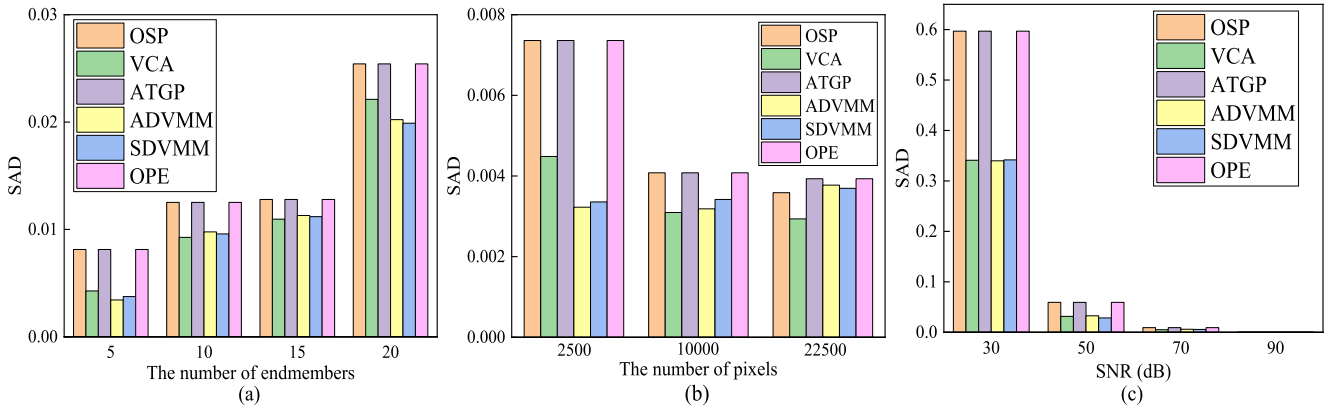
Fig. 2. SAD obtained by different endmember extraction algorithms on synthetic data (containing pure pixels) under (a) different number of endmembers, (b) different number of pixels, and (c) different levels of noise.

we only use them as comparison methods to evaluate the performance of our proposed OPA method on synthetic data.

Let $\hat{\mathbf{m}}_i$ denote the estimated endmember signature by OPE and $\mathbf{m}_i$ are the true endmember signature. We use the spectral angle distance (SAD) to evaluate the estimated endmembers with the ground-truth ones used in our experiments by

$$ \text{SAD} = \sqrt{\frac{1}{c}\sum_{i=1}^{c}\left[\arccos\left(\frac{\hat{\mathbf{m}}_i^T \mathbf{m}_i}{\|\hat{\mathbf{m}}_i\|_2 \|\mathbf{m}_i\|_2}\right)\right]^2} \qquad (23) $$

where c is the number of endmembers. Let $\hat{\mathbf{A}}$ denote the estimated abundance matrix by OPA and $\mathbf{A}$ be the true abundance matrix. Simultaneously $\hat{a}_{ij}$ and a_{ij} are the ijth element of $\hat{\mathbf{A}}$ and $\mathbf{A}$, respectively. We use the root-mean-square error (RMSE) to evaluate the perforamnce of OPA in abundance estimation by

$$ \text{RMSE} = \sqrt{\frac{1}{N}\frac{1}{c}\sum_{i=1}^{N}\sum_{j=1}^{c}(\hat{a}_{ij} - a_{ij})^2} \qquad (24) $$

where N is the number of pixels of HSI dataset. Especially, OPE and OPA get the most accurate results in endmember extraction and abundance estimation when the values of SAD and RMSE are zero.

A. Synthetic Data

The synthetic data are created based on the USGS library. We generate synthetic data based on the linear mixture model [12], [39]. We consider two big families of synthetic data, i.e., with and without pure pixel assumption, to evaluate the effectiveness and efficiency of our proposed OPE and OPA algorithms in extracting endmembers and estimating the abundance of each endmember at every pixel. Each family of synthetic data has three types.

1) Synthetic data with 50×50 pixels and 70-dB noise under different numbers of endmembers, i.e., 5, 10, 15, and 20. We use this type of synthetic data to evaluate the influence of the number of endmembers on the performance of different algorithms in endmember extraction and abundance estimation.

2) Synthetic data with five endmembers and 70-dB noise under a different number of pixels, i.e., 2500, 10 000, and 22 500. The synthetic data are mainly used to exploit the impact of the number of pixels on the performance of different algorithms in the tasks of endmember extraction and abundance estimation.

3) Synthetic data with 50×50 pixels and five endmembers under different levels of noise, i.e., 30, 50, 70, and 90 dB. The synthetic data are designed to evaluate the robustness of different algorithms in the endmember extraction and abundance estimation.

1) Synthetic Data With Pure Pixel Assumption: We employ SAD to evaluate the performance of different algorithms in extracting endmember on three types of synthetic data. Fig. 2 shows the results of SAD obtained by different endmember extraction algorithms on different types of synthetic data. From Fig. 2(a), we see that the values of SAD of all algorithms will increase with the number of endmembers, but our proposed OPE algorithm keeps almost the same accuracy as other popular algorithms. From Fig. 2(b), we observe that the values of SAD will be smaller when the number of pixels increases, and our proposed OPE algorithm also has good performance in endmember extraction compared to other algorithms. From Fig. 2(c), we also observe that our proposed OPE algorithm has good robustness to noise. Therefore, our proposed OPE algorithm is a very effective endmember extraction algorithm.

We employ RMSE as the performance metric and use SUnSPI, CSUnL0, and FCLS as comparison methods to evaluate the performance of the proposed OPA method on different types of synthetic data in the task of abundance estimation, and the results of RMSE are presented in Tables I–III. From Table I, we see that OPA-based methods get more accurate results compared with SUnSPI and CSUnL0, and almost all OPA-based methods are better than the corresponding FCLS-based methods. From Table II, we observe that all OPA-based methods not only are better than SUnSPI, CSUnL0, and the corresponding FCLS-based methods but also obtain the most accurate results in terms of RMSE (its value is very close to zero). From Table III, we obtain that all OPA-based methods have better robustness compared to SUnSPI, CSUnL0, and the

TABLE I
RMSE Obtained by Different Abundance Estimation Algorithms on Synthetic Data (Containing Pure Pixels) With 2500 Pixels and 70-dB Noise Under a Different Number of Endmembers

The number of endmembers	SUnSPI	CSUnL0	OSP		VCA		ATGP		ADVMM		SDVMM		OPE	
			FCLS	OPA	FCLS	OPA	FCLS	OPA	FCLS	OPA	FCLS	OPA	FCLS	OPA
5	0.09094	0.01448	0.11106	**0.00001**	0.12256	**0.00001**	0.11106	**0.00001**	0.12390	**0.00001**	0.12313	**0.00002**	0.12365	**0.00003**
10	0.03504	0.01953	**0.04624**	0.09112	0.04105	**0.01059**	0.04624	**0.09112**	0.04898	**0.00946**	0.04614	**0.00198**	0.04619	**0.00248**
15	0.02111	0.01233	0.03348	**0.00532**	0.03339	**0.00222**	0.03348	**0.00532**	0.03362	**0.00363**	0.03463	**0.00051**	0.03233	**0.00332**
20	0.01774	0.01240	0.02994	**0.00124**	0.03001	**0.00421**	0.02994	**0.00124**	0.03059	**0.02630**	0.02918	**0.00442**	0.03071	**0.00771**

TABLE II
RMSE Obtained by Different Abundance Estimation Algorithms on Synthetic Data (Containing Pure Pixels) With Five Endmembers and 70-dB Noise Under a Different Number of Pixels

The number of pixels	SUnSPI	CSUnL0	OSP		VCA		ATGP		ADVMM		SDVMM		OPE	
			FCLS	OPA	FCLS	OPA	FCLS	OPA	FCLS	OPA	FCLS	OPA	FCLS	OPA
2500	0.11246	0.04831	0.12312	**0.00008**	0.09580	**0.00002**	0.12312	**0.00008**	0.09429	**0.00002**	0.12348	**0.00005**	0.09376	**0.00002**
10000	0.09034	0.00132	0.10883	**0**	0.12182	**0**	0.10883	**0**	0.07686	**0**	0.10886	**0**	0.12198	**0**
22500	0.07992	0.00238	0.07595	**0**	0.11978	**0**	0.10733	**0**	0.07696	**0**	0.11990	**0**	0.12015	**0**

TABLE III
RMSE Obtained by Different Abundance Estimation Algorithms on Synthetic Data (Containing Pure Pixels) With Five Endmembers and 2500 Pixels Under Different Levels of Noise

SNR (dB)	SUnSPI	CSUnL0	OSP		VCA		ATGP		ADVMM		SDVMM		OPE	
			FCLS	OPA	FCLS	OPA	FCLS	OPA	FCLS	OPA	FCLS	OPA	FCLS	OPA
30	0.11082	0.00341	0.11292	**0.00186**	0.11225	**0.00239**	0.11292	**0.00186**	0.11195	**0.00197**	0.11453	**0.00187**	0.11122	**0.00203**
50	0.09310	0.00689	0.10919	**0.00022**	0.09352	**0.00018**	0.10919	**0.00022**	0.10825	**0.00019**	0.07771	**0.00014**	0.10799	**0.00025**
70	0.08796	0.01683	0.12255	**0.00013**	0.12297	**0.00001**	0.12255	**0.00013**	0.12312	**0.00004**	0.11111	**0.00001**	0.09540	**0.00013**
90	0.07941	0.00652	0.08021	**0**	0.12193	**0**	0.08021	**0**	0.11140	**0**	0.10865	**0**	0.07659	**0**

corresponding FCLS-based methods. The above three results reflect that our proposed OPA algorithm is very effective in the task of abundance estimation.

In addition to evaluating the effectiveness of our proposed OPE and OPA algorithms, we also test their efficiency on three different types of synthetic data. Fig. 3 shows the processing time of different endmember extraction algorithms on different types of synthetic data. From Fig. 3(a), we see that OSP and ATGP need more time to complete endmember extraction with the increase of the number of endmembers, but VCA, ADVMM, and SDVMM keep a stable state. Most importantly, our proposed OPE algorithm requires less time on any number of endmembers compared to other algorithms. From Fig. 3(b), we observe that all algorithms both need more time when the number of pixels increases. Similarly, the efficiency of our proposed OPE algorithm is the highest compared to other algorithms. From Fig. 3(c), we obtain that almost all algorithms both maintain a relatively stable state. However, our proposed OPE algorithm is still the most efficient. Moreover, we consider the results of all endmember signatures extracted by different endmember extraction algorithms, i.e., VCA, ATGP, ADVMM, SDVMM, and OPE, to evaluate the efficiency of SunSPI, CSUnL0, FCLS, and

our proposed OPA algorithms in the task of abundance estimation, and Fig. 4 shows their processing time on different types of synthetic data. From Fig. 4(a) and (b), we see that the FCLS-based algorithms (considering different endmember extraction results) need more time to complete abundance estimation with the growth of the number of endmembers and pixels, but OPA-based algorithms keep a stable state on any number of endmembers and pixels and require less time in the task of abundance estimation compared to FCLS-based algorithms. Moreover, SunSPI and CSUnL0 need more time compared with OPA-based algorithms and FCLS-based algorithms. From Fig. 4(c), we infer that all algorithms both maintain a relatively stable state on any noise levels and OPA-based algorithms are still the most efficient compared with SUnSPI, CSUnL0, and the corresponding FCLS-based algorithms. To sum up, our proposed OPE and OPA algorithms for endmember extraction and abundance estimation are not only very effective but also very efficient.

2) Synthetic Data Without Pure Pixel Assumption: To evaluate the performance of the proposed OPE algorithm in the task of endmember extraction on synthetic data without pure pixels, we employ SAD as the performance metric and conduct some related experiments on three different types of

Fig. 3. Processing time (seconds) of different endmember extraction algorithms on synthetic data (containing pure pixels) under (a) different number of endmembers, (b) different number of pixels, and (c) different levels of noise.

Fig. 4. Processing time (seconds) of different abundance estimation algorithms on synthetic data (containing pure pixels) under (a) different number of endmembers, (b) different number of pixels, and (c) different levels of noise.

Fig. 5. SAD obtained by different endmember extraction algorithms on synthetic data (maximum purity of 0.8) under (a) different number of endmembers, (b) different number of pixels, and (c) different levels of noise.

synthetic data. An accurate endmember extraction algorithm has a small value of SAD. Fig. 5 shows the values of SAD obtained by different endmember extraction algorithms on three different types of synthetic data. From Fig. 5(a), we see that the increase of the number of endmembers will reduce the accuracy of all algorithms in endmember extraction. From Fig. 5(b), we observe that the performance of all endmember extraction algorithms basically maintains a stable state with the increment of the number of pixels. From Fig. 5(c), we observe that synthetic data with 50-dB noise will affect the value of

SAD compared to synthetic data with other noise levels. The above three results show that our proposed OPE algorithm is not the best algorithm, but it maintains almost the same accuracy as other algorithms. Therefore, our proposed OPE algorithm is an effective endmember extraction algorithm.

We employ RMSE as the performance metrics to test the performance of SUnSPI, CSUnL0, FCLS, and OPA on three different types of synthetic data without pure pixels. The related results of RMSE are presented in Tables IV–VI. From Table IV, we see that all OPA-based algorithms almost

TABLE IV
RMSE OBTAINED BY DIFFERENT ABUNDANCE ESTIMATION ALGORITHMS ON SYNTHETIC DATA (MAXIMUM PURITY OF 0.8)
WITH 2500 PIXELS AND 70-dB NOISE UNDER A DIFFERENT NUMBER OF ENDMEMBERS

The number of endmembers	SUnSPI	CSUnL0	OSP		VCA		ATGP		ADVMM		SDVMM		OPE	
			FCLS	OPA	FCLS	OPA	FCLS	OPA	FCLS	OPA	FCLS	OPA	FCLS	OPA
5	0.05308	0	0.14065	**0.01061**	0.13302	**0.00426**	0.14065	**0.01061**	0.14743	**0.00320**	0.12222	**0.01122**	0.13292	**0.00413**
10	0.03093	0.02049	0.07898	**0.00528**	0.08014	**0.02595**	0.07898	**0.00528**	0.07942	**0.00396**	0.06853	**0.01613**	0.07377	**0.02014**
15	0.01691	0.01168	0.05155	**0.01907**	0.06014	**0.01689**	0.05155	**0.01907**	0.06014	**0.03853**	0.05040	**0.01177**	0.05249	**0.01564**
20	0.00974	0.00593	0.04541	**0.01956**	0.05001	**0.01201**	0.04541	**0.01956**	0.04751	**0.01093**	0.04504	**0.05630**	0.04941	**0.01562**

TABLE V
RMSE OBTAINED BY DIFFERENT ABUNDANCE ESTIMATION ALGORITHMS ON SYNTHETIC DATA (MAXIMUM PURITY OF 0.8)
WITH FIVE ENDMEMBERS AND 70-dB NOISE UNDER A DIFFERENT NUMBER OF PIXELS

The number of pixels	SUnSPI	CSUnL0	OSP		VCA		ATGP		ADVMM		SDVMM		OPE	
			FCLS	OPA	FCLS	OPA	FCLS	OPA	FCLS	OPA	FCLS	OPA	FCLS	OPA
2500	0.10287	0.0001	0.15327	**0.00937**	0.13493	**0.01732**	0.15327	**0.00937**	0.14755	**0.00315**	0.15006	**0.00337**	0.13877	**0.01003**
10000	0.10127	0.05243	0.15072	**0.00470**	0.15440	**0.00416**	0.15072	**0.00470**	0.13351	**0.00350**	0.14776	**0.00429**	0.14815	**0.00347**
22500	0.09462	0.05262	0.14990	**0.00526**	0.12503	**0.00560**	0.11469	**0.00858**	0.13903	**0.00322**	0.11409	**0.00847**	0.15531	**0.00928**

TABLE VI
RMSE OBTAINED BY DIFFERENT ABUNDANCE ESTIMATION ALGORITHMS ON SYNTHETIC DATA (MAXIMUM PURITY OF 0.8)
WITH FIVE ENDMEMBERS AND 2500 PIXELS UNDER DIFFERENT LEVELS OF NOISE

SNR (dB)	SUnSPI	CSUnL0	OSP		VCA		ATGP		ADVMM		SDVMM		OPE	
			FCLS	OPA	FCLS	OPA	FCLS	OPA	FCLS	OPA	FCLS	OPA	FCLS	OPA
30	0.08959	0.05321	0.19535	**0.10531**	0.19271	**0.06483**	0.19535	**0.10531**	0.15900	**0.12292**	0.14197	**0.21858**	0.16270	**0.04545**
50	0.07811	0.05501	0.14459	**0.00682**	0.14809	**0.02228**	0.14459	**0.00682**	0.13449	**0.00342**	0.09217	**0.00656**	0.15385	**0.01478**
70	0.10233	0.05154	0.12348	**0.00440**	0.10606	**0.02266**	0.12348	**0.0044**	0.12472	**0.00353**	0.09801	**0.00441**	0.11928	**0.00433**
90	0.07665	0.05411	0.13686	**0.00610**	0.16143	**0.00927**	0.13686	**0.0061**	0.11537	**0.00323**	0.13325	**0.00445**	0.14313	**0.00354**

get more accurate results compared with the corresponding FCLS-based algorithms. Moreover, SUnSPI is better than OPA-based algorithms on synthetic data with 20 endmembers, and CSUnL0 is better than OPA-based algorithms on synthetic data with 5 and 20 endmembers. From Table V, we obtain that OPA-based algorithms are better than the corresponding FCLS-based algorithms in the task of abundance estimation based on different endmember extraction results. Moreover, OPA-based algorithms are generally better than SUnSPI and CSUnL0. From Table VI, we observe that OPA-based algorithms have smaller RMSE values compared to SUnSPI, CSUnL0, and the corresponding FCLS-based algorithms, and they get more accurate results than SUnSPI, CSUnL0, and FCLS-based algorithms in abundance estimation. Moreover, although the advantages of OPA over FCLS are not obvious on synthetic data without pure pixels, OPA-based algorithms are still better than SUnSPI, CSUnL0, and the corresponding FCLS-based algorithms under any number of endmembers, any number of pixels, and any noise levels. Therefore, our proposed OPA algorithm is very effective in the task

of estimating the abundance of each endmember at every pixel.

Fig. 6 shows the efficiency of different endmember extraction algorithms. In terms of endmember extraction, from Fig. 6(a), we see that OSP and ATGP require more time when the number of endmembers increases, ADVMM fluctuates up and down, and VCA, SDVMM, and our proposed OPE algorithm keep a stable state. Moreover, compared to other endmember extraction algorithms, our proposed OPE algorithm is the most efficient endmember extraction algorithm at any number of endmembers. From Fig. 6(b), we see that all endmember extraction algorithms both require more time with the growth of the number of pixels, but OPE still has the highest efficiency compared to other algorithms at any number of pixels. From Fig. 6(c), we obtain that all algorithms both maintain a stable state at any noise levels, but our proposed OPE algorithm is still the most efficient. In addition, we use all results of different endmember extraction algorithms to evaluate the efficiency of SUnSPI, CSUnL0, OPA-based algorithms, and FCLS-based algorithms in abundance estimation,
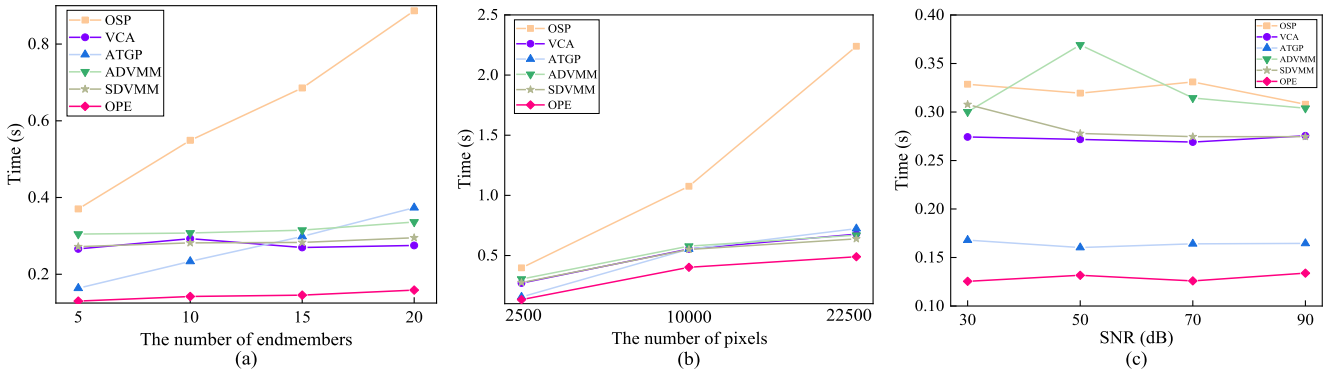
Fig. 6. Processing time (seconds) of different endmember extraction algorithms on synthetic data (maximum purity of 0.8) under (a) different number of endmembers, (b) different number of pixels, and (c) different levels of noise.
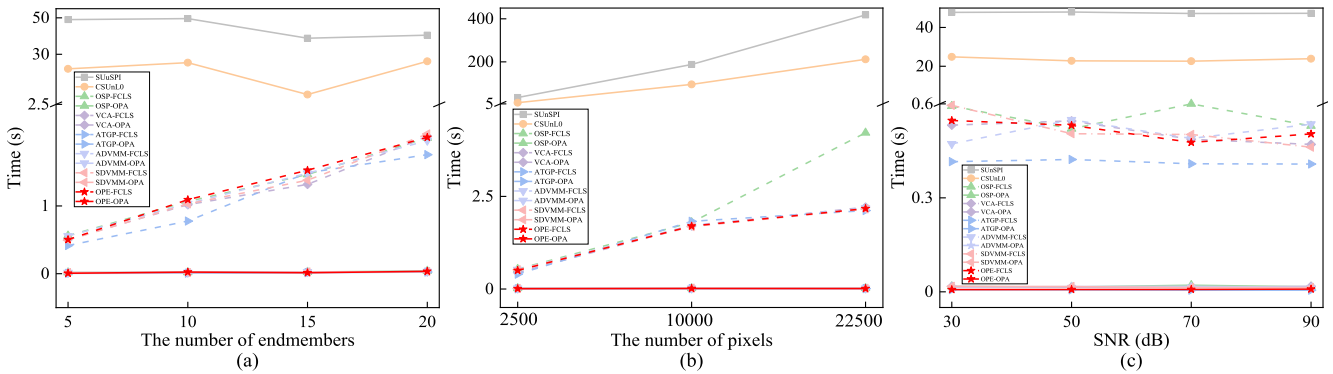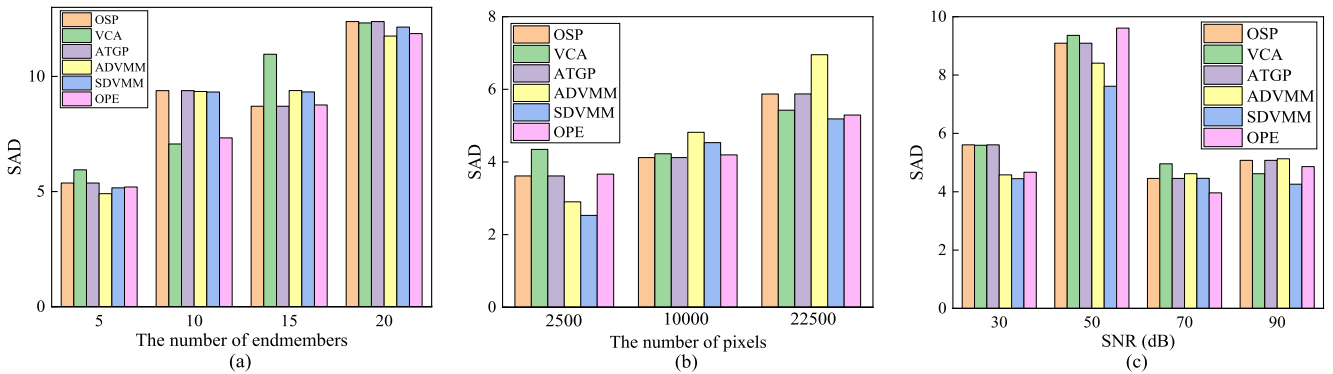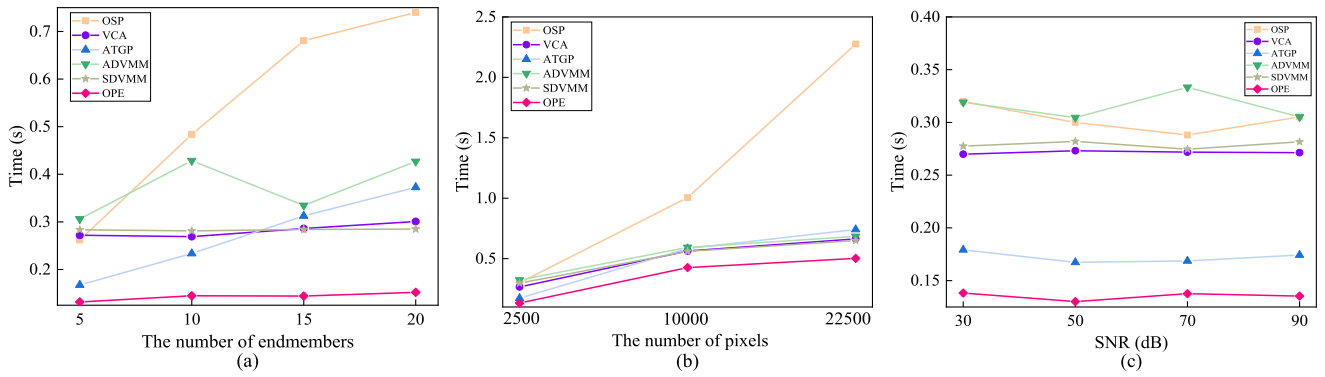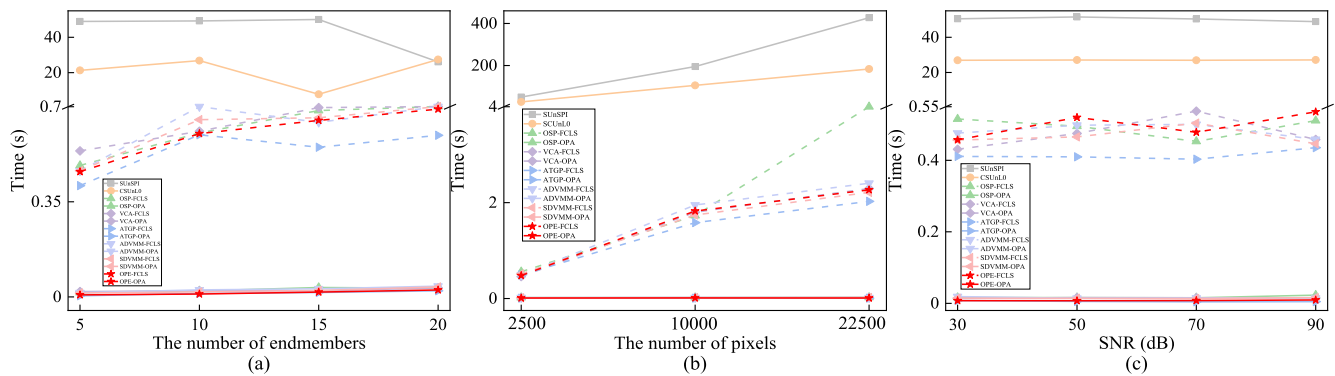

Fig. 7. Processing time (seconds) of different abundance estimation algorithms on synthetic data (maximum purity of 0.8) under (a) different number of endmembers, (b) different number of pixels, and (c) different levels of noise.
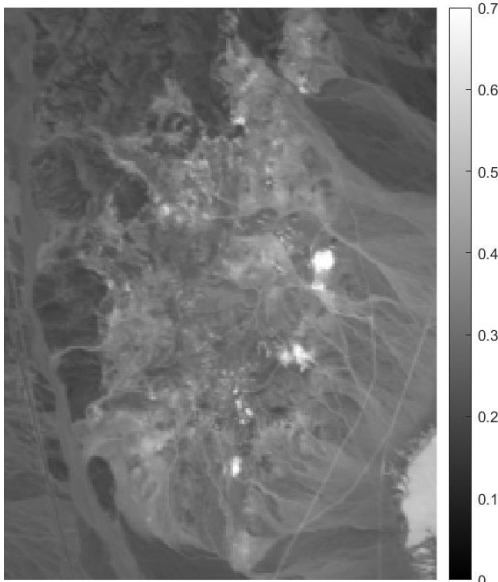
Fig. 8. 30th band of the considered AVIRIS Cuprite subimage.

and their processing time is presented in Fig. 7. From Fig. 7(a), we see that the processing time of SUnSPI, CSUnL0, and FCLS-based algorithms in abundance estimation fluctuates up and down with the growth of the number of endmembers. Moreover, we also see that OPA-based algorithms keep a stable state at any number of endmembers and require less time compared to SUnSPI, CSUnL0, and the corresponding FCLS-based algorithms. From Fig. 7(b), we observe that the processing time of SUnSPI, CSUnL0, and FCLS-based algorithms in abundance estimation will increase with the growth of the number of pixels. Simultaneously, OPA-based algorithms still keep a stable state and have the highest efficiency at any number of pixels compared to SUnSPI, CSUnL0, and the corresponding FCLS-based algorithms. From Fig. 7(c), we obtain that SUnSPI, CSUnL0, FCLS-based algorithms, and OPA-based algorithms have a relatively stable state, but OPA-based algorithms have the lowest processing time at any noise levels in abundance estimation. To sum up, our proposed OPE and OPA algorithms for endmember extraction and abundance estimation are not only very effective but also very efficient.

B. Real Cuprite Data

We use the well-known AVIRIS Cuprite data[2] to evaluate the effectiveness of efficiency of our proposed OPE and OPA algorithms in the endmember extraction and abundance estimation. The portion used in experiments corresponds to

[2]http://aviris.jpl.nasa.gov/html/aviris.freedata.html

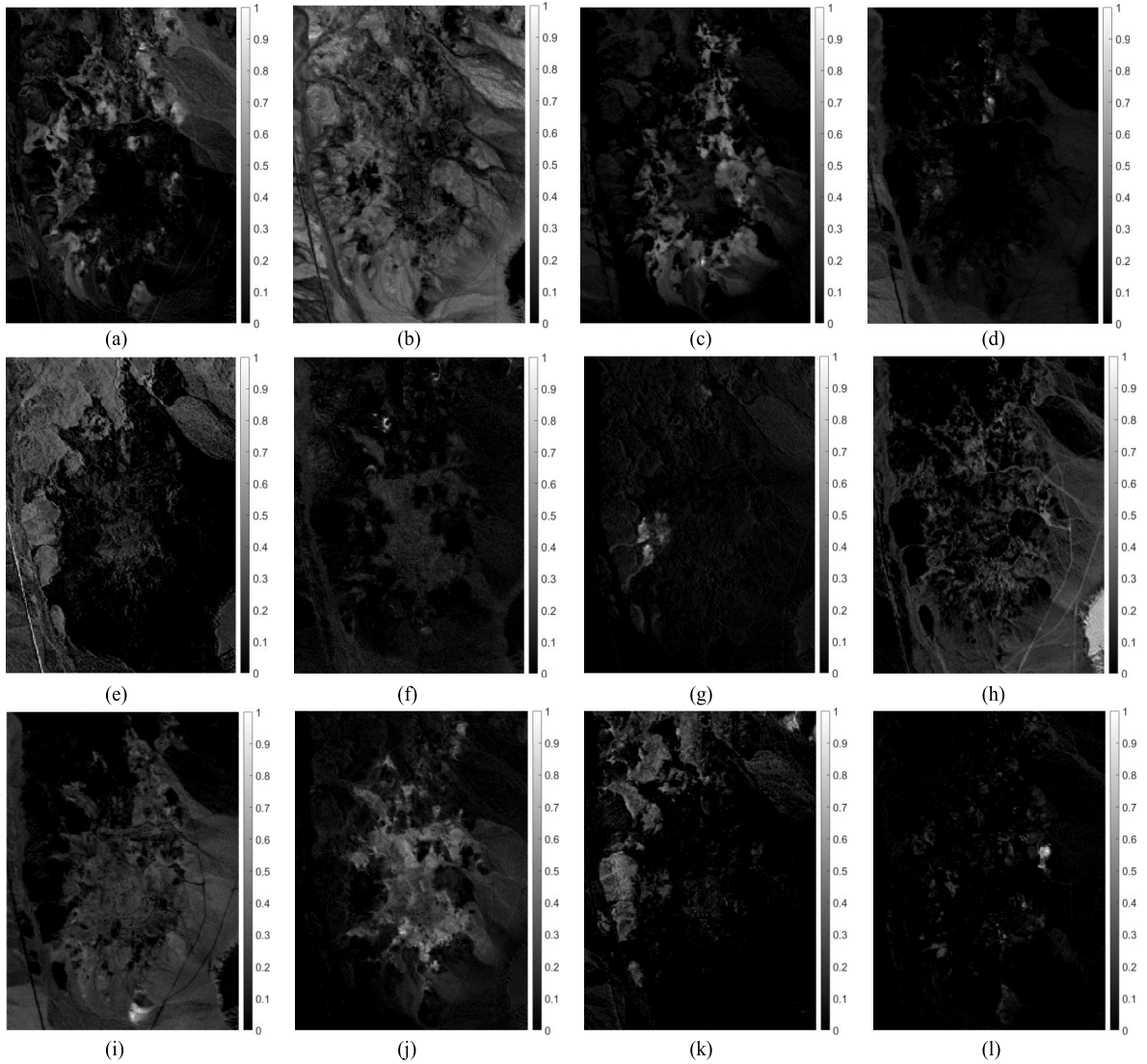Fig. 9. Abundance fraction maps estimated by our proposed OPA algorithm from the Cuprite subimage. (a) Montmorillonite SCa-2.b. (b) Jarosite WS368 Pb #1. (c) Montmorillonite CM20. (d) Nontronite SWa-1.a. (e) Montmorillonite + Illi CM37 #1. (f) Nontronite NG-1.a. (g) Kaolin/Smect H89-FR-5 30K #1. (h) Montmorillonite + Illi CM37 #2. (i) Kaolin/Smect H89-FR-5 30K #2. (j) Kaolin/Smect H89-FR-5 30K #3. (k) Jarosite WS368 Pb #2. (l) Chalcedony CU91-6A.

TABLE VII

SAD AND THE PROCESSING TIME OBTAINED BY DIFFERENT ENDMEMBER EXTRACTION ALGORITHMS ON CUPRITE SUBIMAGE

Metrics	OSP	VCA	ATGP	ADVMM	SDVMM	OPE
SAD	4.3723	4.8400	4.3723	4.6830	**4.3461**	4.4785
Time	5.9829	**3.6465**	4.3723	4.07224	3.9454	4.0650

a 250×191 pixel subset of Cuprite data, with 224 spectral bands in the range 0.4–2.5 μm and a nominal spectral resolution of 10 nm. Due to water absorption and low SNR, we remove bands 1–6, 105–115, 150–170, and 222–224 and retain 183 spectral bands in our experiments. Fig. 8 shows the 30th band of Cuprite subimage for this experiment.

We employ SAD to evaluate the effectiveness of our proposed OPE algorithm on the Cuprite subimage. Table VII shows the values of SAD obtained by different endmember extraction algorithms and their processing time. From Table VII, we see that our proposed OPE endmember extraction algorithm is not optimal or the fastest, but it gets almost the same results as other algorithms. Therefore, our proposed OPE algorithm is an effective endmember extraction algorithm.

Moreover, Fig. 9 shows the abundance maps obtained by our proposed OPA algorithm. Simultaneously, we also consider the reconstruction error (RSE) to evaluate the performance of our proposed OPA algorithm in abundance estimation by

$$\text{RSE} = \frac{||\mathbf{Y} - \hat{\mathbf{Y}}||_F}{||\mathbf{Y}||_2} \qquad (25)$$

where $||\cdot||_F$ indicates the Frobenius norm, $||\cdot||_2$ represents the L_2-norm, $\mathbf{Y}$ is the real Cuprite subimage dataset, and $\hat{\mathbf{Y}}$ is the estimated Cuprite subimage dataset. The related results of

TABLE VIII
RSE AND THE PROCESSING TIME OF DIFFERENT ABUNDANCE ESTIMATION ALGORITHMS ON CUPRITE SUBIMAGE

Metrics	OSP		VCA		ATGP		ADVMM		SDVMM		OPE	
	FCLS	OPA	FCLS	OPA	FCLS	OPA	FCLS	OPA	FCLS	OPA	FCLS	OPA
RSE	**0.50515**	0.52086	**0.57674**	0.60858	**0.50515**	0.52086	**0.52792**	0.54824	**0.50978**	0.52063	**0.51160**	0.54986
Time	11.37464	**0.03294**	10.11134	**0.02912**	10.21231	**0.02803**	10.47131	**0.03055**	10.11796	**0.02843**	10.05766	**0.02345**

RSE and the processing time of FCLS and OPA in abundance estimation based on different endmember extraction algorithms are presented in Table VIII. From Table VIII, we see that OPA-based algorithms may not be optimal in abundance estimation, but they get almost the same results as FCLS-based algorithms. Simultaneously, OPA-based algorithms are always faster than FCLS-based algorithms in abundance estimation on the Cuprite subimage, and they are about 500 times faster than FCLS-based algorithms. The results obtained by OPA on synthetic data and real Cuprite data show that our proposed OPA algorithm can be considered as a new and fast abundance estimation algorithm to replace FCLS.

IV. CONCLUSION AND FUTURE WORK

We have developed a new and fast orthogonal projection framework to complete the overall hyperspectral unmixing task, i.e., extracting endmembers and estimating the abundance of each endmember at every pixel. We refer to the proposed algorithms for endmember extraction and abundance estimation as OPE and OPA, respectively. The proposed OPE algorithm completes endmember extraction by searching for the orthogonal vector of endmembers that have been extracted and chooses the pixels with the maximum projection to the orthogonal vector as endmembers. The proposed OPA algorithm computes a diagonal matrix whose multiplication with the endmember matrix is not only a square matrix but also a diagonal matrix and then exploits some specific operations of matrices to estimate the abundance of each endmember at every pixel. Both OPE and OPA only require fast matrix operations, and they have high efficiency in the tasks of endmember extraction and abundance estimation, respectively. Simultaneously, the proposed OPE and OPA algorithms have been evaluated on synthetic data and real Cuprite data, and the experimental results show that they are not only very efficient but also very effective in the endmember extraction and abundance estimation.

As with any new approach, there are some unresolved issues that may present challenges over time. Especially, we assume that HSI is represented by the multiplication of the endmember matrix and the abundance matrix. In our future work, we will investigate how to extract endmembers and estimate abundances from HSIs that are nonlinearly mixed.

REFERENCES

[1] M. E. Paoletti, J. M. Haut, X. Tao, J. Plaza, and A. Plaza, "FLOP-reduction through memory allocations within CNN for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 5938–5952, Jul. 2021.

[2] X. Lu, B. Wang, X. Zheng, and X. Li, "Exploring models and data for remote sensing image caption generation," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 2183–2195, Apr. 2018.

[3] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. M. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.

[4] B. Pan, Z. Shi, and X. Xu, "MugNet: Deep learning for hyperspectral image classification using limited samples," *ISPRS J. Photogramm. Remote Sens.*, vol. 145, pp. 108–119, Nov. 2018.

[5] J. C. Harsanyi and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 4, pp. 779–785, Jul. 1994.

[6] X. Tao, T. Cui, and P. Ren, "Cofactor-based efficient endmember extraction for green algae area estimation," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 6, pp. 849–853, Jun. 2019.

[7] X. Tao, T. Cui, A. Plaza, and P. Ren, "Simultaneously counting and extracting endmembers in a hyperspectral image based on divergent subsets," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 12, pp. 8952–8966, Dec. 2020.

[8] J. Li, J. M. Bioucas-Dias, A. Plaza, and L. Liu, "Robust collaborative nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6076–6090, Oct. 2016.

[9] X. Xu, Z. Shi, and B. Pan, "A supervised abundance estimation method for hyperspectral unmixing," *Remote Sens. Lett.*, vol. 9, no. 4, pp. 383–392, Apr. 2018.

[10] A. M. Ambikapathi, T. H. Chan, C. Y. Chi, and K. Keizer, "Hyperspectral data geometry-based estimation of number of endmembers using p-norm-based pure pixel identification algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2753–2769, May 2013.

[11] C.-I. Chang, "Orthogonal subspace projection (OSP) revisited: A comprehensive study and analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 502–518, Mar. 2005.

[12] J. M. P. Nascimento and J. M. Bioucas-Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, Apr. 2005.

[13] M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data," *Proc. SPIE*, vol. 3753, pp. 266–275, Oct. 1999.

[14] H. Ren and C.-I. Chang, "Automatic spectral target recognition in hyperspectral imagery," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1232–1249, Oct. 2003.

[15] T.-H. Chan, J.-Y. Liou, A. Ambikapathi, W.-K. Ma, and C.-Y. Chi, "Fast algorithms for robust hyperspectral endmember extraction based on worst-case simplex volume maximization," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2012, pp. 1237–1240.

[16] L. Drumetz, J. Chanussot, C. Jutten, W.-K. Ma, and A. Iwasaki, "Spectral variability aware blind hyperspectral image unmixing based on convex geometry," *IEEE Trans. Image Process.*, vol. 29, pp. 4568–4582, 2020.

[17] S. Das, A. Routray, and A. K. Deb, "Convex set based abundance constrained unmixing of hyperspectral image," in *Proc. 14th IEEE India Council Int. Conf. (INDICON)*, Dec. 2017, pp. 1–5.

[18] S. Das, S. Chakraborty, A. Routray, and A. K. Deb, "Fast linear unmixing of hyperspectral image by slow feature analysis and simplex volume ratio approach," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2019, pp. 560–563.

[19] T.-H. Chan, C.-Y. Chi, Y.-M. Huang, and W.-K. Ma, "A convex analysis-based minimum-volume enclosing simplex algorithm for hyperspectral unmixing," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4418–4432, Nov. 2009.

[20] J. Li and J. M. Bioucas-Dias, "Minimum volume simplex analysis: A fast algorithm to unmix hyperspectral data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, vol. 3, 2008, pp. III–250.

[21] J. Li, A. Agathos, D. Zaharie, J. M. Bioucas-Dias, A. Plaza, and X. Li, "Minimum volume simplex analysis: A fast algorithm for linear hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 9, pp. 5067–5082, Sep. 2015.

[22] S. Zhang, A. Agathos, and J. Li, "Robust minimum volume simplex analysis for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 11, pp. 6431–6439, Nov. 2017.

[23] A. Marinoni and P. Gamba, "A novel approach for efficient p-linear hyperspectral unmixing," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 6, pp. 1156–1168, Sep. 2015.

[24] M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2014–2039, Aug. 2011.

[25] W. Tang, Z. Shi, Y. Wu, and C. Zhang, "Sparse unmixing of hyperspectral data using spectral *a priori* information," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 2, pp. 770–783, Feb. 2015.

[26] Z. Shi, T. Shi, M. Zhou, and X. Xu, "Collaborative sparse hyperspectral unmixing using l_0 norm," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5495–5508, Sep. 2018.

[27] X. Xu and Z. Shi, "Multi-objective based spectral unmixing for hyperspectral images," *ISPRS J. Photogramm. Remote Sens.*, vol. 124, pp. 54–69, Feb. 2017.

[28] S. Das, A. Routray, and A. K. Deb, "Sparsity measure based library aided unmixing of hyperspectral image," *IET Image Process.*, vol. 13, no. 12, pp. 2077–2085, Oct. 2019.

[29] M. Tang, L. Gao, A. Marinoni, P. Gamba, and B. Zhang, "Integrating spatial information in the normalized P-linear algorithm for nonlinear hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 4, pp. 1179–1190, Apr. 2018.

[30] M. Tang, B. Zhang, A. Marinoni, L. Gao, and P. Gamba, "Multiharmonic postnonlinear mixing model for hyperspectral nonlinear unmixing," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 11, pp. 1765–1769, Nov. 2018.

[31] L. Drumetz, M.-A. Veganzones, S. Henrot, R. Phlypo, J. Chanussot, and C. Jutten, "Blind hyperspectral unmixing using an extended linear mixing model to address spectral variability," *IEEE Trans. Image Process.*, vol. 25, no. 8, pp. 3890–3905, Aug. 2016.

[32] Y. Altmann, A. Halimi, N. Dobigeon, and J.-Y. Tourneret, "Supervised nonlinear spectral unmixing using a postnonlinear mixing model for hyperspectral imagery," *IEEE Trans. Image Process.*, vol. 21, no. 6, pp. 3017–3025, Jun. 2012.

[33] D. C. Heinz and C.-I. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar. 2001.

[34] J. M. Bioucas-Dias and M. A. T. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *Proc. 2nd Workshop Hyperspectral Image Signal Process., Evol. Remote Sens.*, Jun. 2010, pp. 1–4.

[35] P.-A. Thouvenin, N. Dobigeon, and J.-Y. Tourneret, "Hyperspectral unmixing with spectral variability using a perturbed linear mixing model," *IEEE Trans. Signal Process.*, vol. 64, no. 2, pp. 525–538, Jan. 2016.

[36] D. Hong, N. Yokoya, J. Chanussot, and X. X. Zhu, "An augmented linear mixing model to address spectral variability for hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1923–1938, Apr. 2019.

[37] T. Zhi, B. Yang, Z. Chen, and B. Wang, "Nonnegative matrix factorization with constraints on endmember and abundance for hyperspectral unmixing," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2017, pp. 1149–1152.

[38] Y. Yu and W. Sun, "Minimum distance constrained non-negative matrix factorization for the endmember extraction of hyperspectral images," *Proc. SPIE*, vol. 6790, Nov. 2007, Art. no. 679015.

[39] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 765–777, Mar. 2007.

[40] J. M. Bioucas-Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sensing*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.

Escuela Politecnica
Av. de la Universidad, S/N, 10003
Caceres, Spain
Phone: 0034927257000. Ext. 51662
Email: aplaza,juanmariohaut{@unex.es}

Dr. Antonio Plaza Miguel y Dr. Juan Mario Haut Hurtado como directores de la tesis titulada "Procesamiento Eficiente y Profundo de Imágenes Hiperespectrales y Aplicación en Tareas de Desmezclado Espectral", certifican el factor de impacto y la categorización de la siguiente publicación, incluída en la tesis doctoral. Del mismo modo, se especifica la aportación del doctorado.

Antonio Plaza Miguel PhD, and Juan Mario Haut Hurtado PhD, as directors of the Phd thesis titled "Efficient and Deep Processing of Hyperspectral Images and Application in Spectral Unmixing Tasks", certify the impact factor and the categorization of the following publication, included in the doctoral thesis. In the same way, the contribution of the doctorate is specified.

Artículo / Paper

Autores/Authors: **X. Tao**, M. E. Paoletti, L. Han, Z. Wu, P. Ren, J. Plaza, A. Plaza, and J. M. Haut.
Title: A New Deep Convolutional Neural Network for Effective Hyperspectral Unmixing.
Journal: IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15, pp. 6999-7012, August 2022.
DOI: 10.1109/JSTARS.2022.3200733. Impact factor 2021: 4.715. Q2

Abstract: Most traditional hyperspectral unmixing methods require the results of other endmember extraction algorithms to complete the abundance estimation step. Due to the impressive learning and data fitting capabilities of convolutional neural networks (CNNs), deep learning (DL)-based hyperspectral unmixing technologies have rapidly developed in the literature. According to the procedure used to combine different layers (i.e., fully connected layers, convolution layers, and activation layers), these techniques are mainly divided into three main categories, i.e., those based on autoencoder networks, convolutional neural networks, and convolutional autoencoder networks. They usually extract the weight and output of a specific activation layer as endmember signatures and abundance maps, respectively. Moreover, most existing DL-based unmixing approaches usually use 2-D CNNs to learn the features contained in hyperspectral images, and very few approaches employ 3-D CNNs to extract spectral and spatialinformation. However, 2-D CNN-based techniques cannot capture good discriminative feature maps from the spectral viewpoint, and 3-D CNN-based techniques usually have high computational overload. In this work, to further exploit the feature extraction capability of CNNs, we combine 3-D and 2-D convolutions to propose a cross-convolution unmixing network (CrossCUN) for hyperspectral unmixing. Simultaneously, to better illustrate the improvements of our proposed CrossCUN, we also build the corresponding 2-D convolution unmixing network (2-DCUN) and 3-D convolution unmixing network (3-DCUN). We evaluate the performance of our newly developed networks on two types of synthetic datasets and three real hyperspectral images. Experimental results show that the proposed networks not only obtain better results than other DL-based unmixing methods but also do not require any prior knowledge (e.g., the results of other endmember extraction algorithms) to estimate the abundance maps.

Contribución del doctorado: Planteamiento de la hipótesis, desarrollo práctico, análisis y discusión de los resultados, elaboración y escritura del manuscrito.

Firma / Signature
Marzo / Marzo, 2023

Antonio Plaza Miguel   Juan Mario Haut Hurtado

A New Deep Convolutional Network for Effective Hyperspectral Unmixing

Xuanwen Tao ⓘ, *Student Member, IEEE*, Mercedes E. Paoletti ⓘ, *Senior Member, IEEE*, Lirong Han ⓘ,
Zhaoyue Wu ⓘ, Peng Ren ⓘ, *Senior Member, IEEE*, Javier Plaza ⓘ, *Senior Member, IEEE*,
Antonio Plaza ⓘ, *Fellow, IEEE*, and Juan M. Haut ⓘ, *Senior Member, IEEE*

Abstract—Hyperspectral unmixing extracts pure spectral constituents (endmembers) and their corresponding abundance fractions from remotely sensed scenes. Most traditional hyperspectral unmixing methods require the results of other endmember extraction algorithms to complete the abundance estimation step. Due to the impressive learning and data fitting capabilities of convolutional neural networks (CNNs), deep learning (DL)-based hyperspectral unmixing technologies have rapidly developed in the literature. According to the procedure used to combine different layers (i.e., fully connected layers, convolution layers, and activation layers), these techniques are mainly divided into three main categories, i.e., those based on autoencoder networks, convolutional neural networks, and convolutional autoencoder networks. They usually extract the weight and output of a specific activation layer as endmember signatures and abundance maps, respectively. Moreover, most existing DL-based unmixing approaches usually use 2-D CNNs to learn the features contained in hyperspectral images, and very few approaches employ 3-D CNNs to extract spectral and spatial information. However, 2-D CNN-based techniques cannot capture good discriminative feature maps from the spectral viewpoint, and 3-D CNN-based techniques usually have high computational overload. In this work, to further exploit the feature extraction capability of CNNs, we combine 3- and 2-D convolutions to propose a cross-convolution unmixing network (CrossCUN) for hyperspectral unmixing. Simultaneously, to better illustrate the improvements of our proposed CrossCUN, we also build the corresponding 2-D convolution unmixing network (2-DCUN) and 3-D convolution unmixing network (3-DCUN). We evaluate the performance of our newly developed networks on two types of synthetic datasets and three real hyperspectral images. Experimental results show that the proposed networks not only obtain better results than other DL-based unmixing methods but also do not require any prior knowledge (e.g., the results of other endmember extraction algorithms) to estimate the abundance maps.

Index Terms—Convolutional neural networks (CNNs), cross convolution, deep learning, hyperspectral unmixing.

I. INTRODUCTION

HYPERSPECTRAL images (HSIs) have been widely studied and applied in different tasks with the aim of processing and analyzing their sheer amount of information. This has been done through various techniques, such as image classification [1], [2], [3], [4], [5], data fusion [6], [7], [8], target detection [9], [10], [11], anomaly detection [12], [13], [14], [15], data denoising [16], [17], [18], and so on. In this context, hyperspectral unmixing [19], [20], [21], [22], [23] is one of the most important applications of HSI data processing. It aims at addressing the problem of low spatial resolution HSIs, which usually contain many mixed pixels. These pixels pose a significant challenge for identifying the materials that compose the considered scenes. To deal with this problem, hyperspectral unmixing techniques decompose mixed pixels into pure materials (endmembers) and their corresponding abundance maps at each pixel. The linear (LMM) [24], [25], [26], [27], [28], [29], [30], [31] and nonlinear mixing model (NLMM) [32], [33], [34] are the two most widely used models for hyperspectral unmixing. The LMM assumes that the observed pixel spectrum is a linear combination of endmember signatures, where the weights conform the corresponding abundance maps that need to be estimated. Due to the simplicity and effectiveness of the LMM, significant efforts to develop hyperspectral unmixing approaches based on the LMM have been conducted in the literature. These hyperspectral unmixing approaches are mainly used to complete two significant tasks, i.e., endmember extraction and abundance estimation.

Traditional endmember extraction methods include vertex component analysis (VCA), automatic target generation process (ATGP), negative abundance-oriented (NABO), and alternating decoupled volume max-min (ADVMM), among others. VCA [26] iteratively projects the HSI data to one direction, which is orthogonal to the subspace formed by the determined endmembers. As a result, the pixel with the extreme projection is extracted as the new endmember. ATGP [29] uses the notion of orthogonal subspace projection to extract endmember signatures. NABO

[30] considers the pixels outside the hull as alternative candidate endmembers to complete the unmixing task. ADVMM [31] addresses the worst case simplex volume maximization problem by alternating optimization. In addition to these approaches, within the current literature, there are also other interesting probabilistic techniques, developed as an attempt to address some of the shortcomings faced by the previous methods [35], [36]. Nevertheless, these methods are computationally very expensive. In contrast, the maximum distance analysis (MDA) [37], [38] has proven to be a simple but effective unmixing method that does not require any prior knowledge about the number of endmembers. However, the above methods both focus on endmember identification and need to be combined with other abundance estimation methods to perform the abundance estimation task.

Traditional abundance estimation methods include fully constrained least squares (FCLS) [39], sparse unmixing by variable splitting and augmented Lagrangian (SUnSAL) [40], spatial group sparsity regularized nonnegative matrix factorization (SGSNMF) [41], minimum volume-constrained nonnegative matrix factorization (MVCNMF) [27], minimum-volume enclosing simplex (MVES) [24], and robust collaborative nonnegative matrix factorization (R-CoNMF) [42], among others. SUnSAL utilizes the alternating direction method of multipliers to decompose a difficult problem into a sequence of simpler ones. SGSNMF incorporates the group-structured prior information of hyperspectral images into the nonnegative matrix factorization optimization to complete abundance estimation, where the data are organized into spatial groups. MVCNMF includes a volume constraint into the nonnegative matrix factorization formulation to integrate least-squares analysis and the convex geometry model to perform the unmixing task. MVES incorporates convex analysis and Craig's criterion to enhance the performance of spectral unmixing. It is worth noting that FCLS, SUnSAL, and SGSNMF need the results of endmember extraction from other methods to complete abundance estimation, and MVCNMF and MVES can simultaneously complete endmember extraction and abundance estimation. Compared with the above methods, R-CoNMF can simultaneously complete the three tasks, i.e., determining the number of endmembers, extracting endmember signatures, and estimating the abundance maps. Especially, a different advantage of R-CoNMF compared to other abundance estimation methods is that it completes abundance estimation without any prior knowledge of the number of endmembers.

A. Deep Learning-Based Methods for Hyperspectral Unmixing

Recently, based on their impressive learning and data fitting capabilities, some DL-based unmixing methods have been proposed in the literature [43], [44], [45], [46], [47], [48], [49], [50]. Depending on how different layers are combined, these DL-based unmixing methods are mainly divided into three groups, i.e., autoencoder networks, convolutional neural networks, and convolutional autoencoder networks. These three types of unmixing methods usually extract the endmember signatures and the corresponding abundance maps from the weights and outputs of a specific activation layer, respectively.

Regarding the first group, unmixing methods based on autoencoder networks mainly use fully connected layers and activation layers. In [43], Guo et al. proposed an autoencoder cascade framework that integrates a denoising autoencoder and a nonnegative sparse autoencoder in order to conduct data denoising and endmember estimation, respectively. In this sense, the autoencoder cascade framework is the first attempt to apply DL-based methods to spectral unmixing, improving the performance of unmixing by considering the observation noise and a sparsity prior. Nevertheless, its performance may be affected due to the fact that the same encoder and decoder is used. To overcome this limitation, in [44], Qu et al. proposed a so-called untied denoising autoencoder with sparsity (uDAS) method to address the unsupervised unmixing problem for HSIs with high noise levels. uDAS incorporates denoising capabilities in the form of a constraint to reduce the reconstruction errors and introduces a novel backpropagation method that uses simple matrix operations to effectively force the endmembers to be nonnegative (while the abundance vector is constrained to sum to one). However, both the autoencoder cascade framework and uDAS ignore spatial information and aim at addressing LMM-based problems.

On the other hand, unmixing methods based on CNNs mainly contain convolutional and activation layers. For instance, in [45], Palsson et al. proposed a novel spectral and spatial linear mixture model by developing an associated estimation method based on a convolutional neural network autoencoder unmixing (CNNAEU) to perform the overall unmixing task, i.e., endmember extraction and abundance estimation. CNNAEU extends the LMM to consider the situation where neighboring pixels take part in the pixel reconstruction. As a result, it represents the first attempt to use a CNN to directly exploit the spatial correlation in HSIs for spectral unmixing. Nevertheless, CNNAEU requires a significant number of training samples to adequately fit a large number of learnable parameters. In this sense, the network quickly tends to overfit. In addition, CNNAEU is designed for LMM-based problems. In [46], Gao et al. proposed a cycle-consistency unmixing network (Cycu-Net) which learns two cascaded autoencoders (in an end-to-end manner) to more effectively improve unmixing performance. Cycu-Net designs a cycle-consistency strategy and introduces a new self-perception loss containing two spectral reconstruction terms and an abundance reconstruction term to further refine the unmixing process. However, it ignores the challenges introduced by the high spectral dimensionality of the data.

Finally, unmixing methods based on convolutional autoencoder networks mainly use fully connected layers, convolutional layers, and activation layers. For instance, in [47], Yasiru et al. developed a convolutional autoencoder (CAE) architecture for spectral unmixing. CAE captures the spatial distribution in the original data through convolution filters, which are used to parameterize the spectral features. The encoder and decoder parts of CAE are used to translate the feature space into a latent space representation and to reconstruct the input from the latent space, respectively. However, CAE is based on the LMM and is generally unsuitable for nonlinear unmixing problems. In [48], Qi et al. proposed a deep spectral convolution network with spectral library (SCSL) support to conduct spectral unmixing.

SCNL uses convolutional layers and fully connected layers to capture features and estimate abundance maps, respectively. Moreover, it exploits another fully connected layer to reconstruct the original data. Although SCNL reduces complexity and improves the utilization of the network, it heavily relies on the quality of the spectral library.

B. Contributions of This Work

As mentioned, existing unmixing methods based on DL architectures usually adopt autoencoder networks, CNNs, and convolutional autoencoder networks to perform spectral unmixing. The methods based on autoencoder networks are the most commonly used, whilst those involving CNNs and convolutional autoencoder networks are quite rare. Moreover, most existing unmixing methods usually employ 2-D convolution kernels to extract features, but 3-D CNNs have the ability to capture spectral and spatial features. Nevertheless, only using 2-D-pure CNNs or 3-D-pure CNNs has some disadvantages, such as the inability to capture information on the relationship between the channels or the need to design a very complex model. For instance, by only using 2-D kernels, the model cannot learn good discriminative feature maps from the spectral information. Similarly, a deep 3-D CNN is significantly more complex and expensive in computational terms. At the same time, it can obtain worse results for classes with similar textures on many spectral bands.

To address the aforementioned drawbacks and limitations faced by existing DL-based methods, in this article we combine 3- and 2-D convolution to propose a new cross convolution unmixing network (CrossCUN) for hyperspectral unmixing. The proposed CrossCUN can not only directly complete the abundance estimation task without any prior knowledge about endmember extraction, but also get more accurate results than traditional unmixing methods. Notably, most existing DL-based methods usually use autoencoder networks or 2-D convolutional networks to complete abundance estimation, and our proposed CrossCUN is the first one to exploit cross convolution—that combines 2- and 3-D convolutions—to perform abundance estimation. Our newly proposed CrossCUN shows the potential of cross convolution compared to using 2- or 3-D convolution alone in terms of abundance estimation, and obtains better results than other traditional and DL-based methods.

The rest of this article is organized as follows. Section II introduces the newly proposed CrossCUN and the corresponding 2-D and 3-D convolution networks, i.e., 2-DCUN and 3-DCUN. In Section III, we test the performance of different unmixing methods in abundance estimation on synthetic data and three real HSIs. Finally, Section IV, concludes this article.

II. Proposed Framework

Our proposed CrossCUN combines 2- and 3-D convolutions and aims at estimating the abundance fraction of each endmember at each pixel in the scene. Let $\mathbf{Y} \in \mathbb{R}^{Nx \times Ny \times Nb}$ denote HSI data, where Nx is the width, Ny is the height, and Nb is the number of spectral bands. Since the 3-D CNN is more computationally complex and hyperspectral pixels contain mixed land cover classes (with the subsequent inter-class similarity and intra-class variability), we employ principal component analysis (PCA) to remove the spectral redundancy on the proposed CrossCUN. PCA reduces the number of bands from Nb to D, maintaining the same spatial dimensions whilst keeping the most relevant spectral information. In this sense, let $\mathbf{X} \in \mathbb{R}^{Nx \times Ny \times D}$ denote the reduced data by PCA (in our work, D is empirically set to 15).

We build 3-D patches to further improve the performance of 2-DCUN, 3-DCUN, and CrossCUN when dealing with the unmixing task. In this context, let $\mathbf{P} \in \mathbb{R}^{S \times S \times D}$ denote the 3-D patches, whilst $S \times S$ denotes the window size. The number of created 3-D patches by these three networks is $(Nx - S + 1) \times (Ny - S + 1)$. We assume that $P_{a,b}$ is the 3-D patch and (a, b) is its center pixel. In this sense, the width and the height of the 3-D patch $P_{a,b}$ are from $a - (S-1)/2$ to $a + (S-1)/2$ and $b - (S-1)/2$ to $b + (S-1)/2$, respectively. In our work, S is empirically set to 9.

On the one hand, our proposed 2-DCUN and CrossCUN involve 2-D convolutions, and their input patches are convolved with 2-D kernels. 2-D convolution computes the sum of the dot products between the input patch and the 2-D kernel to obtain the corresponding results, and the kernel spans the input data to cover the entire spatial dimension. The convolved features use an activation function to introduce nonlinearity in the model. In 2-D convolution, the activation value $v_{i,j}^{x,y}$ at spatial position (x, y) in the jth feature map of the ith layer is computed by

$$v_{i,j}^{x,y} = \varphi \left(b_{i,j} + \sum_{\tau=1}^{d_{i-1}} \sum_{\beta=-\rho}^{\rho} \sum_{\alpha=-\gamma}^{\gamma} w_{i,j,\tau}^{\alpha,\beta} \times v_{i-1,\tau}^{x+\alpha,y+\beta} \right) \quad (1)$$

where φ denotes the activation function, $w_{i,j}$ is the value of the weight parameter for the jth feature map of the ith layer, $b_{i,j}$ is the bias for the jth feature map of the ith layer, d_{i-1} is the number of feature maps in the $(i-1)$th layer, and $2\rho + 1$ and $2\gamma + 1$ are the width and height of kernel, respectively.

On the other hand, 3-DCUN and the proposed CrossCUN first employ 3-D convolution to convolve the input patches with 3-D kernels. Similarly, 3-D convolution also computes the sum of the dot products between the input patch and the 3-D kernel to obtain the corresponding results. We then use an activation function to learn the nonlinear features. In this regard, the activation value $v_{i,j}^{x,y,z}$ at spatial position (x, y, z) in the jth feature map of the ith layer is obtained by

$$v_{i,j}^{x,y,z} = \varphi \left(b_{i,j} + \sum_{\tau=1}^{d_{i-1}} \sum_{\lambda=-\eta}^{\eta} \sum_{\beta=-\rho}^{\rho} \sum_{\alpha=-\gamma}^{\gamma} w_{i,j,\tau}^{\alpha,\beta,\lambda} \times v_{i-1,\tau}^{x+\alpha,y+\beta,z+\lambda} \right)$$
$$(2)$$

where $2\eta + 1$ is the depth of the kernel.

The 2-DCUN, 3-DCUN frameworks and our proposed Cross-CUN are graphically illustrated in Figs. 1–3. From Fig. 1, we see that 2-DCUN contains three 2-D convolutional layers, one flatten layer, and one dense layer. The kernel sizes of the three 2-D convolutional layers are 3×3 ($f_1^1 = f_1^2 = 3, f_2^1 = f_2^2 = 3, f_3^1 = f_3^2 = 3$, where f_j^i denotes the kernel size of the ith dimension in the jth convolution), and the number of kernels
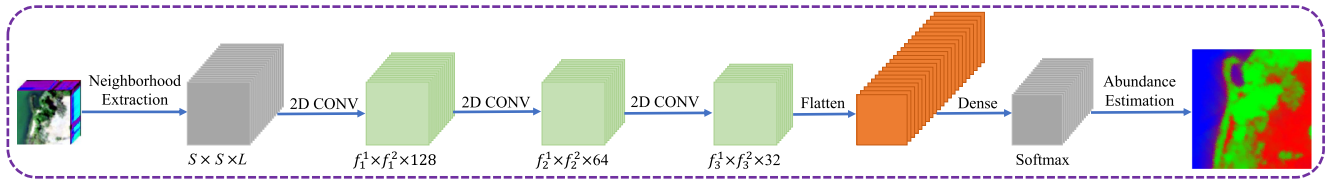
Fig. 1. Two-dimensional convolution network for spectral unmixing.
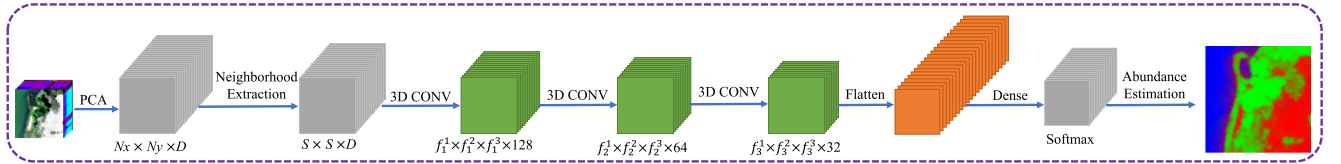
Fig. 2. Three-dimensional convolution network for spectral unmixing.
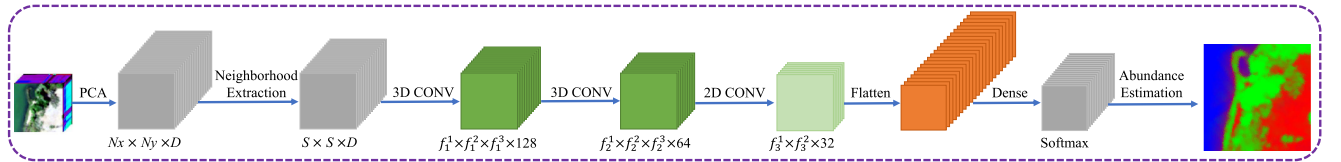
Fig. 3. Graphical overview of our proposed CrossCUN for spectral unmixing.

(i.e., the depth) of the three 2-D convolutional layers are 128, 64, and 32, respectively. Furthermore, from Fig. 2, we observe that 3-DCUN includes three 3-D convolutional layers, one flatten layer, and one dense layer. The kernel sizes of three 3-D convolutional layers are $3 \times 3 \times 7$ (in particular $f_1^1 = f_1^2 = 3$ and $f_1^3 = 7$), $3 \times 3 \times 5$ (where $f_2^1 = f_2^2 = 3$ and $f_2^3 = 5$), and $3 \times 3 \times 3$ (setting $f_3^1 = f_3^2 = f_3^3 = 3$), respectively. Moreover, the number of kernels of the three 3-D convolutional layers is 128, 64, and 32, respectively. Finally, from Fig. 3, we see that our proposed CrossCUN has two 3-D convolutional layers, one 2-D convolutional layer, one flatten layer, and one dense layer. The kernel sizes of the two 3-D convolutional layers are $3 \times 3 \times 7$ (with $f_1^1 = f_1^2 = 3$ and $f_1^3 = 7$), $3 \times 3 \times 5$ (where $f_2^1 = f_2^2 = 3$ and $f_2^3 = 5$). The kernel size of the 2-D convolutional layer is $f_3^1 = f_3^2 = 3$. Especially, applying a 2-D convolutional layer after the 3-D convolutional layer and before the flatten layer can better distinguish spatial information in different spectral bands without losing a large amount of spectral information, which is very significant for HSI data analysis. Similar to existing unmixing methods, the estimated abundance maps need to satisfy two constraints, i.e., abundance sum-to-one constraint (ASC) and abundance nonnegativity constraint (ANC). We use SoftMax as the activation function of the last dense layer in 2-DCUN, 3-DCUN, and CrossCUN, which is computed by

$$\hat{a}_{ij} = \frac{e^{z_j}}{\sum_{j=1}^{C} e^{z_j}} \tag{3}$$

where $\hat{a}_{ij}$ is the estimated abundance of the jth endmember on the ith pixel, z_j is the output value of the dense layer, and C is the total number of endmembers. We use cross-entropy as the loss function to train the 2-DCUN, 3-DCUN, and CrossCUN,

TABLE I
CONFIGURATION OF 2-DCUN ON SAMSON DATA

Layer type	Output shape	Parameter
input_1 (InputLayer)	(9,9,156)	0
conv2d_1 (CONV 2-D)	(7,7,128)	179 840
dropout_1 (Dropout)	(7,7,128)	0
conv2d_2 (CONV 2-D)	(5,5,64)	73 792
dropout_2 (Dropout)	(5,5,64)	0
conv2d_3 (CONV 2-D)	(3,3,32)	18 464
dropout_3 (Dropout)	(3,3,32)	0
flatten_1 (Flatten)	288	0
dense_1 (Dense)	3	867
Total trainable parameter: 272,963		

which is computed as follows:

$$L = -\frac{1}{N} \sum_{i=1}^{N} [\mathbf{a}_i \log \hat{\mathbf{a}}_i] \tag{4}$$

where $\hat{\mathbf{a}}_i$ denotes the abundance estimated for the ith pixel, $\mathbf{a}_i$ is the real abundance vector of the ith pixel, and N is the number of pixels. Tables I– III show a detailed description of 2-DCUN, 3-DCUN, and our proposed CrossCUN in terms of layer types, output types, and the number of parameters on a specific HSI dataset (Samson data, described in the following section). From Tables I–III, we can see that 2-DCUN requires the fewest number of parameters, and 3-DCUN and CrossCUN have the same number of parameters. Although CrossCUN does not

TABLE II
CONFIGURATION OF 3-DCUN ON SAMSON DATA

Layer type	Output shape	Parameter
input_1 (InputLayer)	(9,9,13,1)	0
conv3d_1 (CONV 3-D)	(7,7,7,128)	8192
dropout_1 (Dropout)	(7,7,7,128)	0
conv3d_2 (CONV 3-D)	(5,5,3,64)	368 704
dropout_2 (Dropout)	(5,5,3,64)	0
conv3d_3 (CONV 3-D)	(3,3,1,32)	55 328
dropout_3 (Dropout)	(3,3,1,32)	0
flatten_1 (Flatten)	288	0
dense_1 (Dense)	3	867

Total trainable parameter: 433,091

TABLE III
CONFIGURATION OF OUR PROPOSED CrossCUN ON SAMSON DATA

Layer type	Output shape	Parameter
input_1 (InputLayer)	(9,9,13,1)	0
conv3d_1 (CONV 3-D)	(7,7,7,128)	8192
dropout_1 (Dropout)	(7,7,7,128)	0
conv3d_2 (CONV 3-D)	(5,5,3,64)	368 704
dropout_2 (Dropout)	(5,5,3,64)	0
reshape_1 (Reshape)	(5,5,192)	0
conv2d_1 (CONV 2-D)	(3,3,32)	55 328
dropout_3 (Dropout)	(3,3,32)	0
flatten_1 (Flatten)	288	0
dense_1 (Dense)	3	867

Total trainable parameter: 433,091

need the fewest number of parameters, it obtains more accurate results on experimental data. Moreover, it is worth noting that the number of nodes of the last dense layer in the three networks is 3, which is equal to the number of endmembers in Samson data.

III. EXPERIMENTS

To assess the performance of 2-DCUN, 3-DCUN, and our proposed CrossCUN in terms of abundance estimation, two big families of synthetic data, i.e., with and without pure pixels, and three real HSIs (Samson, Jasper, and Urban scenes) are adopted in our experiments. We use the following:

1) *minimum-volume enclosing simplex* (MVES) [24];
2) *spatial group sparsity regularized nonnegative matrix factorization* (SGSNMF) [41];
3) *robust collaborative nonnegative matrix factorization* (R-CoNMF) [42];
4) *convolutional autoencoder* (CAE) [47];

5) *untied denoising autoencoder with sparsity* (uDAS) as comparison methods to test the performance of different unmixing methods in abundance estimation.

Especially, the initial endmember results required by SGSNMF are extracted by VCA. We employ the root-mean-square error (rmse) and the mean rmse as metrics to test the performance of 2-DCUN, 3-DCUN, and CrossCUN. RMSE is computed by

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{a}_{ij} - a_{ij})^2} \qquad (5)$$

where N is the number of pixels, and $\hat{a}_{ij}$ and a_{ij} are the estimated and real abundance fractions of the jth endmember on the ith pixel. The mean rmse is computed by

$$\mathrm{Mean\ RMSE} = \frac{1}{C}\sum_{j=1}^{C}\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{a}_{ij} - a_{ij})^2} \qquad (6)$$

where C is the number of endmembers.

A. Data Descriptions

1) Synthetic Dataset: To evaluate the impact of pixel purity, the number of endmembers, and the level of noise on abundance estimation, we generate two big families of synthetic data, i.e., with and without pure pixels. Each kind of synthetic data comprises two types: 1) synthetic data with 2500 pixels and 20dB noise with different numbers of endmembers, i.e., 5, 10, 15, and 20; and 2) synthetic data with 2500 pixels and 5 endmembers under different levels of noise, i.e., 10, 20, 30, and 40 dB.

2) Samson Dataset: The Samson dataset was captured by the SAMSON sensor, and it is one of the most widely used datasets for evaluating hyperspectral unmixing algorithms. The scene contains three endmembers, i.e., soil, tree, and water. The original Samson data has 952×952 pixels and 156 bands covering the wavelengths from 401 to 889 nm. In our experiments, to reduce the computational burden, we start from the (252,332)-th pixel and utilize a region of 95×95 pixels to validate the performance of different unmixing methods in estimating abundance maps. The false color composition and the corresponding spectral of Samson are shown in Fig. 4(a) and (d), respectively.

3) Jasper Dataset: The Jasper dataset was obtained by the airborne visible/infrared imaging spectrometer (AVIRIS) of NASA's Jet Propulsion Laboratory (JPL). The original Jasper data contains 512×614 pixels recorded in 224 bands, ranging from 380 to 2500 nm. It has four endmembers, i.e., tree, water, soil, and road. Since the original Jasper data are too complex and will bring high computational overload, we start from the (105 269)th pixel and keep 100×100 pixels in our experiments. In addition, due to dense water vapor and atmospheric artifacts, we remove bands 1–3, 108–112, 154–166, and 220–224 and retain 198 bands in our experiments. The false color composition and the corresponding spectral of Jasper are shown in Fig. 4(b) and (e), respectively.
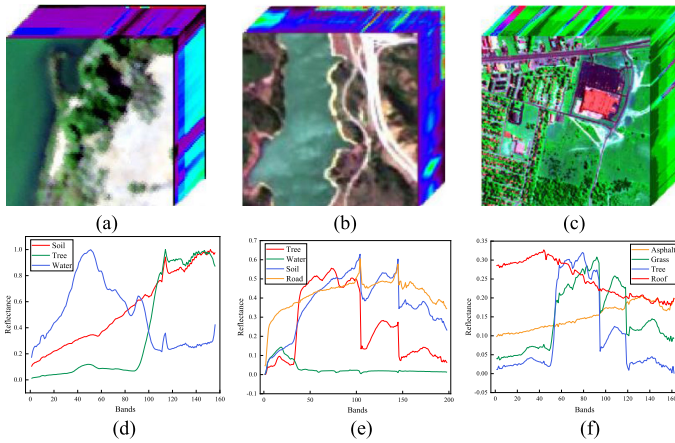
Fig. 4. Proposed models have been evaluated on three different hyperspectral scenes. (a) Samson, (b) Jasper, (c) Urban, (d) endmember spectrum of Samson, (e) eEndmember spectrum of Jasper, and (f) endmember spectrum of Urban.

4) Urban Dataset: The Urban dataset contains four different endmembers, i.e., asphalt, grass, tree, and roof. It comprises 307×307 pixels ranging from 400 to 2500 nm. Due to dense water vapor and atmospheric effects, we remove the bands 1–4, 78, 87, 101–111, 136–153, 198–210 and retain 162 bands for our experiments. The false color composition and the corresponding spectral signatures of Urban are shown in Fig. 4(c) and (f), respectively.

B. Implementation Details and Performance Metrics

In our work, we empirically set the number of layers and randomly select 80% pixels and run 50 epochs for training 2-DCUN, 3-DCUN, and our proposed CrossCUN. Especially, for randomly selecting 80% pixels, we first determine the number of pixels (e.g., N) and then randomly generate $N \times 80\%$ non-repeating numbers between 1 and N. Next, we use the $N \times 80\%$ numbers to extract the corresponding data and abundances from the original experimental data and abundances to ensure randomness. Finally, we use the extracted data and abundances to train our proposed network. The parameters that need to be adjusted in our proposed CrossCUN are the learning rate, input size, input dimension, dropout rate, and weight decay. We perform a parameter sensitivity analysis on Samson data and report their results in Fig. 5. From Fig. 5(a), we see that the mean rmse increases as the learning rate grows, and the mean rmse reaches the lowest value when the learning rate is 0.0001. Therefore, we choose 0.0001 as the final training rate in our proposed CrossCUN. From Fig. 5(b), the mean rmse achieves the lowest value when the input size is 9×9. In this sense, we choose 9×9 as the final input size to train our proposed CrossCUN. From Fig. 5(c), we can see that the value of the mean rmse fluctuates as the input dimension increases, and the mean rmse reaches an optimal value when the input dimension is 13. Therefore, we consider 13 as the final input dimension in our proposed CrossCUN. From Fig. 5(d), we also can observe that the mean rmse fluctuates as the dropout rate grows, and reaches the lowest value when the dropout rate is 0.03. Therefore, we choose 0.03 as the final dropout rate to train our proposed CrossCUN. Fig. 5(e)

shows that our method obtains the best results in terms of mean rmse when the weight decay is 0.0001. As a result, 0.0001 is considered the final weight decay in our experiments.

C. Experiments With Synthetic Data

We generate two big families of synthetic data, i.e., with and without pure pixels, to evaluate the performance of different unmixing methods in estimating abundance maps, and retrain the network when synthetic data have different numbers of endmembers or different levels of noise. Especially, considering that synthetic data include many endmembers, we only report the mean rmse results of different unmixing methods on the two big families of synthetic data. Table IV presents the mean rmse results obtained by different unmixing methods on synthetic data with pure pixels under different numbers of endmembers. From Table IV, we see that 2-DCUN, 3-DCUN, and CrossCUN both get better results than other methods and CrossCUN obtains the most accurate results on synthetic data with any number of endmembers. The results shown in Table IV reflect that the number of endmembers almost has no effect on the performance of our proposed CrossCUN in abundance estimation. Table V displays the mean rmse results of different unmixing methods on synthetic data with pure pixels under different levels of noise, i.e., 10, 20, 30, and 40 dB. From Table V, we observe that our proposed CrossCUN obtains the best results on synthetic data with 10 and 20 dB, and R-CoNMF achieves the best results on synthetic data with 30 and 40dB. From Table V, we also see that although our proposed CrossCUN does not achieve the best results on all synthetic data (the gap between it and R-CoNMF is very small). Therefore, our proposed CrossCUN can be used as an effective abundance estimation method.

Table VI shows the mean rmse results obtained by different unmixing methods on synthetic data without pure pixels using different numbers of endmembers. From Table VI, we observe that MVES and R-CoNMF exhibit better results than other methods. Moreover, 2-DCUN, 3-DCUN, and CrossCUN achieve more accurate results than MVES and R-CoNMF, and CrossCUN obtains the best results in terms of abundance estimation on synthetic data with any number of endmembers. Table VII presents the mean rmse results obtained by different abundance estimation methods on the synthetic data without pure pixels under different levels of noise, i.e., 10, 20, 30, and 40 dB. From Table VII, we observe that our proposed CrossCUN is better than other methods on synthetic data with 10 and 20 dB, and R-CoNMF obtains the most accurate results compared with other methods on synthetic data with 30 and 40dB (but the gap between R-CoNMF and our proposed CrossCUN is small). Therefore, our proposed CrossCUN is also very effective in terms of abundance estimation.

D. Experiments With Samson Data

Table VIII quantitatively compares the performance of different unmixing methods in the task of estimating abundance maps on Samson data, reporting the rmse and mean rmse scores obtained by them. From Table VIII, we see that SGSNMF obtains the best results in the task of estimating the abundance

TABLE IV
MEAN RMSE RESULTS OF DIFFERENT UNMIXING METHODS ON SYNTHETIC DATA (CONTAINING PURE PIXELS) WITH 2500 PIXELS AND 20 dB NOISE USING DIFFERENT NUMBERS OF ENDMEMBERS

Number of endmembers	Methods							
	MVES	SGSNMF	R-CoNMF	CAE	uDAS	2-DCUN	3-DCUN	CrossCUN
5	0.0947	0.2239	0.1041	0.2474	0.1471	0.0686	0.0632	**0.0617**
10	0.0652	0.1495	0.0666	0.1268	0.0919	0.0452	0.0403	**0.0400**
15	0.0633	0.0976	0.0900	0.0860	0.1124	0.0436	0.0380	**0.0380**
20	0.0500	0.0878	0.0841	0.0644	0.0894	0.0372	0.0329	**0.0328**

Best results are highlighted in bold.

TABLE V
MEAN RMSE RESULTS OF DIFFERENT UNMIXING METHODS ON SYNTHETIC DATA (CONTAINING PURE PIXELS) WITH 2500 PIXELS AND FIVE ENDMEMBERS USING DIFFERENT LEVELS OF NOISE

SNR (dB)	Methods							
	MVES	SGSNMF	R-CoNMF	CAE	uDAS	2-DCUN	3-DCUN	CrossCUN
10	0.1473	0.1770	0.1337	0.2467	0.1692	0.0675	0.0649	**0.0647**
20	0.0882	0.2453	0.0600	0.2454	0.0676	0.0530	0.0517	**0.0504**
30	0.0499	0.2169	**0.0362**	0.2436	0.1414	0.0648	0.0455	0.0432
40	0.0287	0.2855	**0.0099**	0.2485	0.1268	0.0512	0.0408	0.0391

Best results are highlighted in bold.

TABLE VI
MEAN RMSE RESULTS OF DIFFERENT UNMIXING METHODS ON SYNTHETIC DATA (MAXIMUM PURITY OF 0.8) WITH 2500 PIXELS AND 20dB NOISE UNDER DIFFERENT NUMBERS OF ENDMEMBERS

Number of endmembers	Methods							
	MVES	SGSNMF	R-CoNMF	CAE	uDAS	2-DCUN	3-DCUN	CrossCUN
5	0.0696	0.1953	0.0506	0.2455	0.0601	0.0483	0.0461	**0.0457**
10	0.0774	0.1615	0.0821	0.1285	0.1276	0.0525	0.0468	**0.0464**
15	0.0626	0.1114	0.0853	0.0859	0.0970	0.0427	0.0377	**0.0374**
20	0.0472	0.0732	0.0801	0.0647	0.0895	0.0365	0.0316	**0.0315**

Best results are highlighted in bold.
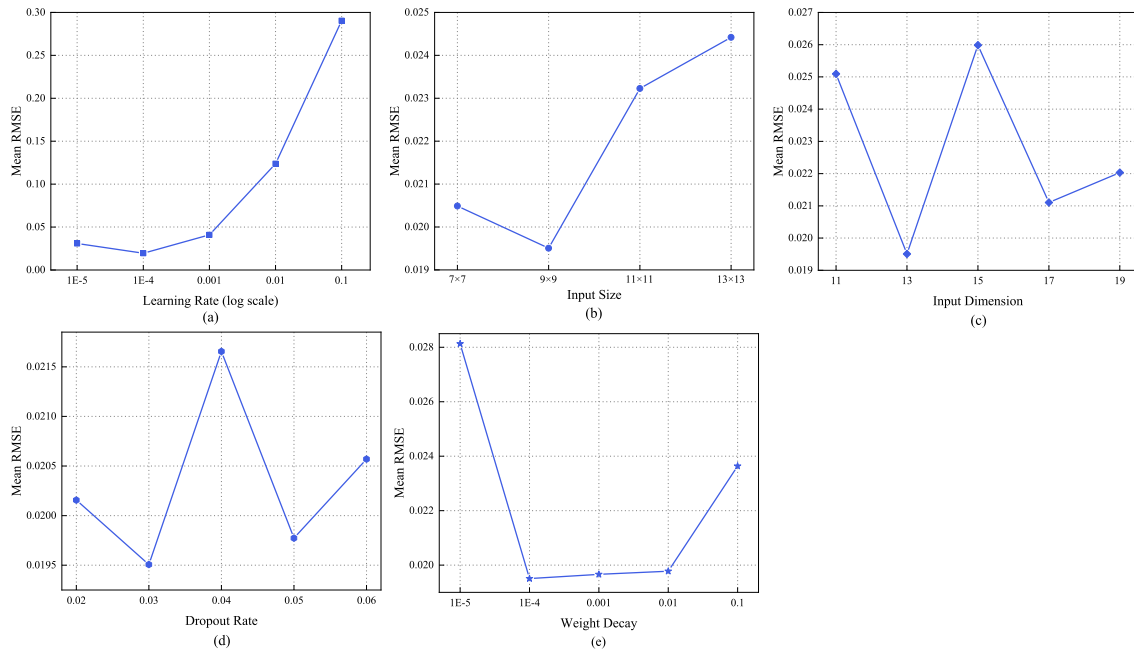
Fig. 5. Quantitative analysis of different (a) learning rates, (b) input sizes, (c) input dimensions, (d) dropout rates, and (e) weight decays for our proposed CrossCUN on Samson dataset.

TABLE VII
MEAN RMSE RESULTS OF DIFFERENT UNMIXING METHODS ON SYNTHETIC DATA (MAXIMUM PURITY OF 0.8) WITH 2500 PIXELS AND FIVE ENDMEMBERS UNDER DIFFERENT LEVELS OF NOISE

SNR (dB)	Methods							
	MVES	SGSNMF	R-CoNMF	CAE	uDAS	2-DCUN	3-DCUN	CrossCUN
10	0.1335	0.1732	0.1575	0.2480	0.1571	0.0895	0.0783	**0.0761**
20	0.0701	0.2106	0.0547	0.2442	0.0753	0.0498	0.0468	**0.0455**
30	0.0355	0.2179	**0.0208**	0.2481	0.0501	0.0484	0.0418	0.0403
40	0.0063	0.2206	**0.0059**	0.2484	0.0394	0.0447	0.0386	0.0380

Best results are highlighted in bold.

maps of soil, tree, and water compared with the other two traditional unmixing methods, i.e., MVES and R-CoNMF. Moreover, CAE is the best abundance estimation method compared with uDAS. From Table VIII, we also find that 2-DCU, 3-DCUN, and CrossCUN are better than the best traditional and DL-based methods, i.e., SGSNMF and CAE, and our proposed CrossCUN achieves better abundance estimation results compared with 2-DCUN and 3-DCUN. This fact reveals that our proposed CrossCUN is not only very effective for abundance estimation purposes but also improves the results obtained by the corresponding 2- and 3-D convolution networks, i.e., 2- and 3-DCUN.

Fig. 6 presents the abundance maps obtained by different unmixing methods from the Samson data. From Fig. 6, we observe that SGSNMF and CAE provide results that are more similar to the ground-truth compared with other traditional and DL-based unmixing methods. We also find that 2-DCUN, 3-DCUN, and our proposed CrossCUN are closer to the ground-truth compared with all methods, and CrossCUN is slightly better than 2- and 3-DCUN.

E. Experiments With Jasper Data

Table IX shows the results of rmse and mean rmse obtained by different unmixing methods and reports their performance evaluation in abundance estimation on Jasper data. From Table IX, we observe that SGSNMF is the best method in the task of estimating the abundance maps of three, water, soil, and road compared with another two traditional unmixing methods, i.e., MVES and R-CoNMF. CAE obtains the best results in

TABLE VIII
RMSE and Mean RMSE Results of Different Unmixing Methods on Samson Data

Mineral		Methods							
		MVES	SGSNMF	R-CoNMF	CAE	uDAS	2-DCUN	3-DCUN	CrossCUN
RMSE	Soil	0.2326	0.1784	0.2584	0.1156	0.2598	0.0491	0.0354	**0.0234**
	Tree	0.3400	0.2448	0.2729	0.1186	0.2509	0.0489	0.0366	**0.0235**
	Water	0.3916	0.3579	0.4125	0.0667	0.4178	0.0238	0.0254	**0.0116**
Mean RMSE		0.3214	0.2604	0.3146	0.1003	0.3095	0.0406	0.0325	**0.0195**

Best results are highlighted in bold.

Fig. 6. Ground-truth and estimated abundances obtained by different unmixing methods for each endmember material on Samson dataset.

TABLE IX
RMSE and Mean RMSE Results of Different Unmixing Methods on Jasper Data

Mineral		Methods							
		MVES	SGSNMF	R-CoNMF	CAE	uDAS	2-DCUN	3-DCUN	CrossCUN
RMSE	Tree	0.2156	0.1937	0.1790	0.1311	0.1587	0.0672	0.0385	**0.0292**
	Water	0.2650	0.1264	0.2181	0.1285	0.1918	0.0288	0.0267	**0.0213**
	Soil	0.2194	0.1697	0.1508	0.2225	0.1315	0.0851	0.0437	**0.0320**
	Road	0.1252	0.1719	0.1281	0.2276	0.1150	0.0685	0.0361	**0.0252**
Mean RMSE		0.2063	0.1654	0.1690	0.1774	0.1492	0.0624	0.0363	**0.0269**

Best results are highlighted in bold.

estimating the abundance maps of tree and water, and uDAS obtains the best abundance estimations for soil and road. In fact, uDAS is the best among DL-based methods in terms of the value of the mean rmse. In addition, from Table IX we also see that 2-DCUN, 3-DCUN, and CrossCUN are not only

better than the best traditional and DL-based methods, i.e., SGSNMF and uDAS, but also can get quite accurate results when estimating the abundance maps of all minerals. Simultaneously, CrossCUN is better than 2- and 3-DCUN. Therefore, CrossCUN is not only very effective in abundance estimation but also has
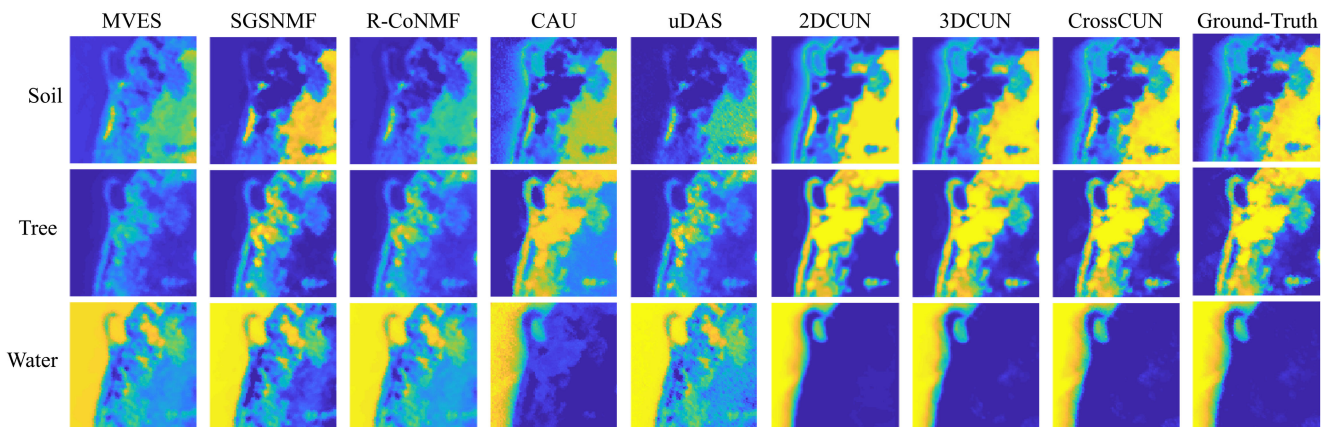
Fig. 7.   Ground-truth and estimated abundances obtained by different unmixing methods for each endmember material on Jasper dataset.

TABLE X
RMSE AND MEAN RMSE RESULTS OF DIFFERENT UNMIXING METHODS ON URBAN DATA

Mineral		Methods							
		MVES	SGSNMF	R-CoNMF	CAE	uDAS	2-DCUN	3-DCUN	CrossCUN
RMSE	Asphalt	0.2886	0.2870	0.3307	0.1745	0.3280	0.0844	0.0327	**0.0247**
	Grass	0.3341	0.4452	0.4024	0.4409	0.4241	0.0956	0.0390	**0.0293**
	Tree	0.2991	0.2558	0.3045	0.2884	0.3204	0.0725	0.0349	**0.0231**
	Roof	0.1849	0.1899	0.1974	0.2254	0.2089	0.0602	0.0275	**0.0185**
Mean RMSE		0.2767	0.2945	0.3088	0.2823	0.3203	0.0782	0.0335	**0.0239**

Best results are highlighted in bold.

improvements compared with traditional and DL-based methods, and the corresponding 2- and 3-D convolution networks, i.e., 2- and 3-DCUN.

The results obtained by different unmixing methods on Jasper data are presented in Fig. 7. From Fig. 7, we see that MVES and CAE provide the worst results compared with the other traditional and DL-based methods, while 2-DCUN, 3-DCUN, and our proposed CrossCUN are more similar to ground-truth than all other methods. This fact, together with the results in Table IX, leads us to conclude that our proposed CrossCUN is not only an effective abundance estimation method but also outperforms 2- and 3-DCUN.

F. Experiments With Urban Data

Table X shows the rmse and mean rmse values obtained by different unmixing methods and reports their performance evaluation in abundance estimation on Urban data. From Table X, we can conclude that the three traditional unmixing methods, i.e., MVES, SGSNMF, and R-CONMF, and the two DL-based methods i.e., CAE and uDAS cannot obtain ideal results in abundance estimation. Clearly, 2-DCUN, 3-DCUN, and CrossCUN obtain better results in terms of rmse and mean rmse compared with all traditional and DL-based unmixing methods. As in previous experiments, CrossCUN is better than 2- and 3-DCUN.

Fig. 8 reports the abundance maps obtained by different unmixing methods on Urban data. From Fig. 8, we see that 2-DCUN, 3-DCUN, and our proposed CrossCUN are all close to the ground-truth. After analyzing the results in Table X and Fig. 8, we can conclude that CrossCUN not only gets the best results in abundance estimation compared with other traditional and DL-based unmixing methods, but is also more effective than 2- and 3-DCUN.
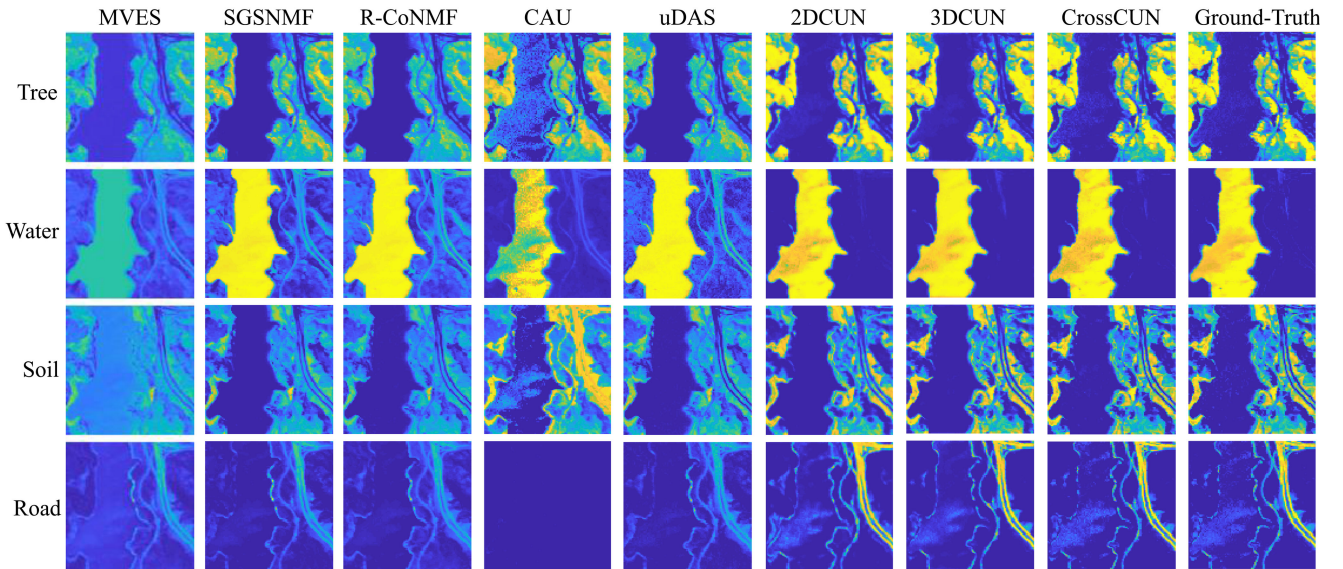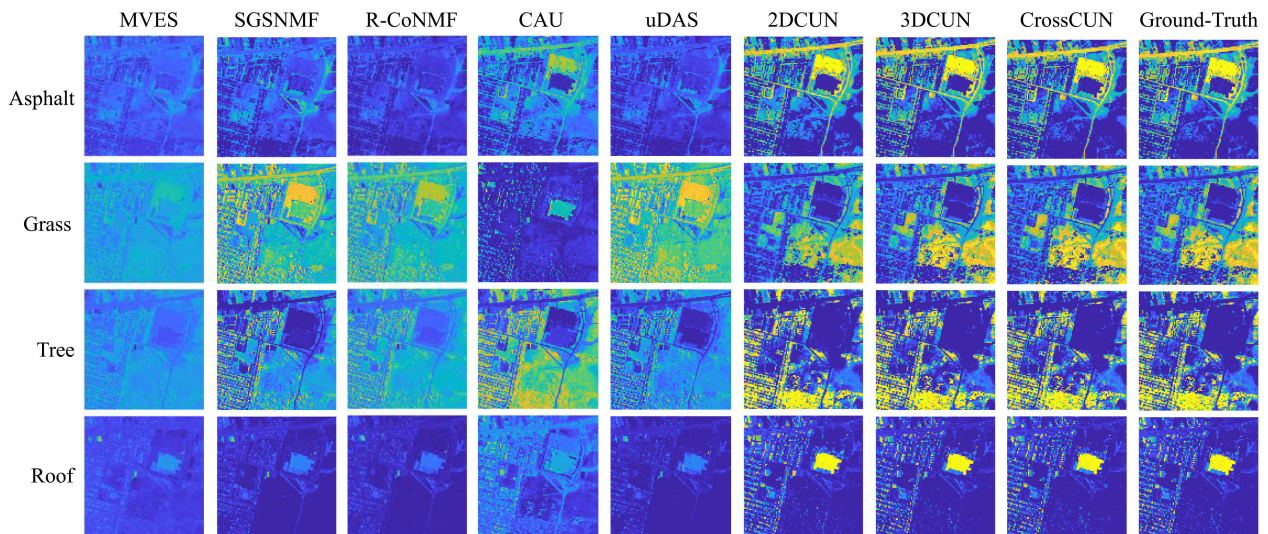
Fig. 8.   Ground-truth and estimated abundances obtained by different unmixing methods for each endmember material on Urban dataset.

TABLE XI
AVERAGE RUNNING TIME (INCLUDING TRAINING AND TESTING TIME) OF DIFFERENT UNMIXING METHODS FOR ALL PIXELS ON SYNTHETIC DATA WITH
DIFFERENT NUMBERS OF ENDMEMBERS

Number of endmembers	Methods							
	MVES	SGSNMF	R-CoNMF	CAE	uDAS	2-DCUN	3-DCUN	CrossCUN
5	**2.1387**	10.4748	4.3838	11.0584	6.2217	15.6695	22.6058	23.2870
10	56.6434	16.3271	**5.7390**	11.0626	27.8004	16.0360	23.0618	23.1512
15	340.6189	20.1648	**6.5037**	11.1522	50.3769	15.9557	23.2579	23.1756
20	676.8515	24.1815	**7.5456**	11.1313	53.6579	16.6114	23.1987	23.5684

Best results are highlighted in bold.

TABLE XII
AVERAGE RUNNING TIME (INCLUDING TRAINING AND TESTING TIME) OF DIFFERENT UNMIXING METHODS FOR ALL PIXELS ON SYNTHETIC DATA WITH
DIFFERENT LEVELS OF NOISE

SNR (dB)	Methods							
	MVES	SGSNMF	R-CoNMF	CAE	uDAS	2-DCUN	3-DCUN	CrossCUN
10	**1.7187**	7.8009	3.9688	10.7874	44.0712	16.6987	22.9547	22.4610
20	**1.7969**	7.2501	4.2816	10.9256	4.0456	15.6064	23.0554	22.7529
30	**1.9807**	9.9005	4.2605	10.7421	21.0370	15.8768	22.9798	22.1205
40	**2.4055**	8.0369	4.7387	10.8124	6.0168	15.5082	23.1463	22.7488

Best results are highlighted in bold.

TABLE XIII
RUNNING TIME (INCLUDING TRAINING AND TESTING TIME) OF DIFFERENT UNMIXING METHODS FOR ALL PIXELS ON DIFFERENT DATASETS

Datasets	Methods							
	MVES	SGSNMF	R-CoNMF	CAE	uDAS	2-DCUN	3-DCUN	CrossCUN
Samson	**1.3309**	20.9016	3.7967	42.9406	8.4782	45.4194	74.2941	79.8372
Jasper	**1.3629**	26.3976	3.5857	59.8035	174.4122	57.2994	88.4914	83.2869
Urban	**1.5257**	189.6412	30.4376	454.7000	1222.8033	577.2054	791.3201	769.4406

Best results are highlighted in bold.

TABLE XIV
MEAN RMSE RESULTS OF 2-DCUN, 3-DCUN, AND CrossCUN ON THE CONSIDERED DATASETS USING DIFFERENT PERCENTAGES OF THE AVAILABLE TRAINING SAMPLES

Training samples (%)	Samson			Jasper			Urban		
	2DCUN	3DCUN	CrossCUN	2-DCUN	3-DCUN	CrossCUN	2-DCUN	3-DCUN	CrossCUN
20	0.0558	0.0475	**0.0447**	0.1294	0.0529	**0.0480**	0.0925	0.0498	**0.0417**
40	0.0464	0.0405	**0.0327**	0.1153	0.0398	**0.0385**	0.0860	0.0424	**0.0352**
60	0.0437	0.0360	**0.0262**	0.1248	0.0374	**0.0320**	0.0798	0.0378	**0.0268**
80	0.0406	0.0325	**0.0195**	0.0624	0.0363	**0.0269**	0.0782	0.0335	**0.0239**

Best results are highlighted in bold.

TABLE XV
SUMMARY OF THE MEAN RMSE RESULTS OBTAINED BY DIFFERENT UNMIXING METHODS ON THE CONSIDERED DATASETS (ALL METHODS TRAINED USING 80% OF THE AVAILABLE TRAINING SAMPLES)

Datasets	Methods							
	MVES	SGSNMF	R-CoNMF	CAE	uDAS	2-DCUN	3-DCUN	CrossCUN
Samson	0.3214	0.2604	0.3146	0.1003	0.3095	0.0406	0.0325	**0.0195**
Jasper	0.2063	0.1654	0.1690	0.1774	0.1492	0.0624	0.0363	**0.0269**
Urban	0.2767	0.2945	0.3088	0.2823	0.3203	0.0782	0.0335	**0.0239**

Best results are highlighted in bold.

G. Computational Cost

We conducted all experiments on a computer with 2.6-GHz Intel Core i7 CPU and 16 GB of memory (NVIDIA GeForce RTX 2060 GPU) and evaluated the computational cost of different unmixing methods using different datasets. By analyzing the experimental results on synthetic data, we find that the existence of pure pixels does not significantly affect the efficiency of different unmixing methods. In this sense, we report the results of the average running time (including training and testing time) of different unmixing methods on synthetic data with and without pure pixels. Table XI displays the average running time of different unmixing methods on synthetic data under different numbers of endmembers. From Table XI, we see that MVES achieves the highest efficiency on synthetic data with five endmembers, and the number of endmembers has a greater impact on MVES. In addition, R-CoNMF is the fastest on synthetic data with 10, 15, and 20 endmembers, and it is stable on synthetic data with any number of endmembers. From Table XI, we also find

that 2-DCUN is faster than 3-DCUN and CrossCUN, and the efficiency of 3-DCUN and CrossCUN is similar. Although our proposed CrossCUN is not faster than MVES and R-CoNMF, the time gap between them is not significant. Moreover, the results in Tables IV and VI show that CrossCUN is more effective than MVES and R-CoNMF. Therefore, our proposed CrossCUN is very effective and efficient for abundance estimation. Table XII shows the results of different unmixing methods on synthetic data under different levels of noise, 10, 20, 30, and 40 dB. The experimental results in Table XII indicate that MVES is the fastest method on synthetic data with any level of noise. Moreover, we find that 2-DCUN achieves higher efficiency than 3-DCUN and CrossCUN, and the running time of 3-DCUN and CrossCUN is similar. Although our proposed CrossCUN does not have higher efficiency than MVES, the results in Tables V and VII show that CrossCUN is more effective than MVES. Therefore, our proposed CrossCUN is an effective and efficient abundance estimation method.

Table XIII shows the running time of different unmixing methods on the three real hyperspectral datasets, i.e., Samson, Jasper, and Urban. From Table XIII, we see that MVES is the most efficient among the considered methods. Although MVES is faster than CrossCUN, the results in Tables VIII–X show that CrossCUN is more effective than MVES. Moreover, the computational cost of CrossCUN is acceptable for abundance estimation.

H. Sensitivity to the Number of Training Samples

Although the proposed CrossCUN is a supervised unmixing method, it obtains accurate abundance estimation results without requiring a lot of training samples. To illustrate this point, we choose different training set sizes (i.e., 20%, 40%, 60%, and 80% of the available labeled samples) to evaluate the performance of 2-DCUN, 3-DCUN, and CrossCUN on different datasets. The related mean rmse results are reported in Table XIV. From Table XIV, we see that: 1) more training pixels will lead to more accurate results (as expected); and 2) CrossCUN provides more accurate results than 2- and 3-DCUN with limited training samples. For clarity, we summarize the mean rmse results of different unmixing methods in Table XV (using 80% of the available samples). After analyzing the results of Tables XIV and XV, we can conclude that 2-DCUN, 3-DCUN, and CrossCUN outperform other compared methods even when the training pixels are very few, and CrossCUN obtains the best results compared to all other methods.

IV. CONCLUSION

In this work, a new cross convolution unmixing network (CrossCUN) has been introduced. The proposed CrossCUN can not only get more accurate results than DL-based methods but also directly complete the abundance estimation results without any prior results about endmember extraction, as opposed to traditional unmixing methods. We evaluated the performance of the newly proposed CrossCUN in terms of abundance estimation on synthetic data and three real HSIs, and our experiments reveal that CrossCUN outperforms the competitors in most cases.

As with any new approach, there are some unresolved issues that may present challenges over time. Similar to other DL-based methods, our proposed CrossCUN is a supervised method that requires labeled samples to train the network. Although the process of collecting training data for unmixing applications is costly, this can be done using image data at multiple resolutions or even field visits. In future work, we will work on designing a new unsupervised network to directly complete abundance estimation.

REFERENCES

[1] J. M. Haut, M. E. Paoletti, J. Plaza, J. Li, and A. Plaza, "Active learning with convolutional neural networks for hyperspectral image classification using a new Bayesian approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 11, pp. 6440–6461, Nov. 2018.

[2] M. E. Paoletti, J. M. Haut, J. Plaza, and A. Plaza, "A new deep convolutional neural network for fast hyperspectral image classification," *ISPRS J. Photogrammetry Remote Sens.*, vol. 145, pp. 120–147, 2018.

[3] Q. Zhu et al., "A spectral-spatial-dependent global learning framework for insufficient and imbalanced hyperspectral image classification," *IEEE Trans. Cybern.*, to be published, doi: 10.1109/TCYB.2021.3070577.

[4] K.-K. Huang, C.-X. Ren, H. Liu, Z.-R. Lai, Y.-F. Yu, and D.-Q. Dai, "Hyperspectral image classification via discriminant Gabor ensemble filter," *IEEE Trans. Cybern.*, vol. 52, no. 8, pp. 8352–8365, Aug. 2022.

[5] Z. Gong, W. Hu, X. Du, P. Zhong, and P. Hu, "Deep manifold embedding for hyperspectral image classification," *IEEE Trans. Cybern.*, to be published, doi: 10.1109/TCYB.2021.3069790.

[6] D. Hong, N. Yokoya, J. Chanussot, and X. X. Zhu, "CoSpace: Common subspace learning from hyperspectral-multispectral correspondences," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4349–4359, Jul. 2019.

[7] Y. Xu, Z. Wu, J. Chanussot, P. Comon, and Z. Wei, "Nonlocal coupled tensor CP decomposition for hyperspectral and multispectral image fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, pp. 348–362, Jan. 2020.

[8] R. Dian, S. Li, L. Fang, T. Lu, and J. M. Bioucas-Dias, "Nonlocal sparse tensor factorization for semiblind hyperspectral and multispectral image fusion," *IEEE Trans. Cybern.*, vol. 50, no. 10, pp. 4469–4480, Oct. 2020.

[9] H. V. Nguyen, A. Banerjee, and R. Chellappa, "Tracking via object reflectance using a hyperspectral video camera," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*, 2010, pp. 44–51.

[10] Y. Fu, Y. Zheng, I. Sato, and Y. Sato, "Exploiting spectral-spatial correlation for coded hyperspectral image restoration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 3727–3736.

[11] J. Zhang, M. Z. A. Bhuiyan, X. Yang, A. K. Singh, D. F. Hsu, and E. Luo, "Trustworthy target tracking with collaborative deep reinforcement learning in EdgeAi-Aided IoT," *IEEE Trans. Ind. Inform.*, vol. 18, no. 2, pp. 1301–1309, Feb. 2022.

[12] H. Su, Z. Wu, Q. Du, and P. Du, "Hyperspectral anomaly detection using collaborative representation with outlier removal," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 12, pp. 5029–5038, Dec. 2018.

[13] H. Su, Z. Wu, A.-X. Zhu, and Q. Du, "Low rank and collaborative representation for hyperspectral anomaly detection via robust dictionary construction," *ISPRS J. Photogrammetry Remote Sens.*, vol. 169, pp. 195–211, 2020.

[14] Y. Yuan, D. Ma, and Q. Wang, "Hyperspectral anomaly detection by graph pixel selection," *IEEE Trans. Cybern.*, vol. 46, no. 12, pp. 3123–3134, Dec. 2016.

[15] W. Xie, X. Zhang, Y. Li, J. Lei, J. Li, and Q. Du, "Weakly supervised low-rank representation for hyperspectral anomaly detection," *IEEE Trans. Cybern.*, vol. 51, no. 8, pp. 3889–3900, Aug. 2021.

[16] W. He, Q. Yao, C. Li, N. Yokoya, and Q. Zhao, "Non-local meets global: An integrated paradigm for hyperspectral denoising," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 6861–6870.

[17] A. Maffei, J. M. Haut, M. E. Paoletti, J. Plaza, L. Bruzzone, and A. Plaza, "A single model CNN for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 4, pp. 2516–2529, Apr. 2020.

[18] Y. Li, X. Luo, N. Wu, and X. Dong, "The application of semisupervised attentional generative adversarial networks in desert seismic data denoising," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[19] B. Pan, Z. Shi, Z. An, Z. Jiang, and Y. Ma, "A novel spectral-unmixing-based green algae area estimation method for GOCI data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 2, pp. 437–449, Feb. 2017.

[20] X.-R. Feng, H.-C. Li, J. Li, Q. Du, A. Plaza, and W. J. Emery, "Hyperspectral unmixing using sparsity-constrained deep nonnegative matrix factorization with total variation," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 6245–6257, Oct. 2018.

[21] X. Xu, B. Pan, Z. Chen, Z. Shi, and T. Li, "Simultaneously multiobjective sparse unmixing and library pruning for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 4, pp. 3383–3395, Apr. 2021.

[22] F. Li, S. Zhang, B. Liang, C. Deng, C. Xu, and S. Wang, "Hyperspectral sparse unmixing with spectral-spatial low-rank constraint," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 6119–6130, 2021.

[23] B. Rasti and B. Koirala, "SUnCNN: Sparse unmixing using unsupervised convolutional neural network," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[24] T.-H. Chan, C.-Y. Chi, Y.-M. Huang, and W.-K. Ma, "A convex analysis-based minimum-volume enclosing simplex algorithm for hyperspectral unmixing," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4418–4432, Nov. 2009.

[25] J. Li, A. Agathos, D. Zaharie, J. M. Bioucas-Dias, A. Plaza, and X. Li, "Minimum volume simplex analysis: A fast algorithm for linear hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 9, pp. 5067–5082, Sep. 2015.

[26] J. M. P. Nascimento and J. M. B. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, Apr. 2005.

[27] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 765-777, Mar. 2007.

[28] W. Tang, Z. Shi, Y. Wu, and C. Zhang, "Sparse unmixing of hyperspectral data using spectral a priori information," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 2, pp. 770–783, Feb. 2015.

[29] H. Ren and C.-I. Chang, "Automatic spectral target recognition in hyperspectral imagery," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1232–1249, Oct. 2003.

[30] R. Marrero et al., "A novel negative abundance-oriented hyperspectral unmixing algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3772–3790, Jul. 2015.

[31] T.-H. Chan, J.-Y. Liou, A. Ambikapathi, W.-K. Ma, and C.-Y. Chi, "Fast algorithms for robust hyperspectral endmember extraction based on worst-case simplex volume maximization," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2012, pp. 1237–1240.

[32] C. Févotte and N. Dobigeon, "Nonlinear hyperspectral unmixing with robust nonnegative matrix factorization," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 4810–4819, Dec. 2015.

[33] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4153–4162, Nov. 2011.

[34] N. Yokoya, J. Chanussot, and A. Iwasaki, "Nonlinear unmixing of hyperspectral data using semi-nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 2, pp. 1430–1437, Feb. 2014.

[35] R. Fernandez-Beltran, A. Plaza, J. Plaza, and F. Pla, "Hyperspectral unmixing based on dual-depth sparse probabilistic latent semantic analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 11, pp. 6344–6360, Nov. 2018.

[36] J. A. G. Jaramago, M. E. Paoletti, J. M. Haut, R. Fernandez-Beltran, A. Plaza, and J. Plaza, "GPU parallel implementation of dual-depth sparse probabilistic latent semantic analysis for hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 9, pp. 3156–3167, Sep. 2019.

[37] X. Tao, M. E. Paoletti, J. M. Haut, P. Ren, J. Plaza, and A. Plaza, "Endmember estimation with maximum distance analysis," *Remote Sens.*, vol. 13, no. 4, 2021, Art. no. 713.

[38] X. Tao et al., "Endmember estimation from hyperspectral images using geometric distances," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[39] D. C. Heinz and C. I. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar. 2001.

[40] J. M. Bioucas-Dias and M. A. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *Proc. 2nd Workshop Hyperspectral Image Signal Process. Evol. Remote Sens.*, 2010, pp. 1–4.

[41] X. Wang, Y. Zhong, L. Zhang, and Y. Xu, "Spatial group sparsity regularized nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 11, pp. 6287–6304, Nov. 2017.

[42] J. Li, J. M. Bioucas-Dias, A. Plaza, and L. Liu, "Robust collaborative nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6076–6090, Oct. 2016.

[43] R. Guo, W. Wang, and H. Qi, "Hyperspectral image unmixing using autoencoder cascade," in *Proc. 7th Workshop Hyperspectral Image Signal Process. Evol. Remote Sens.*, 2015, pp. 1–4.

[44] Y. Qu and H. Qi, "uDAS: An untied denoising autoencoder with sparsity for spectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 3, pp. 1698–1712, Mar. 2019.

[45] B. Palsson, M. O. Ulfarsson, and J. R. Sveinsson, "Convolutional autoencoder for spectral-spatial hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 1, pp. 535–549, Jan. 2021.

[46] L. Gao, Z. Han, D. Hong, B. Zhang, and J. Chanussot, "CyCU-Net: Cycle-consistency unmixing network by learning cascaded autoencoders," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–14, 2022.

[47] Y. Ranasinghe et al., "Convolutional autoencoder for blind hyperspectral image unmixing," in *Proc. IEEE 15th Int. Conf. Ind. Inf. Syst.*, 2020, pp. 174–179, doi: 10.1109/ICIIS51140.2020.9342727.

[48] L. Qi, J. Li, Y. Wang, M. Lei, and X. Gao, "Deep spectral convolution network for hyperspectral image unmixing with spectral library," *Signal Process.*, vol. 176, 2020, Art. no. 107672.

[49] D. Hong, J. Chanussot, N. Yokoya, U. Heiden, W. Heldens, and X. X. Zhu, "WU-Net: A. weakly-supervised unmixing network for remotely sensed hyperspectral imagery," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2019, pp. 373–376.

[50] Z. Han, D. Hong, L. Gao, B. Zhang, and J. Chanussot, "Deep half-siamese networks for hyperspectral unmixing," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 11, pp. 1996–2000, Nov. 2021.

Escuela Politecnica
Av. de la Universidad, S/N, 10003
Caceres, Spain
Phone: 0034927257000. Ext. 51662
Email: aplaza,juanmariohaut{@unex.es}

Dr. Antonio Plaza Miguel y Dr. Juan Mario Haut Hurtado como directores de la tesis titulada "Procesamiento Eficiente y Profundo de Imágenes Hiperespectrales y Aplicación en Tareas de Desmezclado Espectral", certifican el factor de impacto y la categorización de la siguiente publicación, incluída en la tesis doctoral. Del mismo modo, se especifica la aportación del doctorado.

Antonio Plaza Miguel PhD, and Juan Mario Haut Hurtado PhD, as directors of the Phd thesis titled "Efficient and Deep Processing of Hyperspectral Images and Application in Spectral Unmixing Tasks", certify the impact factor and the categorization of the following publication, included in the doctoral thesis. In the same way, the contribution of the doctorate is specified.

Artículo / Paper

Autores/Authors: **X. Tao**, T. Cui, A. Plaza, and P. Ren.
Title: Simultaneously Counting and Extracting Endmembers in a Hyperspectral Image Based on Divergent Subsets.
Journal: IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no 12, pp. 8952-8966, December 2020.
DOI: 10.1109/TGRS.2020.2992542.
Impact factor 2021: 8.125. Q1

Abstract: Most existing endmember extraction techniques require prior knowledge about the number of endmembers in a hyperspectral image. The number of endmembers is normally estimated by a separate procedure, whose accuracy has a large influence on the endmember extraction performance. In order to bridge the two seemingly independent but, in fact, highly correlated procedures, we develop a new endmember estimation strategy that simultaneously counts and extracts endmembers. We consider a hyperspectral image as a hyperspectral pixel set and define the subset of pixels that are most different from one another as the divergent subset (DS) of the hyperspectral pixel set. The DS is characterized by the condition that any additional pixel would increase the likeness within the DS and, thus, reduce its divergent degree. We use the DS as the endmember set, with the number of endmembers being the subset cardinality. To render a practical computation scheme for identifying the DS, we reformulate it in terms of a quadratic optimization problem with a numerical solution. In addition to operating as an endmember estimation algorithm by itself, the DS method can also co-operate with existing endmember extraction techniques by transforming them into a novel and more effective schemes. Experimental results validate the effectiveness of the DS methodology in simultaneously counting and extracting endmembers not only as an individual algorithm but also as a foundation algorithm for improving existing methods.

Contribución del doctorado: Planteamiento de la hipótesis, desarrollo práctico, análisis y discusión de los resultados, elaboración y escritura del manuscrito.

Firma / Signature
Marzo / Marzo, 2023

Antonio Plaza Miguel   Juan Mario Haut Hurtado

Simultaneously Counting and Extracting Endmembers in a Hyperspectral Image Based on Divergent Subsets

Xuanwen Tao, *Student Member, IEEE*, Tingwei Cui, Antonio Plaza, *Fellow, IEEE*, and Peng Ren, *Senior Member, IEEE*

Abstract—Most existing endmember extraction techniques require prior knowledge about the number of endmembers in a hyperspectral image. The number of endmembers is normally estimated by a separate procedure, whose accuracy has a large influence on the endmember extraction performance. In order to bridge the two seemingly independent but, in fact, highly correlated procedures, we develop a new endmember estimation strategy that simultaneously counts and extracts endmembers. We consider a hyperspectral image as a hyperspectral pixel set and define the subset of pixels that are most different from one another as the divergent subset (DS) of the hyperspectral pixel set. The DS is characterized by the condition that any additional pixel would increase the likeness within the DS and, thus, reduce its divergent degree. We use the DS as the endmember set, with the number of endmembers being the subset cardinality. To render a practical computation scheme for identifying the DS, we reformulate it in terms of a quadratic optimization problem with a numerical solution. In addition to operating as an endmember estimation algorithm by itself, the DS method can also co-operate with existing endmember extraction techniques by transforming them into a novel and more effective schemes. Experimental results validate the effectiveness of the DS methodology in simultaneously counting and extracting endmembers not only as an individual algorithm but also as a foundation algorithm for improving existing methods. Our full code is released for public evaluation.[1]

Index Terms—Divergent subset (DS), endmember estimation, hyperspectral image, spectral unmixing.

Xuanwen Tao and Peng Ren are with the College of Oceanography and Space Informatics, China University of Petroleum (East China), Qingdao 266580, China (e-mail: txw_upc@126.com; pengren@upc.edu.cn).

Tingwei Cui is with the School of Atmospheric Sciences, Sun Yat-sen University, Zhuhai 519000, China, also with the Guangdong Province Key Laboratory for Climate Change and Natural Disaster Studies, Sun Yat-sen University, Zhuhai 519000, China, and also with the Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Sun Yat-sen University, Zhuhai 519000, China (e-mail: cuitw@mail.sysu.edu.cn).

Antonio Plaza is with the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, University of Extremadura, E-10071 Caceres, Spain (e-mail: aplaza@unex.es).

Color versions of one or more of the figures in this article are available online at https://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2020.2992542

[1]https://github.com/xuanwentao/DivergentSubset

I. INTRODUCTION

REMOTELY sensed hyperspectral images [1] provide detailed information about different materials on the surface of the earth. The detailed spectral information provided by these images has been widely utilized for object classification [2]–[4], target identification [5]–[7], change detection [8]–[10], and so on. One common problem associated with hyperspectral images is low spatial resolution, which inevitably leads to mixed pixels. A mixed pixel in a hyperspectral image covers several types of materials, and the corresponding spectrum is a mixture of several ground cover spectra known as endmembers. The mixed pixel phenomenon complicates the process of classifying one hyperspectral pixel into one unique material type. To address the ambiguity, spectral unmixing techniques [11]–[13] have been comprehensively investigated for the purpose of extracting endmembers. These techniques decompose each mixed pixel into a proportional composition of endmembers. The constituent proportion with respect to different types of materials [14] for each pixel is referred to as the abundance.

One fundamental requirement for effectively performing spectral unmixing is to accurately identify the endmembers of a hyperspectral image. This is normally achieved by two steps. First, the number of endmembers, i.e., the number of types of pure pixels (possibly with noise) in the image, is determined. Second, according to the number, the endmembers are extracted from the hyperspectral image. We refer to the overall procedures regarding estimating the number of endmembers and endmember extraction as endmember estimation [15] in this article. In Section I-A, we review the literature about estimating the number of endmembers and endmember extraction. In Section I-B, we briefly describe the major contributions of this article in terms of presenting a new endmember estimation framework that simultaneously determines the number of endmembers and extracts their spectral signatures.

A. Literature Review

Determining the number of endmembers [16], [17] is the first step of endmember estimation. In the literature, the number of endmembers is normally regarded as prior knowledge. However, such prior knowledge is almost unavailable in an arbitrary hyperspectral image. In this scenario, if the number

of endmembers is not accurately determined, most existing endmember extraction techniques cannot operate properly. The limitation regarding an unknown number of endmembers hinders most existing endmember extraction techniques from an operational viewpoint. To address the limitation, various algorithms have been developed for estimating the number of endmembers. Particularly, information theory-based algorithms, eigenvalue thresholding algorithms, and geometry characterization methods are three main families of techniques for estimating the number of endmembers. The first family includes information-theoretic criteria based on minimum description length [18], Akaike's information criterion [19], the Bayesian information criterion [20], and so on. Different models have been developed for encoding a negative data log-likelihood term, based on which an accurate estimation of the number of endmembers is expected to be obtained when the model achieves a global optimum. These strategies depend on the empirical configuration of specific mixed models or likelihood functions, and inappropriate configurations would cause estimation errors of the number of endmembers. The second family is related to the eigenvalue thresholding schemes [21]–[23]. A threshold is applied to eigendecomposition results from subspace analysis, such as principal component analysis (PCA) [24], hyperspectral signal subspace by minimum error (Hysime) [25], and the so-called Harsanyi–Farrand–Chang (HFC) analysis [26]. The PCA-based approaches aim to characterize a cutoff gap between the eigenvalues caused by signals and noise. However, the variation between the two eigenvalues may be insignificant, possibly resulting in an incorrect estimation of the number of endmembers. The Hysime approach conducts spectrum noise and the noise covariance estimation, which requires high computational complexity. The HFC approach requires a constant false alarm rate, which has an influence on the estimated number of endmembers. The third family is characterized by the geometry-based estimation of the number of endmembers (GENE), which includes the convex hull (GENE-CH) algorithm, the affine hull (GENE-AH) algorithm [27], and so on. They assume that data samples should lie within a convex hull (CH) or an affine hull (AH) with the endmember signatures as vertices. GENE algorithms operate along with an endmember extraction algorithm (EEA). In this scenario, a maximum hull volume would stop the EEA from extracting the next endmember signature. Therefore, the GENE algorithms depend on the effectiveness of the EEA used, and different EEAs cause different accuracies for counting endmembers.

Once the number of endmembers has been determined, the subsequent processing step for endmember estimation is to extract the determined number of endmember signatures [28]. The N-FINDR algorithm and its derivatives, such as cofactor NFINDR [29], geometric distance constrained N-FINDR [30], and locality preserving N-FINDR [31], are representative methods for endmember extraction. The N-FINDR-based algorithms are devised to search for a simplex with the greatest volume. The pixels that form the simplex are endmembers. The N-FINDR algorithms repeatedly extract endmembers, and the initial endmember matrix chosen from all pixel combinations

is random. The simplex growing algorithm (SGA) [32] seeks an endmember convex geometry with the minimum volume, and it also requires specific initialization. To avoid repeated identification of an endmember, the automatic target generation process (ATGP) [33] is proposed to extract endmembers by utilizing the notion of orthogonal subspace projection. The alternating decoupled volume max–min (ADVMM) [34] and successive decoupled volume max–min (SDVMM) [34] are developed for addressing the worst case simplex volume maximization problem by alternating optimization and by successive optimization, respectively. Specially, alternating volume maximization (AVMAX) [35] and successive volume maximization (SVMAX) [35] have been exploited to characterize more comprehensive criteria. Negative abundance-oriented (NABO) [36] uses the pixels outside the hull as alternative candidate endmembers. Standard least-square optimization is exploited to achieve estimates of endmember combinations following the geometrical properties of hyperspectral signatures [37]. The improved quantum-behaved particle swarm optimization (IQPSO) [38] is a novel optimization algorithm for hyperspectral endmember extraction. It is the first solid work to construct quantum mechanics-driven evolutionary computing algorithms, surpassing state-of-the-art evolutionary computing EEA. In this scenario, the accurate number of endmembers is necessary for effective endmember extraction [39].

B. Contributions

As reviewed in Section I-A, existing studies for estimating the number of endmembers and for extracting endmembers are performed as two independent topics in the literature. Most EEAs assume the known number of endmembers, but the estimation algorithms of the number of endmembers tend to be characterized by independent methods that may not seamlessly benefit the EEAs. One overall goal of this article is to bridge these two seemingly independent but, in fact, highly related procedures in terms of establishing a unifying endmember estimation framework. The major contributions of this article are threefold.

1) We develop an overall endmember estimation framework that simultaneously determines the number of endmembers and extracts their endmember signatures. Specifically, we define a new concept, i.e., the divergent subset (DS), the computation of which extracts endmembers without prior knowledge of the number of endmembers. In addition, the cardinality of the DS is naturally in accordance with the number of endmembers. Our DS method is highly motivated by the concept of the dominant set [40]–[42], which aims at extracting a homogeneous cluster from a set. However, a DS and a dominant set are defined in opposite ways. Specifically, the DS seeks the largest difference, but the dominant set characterizes the greatest homogeneity. Be that as it may, the motivations and computational techniques concerning the dominant set provide a tractable vehicle for modeling our DS. To the best of our knowledge, we are among the first to develop an automatic framework that simultaneously counts and extracts endmembers.

2) The DS framework provides a strategy for transforming existing endmember extraction methods into a novel and more accurate ones that do not require the knowledge of the number of endmembers. To this end, the DS co-operates with one arbitrary existing endmember extraction method. As reviewed in Section I-A, endmember extraction normally requires the knowledge of the number of endmembers. We artificially set the number of endmembers to be a value that is reasonably larger than the actual number of endmembers and, subsequently, perform the endmember extraction. In this scenario, the real endmembers are a subset of the extracted pixels, which are redundant and with less discriminativeness. We then apply the DS computation to the extracted pixel set. The DS finely prunes the coarsely extracted pixels, resulting in accurate endmembers with their number automatically determined. Therefore, the DS framework not only is capable of simultaneously counting and extracting endmembers by itself but also can be considered as a foundation stone for developing novel endmember extraction schemes based on arbitrary existing endmember extraction methods. The resulting endmember extraction schemes have no requirement of knowing the number of endmembers but exhibit the potential to provide better accuracy than the original endmember extraction method.

3) The DS has the property that it encompasses pixels with the largest unlikeness and any additional pixel would decrease its unlikeness. Such property renders the DS robust in both endmember counting and endmember extraction. Experiments validate the advantages of the DS over state-of-the-art endmember counting methods and endmember extraction methods.

C. Paper Outline

The rest of this article is organized as follows. In Section II, we present a series of definitions and then give a recursive definition of a DS. In Section III, we explain the possibility of using a DS to characterize endmembers. In Section IV, we introduce a closed-form formulation of a DS. In Section V, we describe how to numerically compute the DS. In Section VI, we describe how to employ the DS to transform existing endmember extraction methods into new effective schemes. In Section VII, we empirically evaluate the performance of the DS method in extracting and counting endmembers on both synthetic and real data. In Section VIII, we conclude this article with some remarks and hints at plausible future research lines.

II. RECURSIVE DEFINITION OF A DS

We consider a hyperspectral image with L-band and N pixels as a hyperspectral pixel set $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n, \ldots, \mathbf{x}_N] \in \mathbb{R}^{L \times N}$, in which the nth pixel $\mathbf{x}_n = [x_{1n}, x_{2n}, \ldots, x_{Ln}]^{\mathrm{T}} \in \mathbb{R}^{L \times 1}$ is an L-dimensional vector representing its L-band spectrum. Let $\mathbf{D} \in \mathbb{R}^{N \times N}$ denote the feature distance matrix, and the (i, j)th element d_{ij} of $\mathbf{D}$ represents the feature distance

between the ith and jth hyperspectral pixels in a feature space. The feature distance d_{ij} is computed by the Euclidean distance

$$
\begin{aligned}
d_{ij} &= \parallel \mathbf{x}_i - \mathbf{x}_j \parallel_2 \\
&= \sqrt{(x_{1i} - x_{1j})^2 + \cdots + (x_{Li} - x_{Lj})^2}.
\end{aligned} \tag{1}
$$

Let $V = \{1, 2, \ldots, N\}$ denote the index set for all pixels in the hyperspectral image. We give the following definitions regarding a subset $E \in V$ and derive the most important concept in our work, i.e., the DS, at the end of this section.

Definition 1: We define the intrafeature distance $\overline{d}_E(j)$ for a pixel $j \in E$ with respect to the subset E as follows:

$$
\overline{d}_E(j) = \frac{1}{|E|} \sum_{i \in E} d_{ij}. \tag{2}
$$

It is easy to observe from (2) that $\overline{d}_{\{j\}}(j) = 0$ because $d_{jj} = 0$ represents a self-feature distance of the pixel j. The intrafeature distance $\overline{d}_E(j)$ measures the average feature distance between the pixel $j \in E$ and the rest of pixels in E. It also reflects how different the pixel $j \in E$ is from the rest of pixels in E. A large value of the intrafeature distance $\overline{d}_E(j)$ implies that the pixel $j \in E$ is generally very different from the rest of pixels in E.

Definition 2: For the pixels $j \in E$ and $i \notin E$, we define the incremental feature distance $\Delta d_E^j(i)$ of the feature distance d_{ij} over the intrafeature distance $\overline{d}_E(j)$ as follows:

$$
\Delta d_E^j(i) = d_{ij} - \overline{d}_E(j). \tag{3}
$$

The incremental feature distance $\Delta d_E^j(i)$ reflects a relative feature distance between the pixel $i \notin E$ and the pixel set E, with the pixel $j \in E$ being an intermediary between them. It can be either positive or negative. A positive value of $\Delta d_E^j(i)$ implies that the feature distance between the pixel $i \notin E$ and the pixel $j \in E$ is greater than the average feature distance between the pixel $j \in E$ and those in E. In other words, the difference between the pixels i and j is greater than the general difference between the pixel j and those in E. A negative value of $\Delta d_E^j(i)$ implies that the feature distance between the pixel $i \notin E$ and the pixel $j \in E$ is shorter than the average feature distance between the pixel $j \in E$ and those in E. That is, the difference between the pixels i and j is less than the general difference between the pixel j and those in E.

Definition 3: We define the divergent grade $d_E(i)$ of the pixel $i \in E$ with respect to the subset E as follows:

$$
d_E(i) = \begin{cases} 1, & \text{if } |E| = 1 \\ \sum_{j \in E \setminus i} \Delta d_{E \setminus i}^j(i) d_{E \setminus i}(j), & \text{otherwise.} \end{cases} \tag{4}
$$

The divergent grade measures, in the feature space, how distant one pixel is from all the pixels in the set. It can be either positive or negative. A positive value of the divergent grade $d_E(i)$ reflects that the pixel i is quite different from most pixels in E. On the other hand, a negative value of the divergent grade $d_E(i)$ indicates the indiscriminate resemblance of the pixel i to the subset E.

Definition 4: We define the divergent degree of the pixel subset E as follows:

$$
\mathcal{D}(E) = \sum_{i \in E} d_E(i). \tag{5}
$$

The divergent degree measures, in the feature space, how far apart each pixel in E is from one another. It reflects the unlikeness within a set. A big divergent degree value reflects that any pair of pixels in E is very distant in the feature space, i.e., they are quite different. On the other hand, a small divergent degree implies that in E, some of the pixels are not so far apart from another, such that they tend to be close to one another in the feature space, i.e., at least some of the pixels are very similar.

Proposition 1: For a nonempty set V in which the pixel feature distances are not unique, there exists a subset $\hat{E} \subseteq V$ that satisfies the following two conditions.

1) $\forall j \in \hat{E}, \ d_{\hat{E}}(j) > 0$.
2) $\forall i \notin \hat{E}, \ d_{\hat{E} \bigcup i}(i) < 0$.

Proof: The proof of the proposition requires a Lagrangian formulation. The evidence for the proposition can then be identified via the Karush–Kuhn–Tucker (KKT) conditions. The details can be found within the proof of Proposition 2 in Section IV.

Definition 5: For a set V, we define the subset which satisfies the two conditions in Proposition 2 as the DS $\hat{E}$ of the set V. Equivalently, a DS is the one that has the largest divergent degree.

The two conditions in Proposition 2 describe the two main properties of a DS $\hat{E}$. The first condition favors the internal unlikeness of pixels within $\hat{E}$. It requires that each pixel should be far apart from any other pixels within $\hat{E}$. The second condition excludes external likeness. The incorporation of an additional pixel outside $\hat{E}$ into $\hat{E}$ would increase its likeness in terms of introducing a negative divergent grade. Therefore, the additional pixel would decrease the divergence degree and should be excluded, and we use the DS $\hat{E}$ as the endmember set, and its number is the number of endmembers.

III. DS of Pixels as Endmembers

The goal of endmember extraction from a hyperspectral image is to identify a subset of pixels that are most different from one another. The divergent degree provides a meaningful tool for measuring the unlikeness of pixels within a subset, with respect to pixels outside the subset. By Definition 5, a DS represents a group of pixels with the largest divergent degree. We, thus, consider the DS $\hat{E}$ as the endmembers of the hyperspectral image. Furthermore, once a DS is determined, the number of endmembers can be accordingly determined in terms of its cardinality.

However, it is almost intractable to compute a DS based on Proposition 2 and Definition 5. Specifically, according to (4), recursively computing the divergent grade $d_E(i)$ of the pixel $i \in E$ with respect to the subset E is computationally exhaustive. To render a practical computation scheme, we introduce a closed-form formulation of the DS in Section IV.

IV. Closed-Form Formulation of the DS

We reformulate the concept of DS in terms of Lemma 2 and Proposition 2 in this section.

Lemma 1: The divergent grade $d_{\hat{E} \bigcup i}(i)$ can be computed as follows:

$$d_{\hat{E} \bigcup i}(i) = \sum_{k \in \hat{E}} (d_{ik} - d_{jk}) d_{\hat{E}}(k). \tag{6}$$

Proof: The computing method for $d_{\hat{E} \bigcup i}(i)$ requires elementary properties of the determinant. The detailed proof can be found in [41, p. 98].

Proposition 2: The DS can be computed in terms of solving the following quadratic optimization problem:

$$\hat{\mathbf{y}} = \arg\max_{\mathbf{y}} \frac{1}{2}\mathbf{y}^{\mathrm{T}}\mathbf{D}\mathbf{y}$$

$$\text{s. t.} \quad y_i \geq 0, \quad \sum_{i=1}^{N} y_i = 1. \tag{7}$$

The nonzero elements in the optimal $\hat{\mathbf{y}} \in \mathbb{R}^{N \times 1}$ indicate the DS.

Proof: We commence by the following setting:

$$\hat{y}_k = \begin{cases} \dfrac{d_{\hat{E}}(k)}{\mathcal{D}(\hat{E})}, & \text{if } k \in \hat{E} \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

According to Definition 5, we have $\mathcal{D}(\hat{E}) = \sum_{k \in \hat{E}} d_{\hat{E}}(k)$. According to condition 1 of Proposition 2, for $k \in \hat{E}$, $d_{\hat{E}}(k) > 0$. Therefore, the setting (8) plays the same role with the conditions $y_i \geq 0$ and $\sum_{i=1}^{N} y_i = 1$ in the optimization problem (7). It follows condition 1 of Proposition 2.

According to the second condition of Proposition 2, for $i \notin \hat{E}$, $d_{\hat{E} \bigcup i}(i) < 0$. Therefore, we have

$$\frac{d_{\hat{E} \bigcup i}(i)}{\mathcal{D}(\hat{E})} < 0. \tag{9}$$

According to Lemma 1 and the setting (8), we have

$$\begin{aligned} \frac{d_{\hat{E} \bigcup i}(i)}{\mathcal{D}(\hat{E})} &= \frac{\sum_{k \in \hat{E}}(d_{ik} - d_{jk})d_{\hat{E}}(k)}{\mathcal{D}(\hat{E})} \\ &= \sum_{k \in \hat{E}}(d_{ik}\frac{d_{\hat{E}}(k)}{\mathcal{D}(\hat{E})} + d_{jk}\frac{d_{\hat{E}(k)}}{\mathcal{D}(\hat{E})}) \\ &= \sum_{k \in \hat{E}} d_{ik}\hat{y}_k - \sum_{k \in \hat{E}} d_{jk}\hat{y}_k \\ &= (\mathbf{D}\hat{\mathbf{y}})_i - (\mathbf{D}\hat{\mathbf{y}})_j \end{aligned} \tag{10}$$

where $(\mathbf{D}\hat{\mathbf{y}})_i$ and $(\mathbf{D}\hat{\mathbf{y}})_j$ represent the ith and jth elements of $\mathbf{D}\hat{\mathbf{y}}$, respectively.

Based on (9) and (10), we have the relation between one pixel $i \notin \hat{E}$ and one pixel $j \in \hat{E}$ formulated by using $\mathbf{D}$ and $\hat{\mathbf{y}}$ as follows:

$$(\mathbf{D}\hat{\mathbf{y}})_i - (\mathbf{D}\hat{\mathbf{y}})_j < 0. \tag{11}$$

The relation in (11) follows condition 2 of Proposition 2. Therefore, the relations in (8) and (11) satisfy the two conditions of Proposition 2, and they conceptually characterize a DS. We then investigate whether the optimization problem (7) leads to the same relations as (8) and (11).

As discussed earlier, the conditions $y_i \geq 0$ and $\sum_{i=1}^{N} y_i = 1$ for the optimization problem (7) are in accordance with the

setting (8). Therefore, they follow condition 1 of Proposition 2. Furthermore, the Lagrangian function for the optimization problem (7) is given as

$$L(\hat{\mathbf{y}}, \lambda, \gamma) = -\frac{1}{2}\hat{\mathbf{y}}^{\mathrm{T}}\mathbf{D}\hat{\mathbf{y}} + \lambda(\sum_{k \in \hat{E}} \hat{y}_k - 1) - \sum_{k \in \hat{E}} \gamma_k \hat{y}_k. \quad (12)$$

The KKT conditions for the optimization problem in (7) require

$$\frac{\partial L}{\partial \hat{y}_k} = -(\mathbf{D}\hat{\mathbf{y}})_k + \lambda - \gamma_k = 0$$
$$\hat{y}_k \gamma_k = 0$$
$$\gamma_k \geq 0. \quad (13)$$

Based on (13), we have

$$(\mathbf{D}\hat{\mathbf{y}})_k = \lambda - \gamma_k. \quad (14)$$

In addition, the complementary slackness of the KKT conditions has specific requirements for the term $\hat{y}_k \gamma_k = 0$ as follows:

$$\forall k \text{ at } \hat{y}_k = 0, \quad \text{it is required that } \gamma_k > 0$$
$$\forall k \text{ at } \hat{y}_k > 0, \quad \text{it is required that } \gamma_k = 0.$$

The proposition claims that one nonzero $\hat{y}_k > 0$ indicates an endmember. According to the complementary slackness, we have

$$\gamma_i > 0 \text{ for } i \notin \hat{E}$$
$$\gamma_j = 0 \text{ for } j \in \hat{E}. \quad (15)$$

Based on (14) and (15), we have

$$(\mathbf{D}\hat{\mathbf{y}})_i - (\mathbf{D}\hat{\mathbf{y}})_j = (\lambda - \gamma_i) - (\lambda - \gamma_j)$$
$$= -\gamma_i$$
$$< 0. \quad (16)$$

We observe that the relation in (16), which is derived from the closed-form (7), is consistent with the relation in (11), which follows condition 2 of Proposition 2. This proves that the closed-form (7) and Definition 5 provide equivalent formulations for characterizing a DS. This completes the proof.

The relation in (16) is obtained subject to the KKT conditions that are the necessary conditions of the closed-form formulation (7) of a DS. The necessity proves the existence of a DS, which is claimed in Proposition 2.

It is worth noting that the solution y_i of the optimization problem (7) not only extracts the endmember signatures in terms of its nonzero elements but also indicates the number of endmembers by counting the nonzero elements. The number of nonzero elements in $\mathbf{y}$ is the cardinality of the DS and is equal to the number of endmembers. Therefore, our DS method simultaneously extracts endmember signatures and determines their number.

Though Proposition 2 provides a principled manner for formulating a DS, it is still difficult for us to obtain an analytical solution. We will describe how to numerically compute the optimization problem (7) in Section V.

V. NUMERICAL COMPUTATION OF THE DS

We follow the strategy of replicator dynamics in evolutionary game theory [43] to numerically compute the solution for the optimization problem (7) as follows:

$$y_i^{\mathrm{new}} = y_i \frac{(\mathbf{D})_i^{\mathrm{T}} \mathbf{y}}{\mathbf{y}^{\mathrm{T}} \mathbf{D} \mathbf{y}} \quad (17)$$

where $(\mathbf{D})_i$ is the ith column of the matrix $\mathbf{D}$. The iterative updates according to (7) make $\mathbf{y}$ converge to an optimum. The convergence of the iterative computation is guaranteed by the Baum–Eagon inequality [44].

VI. IMPROVING ENDMEMBER EXTRACTION TECHNIQUES VIA DSs

When implementing an endmember extraction technique, we normally assume that the number of endmembers M in a hyperspectral image $\mathbf{X}$ should be known as prior information. Based on the information, M endmembers are identified by the endmember extraction technique. If an estimated number of endmembers $M_l < M$ is provided for the endmember extraction technique, at most M_l correct endmembers can be extracted and at least $M - M_l$ endmembers would be ignored. On the other hand, if the endmember extraction technique, e.g., vertex component analysis (VCA) [45], is informed of an estimated number of endmembers $M_g > M$ (in Section VII-A2, we use the DS method to improve the traditional VCA algorithm, and the value of M_g is 50), it extracts M_g candidate endmembers that include the M true endmembers. This observation makes sense for most endmember extraction techniques because they tend to extract a pixel, which is most different from others as an endmember from iteration to iteration such that the M endmember should be picked out before the iteration in which any of the $(M_g - M)$ nonendmembers are extracted. It is also observed that the $(M_g - M)$ extracted nonendmembers are not as discriminative as the M endmembers. The addition of the subset of the $(M_g - M)$ nonendmembers to the subset of M endmembers increases the likeness of the M endmembers and, in turn, reduces their divergent degree.

In the light of these observations, we use the DS as a postprocessing procedure for finely pruning the M_g candidate endmembers extracted by the original endmember extraction technique. Specifically, we consider the candidate endmembers as a set and perform endmember extraction by computing the DS of the candidate endmember set.

In this scenario, a novel improved endmember estimation strategy is established according to the following three steps. First, we assume the number of endmembers to be M_g, which is far greater than the actual number of endmembers M. Second, based on the assumption, we apply an endmember extraction technique, e.g., VCA, to a hyperspectral image to obtain a set of M_g of candidate endmembers. Third, we compute the DS of the candidate endmember set. The DS cardinality is the number of endmembers, and the hyperspectral pixels in DS are the endmembers.

One advantage of the strategy of the endmember extraction improvement based on DSs is that it shares the effectiveness of

an endmember extraction technique in determining individual endmembers and avoids its limitation on the prior knowledge of the number of endmembers. Therefore, the DS methodology provides a foundation stone for improving existing endmember extraction techniques.

VII. Experimental Results

We use both synthetic data and real data to evaluate the performance of the proposed DS methodology in extracting and counting endmembers. The experimental results are compared with eight EEAs, i.e., ADVMM [34], SDVMM [34], AVMAX [35], SVMAX [35], p-norm-based pure pixel identification (TRIP) [27], ATGP [33], NABO [36], and SGA [32]. In addition, we experimentally compare the performance of the DS, Hysime [25], HFC analysis [26], and the AH (GENE-AH) algorithm [27] in estimating the number of endmembers.

We use three metrics, i.e., spectral information divergence (SID), abundance information divergence (AID), and root-mean-square error (RMSE) to evaluate the performance of endmember extraction of the DS method and the comparison algorithms. Let $\mathbf{X}_e = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m, \ldots, \mathbf{x}_M] \in \mathbb{R}^{L \times M}$ denote the true endmember signatures and $\hat{\mathbf{X}} = [\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \ldots, \hat{\mathbf{x}}_m, \ldots, \hat{\mathbf{x}}_M] \in \mathbb{R}^{L \times M}$ be the estimated endmember signatures, where M is the number of endmembers, $\mathbf{x}_m = [x_{1m}, x_{2m}, \ldots, x_{lm}, \ldots, x_{Lm}]^{\mathrm{T}} \in \mathbb{R}^{L \times 1}$, and $\hat{\mathbf{x}}_m = [\hat{x}_{1m}, \hat{x}_{2m}, \ldots, \hat{x}_{lm}, \ldots, \hat{x}_{Lm}]^{\mathrm{T}} \in \mathbb{R}^{L \times 1}$. The probability distribution vector of the mth true endmember signature is given by $\mathbf{p}_m = (\mathbf{x}_m / \sum_{l=1}^{L} x_{jm})$. The vector characterizes the variability of spectral signatures. Let $\mathbf{P}_e = [\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_m, \ldots, \mathbf{p}_M] \in \mathbb{R}^{L \times M}$ be the probability distribution matrix for $\hat{\mathbf{X}}_e$, $\hat{\mathbf{P}} = [\hat{\mathbf{p}}_1, \hat{\mathbf{p}}_2, \ldots, \hat{\mathbf{p}}_m, \ldots, \hat{\mathbf{p}}_M] \in \mathbb{R}^{L \times M}$ represent the probability distribution matrix for $\hat{\mathbf{X}}_e$, p_{lm} be the (i, j)th element of matrix $\mathbf{P}_e$, and $\hat{p}_{lm}$ be the (i, j)th element of matrix $\hat{\mathbf{P}}$. We have the following equation:

$$D(\mathbf{x}_m | \hat{\mathbf{x}}_m) = \sum_{l=1}^{L} p_{lm} \log \frac{p_{lm}}{\hat{p}_{lm}}. \tag{18}$$

The relative entropy between $\mathbf{X}_e$ and $\hat{\mathbf{X}}$ is $D(\mathbf{X}_e | \hat{\mathbf{X}}) = [D(\mathbf{x}_1 | \hat{\mathbf{x}}_1), D(\mathbf{x}_2 | \hat{\mathbf{x}}_2), \ldots, D(\mathbf{x}_M | \hat{\mathbf{x}}_M)] \in \mathbb{R}^{1 \times M}$. The relative entropy is asymmetric with respect to $\mathbf{X}_e$ and $\hat{\mathbf{X}}$. To render a symmetric metric based on the relative entropy, the SID [46] is given by

$$\text{SID} = D(\mathbf{X}_e | \hat{\mathbf{X}}) + D(\hat{\mathbf{X}} | \mathbf{X}_e). \tag{19}$$

In order to make comparisons based on more comprehensive metrics, the abundance $\hat{\mathbf{a}}$ for one pixel $\mathbf{x}$ with respect to the estimated endmembers $\hat{\mathbf{X}}$ is employed, and each element of $\hat{\mathbf{a}} = [\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_M]^{\mathrm{T}} \in \mathbb{R}^{M \times 1}$ represents the abundance with respect to one endmember. Let $\mathbf{q} \in \mathbb{R}^{L \times 1}$ denote additive noise, and the abundance in terms of a linear mixture method [47] is given as follows:

$$\mathbf{x} = \hat{\mathbf{X}}\hat{\mathbf{a}} + \mathbf{q} \tag{20}$$

subject to the nonnegative and sum-to-one constraints

$$\hat{a}_m \geq 0, \ m = 1, 2, \ldots, M; \ \sum_{m=1}^{M} \hat{a}_m = 1. \tag{21}$$

The abundance vectors are used to replace the endmember signatures in SID to derive more comprehensive metrics. Let $\mathbf{A}_e = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n, \ldots, \mathbf{a}_N] \in \mathbb{R}^{M \times N}$ denote the true endmember abundance and $\hat{\mathbf{A}} = [\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \ldots, \hat{\mathbf{a}}_n, \ldots, \hat{\mathbf{a}}_N] \in \mathbb{R}^{M \times N}$ be the estimated endmember abundance. Accordingly, the AID is given as follows:

$$\text{AID} = D(\mathbf{A}_e | \hat{\mathbf{A}}) + D(\hat{\mathbf{A}} | \mathbf{A}_e). \tag{22}$$

The matrix $\mathbf{R}$ is $(\mathbf{A}_e - \hat{\mathbf{A}})$, and R_{ij} is the (i, j)th element of the matrix $\mathbf{R}$. The RMSE is given as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} \sum_{j=1}^{M} R_{ij}^4}{NM}}. \tag{23}$$

The value of zero indicates the best endmember extraction performance in terms of the metrics SID, AID, or RMSE. Therefore, the value of the metric close to zero indicates accurate endmember extraction.

A. Experimental Evaluations on Synthetic Data

The synthetic data are created based on the United States Geological Survey (USGS) mineral spectra library.[2] We generate synthetic data partially based on the linear mixture method by Miao and Qi [48]. The linear mixture method provides two families of synthetic data: 1) synthetic data with no pure pixels and 2) synthetic data with multiple pure pixels for each type of material. In our experiments, we test our methods in three scenarios. In the first scenario, we incorporate one pure pixel for each material type into the synthetic data from family 1. We, thus, create our synthetic data with one pure pixel for each type of material. In the second scenario, we use the synthetic data from family 2 directly as our synthetic data with multiple pure pixels for each type of material. Fig. 1 shows the visualization results of the first and second scenarios. In the third scenario, we add zero-mean Gaussian noise to family 2 as our synthetic data to evaluate the robustness of the DS method to noise.

The original hyperspectral data do not guarantee to maintain a metric space in which the feature distance can be conveniently measured. In order to compute the feature distance in a principled manner, we commence by mapping the original hyperspectral data into a low-dimensional space via PCA. The Euclidean distance is used for measuring the feature distance between pixels in the low-dimensional space, resulting in the feature distance matrix $\mathbf{D}$. We compute the DS indicator $\mathbf{y}$ based on (17). The endmembers are extracted by identifying the nonzero elements in the DS indicator $\mathbf{y}$. Accordingly, the number of nonzero elements is the number of endmembers.

It is worth noting that we use PCA to automatically map the original hyperspectral data into a dimension reduction space to obtain dimension reduction data. We first compute the eigenvalues and eigenvectors of the covariance matrix of the original data. We then compute cumulative contribution rate of the eigenvalues and use the eigenvectors corresponding to the eigenvalues whose the cumulative contribution rate exceeds 99.99% as the projection matrix. We finally use the projection

---

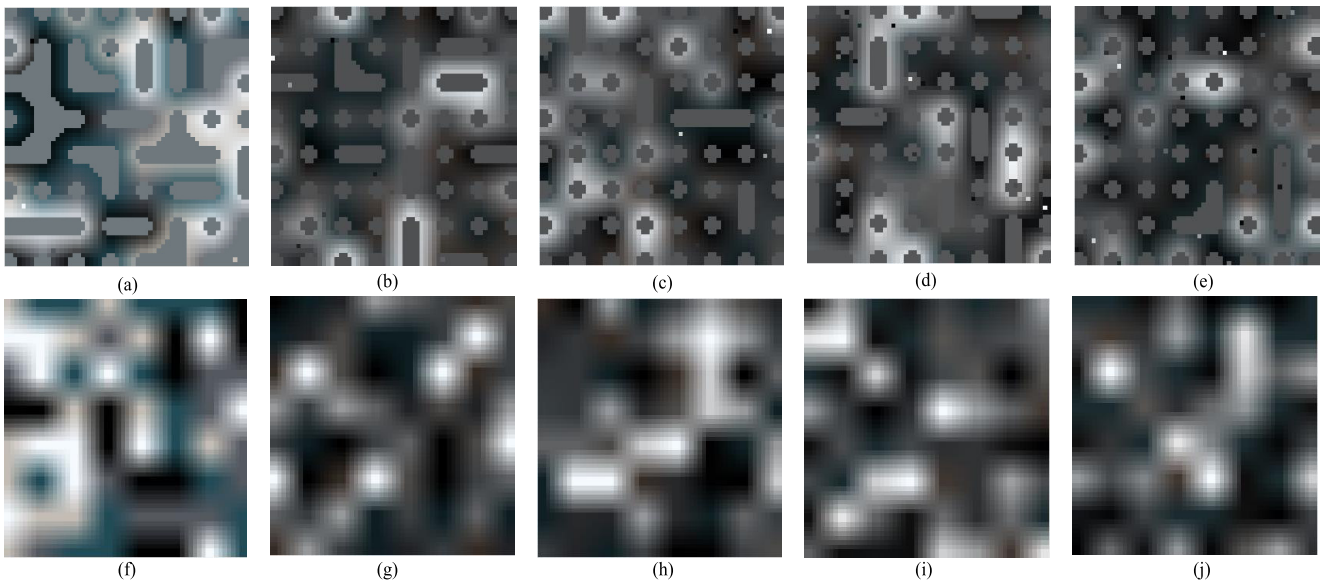[2] https://speclab.cr.usgs.gov/spectral-lib.html

Fig. 1.   Visualization results of synthetic data with different numbers of endmembers, i.e., 5, 10, 15, 20, and 25. (a)–(e) Synthetic data with one pure pixel for each type of material. (f)–(j) Synthetic data with multiple pure pixels for each type of material.
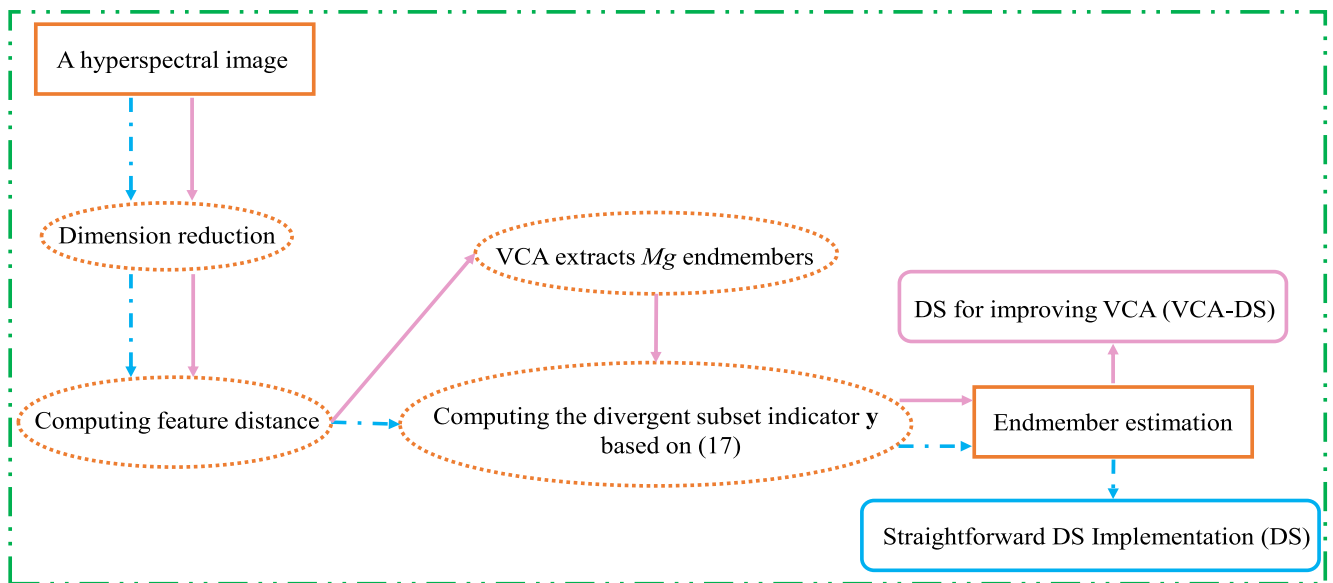

Fig. 2.   Flowchart of two different endmember estimation methods. Dashed dotted line: method of the straightforward DS implementation. Solid line: method of the DS for improving VCA (VCA-DS).

matrix to compute the dimension reduction data. It is also worth noting that more sophisticated data mapping techniques, such as manifold embedding methods, can be applied instead of PCA. Furthermore, more comprehensive distance functions, such as the Manhattan distance and the Mahalanobis distance, can be used rather than the straightforward Euclidean distance. Although the alternative strategies have potentials for improving the effectiveness of feature distance computation, we use the basic PCA and the Euclidean distance in this article because we aim to examine the effectiveness of our DS method even in the situation that the most elementary feature computation strategies are applied. Fig. 2 shows how to complete

the overall endmember estimation task by straightforward DS implementation and the DS for improving VCA.

1) Straightforward DS Implementation: We directly utilize DSs on synthetic data with one pure pixel and multiple pure pixels for each type of material. We obtain the divergent indicator $\mathbf{y}$ based on (17), and the nonzero elements in $\mathbf{y}$ are endmembers. Specially, the endmembers tend to be redundant, and multiple pure pixels might be identified for one type of material. To eliminate the redundancy, we compute the correlation matrix for the preliminary selection of endmembers. The standard correlation greater than a thresholding value (i.e., 0.99 in our experiments) indicates high correlation
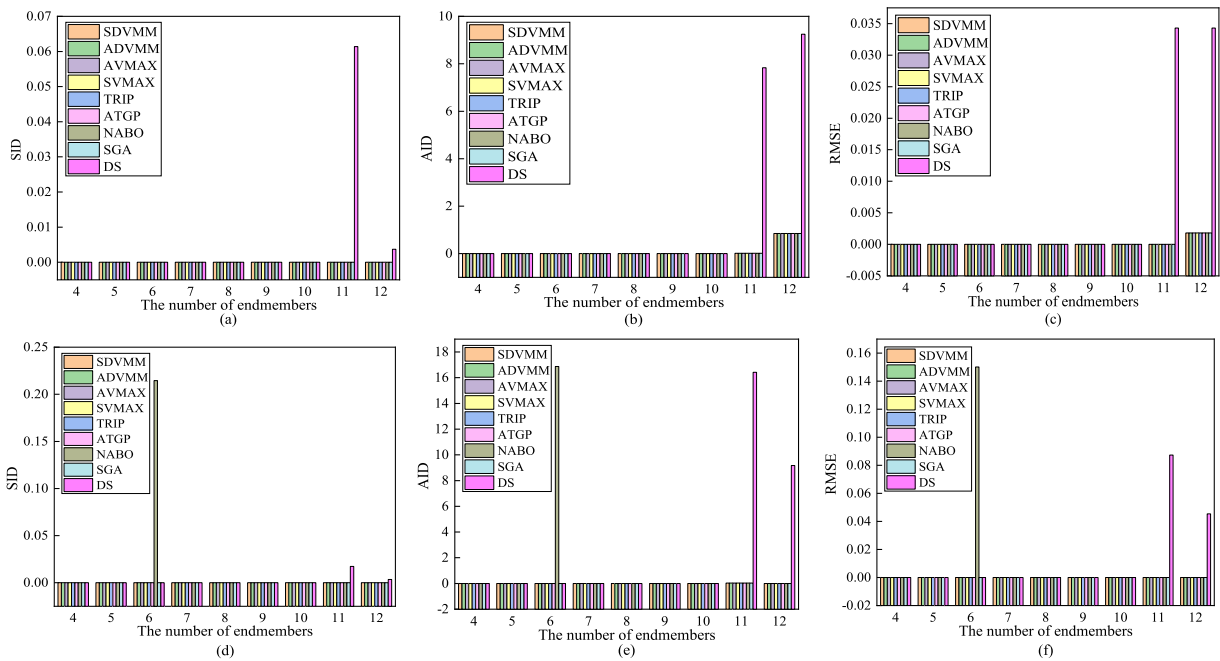
Fig. 3. Performance of the endmember extraction stage. (a)–(c) Synthetic data with one pure pixel for each type of material. (d)–(f) Synthetic data with multiple pure pixels for each type of material.
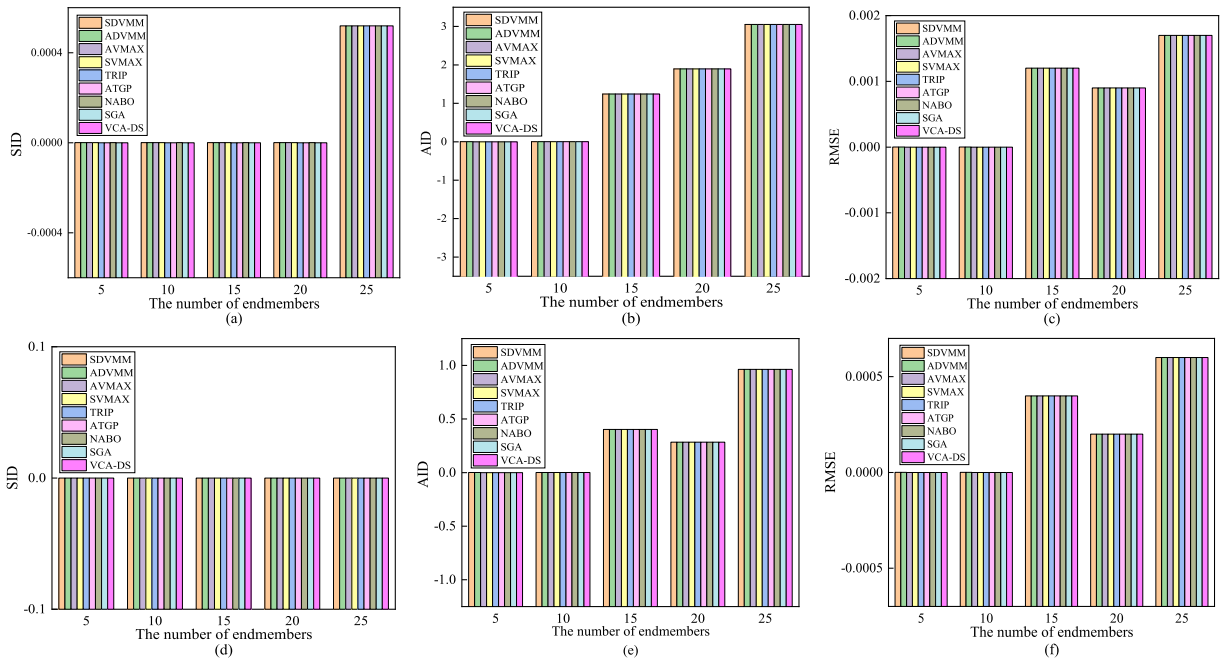
Fig. 4. Performance of extracting endmembers. (a)–(c) Synthetic data with one pure pixel for each type of material. (d)–(f) Synthetic data with multiple pure pixels for each type of material.

between pixels. Pixels that satisfy such high correlation condition in the preliminary selection are considered as multiple pure pixels for one type of material, and only one of them is used as an extracted endmember. We employ the metrics SID, AID, and RMSE to evaluate the endmember extraction performance. Each performance metric indicates an accurate result when its value is close to zero. Fig. 3 illustrates the

accuracy of different methods in extracting endmembers on different synthetic data. The values of the metrics SID, AID, and RMSE are almost close to zero for eight comparison methods, and the NABO method has an error when the number of endmembers is six on synthetic data with multiple pure pixels for each type of material. Though the performance of the straightforward DS implementation keeps accurate for
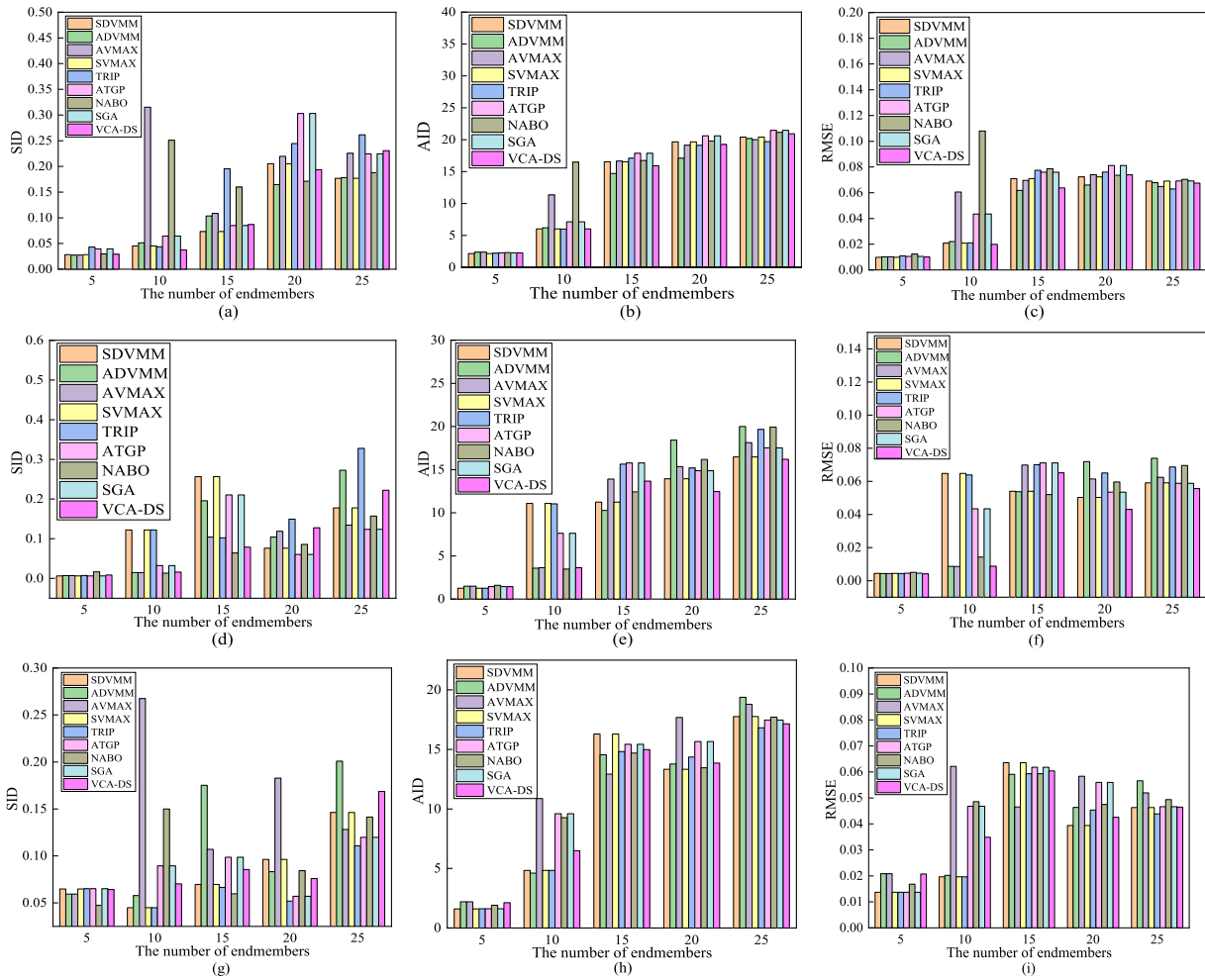
Fig. 5. Performance of extracting endmembers on synthetic data with different noises. (a)–(c) SNR = 10 dB. (d)–(f) SNR = 15 dB. (g)–(i) SNR = 20 dB.

counting a medium number of endmembers, it decreases when the number of endmembers exceeds ten. One possible reason for this inefficiency is that the pairwise relationships encoded by the DS are low order descriptions that are easily contaminated by noisy and interfering observations. To cope with this inefficiency of DSs for counting large numbers of endmembers, we extend our straightforward strategy to that described in Section VII-A2 and present the experimental evaluation in the following sections.

2) DSs for Improving VCA: To neutralize the susceptibility of DSs to noisy and interfering observations, we have proposed an improvement strategy in Section VI. Specifically, we commence by performing VCA on the original data with an assumed number of endmembers 50, which is evidently greater than the actual number of endmembers, and obtain a preliminary endmember extraction result. We then perform DSs on the preliminary result and obtain the divergent indicator **y**, and the nonzero elements in the **y** are endmembers. We refer to the novel strategy as VCA-DS. Fig. 4 shows the results of different methods. We observe that our VCA-DS method is an effective method for endmember extraction, and it is not limited by the large number of endmembers. The reason for SID, AID,

and RMSE exhibiting very small values in Figs. 3 and 4 is that the two types of synthetic data contain no noise, and the endmember spectrums extracted by DS and VCA-DS are the same as the reference spectrums.

Fig. 5 compares the performance of extracting endmembers for different methods at various noise levels. The metrics SID, AID, and RMSE are used to evaluate the sensitivity of different methods in extracting endmembers to various noise (10, 15, and 20 dB). We observe that the VCA-DS method is stable compared to other methods in extracting endmembers when the noise is different.

Fig. 6 shows a qualitative comparison between the endmembers extracted via the VCA-DS method and their corresponding true spectral signatures from the USGS spectra library. From Fig. 6, we observe that the endmember signatures estimated by the VCA-DS method are in good agreement with the reference USGS spectral signatures. The results shown in Fig. 6 reflect that the high accuracy of VCA-DS in endmember extraction. All experimental evaluations about endmember extraction validate that the VCA-DS method is effective in endmember extraction without the prior knowledge of the number of endmembers.
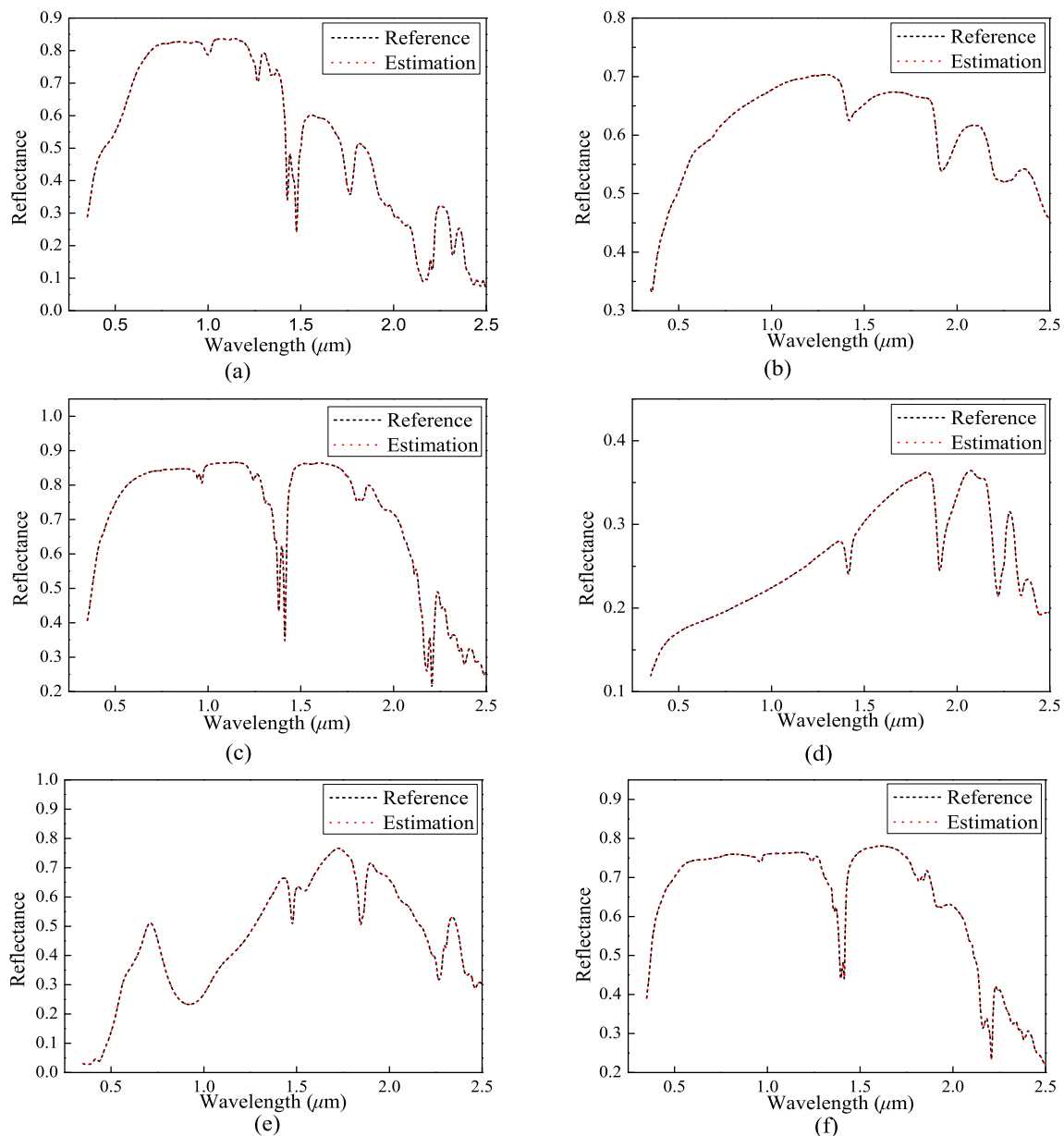
Fig. 6. Reference signatures and estimated endmember signatures by VCA-DS method on synthetic data. (a) Alunite. (b) Chalcedony. (c) Dickite. (d) Illite. (e) Jarosite. (f) Kaolinite.

Fig. 7 reports the performance of the VCA-DS method in counting endmembers on synthetic data with one pure pixel and multiple pure pixels for each type of material. As shown in Fig. 7, the Hysime method has errors when the number of endmembers exceeds 15 and 20 on the synthetic data with one pure pixel and multiple pure pixels for each type of material, respectively. Fig. 7 also reveals that the VCA-DS method and the two traditional methods (HFC and GENE-AH) achieve accurate estimations about the number of endmembers.

Tables I and II show the running average time of different methods in extracting and counting endmembers on different types of synthetic data. NABO and HFC are the fastest methods in extracting and counting endmembers, respectively. The difference in the average running time of different methods is very small. Especially, the VCA-DS method completes the overall endmember estimation task, i.e., extracting and counting endmembers, and other methods only complete one task, i.e., either endmember counting or endmember extraction. Therefore, VCA-DS is an efficient method in both extracting and counting endmembers.

3) Observations: We have empirically evaluated our DS methodology in two different tasks, i.e., endmember extraction and endmember counting. Each comparison method is only suited to one of the two tasks, and in contrast, our DS method simultaneously conducts both tasks. In addition, our DS method achieves comparable results in both tasks with the task-specific state of the art methods.

The derivative of the DS method, i.e., the VCA-DS method, can be viewed from two perspectives. On one hand, VCA can be viewed as a preprocessing step for DSs. It primarily filters

Fig. 7. Performance of counting endmembers. (a) Synthetic data with one pure pixel for each type of material. (b) Synthetic data with multiple pure pixels for each type of material.

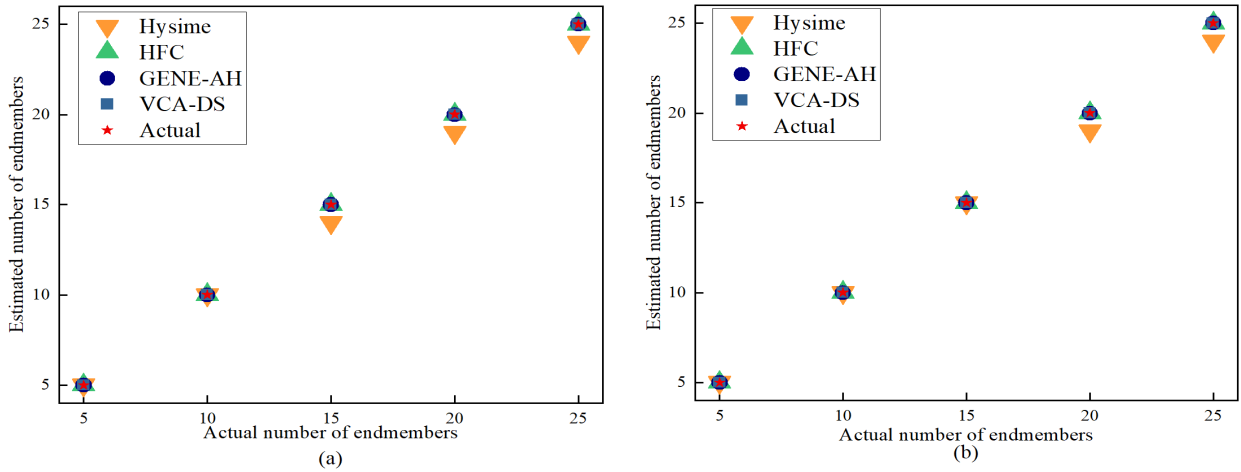TABLE I

AVERAGE RUNNING TIME (SECOND) FOR EXTRACTING ENDMEMBERS ON SYNTHETIC DATA

Data type	Methods								
	SDVMM	ADVMM	AVMAX	SVMAX	TRIP	ATGP	NABO	SGA	VCA-DS
One pure pixel	1.4794	1.4703	1.5394	1.4797	1.4235	1.1476	**1.1003**	1.3824	1.5087
Multiple pure pixels	1.2249	1.2035	1.2120	1.1393	1.0821	0.9724	**0.8612**	1.1366	1.1553
Multiple pure pixels (10db noise)	1.5179	1.5384	1.4853	1.4039	1.4253	1.2495	**1.1283**	1.4616	2.3318
Multiple pure pixels (15db noise)	1.5120	1.5747	1.4695	1.4348	1.4236	1.3104	**1.1463**	1.4792	2.3011
Multiple pure pixels (20db noise)	1.6468	1.7343	1.6237	1.5999	1.5910	1.3606	**1.2940**	1.6055	2.5154

TABLE II

AVERAGE RUNNING TIME (SECOND) FOR COUNTING ENDMEMBERS ON SYNTHETIC DATA

Data type	Methods			
	Hysime	HFC	GENE-AH	VCA-DS
One pure pixel	0.1456	**0.0806**	5.4391	1.5087
Multiple pure pixels	0.1636	**0.0775**	5.5262	1.1553

out the noisy and interweaving pixels that are definitely not endmembers but just interferes with endmember estimation and leaves comparatively clean pixels from which endmembers are simultaneously counted and extracted via the DS method. On the other hand, as described in Section VII-A2, the DS method can be considered as a postprocessing procedure that improves the effectiveness of VCA without the prior knowledge of the number of endmembers. Therefore, VCA and the DS method are complementary to each other, and the

VCA-DS strategy forms an effective framework for simultaneously counting and extracting endmember. It is robust and scalable to large numbers of endmembers.

B. Experimental Evaluations on Real Data

1) GOCI Data: We empirically test the DS method on GOCI data that covers an ocean region with the existence of green algae. Specifically, we test the DS method in terms of computing the abundance with respect to its extracted green algae endmember and accordingly estimating the green algae area.

The GOCI data have eight bands and 1219×1175 pixels. We use the DS method to extract endmembers from the GOCI image. In our experiments, the DS method automatically extracts two endmembers, each of which is denoted by $\hat{\mathbf{x}}_\beta = [x_1, x_2, \ldots, x_L]^T$. In order to determine which endmember represents the green algae, we use the green algae spectrum provided by the China State Oceanic Administration (CSOA) as the reference spectrum $\mathbf{r} = [r_1, r_2, \ldots, r_L]^T$ to identify the green algae endmember. Specifically, we compute the spectral angle $\arccos(\mathbf{r}^T \hat{\mathbf{x}}_\beta / |\mathbf{r}|^{-1/2} |\hat{\mathbf{x}}_\beta|^{-1/2})$ between the spectrum of each extracted endmember and the reference spectrum.
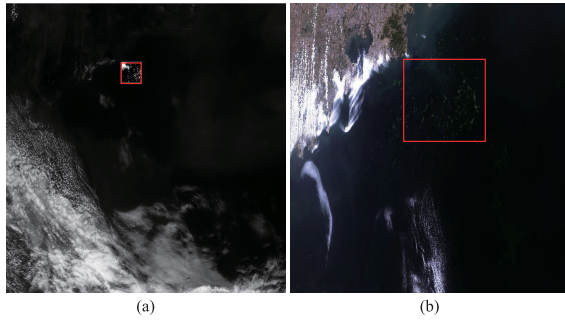
Fig. 8. Investigated ocean region with existence of green algae. (a) GOCI image. (b) HJ-1B image. The GOCI image and HJ-1B image collected on June 29, 2013. The same investigated regions are cropped from both the GOCI image and the HJ-1B images with the red rectangles, separately. The sizes of the cropped patches look different because of the different resolutions of the GOCI image and HJ-1B image.

We consider the endmember with the smallest angle to be the green algae endmember.

We compute the abundance based on (20). We estimate the area B of green algae blooms as follows:

$$B = S^2 \sum_{i=1}^{N} \hat{a}(i) \tag{24}$$

where $\hat{a}(i)$ denotes the green algae abundance of the ith pixel in the GOCI image, and S is the spatial resolution of GOCI image. The spatial resolution of the GOCI image used in our experiments is $R = 500$m. The green algae area is 69.18 km^2. To evaluate the accuracy of the green algae area estimation, we use the HJ-1B data that were collected at a close time to that of the GOCI data as the reference data. We consider that the GOCI and HJ-1B images capture the same scene at the same time. We manually segment the investigated ocean regions from the GOCI image and HJ-1B image by rectangles, as shown in Fig. 8(a) and (b), respectively. Though surrounding the same region, the two rectangles appear in different sizes in Fig. 8 because of the different resolutions of the GOCI image and HJ-1B image.

We used the green algae area 83.75 km^2, which is computed based on the HJ-1B data as the baseline. The area estimated via the DS method is close to the baseline. The results of the green algae area on GOCI and HJ-1B data reflect that the DS method extracts and counts endmembers effectively.

It is worth noting that the green algae area estimated by the DS method and the cofactor N-FINDR method [29] are different. The original assumption of the number of endmembers for the GOCI data with no cropping is four [30], and the cofactor N-FINDR method uses the assumption without considering that the image cropping would reduce the number of endmembers. Benefiting from the effectiveness of the DS method with more accurate endmember estimation, the green algae area estimated by the DS method is more accurate than that by the cofactor N-FINDR method reported in [29].

2) Samson, Jasper Ridge, Urban, and Cuprite Data: We utilize four real data[3] to evaluate the performance of the VCA-DS method in counting endmembers.

[3]http://lesun.weebly.com/hyperspectral-dataset.html

TABLE III
RESULTS OF ESTIMATED NUMBER OF ENDMEMBERS BY
DIFFERENT METHODS ON FOUR REAL DATA

Methods	Real data			
	Samson	Jasper Ridge	Urban	Cuprite
Hysime	43	18	27	13
HFC	5	3	**6**	**12**
GENE-AH	4	**4**	4	18
DS	**3**	**4**	7	**12**
Actual number	3	4	6	12

The first real data is Samson, which contains three endmembers. Samson has 952×952 pixels, each of which has 156 bands covering the wavelengths from 401 to 889 nm. The original image is too large, leading to very expensive in terms of computational cost. In this article, a region of 95×95 pixels is used. The second real data are Jasper Ridge, which has four endmembers. The Jasper Ridge data have 210 bands and 512×614 pixels with wavelengths ranging from 380 to 2500 nm. We crop the original image into a subimage of 100×100 pixels. The position of the first pixel is (105, 269). We remove the bands 1–3, 108–112, 154–166, and 220–224 (due to dense water vapor and atmospheric effects) and keep 198 bands in our experiments. The third real data is Urban that has four endmembers. The Urban data have 307×307 pixels with wavelengths ranging from 400 to 2500 nm. We remove the bands 1–4, 76, 87, 101–111, 136–153, and 198–210 (due to dense water vapor and atmospheric effects), and the modified data have 162 bands. The forth real data are the Cuprite [49]–[51], and it has 12 endmembers and covers the Cuprite in Las Vegas, USA. Cuprite has 224 bands with 250×190 pixels, and the wavelength ranges from 370 to 2480 nm. We remove some noisy bands (1–2 and 221–224) and water absorption bands (104, 113, 148, and 167) from the original 224 bands, and the modified Cuprite data have 188 bands. The abovementioned four real data are shown in Fig. 9.

Table III shows the performance of different algorithms in counting endmembers. According to our experiments, Hysime could not obtain accurate results on any real data set. HFC achieves accurate performance on Urban and Cuprite data. GENE-AH obtains accurate results on the Jasper Ridge data, and the DS method produces accurate results on Samson, Jasper Ridge, and Cuprite data and just has a minor error on the Urban data.

Our VCA-DS method simultaneously counts and extracts endmembers in the real data. Our experiments validate that it extracts accurate endmembers for the real data. Due to the page limitation, we did not illustrate the endmember extraction results here. Interested readers can conduct experiments following our settings. Our full code is released for public evaluation.[4]
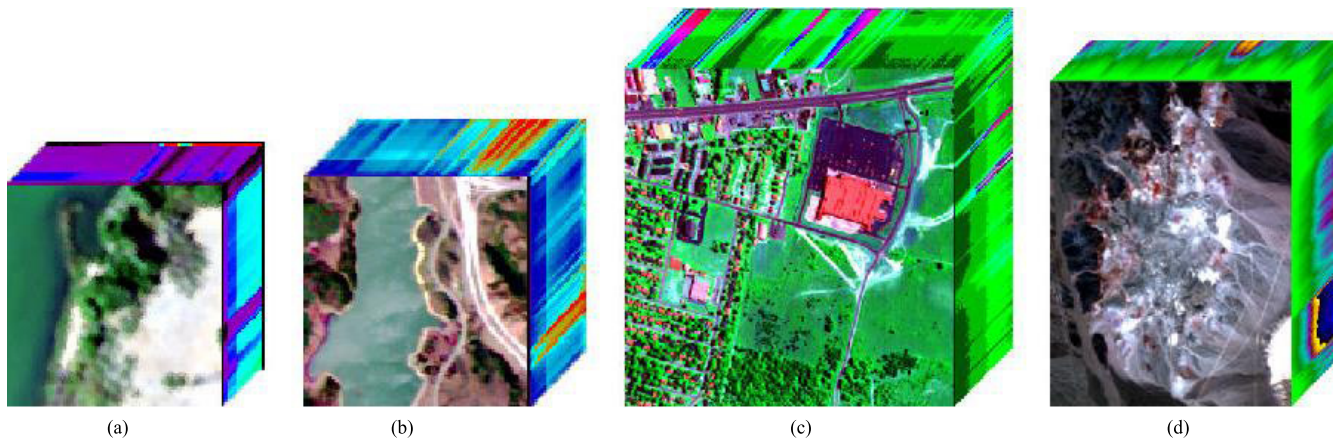
[4]https://github.com/xuanwentao/DivergentSubset

Fig. 9. Four different real data. (a) Samson. (b) Jasper Ridge. (c) Urban. (d) Cuprite.

VIII. CONCLUSION AND FUTURE WORK

We have developed a new endmember estimation methodology, referred to DSs, for simultaneously counting and extracting in a hyperspectral image. The DS has been validated to be capable of counting and extracting a medium number of endmembers. However, due to the limited capability of pairwise characterization in the DS method, it does not scale up to a large number of endmembers in the condition of noisy and interfering observations. To address this limitation, we have developed a new DS-based strategy that first uses an existing endmember extraction technique (e.g., VCA) to extract an arbitrarily large number of candidate endmembers and then uses the DS method to count and extract the actual endmembers in the candidates. This strategy has been validated to be effective to be scaled up to a large number of endmembers, without the requirement of the number of endmembers.

As with any new approach, there are some unresolved issues that may present challenges over time. Specifically, the DS method is one of the first available methods able to simultaneously extract endmember signatures and estimate their number in a hyperspectral scene. One limitation of the DS method is the assumption of the existence of pure pixels in the hyperspectral image. In our future work, we will investigate how to simultaneously determine the number of endmembers and extract the endmember signatures from hyperspectral images without pure pixels.

REFERENCES

[1] G. Wang, G. Dong, H. Li, L. Han, X. Tao, and P. Ren, "Remote sensing image synthesis via graphical generative adversarial networks," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2019, pp. 10027–10030.

[2] M. E. Paoletti, J. M. Haut, X. Tao, J. P. Miguel, and A. Plaza, "A new GPU implementation of support vector machines for fast hyperspectral image classification," *Remote Sens.*, vol. 12, no. 8, p. 1257, 2020.

[3] F. Luo, B. Du, L. Zhang, L. Zhang, and D. Tao, "Feature learning using spatial-spectral hypergraph discriminant analysis for hyperspectral image," *IEEE Trans. Cybern.*, vol. 49, no. 7, pp. 2406–2419, Jul. 2019.

[4] N. He *et al.*, "Feature extraction with multiscale covariance maps for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 755–769, Feb. 2019.

[5] F. Gao, Y. Zhang, J. Wang, J. Sun, E. Yang, and A. Hussain, "Visual attention model based vehicle target detection in synthetic aperture radar images: A novel approach," *Cognit. Comput.*, vol. 7, no. 4, pp. 434–444, Aug. 2015.

[6] F. Gao *et al.*, "Biologically inspired progressive enhancement target detection from heavy cluttered SAR images," *Cognit. Comput.*, vol. 8, no. 5, pp. 955–966, Oct. 2016.

[7] F. Gao, F. Ma, J. Wang, J. Sun, E. Yang, and H. Zhou, "Visual saliency modeling for river detection in high-resolution SAR imagery," *IEEE Access*, vol. 6, no. 1, pp. 1000–1014, 2018.

[8] A. Erturk, M.-D. Iordache, and A. Plaza, "Sparse unmixing-based change detection for multitemporal hyperspectral images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 2, pp. 708–719, Feb. 2016.

[9] A. Erturk, M.-D. Iordache, and A. Plaza, "Sparse unmixing with dictionary pruning for hyperspectral change detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 1, pp. 321–330, Jan. 2017.

[10] B. Du, L. Ru, C. Wu, and L. Zhang, "Unsupervised deep slow feature analysis for change detection in multi-temporal remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 9976–9992, Dec. 2019.

[11] R. Wang, H.-C. Li, A. Pizurica, J. Li, A. Plaza, and W. J. Emery, "Hyperspectral unmixing using double reweighted sparse regression and total variation," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 7, pp. 1146–1150, Jul. 2017.

[12] R. Fernandez-Beltran, A. Plaza, J. Plaza, and F. Pla, "Hyperspectral unmixing based on dual-depth sparse probabilistic latent semantic analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 11, pp. 6344–6360, Nov. 2018.

[13] J. Delgado, G. Martin, J. Plaza, L. I. Jimenez, and A. Plaza, "Fast spatial preprocessing for spectral unmixing of hyperspectral data on graphics processing units," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 2, pp. 952–961, Feb. 2016.

[14] F. Kowkabi, H. Ghassemian, and A. Keshavarz, "A fast spatial-spectral preprocessing module for hyperspectral endmember extraction," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 6, pp. 782–786, Jun. 2016.

[15] L. Drumetz *et al.*, "Hyperspectral local intrinsic dimensionality," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 7, pp. 4063–4078, Jul. 2016.

[16] R. Heylen, M. Parente, and P. Scheunders, "Estimation of the number of endmembers via the hubness phenomenon," in *Proc. 8th Workshop Hyperspectral Image Signal Process., Evol. Remote Sens. (WHISPERS)*, Aug. 2016, pp. 1–4.

[17] Z. Li, J. Li, and G. Wang, "Estimating the number of endmembers from hyperspectral image using noncentral chi-squared distribution model," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2016, pp. 6577–6580.

[18] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, Sep. 1978.

[19] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. AC-19, no. 6, pp. 716–723, Dec. 1974.

[20] M. W. Graham and D. J. Miller, "Unsupervised learning of parsimonious mixtures on large spaces with integrated feature and component selection," *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1289–1303, Apr. 2006.

[21] A. Halimi, P. Honeine, M. Kharouf, C. Richard, and J.-Y. Tourneret, "Estimating the intrinsic dimension of hyperspectral images using a noise-whitened eigengap approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 7, pp. 3811–3821, Jul. 2016.

[22] E. Terreaux, J.-P. Ovarlez, and F. Pascal, "New model order selection in large dimension regime for complex elliptically symmetric noise," in *Proc. 25th Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2017, pp. 1090–1094.

[23] S. Das, A. Routray, and A. K. Deb, "Noise robust estimation of number of endmembers in a hyperspectral image by eigenvalue based gap index," in *Proc. 8th Workshop Hyperspectral Image Signal Processing: Evol. Remote Sens. (WHISPERS)*, Aug. 2016, pp. 1–5.

[24] P. R. Peres-Neto, D. A. Jackson, and K. M. Somers, "How many principal components? Stopping rules for determining the number of non-trivial axes revisited," *Comput. Statist. Data Anal.*, vol. 49, no. 4, pp. 974–997, Jun. 2005.

[25] J. M. Bioucas-Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.

[26] C.-I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 608–619, Mar. 2004.

[27] A. Ambikapathi, T.-H. Chan, C.-Y. Chi, and K. Keizer, "Hyperspectral data geometry-based estimation of number of endmembers using p-norm-based pure pixel identification algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2753–2769, May 2013.

[28] N. Dobigeon, S. Moussaoui, M. Coulon, J.-Y. Tourneret, and A. O. Hero, "Joint Bayesian endmember extraction and linear unmixing for hyperspectral imagery," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4355–4368, Nov. 2009.

[29] X. Tao, T. Cui, and P. Ren, "Cofactor-based efficient endmember extraction for green algae area estimation," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 6, pp. 849–853, Jun. 2019.

[30] B. Pan, Z. Shi, Z. An, Z. Jiang, and Y. Ma, "A novel spectral-unmixing-based green algae area estimation method for GOCI data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 2, pp. 437–449, Feb. 2017.

[31] X. Tao, T. Cui, Z. Yu, and P. Ren, "Locality preserving endmember extraction for estimating green algae area," in *Proc. OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO)*, May 2018, pp. 1–4.

[32] C.-I. Chang, C.-C. Wu, C.-S. Lo, and M.-L. Chang, "Real-time simplex growing algorithms for hyperspectral endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 4, pp. 1834–1850, Apr. 2010.

[33] H. Ren and C.-I. Chang, "Automatic spectral target recognition in hyperspectral imagery," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1232–1249, Oct. 2003.

[34] T.-H. Chan, J.-Y. Liou, A. Ambikapathi, W.-K. Ma, and C.-Y. Chi, "Fast algorithms for robust hyperspectral endmember extraction based on worst-case simplex volume maximization," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2012, pp. 1237–1240.

[35] T.-H. Chan, W.-K. Ma, A. Ambikapathi, and C.-Y. Chi, "A simplex volume maximization framework for hyperspectral endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4177–4193, Nov. 2011.

[36] R. Marrero *et al.*, "A novel negative abundance-oriented hyperspectral unmixing algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3772–3790, Jul. 2015.

[37] A. Marinoni, A. Plaza, and P. Gamba, "A novel preunmixing framework for efficient detection of linear mixtures in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4325–4333, Aug. 2017.

[38] B. Du, Q. Wei, and R. Liu, "An improved quantum-behaved particle swarm optimization for endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 6003–6017, Aug. 2019.

[39] L. Tong, B. Du, R. Liu, and L. Zhang, "An improved multiobjective discrete particle swarm optimization for hyperspectral endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 7872–7882, Oct. 2019.

[40] M. Pavan and M. Pelillo, "Dominant sets and pairwise clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 1, pp. 167–172, Jan. 2007.

[41] M. Pavan, "A new graph-theoretic approach to clustering, with applications to computer vision," Ph.D. dissertation, Univ. Ca' Foscari, Venice, Italy, 2004.

[42] M. Pavan and M. Pelillo, "A new graph-theoretic approach to clustering and segmentation," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2003, p. 1

[43] J. W. Weibull, *Evolutionary Game Theory*, vol. 78, no. 3. Cambridge, MA, USA: MIT Press, 1996, pp. 850–853.

[44] L. E. Baum and J. A. Eagon, "An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology," *Bull. Amer. Math. Soc.*, vol. 73, no. 3, pp. 360–364, May 1967.

[45] J. C. Harsanyi and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 4, pp. 779–785, Jul. 1994.

[46] C.-I. Chang and D. C. Heinz, "Constrained subpixel target detection for remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1144–1159, May 2000.

[47] J. Li, J. M. Bioucas-Dias, A. Plaza, and L. Liu, "Robust collaborative nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6076–6090, Oct. 2016.

[48] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 765–777, Mar. 2007.

[49] F. Zhu, Y. Wang, S. Xiang, B. Fan, and C. Pan, "Structured sparse method for hyperspectral unmixing," *ISPRS J. Photogramm. Remote Sens.*, vol. 88, pp. 101–118, Feb. 2014.

[50] F. Zhu, Y. Wang, B. Fan, G. Meng, and C. Pan, "Effective spectral unmixing via robust representation and learning-based sparsity," 2014, *arXiv:1409.0685*. [Online]. Available: https://arxiv.org/abs/1409.0685

[51] F. Zhu, Y. Wang, B. Fan, G. Meng, S. Xiang, and C. Pan, "Spectral unmixing via data-guided sparsity," 2014, *arXiv:1403.3155*. [Online]. Available: https://arxiv.org/abs/1403.3155

Xuanwen Tao (Student Member, IEEE) received the B.S. degree in electronic information science and technology from Tianjin Chengjian University, Tianjin, China, in 2016. She is currently pursuing the M.Eng. degree in information and communication engineering with the China University of Petroleum (East China), Qingdao, China.

Her research interests include machine learning and hyperspectral data processing.

Tingwei Cui received the B.Eng., M.Eng., and Ph.D. degrees from the Ocean University of China, Qingdao, China, in 2000, 2003, and 2006, respectively.

He is currently a Professor with Sun Yat-sen University, China. Prior to that, he was a Professor with the First Institute of Oceanography, Ministry of Natural Resources of China, Beijing, China. His research interests are ocean optics and ocean color remote sensing.

Antonio Plaza (Fellow, IEEE) received the M.Sc. and Ph.D. degrees in computer engineering from the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, University of Extremadura, Cáceres, Spain, in 1999 and 2002, respectively.

He is currently the Head of the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, University of Extremadura. He is currently one of the top cited authors in Spain and in the University of Extremadura. He has guest edited ten special issues on hyperspectral remote sensing for different journals. His research interests include remotely sensed hyperspectral image analysis, signal processing, and efficient implementations of large-scale scientific problems on high-performance computing architectures, including commodity Beowulf clusters, heterogeneous networks of computers and clouds, and specialized computer architectures, such as field programmable gate arrays or graphical processing units.

Dr. Plaza is a fellow of the IEEE for the contributions to hyperspectral data processing and parallel computing of earth observation data. He was a recipient of a recognition as an Outstanding Associate Editor of IEEE ACCESS in 2017. He was a member of the Editorial Board of the IEEE GEOSCIENCE AND REMOTE SENSING NEWSLETTER from 2011 to 2012 and the *IEEE Geoscience and Remote Sensing Magazine* in 2013. He was also a pixel of the Steering Committee of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING. He is a reviewer of 500 manuscripts for over 50 different journals. He served as the Editor-in-Chief for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING from 2013 to 2017. He is an Associate Editor of IEEE ACCESS. He served as the Director of Education Activities for the IEEE Geoscience and Remote Sensing Society (GRSS) from 2011 to 2012 and the President of the Spanish Chapter of the IEEE GRSS from 2012 to 2016.

Peng Ren (Senior Member, IEEE) received the B.Eng. and M.Eng. degrees in electronic engineering from the Harbin Institute of Technology, Harbin, China, and the Ph.D. degree in computer science from the University of York, York, U.K.

He is currently a Professor with the College of Oceanography and Space Informatics, China University of Petroleum (East China), Qingdao, China. His research interests include remote sensing and machine learning.

Dr. Ren was a recipient of the K. M. Scott Prize from the University of York in 2011 and the Eduardo Caianiello Best Student Paper Award at the 18th International Conference on Image Analysis and Processing in 2015, as one coauthor.

Escuela Politecnica
Av. de la Universidad, S/N, 10003
Caceres, Spain
Phone: 0034927257000. Ext. 51662
Email: aplaza,juanmariohaut{@unex.es}

Dr. Antonio Plaza Miguel y Dr. Juan Mario Haut Hurtado como directores de la tesis titulada "Procesamiento Eficiente y Profundo de Imágenes Hiperespectrales y Aplicación en Tareas de Desmezclado Espectral", certifican el factor de impacto y la categorización de la siguiente publicación, incluída en la tesis doctoral. Del mismo modo, se especifica la aportación del doctorado.

Antonio Plaza Miguel PhD, and Juan Mario Haut Hurtado PhD, as directors of the Phd thesis titled "Efficient and Deep Processing of Hyperspectral Images and Application in Spectral Unmixing Tasks", certify the impact factor and the categorization of the following publication, included in the doctoral thesis. In the same way, the contribution of the doctorate is specified.

Artículo / Paper

Autores/Authors: **X. Tao**, M. E. Paoletti, J. M. Haut, P. Ren, J. Plaza, and A. Plaza.
Title: Endmember Estimation with Maximum Distance Analysis.
Journal: Remote Sensing, vol. 13, no. 4, p. 713, February 2021.
DOI: 10.3390/rs13040713.
Impact factor 2021: 5.349. Q1

Abstract: Endmember estimation plays a key role in hyperspectral image unmixing, often requiring an estimation of the number of endmembers and extracting endmembers. However, most of the existing extraction algorithms require prior knowledge regarding the number of endmembers, being a critical process during unmixing. To bridge this, a new maximum distance analysis (MDA) method is proposed that simultaneously estimates the number and spectral signatures of endmembers without any prior information on the experimental data containing pure pixel spectral signatures and no noise, being based on the assumption that endmembers form a simplex with the greatest volume over all pixel combinations. The simplex includes the farthest pixel point from the coordinate origin in the spectral space, which implies that: (1) the farthest pixel point from any other pixel point must be an endmember, (2) the farthest pixel point from any line must be an endmember, and (3) the farthest pixel point from any plane (or affine hull) must be an endmember. Under this scenario, the farthest pixel point from the coordinate origin is the first endmember, being used to create the aforementioned point, line, plane, and affine hull. The remaining endmembers are extracted by repetitively searching for the pixel points that satisfy the above three assumptions. In addition to behaving as an endmember estimation algorithm by itself, the MDA method can co-operate with existing endmember extraction techniques without the pure pixel assumption via generalizing them into more effective schemes. The conducted experiments validate the effectiveness and efficiency of our method on synthetic and real data.

Contribución del doctorado: Planteamiento de la hipótesis, desarrollo práctico, análisis y discusión de los resultados, elaboración y escritura del manuscrito.

Firma / Signature
Marzo / Marzo, 2023

Antonio Plaza Miguel   Juan Mario Haut Hurtado

Article

Endmember Estimation with Maximum Distance Analysis

Xuanwen Tao [1],* , **Mercedes E. Paoletti [1]** , **Juan M. Haut [2]** , **Peng Ren [3]** , **Javier Plaza [1]** and **Antonio Plaza [1]**

[1] Hyperspectral Computing Laboratory (HyperComp), Department of Computer Technology and Communications, Escuela Politecnica de Caceres, University of Extremadura, Avenida de la Universidad sn, E-10002 Caceres, Spain; mpaoletti@unex.es (M.E.P.); jplaza@unex.es (J.P.); aplaza@unex.es (A.P.)

[2] Department of Communication and Control Systems, Higher Technical School of Computer Engineering, National Distance Education University (UNED), E-28040 Madrid, Spain; juanmariohaut@unex.es

[3] College of Oceanography and Space Informatics, China University of Petroleum (East China), Qingdao 266580, China; pengren@upc.edu.cn

* Correspondence: xtaona@alumnos.unex.es; Tel.: +34-633-688-306

Abstract: Endmember estimation plays a key role in hyperspectral image unmixing, often requiring an estimation of the number of endmembers and extracting endmembers. However, most of the existing extraction algorithms require prior knowledge regarding the number of endmembers, being a critical process during unmixing. To bridge this, a new maximum distance analysis (MDA) method is proposed that simultaneously estimates the number and spectral signatures of endmembers without any prior information on the experimental data containing pure pixel spectral signatures and no noise, being based on the assumption that endmembers form a simplex with the greatest volume over all pixel combinations. The simplex includes the farthest pixel point from the coordinate origin in the spectral space, which implies that: (1) the farthest pixel point from any other pixel point must be an endmember, (2) the farthest pixel point from any line must be an endmember, and (3) the farthest pixel point from any plane (or affine hull) must be an endmember. Under this scenario, the farthest pixel point from the coordinate origin is the first endmember, being used to create the aforementioned point, line, plane, and affine hull. The remaining endmembers are extracted by repetitively searching for the pixel points that satisfy the above three assumptions. In addition to behaving as an endmember estimation algorithm by itself, the MDA method can co-operate with existing endmember extraction techniques without the pure pixel assumption via generalizing them into more effective schemes. The conducted experiments validate the effectiveness and efficiency of our method on synthetic and real data.

Keywords: hyperspectral image; spectral unmixing; maximum distance analysis; endmember extraction; endmember estimation

1. Introduction

Remotely sensed hyperspectral imaging (HSI) [1] provides a significant amount of information regarding different materials on the Earth's surface, capturing their reflectance behavior in the presence of solar radiation (which depends on their chemical composition and physical structure) by measuring the degree of absorption along the wavelengths of the electromagnetic spectrum (usually focused on the visible, near infrared (NIR), and shortwave infrared (SWIR) spectrum [2]), in hundreds of narrow and continuous spectral bands [3]. As a result, each pixel contains the representation of a spectral signature, which is unique for each observed material. However, one typical problem that is associated with HSI data is its low spatial resolution [4], which inevitably leads to mixed pixels. A mixed pixel covers several types of materials, which are usually composed of macroscopic objects, such as water, soil, vegetation, or buildings [5,6], and the corresponding spectrum being a mixture of several ground cover spectra.

A challenge of the mixed pixel phenomenon is how to accurately identify the prime materials that compose it, determining the original spectral signatures and measuring

the proportions in which they are present [7]. These principal constituent spectra are known as endmembers, and obtaining them is a hard and ill-posed problem, where the reflectance pattern is (in general) not known a priori. Additionally, the sensor inaccuracies and atmospheric/light conditions (which introduce an important amount of noise into the spectra), and the high spectral-dimensionality hinders the endmember extraction task [8], due to the high intra-class variability and inter-class similarity of the data [9]. Spectral unmixing is a significant technique for addressing this challenge [10–14]. In the available literature, spectral unmixing techniques [15–17] have been comprehensively investigated for the purpose of extracting endmembers and estimating their corresponding abundances, allowing for the processing of HSI scenes at the sub-pixel scale [18]. In fact, these techniques decompose each mixed pixel into a proportional composition of endmembers, where the constituent proportion with respect to different types of materials [19] for each pixel is defined as the abundance.

A significant challenge for existing spectral unmixing techniques is how to accurately extract endmembers [20–24] from remotely sensed HSI data. This is normally achieved by two seemingly independent, but, in fact, highly correlated procedures: (i) determining the number of endmembers, and (ii) extracting their spectral signatures. Because the accuracy of determining the number of endmembers has large influence on the endmember extraction step, it is natural to consider developing methods to essentially integrate both of the techniques seamlessly into an overall procedure of endmember estimation. This observation forms the motivation of our work.

2. Literature Review

In this paper, we propose the maximum distance analysis (MDA) method as a new unmixing framework that simultaneously counts and extracts endmembers without any prior knowledge. To introduce the proposed method in a detailed and comprehensive way, in Sections 2.1 and 2.2, we review the literature regarding counting and extracting endmembers, respectively. In Section 2.3, we briefly describe the major contributions of our work in terms of presenting a new endmember estimation framework that simultaneously counts the number of endmembers and extracts their spectral signatures.

2.1. Endmember Counting

Estimating the number of endmembers is regarded as the first step of the overall endmember estimation task. The number of endmembers is often unavailable in an arbitrary HSI scene. In this scenario, most of the existing endmember extraction techniques cannot properly extract endmembers if the number of endmembers is not accurately determined. The limitation regarding an unknown number of endmembers hinders most existing endmember extraction techniques from an operational viewpoint. To address this limitation, several algorithms have been developed for counting the number of endmembers. Particularly, information theory-based algorithms, eigenvalue thresholding algorithms, and geometry characterization algorithms are three main families of techniques for estimating the number of endmembers [25]. A brief summary of each family of algorithms is provided below.

- The first family includes some kind of criteria information, such as theoretic criteria that are based on minimum description length [26,27], Akaike's information criterion [28], and Bayesian information criterion [29]. Furthermore, different models have been developed for encoding a negative data log-likelihood term and a penalty term. An accurate estimation of the number of endmembers is expected to be obtained when the model achieves a global optimum. These strategies rely on the empirical configuration of specific mixed models or likelihood functions, and improper configurations will cause estimation errors of the number of endmembers.
- The second family is related to a thresholding scheme, in which a threshold is applied to the eigen-decomposition results from subspace analysis. Eigenvalue thresholding algorithms include principal component analysis (PCA) [30], hyperspectral signal

subspace by minimum error (Hysime) [31], and the so-called Harsanyi-Farrand-Chang (HFC) method [32], coupled with its noise-whitened implementation (NWHFC) [33]. The PCA-based approaches aim to characterize a cutoff gap between the eigenvalues that are caused by signals and noise. However, these approaches will provide an incorrect estimation of the number of endmembers if the variation between the two eigenvalues is negligible. The Hysime approach conducts spectrum noise characterization and noise covariance estimation, and it requires high computational complexity. The HFC approach requires a fixed false alarm probability, which affects the estimated number of endmembers.

- The third family includes the geometry-based estimation of the number of endmembers (GENE), which includes the convex hull (GENE-CH) algorithm and affine hull (GENE-AH) algorithm [34]. Both GENE-CH and the GENE-AH utilize data geometry and exploit the fact that all the observed pixel vectors should lie in the convex hull (CH) and affine hull (AH) of the endmember signatures, respectively. Specifically, GENE algorithms operate along with an endmember extraction algorithm (EEA). In this scenario, a maximum hull volume would stop the EEA from extracting the next endmember signature. Therefore, the GENE algorithms depend on the effectiveness of the EEA used, and different endmember extraction algorithms (EEAs) cause different accuracies for counting endmembers.

2.2. Endmember Extraction

The subsequent step of endmember estimation is to extract endmembers (assuming that the number of endmembers has been previously determined). Linear unmixing assumes that all of the pixels are a linear combination of the endmember signatures present in the scene, weighted by their respective fractional abundances. The related techniques include two families, i.e., statistical algorithms and geometrical algorithms. Statistical algorithms address endmember extraction as an inference problem that is formulated under the Bayesian framework, whereas geometrical algorithms exploit the fact that the spectral vectors (under the linear mixing model) lie in a simplex whose vertices correspond to the endmembers. In our work, we focus on the geometrical approach to endmember extraction. Geometrical algorithms try to find a simplex set whose vertices correspond to the endmembers. As such, we extract endmembers by identifying the vertices of the aforementioned simplex. The main ideas that have been presented in recent research under this framework belong to two different groups. Pure pixel-based algorithms [9] presume that the image scene contains at least one pure pixel per endmember. Even more recently, some algorithms [35] dropped this assumption by assuming that no pure pixel points may be present in real image scenes. In the following, we outline these two types of algorithms.

- If there is at least one pure pixel for per endmember, then EEAs search for the spectral vectors in the data set that corresponds to the vertices of the data simplex. Typical algorithms that are based on the pure pixel assumption are pixel purity index (PPI) [36], N-FINDR [37,38], and vertex component analysis (VCA) [39], among others. The PPI algorithm projects all of the observed pixels onto randomly generated unit-norm vectors, and it records the number of times (i.e., scores) that the value of each projected pixel has an extreme value (either minimum or maximum projection value). Then, the endmembers are those pixels with the highest scores. The N-FINDR algorithm and its derivatives search for a simplex with the greatest volume over all pixel combinations, and the vertices of the simplex correspond to the endmembers. The VCA algorithm iteratively projects data onto a direction orthogonal to the subspace that is spanned by the endmembers already determined, and the new endmember is the pixel with the extreme value of the projected data. It is worth noting that PPI, N-FINDR, and VCA require the number of endmembers in advances to perform endmember extraction.

- If the pure pixel assumption is not satisfied (this is a more realistic scenario, since HSI data are often dominated by highly mixed pixel [9]), the endmember extraction process is a rather intractable task. The difficulty is that the HSI data set may not

contain any endmembers or at least some of them. Some popular algorithms that are implemented under this assumption are the minimum volume constrained non-negative matrix factorization (MVCNMF) [40], minimum-volume enclosing simplex (MVES) [41], minimum volume simplex analysis (MVSA) [42,43], etc. MVCNMF adopts constrained nonnegative matrix factorization, together with a volume-based constraint, to decompose mixed pixels in multispectral and hyperspectral remote sensing images. MVES finds a simplex containing all of the dimensionally-reduced pixels by minimizing the simplex volume subject to the constraint. MVSA fits a minimum volume simplex to the hyperspectral data, constraining the abundance fractions to belong to the probability simplex. Specifically, MVCNMF, MVES, and MVSA also require the number of endmembers when they implement endmember extraction.

2.3. From the Literature to Our Contributions

As reviewed in Sections 2.1 and 2.2, most of the existing techniques for estimating the number of endmembers and extracting their signatures are performed as two independent procedures in the literature. Most existing EEAs require prior knowledge regarding the number of endmembers, but the algorithms for counting endmembers tend to be characterized by independent methods that may not seamlessly benefit the EEAs. To bridge these two seemingly independent (but, in fact, highly related) procedures, we establish a unifying endmember estimation framework. The contributions of our work are two-fold.

- The major contribution of our work is a novel endmember estimation method, referred to as maximum distance analysis (MDA), which accomplishes the overall endmember estimation task (that normally comprises of two independent steps, i.e., determining the number of endmembers and extracting their endmember signatures). Our newly proposed MDA method sequentially identifies pixel points as endmembers that are farthest from specific pixel point, line, plane, and affine hulls. The sequential endmember extraction terminates when the maximum distance between all the pixel points and the affine hull [41,44] formed by all extracted endmembers is zero. Our proposed MDA method does not require knowing the number of endmembers. It only involves straightforward vector-based computations, keeping its computational complexity very low. Moreover, our proposed MDA method does not use dimension reduction.
- A second major contribution of our work is that the proposed MDA method provides an effective way for generalizing most existing EEAs that require prior knowledge of the number of endmembers into more accurate ones without the requirement. We use the MVSA as a special case to demonstrate the generalization strategy that is based on the MDA. Specifically, MVSA requires the number of endmembers and uses the traditional VCA method as an initialization method for extracting endmembers. Therefore, our proposed MDA framework is not only capable of simultaneously counting and extracting endmembers by itself without any prior knowledge, but it also provides a general framework for developing new endmember extraction schemes that are based on arbitrary existing endmember extraction methods without the pure pixel assumption.

3. Maximum Distance Analysis

3.1. Extracting Endmember Signatures

Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_i, \cdots, \mathbf{x}_N] \in \mathbb{R}^{L \times N}$ denote a HSI dataset with L bands and N pixels, where the ith pixel $\mathbf{x}_i = [x_{1i}, x_{2i}, \cdots, x_{Li}]^{\mathrm{T}} \in \mathbb{R}^{L \times 1}$ represents an L-dimensional vector. From a geometric point of view, we know that endmembers are supposed to form a simplex with the greatest volume over all pixel combinations, and they are distributed at the outermost region of all pixels to enclose them. Therefore, if one pixel point is a candidate to be an endmember, then it must satisfy two conditions: (1) the pixel point is an essential component of the simplex with the greatest volume over all pixel combinations and (2) the pixel point is distributed in the outermost region of all pixels. Some of the EEAs utilize the idea of a simplex to search for some pixel points far away from each other as

endmembers, and the operations inevitably render considerable computational overloads and make EEAs inefficient for tackling large-size remote sensing HSI scenes. To address this problem, we give some definitions regarding distances to effectively and efficiently extract endmembers, where the task of extracting endmembers does not require any prior knowledge regarding the number of endmembers.

Definition 1. *We define the farthest pixel point from the coordinate origin in the pixel space as the first endmember.*

The Euclidean distance computes the distance d_{ij} between the ith and jth hyperspectral pixel points in the pixel space

$$
\begin{aligned}
d_{ij} &= \parallel \mathbf{x}_i - \mathbf{x}_j \parallel_2 \\
&= \sqrt{(x_{1i} - x_{1j})^2 + \cdots + (x_{Li} - x_{Lj})^2}
\end{aligned}
\tag{1}
$$

We compute the distance between each pixel point and the coordinate origin by using Equation (1) to obtain the farthest pixel point from the coordinate origin, computing as a result of the first endmember. In this sense, this first point easily satisfies the two conditions to be an endmember. Specifically, the distance calculation in Equation (1) may need to be adapted if the hyperspectral data are nonlinearly mixed or includes noise.

Definition 2. *We define the farthest pixel point from the first extracted endmember as the second endmember.*

We compute the distance between each pixel point and the first extracted endmember through Equation (1) to obtain the farthest pixel point from the first endmember, getting, as a result, the second endmember. We observe that the second extracted endmember also satisfies the two conditions to be considered as endmember.

Definition 3. *We define the farthest pixel point from the line that is defined by the first and second extracted endmembers as the third endmember.*

Assuming that we have acquired the first endmember A and the second endmember B, as shown in Figure 1, we compute the distance $d_{i\_L_{\{AB\}}}$ between any pixel point i and the line $L_{\{AB\}}$ to obtain the third endmember. Subsequently, the angle α between vectors $\overrightarrow{Bi}$ and $\overrightarrow{BA}$ is trigonometrically computed by Equation (2)

$$
\alpha = \arccos \frac{\overrightarrow{Bi} \cdot \overrightarrow{BA}}{\parallel \overrightarrow{Bi} \parallel_2 \parallel \overrightarrow{BA} \parallel_2}
\tag{2}
$$

Figure 1. Schematic diagram illustrating how to extract an endmember (from two previously available ones) by using our maximum distance analysis (MDA) method. The farthest pixel point from the line $L_{\{AB\}}$ is the third endmember. The red dots represent endmembers that have been previously extracted, and the blue dots denote the pixel points in the pixel space.

As shown in Figure 1, the distance $d_{i\_L_{\{AB\}}}$ between any pixel point i and the line $L_{\{AB\}}$ is computed by

$$d_{i\_L_{\{AB\}}} = || \overrightarrow{Bi} ||_2 \sin \alpha \tag{3}$$

In this sense, we obtain the index e_3 of the third endmember by maximizing the distance with

$$e_3 = \underset{i \in \{1,2,\cdots,N\}}{\arg\max} \ d_{i\_L_{\{AB\}}} \tag{4}$$

and we use the index e_3 to extract the third endmember C. Moreover, the third extracted endmember also satisfies the two conditions in order to be considered as an endmember.

Definition 4. *We define the farthest pixel point from the plane that is formed by the first, second, and third extracted endmembers as the fourth endmember.*

Assuming that the three extracted endmembers A, B, and C form a plane $P_{\{ABC\}}$, as shown in Figure 2, we compute the distance $d_{i\_P_{\{ABC\}}}$ between any pixel point i and the plane $P_{\{ABC\}}$ to obtain the fourth endmember. O is the coordinate origin in the pixel space and $\mathbf{n}$ is the normal vector of the plane $P_{\{ABC\}}$. However, in our work, we get the normal vector $\mathbf{n}$ by a unique strategy. First, we get $\mathbf{n}$ by

$$\begin{aligned}
\overrightarrow{OA} \cdot \mathbf{n} &= 0; \\
\overrightarrow{OB} \cdot \mathbf{n} &= 0; \\
\overrightarrow{OC} \cdot \mathbf{n} &= 0;
\end{aligned} \tag{5}$$

In addition, we know

$$\begin{aligned}
\overrightarrow{AB} &= \overrightarrow{OB} - \overrightarrow{OA}; \\
\overrightarrow{BC} &= \overrightarrow{OC} - \overrightarrow{OB}; \\
\overrightarrow{AC} &= \overrightarrow{OC} - \overrightarrow{OA};
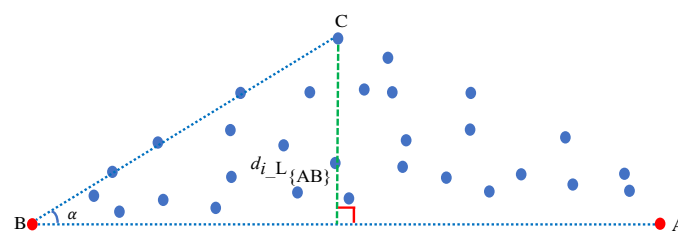\end{aligned} \tag{6}$$
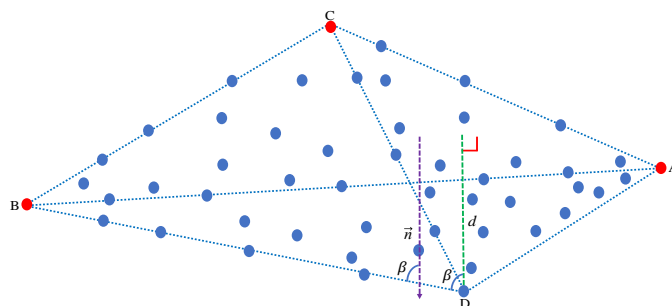
Figure 2. Schematic diagram illustrating how to extract an endmember (from three previously available ones) by using our maximum distance analysis (MDA) method. The farthest pixel point from the plane $P_{\{ABC\}}$ is the fourth endmember.

Based on Equations (5) and (6), we have the expressions

$$\begin{aligned}
(\overrightarrow{OB} - \overrightarrow{OA}) \cdot \mathbf{n} &= \overrightarrow{AB} \cdot \mathbf{n} = 0; \\
(\overrightarrow{OC} - \overrightarrow{OB}) \cdot \mathbf{n} &= \overrightarrow{BC} \cdot \mathbf{n} = 0; \\
(\overrightarrow{OC} - \overrightarrow{OA}) \cdot \mathbf{n} &= \overrightarrow{AC} \cdot \mathbf{n} = 0;
\end{aligned}$$

We know that A, B, and C are three different endmembers, which mean that $\overrightarrow{AB}$, $\overrightarrow{BC}$, $\overrightarrow{AC}$ are not collinear with each other. Therefore, **n** is not only the normal vector of the plane $P_{\{ABC\}}$, but also orthogonal to $\overrightarrow{OA}$, $\overrightarrow{OB}$, $\overrightarrow{OC}$. Therefore, we define matrix $\mathbf{X}_3 = [\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC}] \in \mathbb{R}^{N \times 3}$, where N is the number of pixels. We obtain the normal vector **n** by addressing

$$\mathbf{X}_3^T \mathbf{n} = \mathbf{0} \tag{7}$$

The angle β between vectors $\overrightarrow{Bi}$ and **n** is trigonometrically computed by

$$\beta = \arccos \frac{\overrightarrow{Bi} \cdot \mathbf{n}}{|| \overrightarrow{Bi} ||_2 ||\mathbf{n}||_2} \tag{8}$$

Therefore, the distance $d_{i\_P_{\{ABC\}}}$ is computed by

$$d_{i\_P_{\{ABC\}}} = || \overrightarrow{Bi} ||_2 |\cos \beta| \tag{9}$$

where symbol $| \cdot |$ denotes the absolute value of $\cos \beta$. With this, we obtain the index e_4 of the fourth endmember by

$$e_4 = \underset{i \in \{1,2,\cdots,N\}}{\arg\max} \ d_{i\_P_{\{ABC\}}} \tag{10}$$

Finally, the index e_4 is used to extract the fourth endmember D, which also satisfies the two conditions to be considered as an endmember. Figure 3 graphically shows the process of extracting four endmembers.
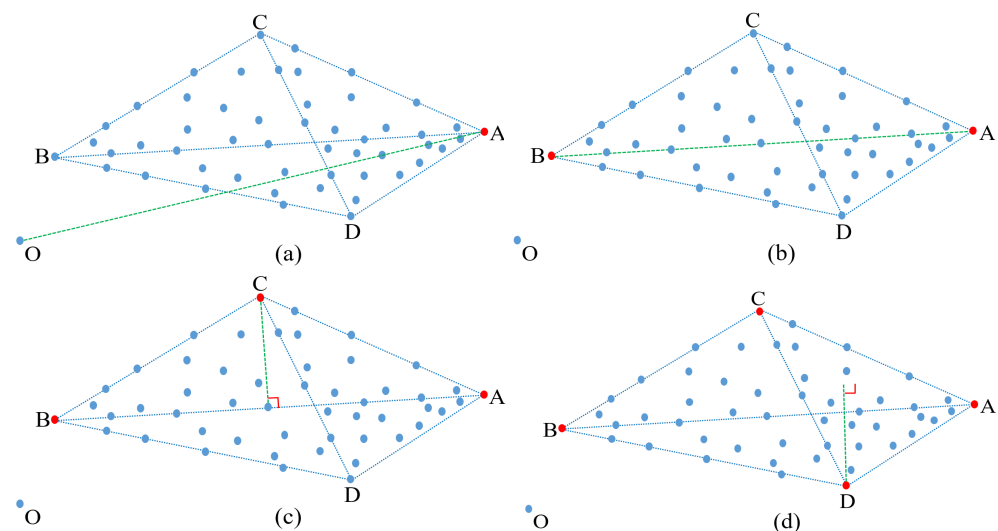
Figure 3. (**a**–**d**) Step-by-step process illustrating how to extract four endmembers by using our maximum distance analysis (MDA) method.

Definition 5. *We define the farthest pixel point from the affine hull* Aff$\{$ABCD$\cdots\}$ *formed by all of the extracted endmembers* $\{A, B, C, D, \cdots\}$ *as the next endmember.*

Assuming that we have extracted endmembers $\{$ABCD$\cdots\}$, and these endmembers form a affine hull Aff$\{$ABCD$\cdots\}$. We have the matrix $\mathbf{X}_{\{ABCD\cdots\}} = \{\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC}, \overrightarrow{OD}, \cdots\} \in \mathbb{R}^{N \times c}$, where N is the number of pixels and c is the number of endmembers that have been extracted. Assuming that **w** is the normal vector of the affine hull

$\text{Aff}\{\text{ABCD}\cdots\}$ and b is the offset or distance between the coordinate origin and affine hull, we compute the normal vector $\mathbf{w}$ by

$$\mathbf{X}^{\text{T}}_{\{\text{ABCD}\cdots\}}\mathbf{w} = \mathbf{0} \tag{11}$$

$\mathbf{x}$ is a pixel point of the affine hull $\text{Aff}\{\text{ABCD}\cdots\}$, and the affine hull is written as Equation (12)

$$\mathbf{w}^{\text{T}}\mathbf{x} + b = 0 \tag{12}$$

In this context, there is a pixel point $\mathbf{x}_i(\vec{x}_i)$, whose distance between it and the affine hull is denoted as d. Additionally, we consider point $\mathbf{x}_j(\vec{x}_j)$ as the projection of $\mathbf{x}_i$ onto the affine hull, where γ is the angle between vectors $\mathbf{w}$ and $\overrightarrow{x_i x_j}$, $\cos\gamma = 1$, and we have

$$\begin{aligned}
\mathbf{w}^{\text{T}} \overrightarrow{x_j x_i} &= w^1(x_i^1 - x_j^1) + w^2(x_i^2 - x_j^2) + \cdots + w^N(x_i^N - x_j^N) \\
&= w^1 x_i^1 + w^2 x_i^2 + \cdots + w^N x_i^N \\
&\quad - (w^1 x_j^1 + w^2 x_j^2 + \cdots + w^N x_j^N)
\end{aligned} \tag{13}$$

and

$$\begin{aligned}
\mathbf{w}\cdot \overrightarrow{x_i x_j} &= ||\mathbf{w}||_2 || \overrightarrow{x_i x_j} ||_2 \cos\gamma \\
&= ||\mathbf{w}||_2 || \overrightarrow{x_i x_j} ||_2 \\
&= ||\mathbf{w}||_2 d
\end{aligned} \tag{14}$$

and also

$$|\mathbf{w}\cdot \overrightarrow{x_i x_j} | = ||\mathbf{w}||_2 d \tag{15}$$

We know that $\mathbf{w}^{\text{T}}\vec{x}_j = -b$ from Equations (12) and (13) is represented as

$$\mathbf{w}^{\text{T}} \overrightarrow{x_j x_i}= w^1 x_i^1 + w^2 x_i^2 + \cdots + w^N x_i^N - (-b) \tag{16}$$

Based on Equations (15) and (16), we have

$$\begin{aligned}
||\mathbf{w}||_2 d &= |w^1 x_i^1 + w^2 x_i^2 + \cdots + w^N x_i^N + b| \\
&= |\mathbf{w}^{\text{T}}\vec{x}_i + b|
\end{aligned} \tag{17}$$

Therefore, the distance between any pixel point $\vec{x}_i$ and the considered affine hull is computed by

$$d = \frac{|\mathbf{w}^{\text{T}}\vec{x}_i + b|}{||\mathbf{w}||_2} \tag{18}$$

Specifically, support vector machines (SVMs) [45] provide the distance calculation that is given by Equation (18) between any pixel point $\vec{x}_i$ and the affine hull. We obtain the index of the next endmember by

$$\hat{e} = \underset{i\in\{1,2,\cdots,N\}}{\arg\max} \frac{|\mathbf{w}^{\text{T}}\vec{x}_i + b|}{||\mathbf{w}||_2} \tag{19}$$

We then use the index $\hat{e}$ to extract the next endmember, which also satisfies the aforementioned two conditions that any pixel point must meet in order to be considered an endmember.

3.2. Estimating the number of endmembers

From the mathematical discussions in the previous subsection, we know that: (1) if one pixel point belongs to one affine hull, the distance between it and the affine hull is zero, and (2) if one point does not belong to one affine hull, the distance between it and the affine hull is greater than zero. Based on the aforementioned concepts, we give the

following definition as the stopping criterion of endmember extraction, also obtaining the final number of endmembers by counting those pixel points that have been extracted as endmembers.

Definition 6. *The stopping criterion of endmember extraction is that the maximum distance between all of the pixel points and affine hull formed by the extracted endmember is zero. Simultaneously, the number of endmembers is determined by counting the number of pixel points that have been extracted.*

To conclude this section, we emphasize that our MDA method extracts endmembers without any previous knowledge regarding the number of endmembers, since it performs both endmember extraction and the estimation of the number of endmembers. It is worth noting that the maximum distance of the stopping criterion in Definition 6 is zero when the experimental data are linearly mixed and do not include noise; otherwise, the maximum distance is greater than zero. In addition, there is a clear difference between the distance between the non-endmember and endmember to the affine hull that formed by all extracted endmembers when the experimental data are non-linearly mixed and include noise. Therefore, we can observe the difference to obtain the maximum distance when the experimental data are non-linearly mixed and include noise, and then determine the number of endmembers.

4. Experiments

We use both synthetic and real hyperspectral data in order to evaluate the performance of our proposed MDA method in the tasks of estimating the number of endmembers and extracting their signatures. We perform the comparison tests using Matlab 2018b on an i5-8300 CPU (2.3 G). We conduct a series of experiments to demonstrate the effectiveness and efficiency of our newly proposed MDA method in counting and extracting endmembers. The endmember extraction capacity of our method is assessed with three endmember extraction algorithms, i.e., minimum-volume enclosing simplex (MVES) [41], vertex component analysis (VCA) [39], and minimum volume constrained nonnegative matrix factorization (MVCNMF) [40]. Additionally, the accuracy of our method in the task of estimating the number of endmembers is assessed with two endmember counting algorithms, i.e., hyperspectral signal subspace by minimum error (Hysime) [31] and Harsanyi–Farrand–Chang (HFC) [32]. We compare the estimated endmembers and abundances with the ground-truth ones that were used in our experiments by the following formulas

$$\phi_M = \frac{\left\| M - \hat{M} \right\|_F}{\|M\|} \tag{20}$$

and

$$\phi_A = \frac{\left\| A - \hat{A} \right\|_F}{\|A\|} \tag{21}$$

where $\|\cdot\|_F$ stands for the Frobenius norm, $\|\cdot\|$ is Euclidean norm, $\hat{M}$ and $\hat{A}$ are the estimated endmember signatures and abundances, and M and A denote the actual endmember signatures and abundances. The reconstruction error is another metric that is considered in our experiments, which is computed as

$$\begin{aligned} \phi_X &= \frac{\left\| X - \hat{X} \right\|_F}{\|X\|} \\ &= \frac{\left\| X - \hat{M}\hat{A} \right\|_F}{\|X\|} \end{aligned} \tag{22}$$

where $\hat{X}$ and X are the estimated and actual spectral signatures, respectively. Additionally, we also use the spectral angle distance (SAD) to evaluate our method by

$$SAD = \sqrt{\frac{1}{c}\sum_{i=1}^{c}\left[\arccos\left(\frac{\hat{\mathbf{m}}_i^{\mathrm{T}}\mathbf{m}_i}{\|\hat{\mathbf{m}}_i\|\|\mathbf{m}_i\|}\right)\right]^2} \tag{23}$$

where c is the number of endmembers, and $\hat{\mathbf{m}}_i$ and $\mathbf{m}_i$ are the ith estimated and actual spectral signatures, respectively.

EEAs get the most accurate results when the values of the above four metrics $\phi_{\mathrm{M}}, \phi_{\mathrm{A}}, \phi_{\mathrm{X}}$ and SAD are zero.

4.1. Synthetic Data

The synthetic data are created based on the United States Geological Survey (USGS) library. The mineral signatures in USGS library both have 224 spectral bands. We generate synthetic data that are based on the linear mixture method [39,40]. Specifically, to improve the unmixing performance, we have removed the low SNR bands as well as the water vapor absorption bands (including bands 1-2, 104-113, 148-167, and 221-224) from the original 224-band data cube, and keep 188 spectral bands in our experiments. We consider four families of synthetic data to evaluate the effectiveness and efficiency of our proposed MDA method in counting and extracting endmembers: (1) synthetic data containing pure mineral signatures and no noise under different number of endmembers; (2) synthetic data containing pure mineral signatures and 2000 pixel points and five endmembers under different noise levles; (3) synthetic data containing nonpure mineral signatures and 2000 pixel points and five endmembers under different noise levels; and, (4) synthetic data, including nonpure spectral signatures with different number of endmembers and different sizes under specific noise (70 dB). We utilize Definitions 1 to 6 to extract the signatures and estimate the number of endmembers. For simplicity, we summarize the steps that were conducted for extracting and counting endmembers with our MDA method:

1. We calculate the distance between every pixel point and coordinate origin in the pixel space and extract the pixel point with the maximum distance as the first endmember.
2. We calculate the distance between every pixel point and the first extracted endmember in the pixel space and extract the pixel point with the maximum distance as the second endmember.
3. We calculate the distance between every pixel point and the line defined by the first and second extracted endmembers, extracting the pixel point with the maximum distance as the third endmember.
4. We calculate the distance between every pixel point and the plane formed by the first, second, and third extracted endmembers, extracting the pixel point with the maximum distance as the fourth endmember.
5. We calculate the distance between every pixel point and the affine hull formed by all the previously extracted endmembers, extracting the pixel point with the maximum distance as the next endmember.
6. We finish the overall endmember extraction procedure when the maximum distance between all pixel points and the affine hull formed by all the previously extracted endmembers is zero. Simultaneously, we determine the number of endmembers by counting the number of pixel points that have been extracted. In the following, we describe the conducted experiments.

4.1.1. Straightforward MDA Algorithm

This experiment uses the first family type of synthetic data to evaluate the proposed MDA algorithm in counting and extracting endmembers. The synthetic data are created by the procedure that was described in [40]. The synthetic data, with 3364 pixel points, have pure spectral signatures and no noise. We use the metrics ϕ_M and SAD to evaluate the endmember extraction performance of our proposed MDA method on synthetic data

containing pure pixels and no noise under different number of endmembers. In addition, due to the fact that MDA can complete the two tasks of extracting endmembers and determining the number of endmembers, we employ VCA as one comparison method of extracting endmember, and Hysime and HFC as the two comparison methods of counting endmembers to evaluate the performance of our proposed MDA method. Table 1 illustrates the performance and efficiency of different methods in extracting their spectral signatures and determining the number of endmembers. Firstly, in terms of extracting endmembers, we observe ϕ_M and SAD from Table 1, and find that both VCA and our proposed MDA method acquire accurate results and they have high efficiency under different number of endmembers. Specifically, our method is slightly faster than VCA. Secondly, in terms of determining the number of endmembers, we find that both HFC and our proposed MDA method obtain accurate results, and Hysime has an error when the number of endmembers 20. In addition, we also observe that Hysime and HFC are faster than our proposed MDA method. However, VCA can only extract endmembers, Hysime and HFC can only determine the number of endmembers, but our proposed MDA method can simultaneously complete the two tasks (extracting endmembers and determining the number of endmembers). Specifically, due to the fact that MDA is faster than VCA, our proposed MDA method extracts endmembers more accurately and quickly as compared to Hysime (or HFC) +VCA. Therefore, our proposed MDA method is an effective and effcienet method in completing the overall endmember estimation task, i.e., determining the number of endmembers and extracting endmembers.

Table 1. Performance and processing time (seconds) for our proposed MDA algorithm on synthetic data containing pure mineral signatures and no noise under different number of endmembers. c and $\hat{c}$ are the actual and estimated number of endmembers, respectively. d denotes the maximum distance between all pixel points and the affine hull formed by all of the extracted endmembers.

	VCA			MDA					Hysime		HFC	
c	ϕ_M	SAD	Time	ϕ_M	SAD	Time	$\hat{c} + d$		$\hat{c}$	Time	$\hat{c}$	Time
5	0	0	0.252	0	0	0.204	5 + 0.1066	6 + 0	5	0.161	5	0.089
10	0	0	0.246	0	0	0.216	10 + 0.0844	11 + 0	10	0.142	10	0.063
15	0	0	0.245	0	0	0.233	15 + 0.0015	16 + 0	15	0.143	15	0.064
20	0	0	0.259	0	0	0.251	20 + 0.0020	21 + 0	19	0.157	20	0.065

4.1.2. MDA for Improving MVSA

As reviewed in Section 2.2, the original MVSA algorithm requires the number of endmembers before performing endmember extraction. Additionally, the original MVSA algorithm needs to utilize the VCA algorithm as an initialization method to extract endmembers. To this end, we embed our method into MVSA, which possibly extracts endmembers from nonpure pixels. Such composition (e.g., combining the MDA and MVSA) forms a new, overall unmixing framework that simultaneously and efficiently determines the number of endmembers and extracts their spectral signatures without any prior knowledge. We refer to this strategy as MDA-MVSA.

We conduct our experiments on synthetic data containing pure and nonpure spectral signatures with different noise levels, respectively. The two types of synthetic data both have 2000 pixel points and five endmembers. The two types of synthetic data are constructed according to the linear model that is given using the procedure described in [39]. Zero-mean white Gaussian noise has been added to the two types of synthetic data, and it is defined by

$$SNR = 10 \log_{10}(\mathbb{E}\|\mathbf{X}\|_F^2 / \mathbb{E}\|\mathbf{N}\|_F^2) \tag{24}$$

where $\mathbf{X}$ and $\mathbf{N}$ are the matrices of pixel points and noise, respectively. We continue to employ the metrics ϕ_M, ϕ_A, ϕ_X, and SAD to evaluate the performance of our proposed

MDA-MVSA method in unmixing. Tables 2 and 3 show the results that were obtained by different methods on synthetic data containing pure and non-pure spectral signatures under different noise levels. Specifically, VCA and sparse unmixing by variable splitting and augmented Lagrangian (Sunsal) [46] are the most typical endmember extraction method and abundance estimation method, respectively. Therefore, we combine VCA and Sunsal as VCA-Sunsal in order to evaluate the performance of different methods. From Tables 2 and 3, we see that: (1) VCA-Sunsal is the fastest method when compared with other methods, but the time gap between it and our proposed MDA-MVSA method is very small. In addition, VCA-Sunsal requires the number of endmembers in endmember extraction compared to our proposed MDA-MVSA method. In this scenario, although VCA-Sunsal is faster than our proposed MDA-MVSA method, the processing of time regarding our proposed MDA-MVSA includes the processing time of determining the number of endmembers and VCA-Sunsal does not include; (2) our proposed MDA-MVSA method obtains better results than VCA-Sunsal, MVCNMF, and MVES in terms of the metrics ϕ_M, ϕ_A, and ϕ_X; and, (3) our proposed MDA-MVSA method and original MVSA acquire almost same results in terms of the metrics ϕ_M, ϕ_A, and ϕ_X, but our proposed MDA-MVSA method completes endmember extraction without requiring the number of endmembers. In addition, the original MVSA uses VCA as initialization to obtain the final results of endmember extraction, and VCA may provide slightly different results in different runs due to its initialization. In this scenario, our proposed MDA-MVSA method is more stable than the original MVSA, and it extracts endmembers without requiring the number of endmembers. Therefore, our proposed MDA-MVSA method improves the performance of original MVSA method, and it is very effective and efficient in extracting endmembers. Figure 4 shows the results of SAD for different methods at various noise levels on the two considered synthetic data scenes. From Figure 4, we see that that our proposed MDA-MVSA method obtains more better results in terms of SAD compared to VCA-Sunsal, MVCNMF, and MVES. We also observe that our proposed MDA-MVSA method and original MVSA acquire the same results. Similarly, due to the fact that original MVSA uses VCA as initialization to complete endmember extraction and requires the number of endmembers when it extracts endmembers, our proposed MDA-MVSA method still improves theoriginal MVSA.

Table 2. Performance and processing time (seconds) for different endmember extraction algorithms on synthetic data with 2000 pixel points and five endmembers and containing pure mineral signatures for the United States Geological Survey (USGS) library under different noise levels.

	VCA-Sunsal				MVCNMF				MVES				MVSA				MDA-MVSA			
dB	ϕ_M	ϕ_A	ϕ_X	Time	ϕ_M	ϕ_A	ϕ_X	Time	ϕ_M	ϕ_A	ϕ_X	Time	ϕ_M	ϕ_A	ϕ_X	Time	ϕ_M	ϕ_A	ϕ_X	Time
30	0.091	0.281	0.012	0.338	0.053	0.171	0.004	10.362	0.051	0.157	0	2.409	0.054	0.127	0	0.530	0.047	0.122	0	0.588
50	0.095	0.322	0.013	0.346	0.056	0.195	0.005	9.443	0.009	0.024	0	2.562	0.005	0.016	0	0.684	0.005	0.016	0	0.582
70	0.188	0.472	0.047	0.350	0.066	0.191	0.005	9.439	0.019	0.038	0	3.385	0.002	0.006	0	0.561	0.002	0.006	0	0.571
90	0.131	0.307	0.014	0.352	0.069	0.178	0.004	10.169	0.009	0.017	0	3.865	0.002	0.005	0	0.562	0.002	0.005	0	0.570
∞	0.186	0.555	0.040	0.346	0.055	0.164	0.003	10.448	0.004	0.009	0	3.714	0.002	0.004	0	0.554	0.002	0.004	0	0.569

Table 3. Performance and processing time (seconds) for different endmember extraction algorithms on synthetic data with 2000 pixel points and five endmembers and containing nonpure mineral signatures (maximum purity of 0.8) for the USGS library under different noise levels.

|	VCA-Sunsal				MVCNMF				MVES				MVSA				MDA-MVSA			
dB	ϕ_M	ϕ_A	ϕ_X	Time	ϕ_M	ϕ_A	ϕ_X	Time	ϕ_M	ϕ_A	ϕ_X	Time	ϕ_M	ϕ_A	ϕ_X	Time	ϕ_M	ϕ_A	ϕ_X	Time
30	0.164	0.496	0.035	0.345	0.057	0.187	0.005	10.200	0.059	0.153	0	2.427	0.056	0.142	0	0.662	0.051	0.140	0	0.696
50	0.206	0.489	0.045	0.355	0.192	0.305	0.004	10.370	0.007	0.022	0	2.528	0.003	0.011	0	0.671	0.003	0.011	0	0.563
70	0.285	0.497	0.111	0.362	0.078	0.190	0.005	9.853	0.010	0.023	0	3.006	0.003	0.005	0	0.568	0.003	0.005	0	0.567
90	0.134	0.351	0.018	0.364	0.084	0.215	0.005	10.306	0.011	0.025	0	3.461	0.003	0.005	0	0.568	0.003	0.005	0	0.571
∞	0.109	0.409	0.017	0.335	0.057	0.180	0.002	10.299	0.002	0.009	0	3.555	0.002	0.005	0	0.551	0.002	0.005	0	0.678
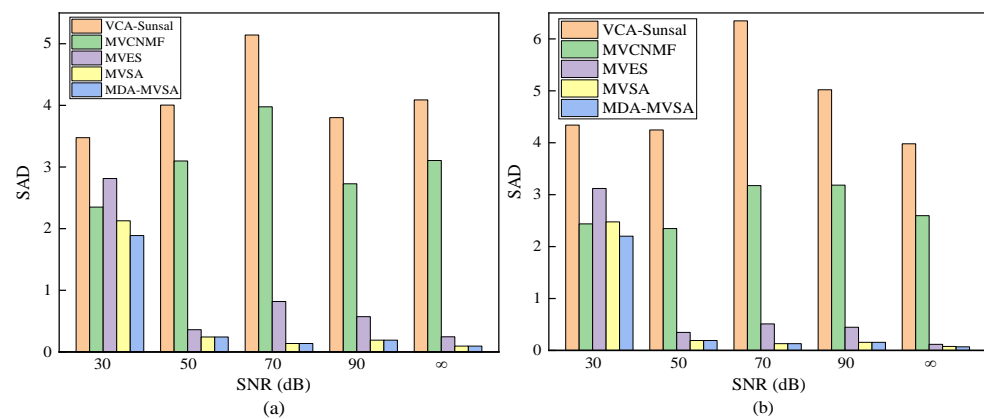
Figure 4. Spectral angle distance (SAD) performance for different endmember extraction algorithms on (**a**) synthetic data with pure pixels under different noise levels and (**b**) synthetic data with nonpure pixels under different noise levels. The two types of synthetic data both have 3364 pixel points and five endmembers under different noise levels.

Table 4 reports the performance of our proposed MDA-MVSA method in counting the endmembers on synthetic data containing pure and non-pure spectral signatures under different noise levels. Specifically, the two types of synthetic data both have 2000 pixel points and five endmembers. From Table 4, we see that the gap between the maximum distance is very obvious when the estimated number of endmembers exceeds the actual number of endmembers. Therefore, our proposed MDA-MVSA method can accurately determine the number of endmembers. Table 5 shows the results (estimated number of endmembers + processing time) of different methods in the task of counting endmembers on the above two types of synthetic data, i.e., containing pure and non-pure spectral signatures under different noise levels. From Table 5, we see that Hysime and our proposed MDA-MVSA method obtain accurate results on synthetic data containing pure and non-pure spectral signatures with different noise levels, and HFC has errors when the actual number of endmembers is fibe on synthetic data containing pure and non-pure spectral signatures with different noise levels. Additionally, HFC is the most effective method in counting endmembers, but the time gap between it and our proposed MDA-MVSA method is very small. In addition, Hysime and HFC can only determine the number of endmembers and our proposed MDA-MVSA method can complete two tasks, i.e., determining the number of endmembers and extracting endmembers. In this scenario, the processing time of Hysime and HFC is the running time of determining the number of endmembers, while the processing time of our proposed MDA-MVSA method is the running time of determining the number of endmembers and extracting endmembers. Therefore, our proposed MDA-MVSA method is also very effective and efficient in determining the number of endmembers.

Table 4. The relationship between the number of endmembers and the maximum distance on synthetic data containing pure and nonpure mineral signatures under different noise levels. The two types of synthetic data both have 2000 pixel points and five endmembers.

dB	The Estimated Number of Endmembers + The Maximum Distance			
	SD with Pure Pixel Points		SD with Nonpure Pixel Points	
30	5 + 0.094	6 + 0.068	5 + 0.086	6 + 0.06
50	5 + 0.048	6 + 0.008	5 + 0.021	6 + 0.007
70	5 + 0.029	6 + 0.001	5 + 0.02	6 + 0.001
90	5 + 0.014	6 + 0	5 + 0.017	6 + 0
∞	5 + 0.014	6 + 0	5 + 0.007	6 + 0

Table 5. The results of estimated number of endmembers + the processing time by different methods on synthetic data containing pure and nonpure mineral signatures under different noise levels. The two types of synthetic data both have 2000 pixel points and five endmembers.

dB	Synthetic Data with Pure Pixel Points			Synthetic Data with Nonpure Pixel Points		
	Hysime	**HFC**	**MDA-MVSA**	**Hysime**	**HFC**	**MDA-MVSA**
30	5 + 0.13	188 + 0.068	5 + 0.588	5 + 0.133	188 + 0.067	5 + 0.696
50	5 + 0.128	5 + 0.067	5 + 0.582	5 + 0.127	5 + 0.067	5 + 0.563
70	5 + 0.13	5 + 0.067	5 + 0.571	5 + 0.126	5 + 0.067	5 + 0.567
90	5 + 0.132	5 + 0.066	5 + 0.57	5 + 0.126	5 + 0.067	5 + 0.571
∞	5 + 0.126	5 + 0.068	5 + 0.569	5 + 0.127	5 + 0.066	5 + 0.678

To take the experimental validation one step further, we conduct some experiments on synthetic data containing non-pure spectral signatures with different numbers of endmembers and sizes under a specific noise level (70 dB) to evaluate the performance of our proposed MDA-MVSA method in extracting and counting endmembers. Table 6 reports the performance and processing time for problems with N = 4000, 8000, and 12,000 pixel points while using different numbers of endmembers (c = 5, 10, 15 and 20) on synthetic data with non-pure spectral signatures. From Table 6, we see that our proposed MDA-MVSA method has good results in terms of the values of ϕ_M, ϕ_A, ϕ_X, SAD, and time over MVES. Specifically, this is a very difficult problem when the number of endmembers is 20. For instance, from Table 6, we see that our method takes about 12 s to perform the computation and in comparison MVES takes around 3.7 h on synthetic data with 12,000 pixel points and 20 endmembers. This fact reflects that our proposed MDA-MVSA method is very efficient for large-scale problems. From Table 6, we also observe that the gap of the maximum distance d is very obvious when the estimated number of endmembers exceeds the actual number of endmembers. Therefore, as compared to MVES, our algorithm is not only a high-speed algorithm, but it also does not require the number of endmembers when it implements endmember extraction.

Table 6. Performance and processing time (seconds) obtained with our method and minimum-volume enclosing simplex (MVES) on synthetic data with N = 4000, 8000, 12,000, SNR = 70 dB, c = 5, 10, 15, and 20, and maximum purity of 0.8. $\hat{c}$ is the estimated number of endmembers, and d is the maximum distance between all pixel points and the affine hull formed by all extracted endmembers.

N	c	MVES					MDA-MVSA					$\hat{c}$ + d	
		ϕ_M	ϕ_A	ϕ_X	SAD	Time	ϕ_M	ϕ_A	ϕ_X	SAD	Time		
4000	5	0.0034	0.0147	0	0.1264	5.4033	0.0005	0.0025	0	0.0286	0.7657	5 + 0.0206	6 + 0.0009
	10	0.0046	0.0146	0	0.1981	158.3557	0.0018	0.0067	0	0.1034	1.4707	10 + 0.0040	11 + 0.0008
	15	0.0335	0.0927	0	4.2413	478.3535	0.0228	0.0755	0	1.6291	2.8269	15 + 0.0012	16 + 0.0008
	20	0.0152	0.0700	0	0.6945	1572.6336	0.0072	0.0316	0	0.2972	3.9038	20 + 0.0017	21 + 0.0008
8000	5	0.0032	0.0061	0	0.2002	3.2270	0.0007	0.0014	0	0.0349	1.0759	5 + 0.0369	6 + 0.0007
	10	0.0086	0.0267	0	0.6264	365.4279	0.0036	0.0068	0	0.0939	1.7829	10 + 0.0037	11 + 0.0006
	15	0.0099	0.0286	0	1.1329	1233.9985	0.0046	0.0124	0	0.3029	4.1594	15 + 0.0008	16 + 0.0012
	20	0.0300	0.0950	0	18.8659	3319.7858	0.0155	0.0515	0	0.7091	7.1124	20 + 0.0012	21 + 0.0010
12,000	5	0.0024	0.0048	0	0.0423	6.2190	0.0004	0.0011	0	0.0257	1.4246	5 + 0.0023	6 + 0.0009
	10	0.0047	0.0110	0	0.3527	534.3258	0.0017	0.0051	0	0.0359	9.5453	10 + 0.0105	11 + 0.0008
	15	0.0314	0.0808	0	1.7409	2651.6650	0.0222	0.0603	0	1.2631	5.9396	15 + 0.0012	16 + 0.0009
	20	0.0278	0.0832	0	1.6324	13,125.5823	0.0178	0.0458	0	0.7527	11.1843	20 + 0.0014	21 + 0.0010

4.2. Real Data

4.2.1. Cuprite Data

We first use the well-known AVIRIS Cuprite data (http://aviris.jpl.nasa.gov/html/aviris.freedata.html (accessed on 15 February 2021).) in order to evaluate the performance of the proposed approach in extracting endmembers. The scene has been widely used to evaluate the performance of endmember extraction algorithms. The portion used in experiments corresponds to a 250 × 191 pixel subset of the scene, with 224 spectral bands in the range 0.4–2.5 μm and nominal spectral resolution of 10 nm. We remove the bands 1–6, 105–115, 150–170, and 222–224 (due to water absorption and low SNR), and retain 183 spectral bands in our experiments. Figure 5 presents the subimage of Cuprite data (displaying only the 30th band) for this experiment.
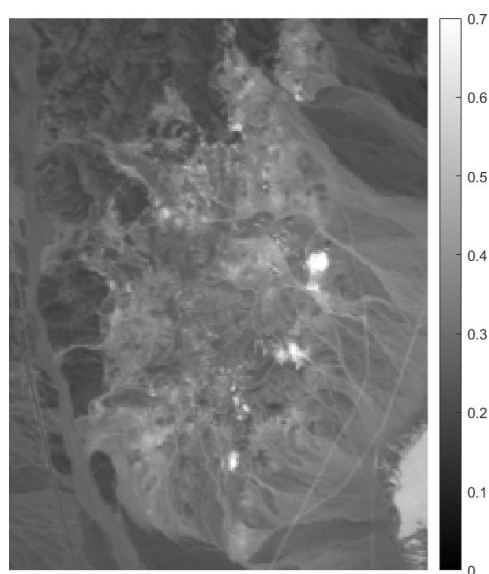
Figure 5. 30th band of the considered airborne visible/infrared imaging spectrometer (AVIRIS) Cuprite subimage.

We employ the metrics ϕ_M, ϕ_X and SAD to evaluate the endmember extraction performance of our method and MVCNMF on the Cuptite subimage. Table 7 shows the processing time and values of ϕ_M, ϕ_X and SAD for our method and MVCNMF. From Table 7, we see that our method gets more accurate results when compared with MVCNMF in terms of the values of the metrics ϕ_M, ϕ_X and SAD. This fact reflects that our method extracts endmembers effectively. From Table 7, we also find that our method takes less than 1 min. to perform the computation, while MVES spends around 20 min. in the considered environment. As a result, we conclude that our proposed method is very efficient for endmember extraction.

Table 7. Performance and processing time (in seconds) obtained by our proposed MDA-minimum volume simplex analysis (MDA-MVSA) method and minimum volume constrained nonnegative matrix factorization (MVCNMF) on the considered Cuprite subimage.

Methods	ϕ_M	ϕ_X	SAD	Time
MDA-MVSA	**0.0747**	**0.4919**	**4.4705**	**34.6597**
MVCNMF	0.0927	0.5519	5.4707	1230.8838

Figure 6 shows the abundance maps obtained by our MDA-MVSA method. The abundance maps that are obtained by MDA-MVSA are identified as mineral maps of Chlorapatite WS423, Jarosite WS368 Pb, Montmorillonite SCa-2.a, Kaolin/Smect KLF511 12%K #1, Kaolin/Smect KLF508 85%K, Nontronite NG-1.a, Montmorillonite SCa-2.b, Nontronite

SWa-1.a, Kaolin/Smect KLF511 12%K #2, Kaolin/Smect H89-FR-5 30K, Buddingtonite GDS85 D-206, Kaolin/Smect KLF511 12%K #3, as shown in Figure 6a–n, respectively. In addition, Figure 7 presents the spectral signatures of the estimated endmembers. From Figure 7, we observe that the endmember signatures that are estimated by our method are in good agreement with the real spectral signatures. The results that are shown in Figure 7 reflect the high accuracy of our method in endmember extraction on the Cuprite subimage.
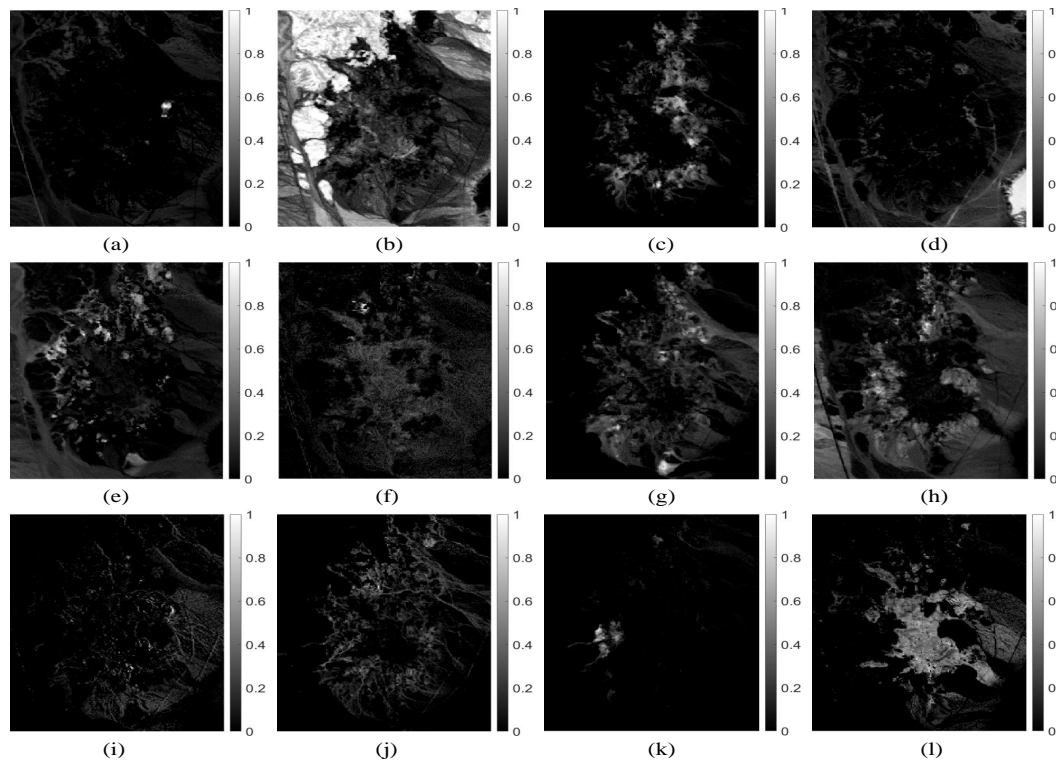

Figure 6. Abundance fraction maps estimated by our MDA-MVSA method from the Cuprite subimage: (**a**) Chlorapatite WS423; (**b**) Jarosite WS368 Pb; (**c**) Montmorillonite SCa-2.a; (**d**) Kaolin/Smect KLF511 12%K #1; (**e**) Kaolin/Smect KLF508 85%K; (**f**) Nontronite NG-1.a; (**g**) Montmorillonite SCa-2.b; (**h**) Nontronite SWa-1.a; (**i**) Kaolin/Smect KLF511 12%K #2; (**j**) Kaolin/Smect H89-FR-5 30K; (**k**) Buddingtonite GDS85 D-206; and, (**l**) Kaolin/Smect KLF511 12%K #3.

4.2.2. Samson Data

We also utilize the Samson real hyperspectral image (http://lesun.weebly.com/hyperspectral-data-set.html (accessed on 15 February 2021).) to evaluate the performance of our method in extracting endmembers. The Samson image comprises 952×952 pixels, each of which has 156 bands that cover the wavelength range from 401 to 889 nm. The original image is quite large, leading to high computational cost for several of the considered algorithms. In this experiment, a subimage with 95×95 pixels is used. The spatial coordinate of the first pixel of the subimage in the full Samson image is (252,332). The Samson data contain three endmembers, i.e., tree, water, and soil (Figure 8a shows the false color composite of the considered subimage).

We employ the metrics ϕ_M, ϕ_A, ϕ_X to evaluate the performance of our method and MVCNMF in extracting endmembers on the considered Samson data. Table 8 shows the performance metrics of different methods that are applied to the Samson subimage, together with the processing time (in seconds) for each method. From Table 8, we conclude that our method exhibits very low values of ϕ_M, ϕ_A, ϕ_X, while our method is faster than MVCNMF. In summary, this reveals that our method effectively and efficiently extracts endmembers from the considered Samson data.
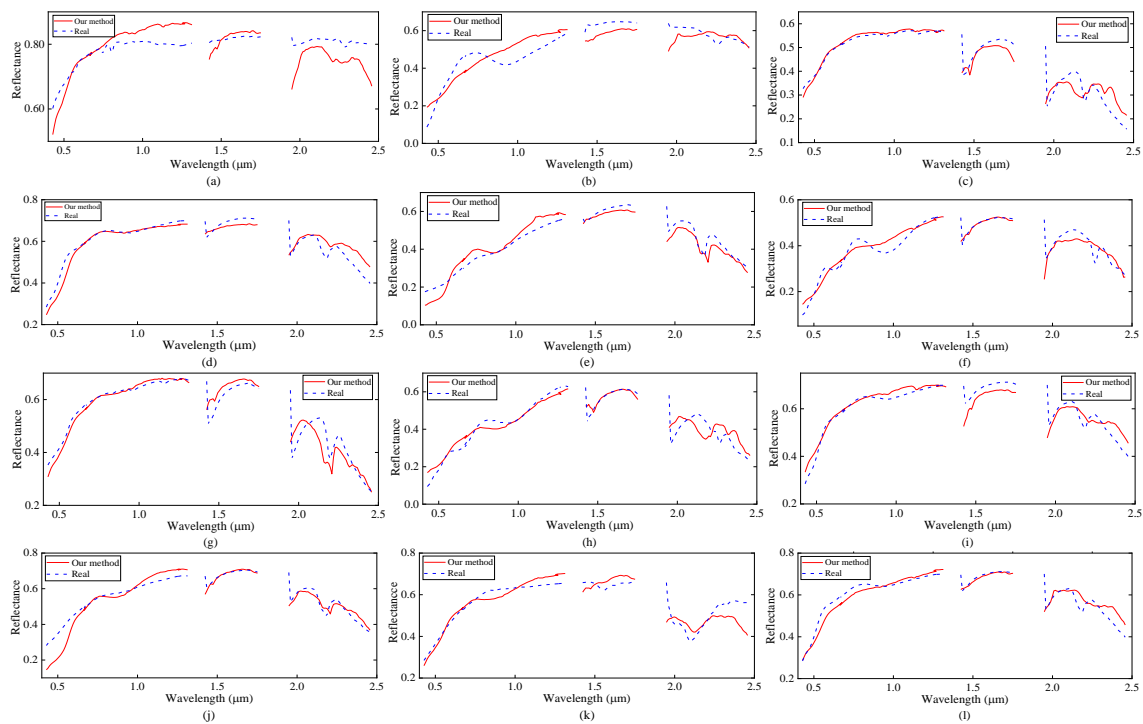
Figure 7. Endmember signatures in the USGS library and the endmember estimates obtained by our MDA-MVSA method from the Cuprite subimage. The corresponding signatures are the following. (**a**) Chlorapatite WS423; (**b**) Jarosite WS368 Pb; (**c**) Montmorillonite SCa-2.a; (**d**) Kaolin/Smect KLF511 12%K #1; (**e**) Kaolin/Smect KLF508 85%K; (**f**) Nontronite NG-1.a; (**g**) Montmorillonite SCa-2.b; (**h**) Nontronite SWa-1.a; (**i**) Kaolin/Smect KLF511 12%K #2; (**j**) Kaolin/Smect H89-FR-5 30K; (**k**) Buddingtonite GDS85 D-206; and, (**l**) Kaolin/Smect KLF511 12%K #3.
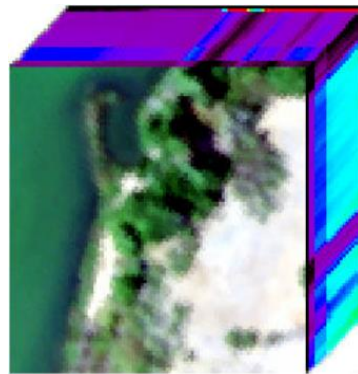
Figure 8. False color composition of the considered Samson subimage.

Table 8. Performance and processing time (seconds) obtained by our proposed MDA-MVSA method and MVCNMF on the considered Samson subimage.

Methods	ϕ_M	ϕ_A	ϕ_X	SAD	Time
MDA-MVSA	**0.0833**	**0.2026**	**1.3961**	**4.3788**	**1.1834**
MVCNMF	0.3007	0.8321	1.5881	11.7210	23.7691

Figure 9 presents the estimated abundance maps by our method. Figure 10 shows a qualitative comparison between the endmembers that are extracted by our method from the Samson subimage and their corresponding reference (ground-truth) signatures. From Figure 10, we conclude that the endmember signatures estimated by our method are in

very good spectral agreement with the reference signatures. The results that are shown in Figures 9 and 10 suggest the high accuracy of our method in extracting endmembers from real data.
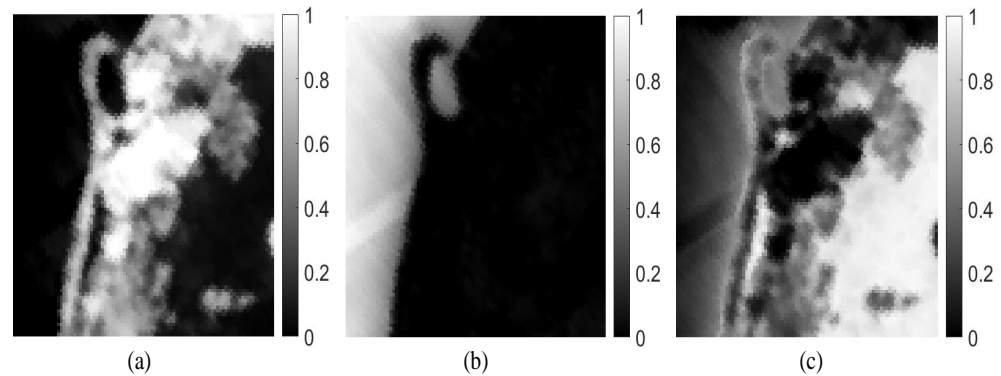

Figure 9. Abundance fraction maps estimated by our MDA-MVSA method from the Samson subimage: (**a**) tree; (**b**) water and (**c**) soil.
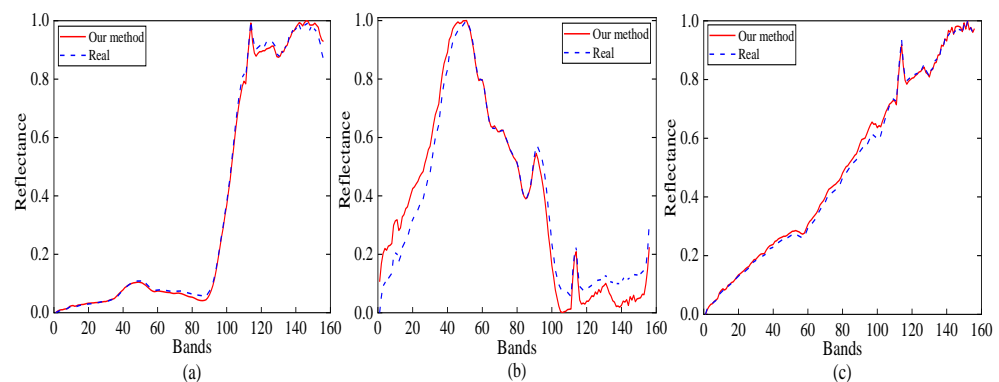

Figure 10. Reference signatures and estimated endmember signatures by our MDA-MVSA method from the Samson subimage: (**a**) tree; (**b**) water; and, (**c**) soil.

5. Conclusions and Future Work

We have developed a new algorithm for estimating the number of endmembers and the corresponding endmember signatures, which is referred to as maximum distance analysis (MDA). Our proposed MDA method extracts endmember signatures without any prior knowledge about the number of endmembers. Moreover, our method cannot obtain the most accurate results of endmember estimation when the experimental data have noise due to the fact that the distance calculation that is shown in Equation (18) is given in the scenario of linear mixing. To address the issue, we embed our proposed MDA algorithm into the original MVSA algorithm to form a new and overall endmember estimation framework. The operation avoids the limitation that the original MVSA algorithm requires prior knowledge on the number of endmembers, and it has high efficiency for the overall endmember estimation task. We refer to this novel strategy as MDA-MVSA. The strategy has been validated while using synthetic data and real data, and it has shown to accurately perform both counting and extracting endmembers, exhibiting very high efficiency in comparison with other traditional endmember extraction methods. Specifically, MDA is usually used on simple scenes (i.e., experimental data containing pure pixels and no noise), and MDA-MVSA combines MDA and MVSA to form a new and effective unmixing framework for complex scenes (i.e., experimental data containing non-pure pixels and noise).

As with any new approach, there are still some unresolved issues that may present challenges over time. Specifically, the distance calculation in our proposed method may need to be adapted if the hyperspectral data are nonlinearly mixed or include noise. In our future

work, we will investigate how to simultaneously determine the number of endmembers and extract their signatures from hyperspectral images when they are nonlinearly mixed.

Author Contributions: Methodology, X.T.; Software, J.M.H.; Writing—review & editing, M.E.P., P.R., J.P. and A.P. The authors have contributed as equally to this work. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Luo, F.; Du, B.; Zhang, L.; Zhang, L.; Tao, D. Feature learning using spatial-spectral hypergraph discriminant analysis for hyperspectral image. *IEEE Trans. Cybern.* **2018**, *49*, 2406–2419. [CrossRef] [PubMed]
2. 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]
3. Goetz, A.F.; Vane, G.; Solomon, J.E.; Rock, B.N. Imaging spectrometry for earth remote sensing. *Science* **1985**, *228*, 1147–1153. [CrossRef]
4. Paoletti, M.; Haut, J.; Plaza, J.; Plaza, A. Deep learning classifiers for hyperspectral imaging: A review. *ISPRS J. Photogramm. Remote Sens.* **2019**, *158*, 279–317. [CrossRef]
5. Keshava, N.; Mustard, J.F. Spectral unmixing. *IEEE Signal Process. Mag.* **2002**, *19*, 44–57. [CrossRef]
6. Gao, F.; You, J.; Wang, J.; Sun, J.; Yang, E.; Zhou, H. A novel target detection method for SAR images based on shadow proposal and saliency analysis. *Neurocomputing* **2017**, *267*, 220–231. [CrossRef]
7. Adams, J.B.; Smith, M.O.; Johnson, P.E. Spectral mixture modeling: A new analysis of rock and soil types at the Viking Lander 1 site. *J. Geophys. Res. Solid Earth* **1986**, *91*, 8098–8112. [CrossRef]
8. Chan, T.H.; Ma, W.K.; Ambikapathi, A.; Chi, C.Y. An optimization perspective onwinter's endmember extraction belief. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Vancouver, BC, Canada, 24–29 July 2011; pp. 1143–1146.
9. 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]
10. 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]
11. Tao, X.; Cui, T.; Plaza, A.; Ren, P. Simultaneously Counting and Extracting Endmembers in a Hyperspectral Image Based on Divergent Subsets. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 8952–8966. [CrossRef]
12. Xu, X.; Pan, B.; Chen, Z.; Shi, Z.; Li, T. Simultaneously Multiobjective Sparse Unmixing and Library Pruning for Hyperspectral Imagery. *IEEE Trans. Geosci. Remote Sens.* **2020**. [CrossRef]
13. Xu, X.; Shi, Z.; Pan, B.; Li, X. A classification-based model for multi-objective hyperspectral sparse unmixing. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 9612–9625. [CrossRef]
14. Geng, X. Study on Hyperspectral Target Detection and Classification. Ph.D. Thesis, Chinese Academy of Sciences, Beijing, China, January 2005.
15. Wang, R.; Li, H.C.; Pizurica, A.; Li, J.; Plaza, A.; Emery, W.J. Hyperspectral unmixing using double reweighted sparse regression and total variation. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1146–1150. [CrossRef]
16. Fernandez-Beltran, R.; Plaza, A.; Plaza, J.; Pla, F. Hyperspectral unmixing based on dual-depth sparse probabilistic latent semantic analysis. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 6344–6360. [CrossRef]
17. Delgado, J.; Martín, G.; Plaza, J.; Jiménez, L.I.; Plaza, A. Fast spatial preprocessing for spectral unmixing of hyperspectral data on graphics processing units. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *9*, 952–961. [CrossRef]
18. Zhang, L.; Du, B.; Zhong, Y. Hybrid detectors based on selective endmembers. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 2633–2646. [CrossRef]
19. Kowkabi, F.; Ghassemian, H.; Keshavarz, A. A fast spatial–spectral preprocessing module for hyperspectral endmember extraction. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 782–786. [CrossRef]
20. Kowkabi, F.; Keshavarz, A. Using spectral Geodesic and spatial Euclidean weights of neighbourhood pixels for hyperspectral Endmember Extraction preprocessing. *ISPRS J. Photogramm. Remote Sens.* **2019**, *158*, 201–218. [CrossRef]
21. Zhang, C.; Qin, Q.; Zhang, T.; Sun, Y.; Chen, C. Endmember extraction from hyperspectral image based on discrete firefly algorithm (EE-DFA). *ISPRS J. Photogramm. Remote Sens.* **2017**, *126*, 108–119. [CrossRef]

22. Sun, W.; Yang, G.; Wu, K.; Li, W.; Zhang, D. Pure endmember extraction using robust kernel archetypoid analysis for hyperspectral imagery. *ISPRS J. Photogramm. Remote Sens.* **2017**, *131*, 147–159. [CrossRef]

23. Du, B.; Wei, Q.; Liu, R. An improved quantum-behaved particle swarm optimization for endmember extraction. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 6003–6017. [CrossRef]

24. Tong, L.; Du, B.; Liu, R.; Zhang, L. An improved multiobjective discrete particle swarm optimization for hyperspectral endmember extraction. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 7872–7882. [CrossRef]

25. Lin, C.H.; Chi, C.Y.; Chen, L.; Miller, D.J.; Wang, Y. Detection of sources in non-negative blind source separation by minimum description length criterion. *IEEE Trans. Neural Networks Learn. Syst.* **2017**, *29*, 4022–4037. [CrossRef]

26. Rissanen, J. Modeling by shortest data description. *Automatica* **1978**, *14*, 465–471. [CrossRef]

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

28. Akaike, H. A new look at the statistical model identification. In *Selected Papers of Hirotugu Akaike*; Springer: Berlin/Heidelberg, Germany, 1974; pp. 716–723.

29. Graham, M.W.; Miller, D.J. Unsupervised learning of parsimonious mixtures on large spaces with integrated feature and component selection. *IEEE Trans. Signal Process.* **2006**, *54*, 1289–1303. [CrossRef]

30. Peres-Neto, P.R.; Jackson, D.A.; Somers, K.M. How many principal components? Stopping rules for determining the number of non-trivial axes revisited. *Comput. Stat. Data Anal.* **2005**, *49*, 974–997. [CrossRef]

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

32. 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]

33. Halimi, A.; Honeine, P.; Kharouf, M.; Richard, C.; Tourneret, J.Y. Estimating the intrinsic dimension of hyperspectral images using a noise-whitened eigengap approach. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 3811–3821. [CrossRef]

34. 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]

35. Parente, M.; Plaza, A. Survey of geometric and statistical unmixing algorithms for hyperspectral images. In Proceedings of the 2010 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, Reykjavik, Iceland, 14–16 June 2010; pp. 1–4.

36. Boardman, J.W.; Kruse, F.A.; Green, R.O. Mapping target signatures via partial unmixing of AVIRIS data. In Proceedings of the Summaries 5th JPL Airborne Earth Science Workshop, Pasadena, CA, USA, 23–26 January 1995; Volume 1, pp. 23–26.

37. Tao, X.; Cui, T.; Yu, Z.; Ren, P. Locality Preserving Endmember Extraction for Estimating Green Algae Area. In Proceedings of the 2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO), Kobe, Japan, 28–31 May 2018; pp. 1–4.

38. Tao, X.; Cui, T.; Ren, P. Cofactor-Based Efficient Endmember Extraction for Green Algae Area Estimation. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 849–853. [CrossRef]

39. Nascimento, J.M.; Dias, J.M. Vertex component analysis: A fast algorithm to unmix hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 898–910. [CrossRef]

40. 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]

41. Chan, T.H.; Chi, C.Y.; Huang, Y.M.; Ma, W.K. A convex analysis-based minimum-volume enclosing simplex algorithm for hyperspectral unmixing. *IEEE Trans. Signal Process.* **2009**, *57*, 4418–4432. [CrossRef]

42. Li, J.; Bioucas-Dias, J.M. Minimum volume simplex analysis: A fast algorithm to unmix hyperspectral data. In Proceedings of the IGARSS 2008-2008 IEEE International Geoscience and Remote Sensing Symposium, Boston, MA, USA, 7–11 July 2008; Volume 3, p. III–250.

43. Li, J.; Agathos, A.; Zaharie, D.; Bioucas-Dias, J.M.; Plaza, A.; Li, X. Minimum volume simplex analysis: A fast algorithm for linear hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 5067–5082.

44. Lin, C.H.; Chi, C.Y.; Wang, Y.H.; Chan, T.H. A fast hyperplane-based minimum-volume enclosing simplex algorithm for blind hyperspectral unmixing. *IEEE Trans. Signal Process.* **2015**, *64*, 1946–1961. [CrossRef]

45. Hearst, M.A.; Dumais, S.T.; Osuna, E.; Platt, J.; Scholkopf, B. Support vector machines. *IEEE Intell. Syst. Their Appl.* **1998**, *13*, 18–28. [CrossRef]

46. Bioucas-Dias, J.M.; Figueiredo, M.A. Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing. In Proceedings of the 2010 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, Reykjavik, Iceland, 14–16 June 2010; pp. 1–4.

Escuela Politecnica
Av. de la Universidad, S/N, 10003
Caceres, Spain
Phone: 0034927257000. Ext. 51662
Email: aplaza,juanmariohaut{@unex.es}

Dr. Antonio Plaza Miguel y Dr. Juan Mario Haut Hurtado como directores de la tesis titulada "Proce-samiento Eficiente y Profundo de Imágenes Hiperespectrales y Aplicación en Tareas de Desmezclado Es-pectral", certifican el factor de impacto y la categorización de la siguiente publicación, incluída en la tesis doctoral. Del mismo modo, se especifica la aportación del doctorado.

Antonio Plaza Miguel PhD, and Juan Mario Haut Hurtado PhD, as directors of the Phd thesis titled "Effi-cient and Deep Processing of Hyperspectral Images and Application in Spectral Unmixing Tasks", certify the impact factor and the categorization of the following publication, included in the doctoral thesis. In the same way, the contribution of the doctorate is specified.

Artículo / Paper

Autores/Authors: **X. Tao**, M. E. Paoletti, J. M. Haut, L. Han, P. Ren, J. Plaza and A. Plaza.
Title: Endmember Estimation From Hyperspectral Images Using Geometric Distances.
Journal: IEEE Geoscience and Remote Sensing Letters, Art. Seq. Number. 5508405, pp. 1-11, August 2021
DOI: 10.1109/LGRS.2021.3102076.
Impact factor 2021: 5.343. Q1

Abstract: Endmember estimation consists of two tasks, that is, determining the number of pure spectral constituents (endmembers) and extracting their spectral signatures. We present a new geometric distance-based method for endmember estimation from hyperspectral images (HSIs), which does not need to know the number of endmembers in advance. Our strategy optimizes the widely used maximum distance analysis (MDA) method from two viewpoints. First, the traditional MDA method performs endmember estimation by computing the maximum distances between any pixel and one specific pixel, line, plane, or affine hull (AH) composed by the endmembers that have been formerly extracted. Instead, our new strategy only requires computing the maximum distance between any pixel and one specific AH. This operation provides a simpler way than MDA to estimate endmembers. Second, our strategy exploits a new distance computation between any pixel and an AH and just needs the normal vector (compared to the traditional MDA method, which uses the normal vector and offset). The new distance computation in our method is much more efficient than that in the traditional MDA method.

Contribución del doctorado: Planteamiento de la hipótesis, desarrollo práctico, análisis y discusión de los resultados, elaboración y escritura del manuscrito.

Firma / Signature
Marzo / Marzo, 2023

Antonio Plaza Miguel   Juan Mario Haut Hurtado

Endmember Estimation From Hyperspectral Images Using Geometric Distances

Xuanwen Tao[ID], *Graduate Student Member, IEEE*, Mercedes E. Paoletti[ID], *Senior Member, IEEE*,
Juan M. Haut[ID], *Senior Member, IEEE*, Lirong Han[ID], Peng Ren[ID], *Senior Member, IEEE*,
Javier Plaza[ID], *Senior Member, IEEE*, and Antonio Plaza[ID], *Fellow, IEEE*

Abstract— Endmember estimation consists of two tasks, that is, determining the number of pure spectral constituents (endmembers) and extracting their spectral signatures. We present a new geometric distance-based method for endmember estimation from hyperspectral images (HSIs), which does not need to know the number of endmembers in advance. Our strategy optimizes the widely used maximum distance analysis (MDA) method from two viewpoints. First, the traditional MDA method performs endmember estimation by computing the maximum distances between any pixel and one specific pixel, line, plane, or affine hull (AH) composed by the endmembers that have been formerly extracted. Instead, our new strategy only requires computing the maximum distance between any pixel and one specific AH. This operation provides a simpler way than MDA to estimate endmembers. Second, our strategy exploits a new distance computation between any pixel and an AH and just needs the normal vector (compared to the traditional MDA method, which uses the normal vector and offset). The new distance computation in our method is much more efficient than that in the traditional MDA method.

Index Terms— Distance computation, endmember estimation, geometric distance, hyperspectral images (HSIs), maximum distance analysis (MDA).

I. INTRODUCTION

HYPERSPECTRAL images (HSIs) have been widely utilized in many applications, such as classification, target identification, change detection, among others. Generally, HSI data contains many mixed pixels because it has low spatial resolution. The spectrum of one mixed pixel is a mixture of many ground cover spectra. Moreover, mixed pixels need to be decomposed into a set of endmembers and their corresponding

proportions at each pixel. To obtain accurate results from the decomposition process, linear spectral unmixing algorithms [1] have been exploited for endmember estimation, which contains two seemingly independent but actually highly related processes: 1) counting endmembers (the number of different types of materials) and 2) according to the previous number, extracting endmembers. A brief review of existing methods is provided below.

1) The algorithms for determining the number of endmembers [2] include hyperspectral signal subspace by minimum error (Hysime), Harsanyi–Farrand–Chang (HFC), geometry-based estimation of number of endmembers-affine hull (GENE-AH), among others. Hysime performs noise covariance estimation and spectrum noise characterization, and it has high computational overloads. The number of endmembers determined by HFC will be affected by a fixed false alarm probability. GENE-AH counts endmembers by exploiting the fact that all pixels should be located in the affine hull (AH) of endmembers. Moreover, GENE-AH needs to use an endmember extraction algorithm (EEA), and *p*-norm-based pure pixel identification (TRIP) is a suitable EEA for GENE-AH.

2) The algorithms for extracting endmembers [3]–[5] include the simplex growing algorithm (SGA), automatic target generation process (ATGP), alternating decoupled volume maxmin (ADVMM) and successive decoupled volume maxmin (SDVMM) [3], alternating volume maximization (AVMAX) and successive volume maximization (SVMAX) [4], negative abundance-oriented (NABO) [5], among many others. SGA searches for the convex geometry of endmembers with the smallest volume and needs specific initialization. ADVMM and SDVMM handle the worst case simplex volume maximization problem by alternating and successive optimizations, respectively. AVMAX and SVMAX take into account more comprehensive criteria to carry out the same process. NABO uses the pixels located outside the hull as alternative candidate endmembers.

All the aforementioned algorithms can only complete one task of endmember estimation, that is, counting or extracting endmembers. Most EEAs in the literature need to know the number of endmembers in advance and cannot operate properly if the number of endmembers is not accurately estimated. To address the limitation, Tao *et al.* [6] developed

a maximum distance analysis (MDA) algorithm that simultaneously estimates the number and signatures of endmembers without any prior knowledge. MDA computes the maximum distance between any pixel and one specific pixel, line, plane, or AH to extract endmembers and finalizes the endmember estimation when the maximum distance is zero. However, MDA has two shortcomings: 1) its procedure is too complex and 2) the maximum distance computation is very expensive for high-dimensional HSIs. To address these limitations, we introduce a new geometric distance-based algorithm derived from MDA and called new MDA (NMDA).

The main difference between our NMDA method and the traditional MDA method can be summarized as follows. The proposed NMDA method first optimizes the process of endmember estimation and yields endmembers by searching for the furthest pixel from one specific AH. Quite opposite, the traditional MDA algorithm obtains endmembers by searching for the furthest pixel from one specific pixel, line, plane, or AH composed by the endmembers that have been formerly extracted. In addition, the newly proposed NMDA algorithm exploits a new distance computation that only requires the normal vector completing the task of endmember estimation. However, the distance computation of the traditional MDA method needs the normal vector and offset. With these issues in mind, there are two major contributions of our newly proposed algorithm: 1) NMDA provides a simpler way than MDA to estimate the endmembers and 2) NMDA is much more efficient than the traditional MDA method. We test our NMDA method in the task of estimating endmembers from synthetic and real HSI data and conclude that our NMDA method is both effective and efficient compared to other state-of-the-art techniques.

II. Methodology

A. Maximum Distance Analysis

Assuming that $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_i, \ldots, \mathbf{y}_N] \in \mathbb{R}^{L \times N}$ is an HSI dataset, and it has L bands and N pixels. The ith pixel $\mathbf{y}_i = [y_{1i}, y_{2i}, \ldots, y_{Li}]^T \in \mathbb{R}^{L \times 1}$ represents an L-dimensional vector. In terms of a geometric point of view, endmembers will form a simplex, which has the largest volume among all pixel combinations. To this end, the MDA algorithm searches for the furthest pixel from one specific pixel, line, plane, or AH. For clarity, we give the main steps of MDA below.

1) MDA extracts the first endmembers by searching for the furthest pixel from the coordinate origin.
2) MDA extracts the second endmembers by searching for the furthest pixel from the first extracted endmember.
3) MDA extracts the third endmembers by searching for the furthest pixel from the line composed by the first and second extracted endmembers.
4) MDA extracts the fourth endmembers by searching for the furthest pixel from the plane composed by the first, second, and third extracted endmembers.
5) MDA extracts the next endmembers by searching for the furthest pixel from the AH composed by all the formerly extracted endmembers.

It is worth mentioning that the furthest pixels in the above steps are acquired by computing the maximum distances
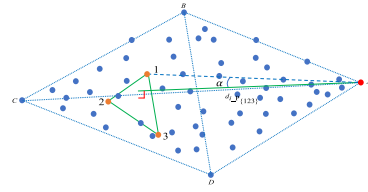

Fig. 1. Schematic using our NMDA method to extract the first endmember. The furthest pixel from the plane $P_{\{123\}}$ is the first endmember. The red dots denote the first endmember that is extracted by our NMDA method, the blue dots represent pixel points, and the yellow dots are randomly selected pixels to form the initial plane $P_{\{123\}}$.

between any pixel and one specific pixel, line, plane, or AH, and the maximum distances are computed by some mathematical method. Specifically, MDA computes the distance between any pixel and one specific AH by

$$d = \frac{\left| \mathbf{w}^T \vec{y}_i + b \right|}{||\mathbf{w}||_2} \tag{1}$$

where $\vec{y}_i$ is a pixel in the pixel space, $\mathbf{w}$ is the normal vector of the AH, and b is the offset. MDA utilizes (1) to compute the maximum distances between all pixels and the specific AH composed by all endmembers extracted by previous steps, and it finalizes the endmember extraction task when the maximum distance is zero.

B. New Maximum Distance Analysis

Our NMDA algorithm can simultaneously estimate the number and signatures of endmembers. It exploits the fact that the furthest pixel from any plane or AH must be an endmember. We randomly choose three pixels (1, 2, and 3) from the pixel space to form the initial plane $P_{\{123\}}$, as shown in Fig. 1. Particularly, random initialization only determines which endmember is extracted first, and it does not affect the final result of unmixing. The vector $\mathbf{n}$ is the normal vector of the plane $P_{\{123\}}$, and the angle α between the two vectors $\vec{1i}$ and $\mathbf{n}$ is trigonometrically computed by

$$\alpha = \arccos \frac{\vec{1i} \cdot \mathbf{n}}{\left\| \vec{1i} \right\|_2 \|\mathbf{n}\|_2}. \tag{2}$$

We compute the distance $d_{i\_P_{\{123\}}}$ between any pixel i and the plane $P_{\{123\}}$ formed by the three pixels 1, 2, and 3 to get the first endmember, and the distance $d_{i\_P_{\{123\}}}$ is computed by

$$d_{i\_P_{\{123\}}} = \left\| \vec{1i} \right\|_2 |\cos \alpha|$$
$$= \frac{\left\| \vec{1i} \cdot \mathbf{n} \right\|}{||\mathbf{n}||_2} \tag{3}$$

where the symbol $|\cdot|$ represents the absolute value. We obtain the index I_1 of the first endmember by

$$I_1 = \arg\max_{i \in \{1,2,\ldots,N\}} d_{i\_P_{\{123\}}} \tag{4}$$

and then we use the index I_1 to extract the first endmember A. Equation (3) offers a new distance computation method, and it
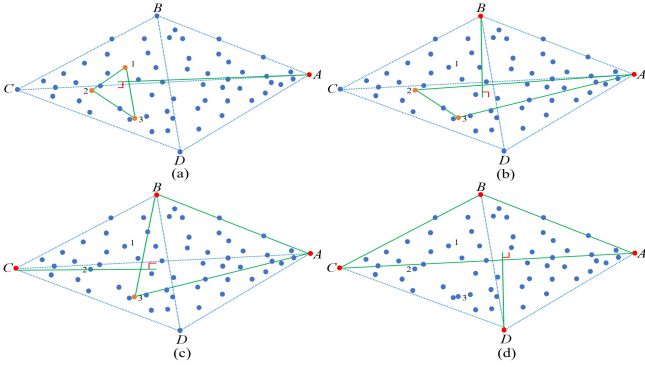
Fig. 2. Red dots denote the endmembers that are extracted by our NMDA method based on different planes. (a)–(d) Process of extracting the first, second, third, and fourth endmembers (A, B, C, and D) based on different planes $P_{\{123\}}$, $P_{\{23A\}}$, $P_{\{3AB\}}$, and $P_{\{ABC\}}$.

can be used to compute the distance between any pixel and one specific plane or AH. Now we only have one endmember A that cannot form a plane or AH, and we use two random pixels (2 and 3) to help the pixel A to form the plane $P_{\{23A\}}$. We compute the distance $d_{i\_P_{\{23A\}}}$ between any pixel i and the plane $P_{\{23A\}}$ to seek for the furthest pixel from the plane $P_{\{23A\}}$ as the second endmember B. Now we have the plane $P_{\{3AB\}}$, and we then compute the distance $d_{i\_P_{\{3AB\}}}$ to obtain the third endmember C. At this point, we have three endmembers that can form a plane, and we update (3) and (4) as follows:

$$
\begin{aligned}
d_{i\_P_{\{ABC\cdots\}}} &= \left\| \overrightarrow{Ai} \right\|_2 |\cos \alpha| \\
&= \frac{\left\| \overrightarrow{Ai} \cdot \mathbf{n} \right\|}{||\mathbf{n}||_2}
\end{aligned}
\tag{5}
$$

and

$$
I = \underset{i \in \{1,2,\ldots,N\}}{\arg \max} \; d_{i\_P_{\{ABC\cdots\}}}
\tag{6}
$$

where α is the angle between the two vectors $\overrightarrow{Ai}$ and $\mathbf{n}$. We extract endmembers by using (5) and (6) to compute the distance between any pixel and the specific plane and finalize the endmember extraction process when the maximum distance is zero. Fig. 2 graphically illustrates the process of extracting four endmembers. Specifically, we note that (5) provides a more efficient way to compute the distance between any pixel and one specific AH by observing (1) and (5). In addition, in Section II-A, we can see that the traditional MDA method needs to compute many distances to complete the endmember estimation, while our method only needs to compute the distance between any pixel and one specific plane to complete such estimation. Therefore, our method provides a simpler way that MDA to complete endmember estimation and does not need to know the number of endmembers in advance.

III. EXPERIMENTAL RESULTS

We evaluate the performance of the proposed NMDA method for the purpose of determining the number of end-members and extracting their signatures on synthetic and

real data. We first compare our proposed NMDA method with some endmember extraction methods, that is, ADVMM, SDVMM, AVMAX, and SVMAX, TRIP, ATGP, NABO, and SGA. We then compare our proposed NMDA method with three counting endmember methods, that is, Hysime, HFC, and GENE-AH. We utilize spectral information divergence (SID), abundance information divergence (AID), and root mean square error (RMSE) to evaluate the effectiveness of different methods in the task of endmember extraction. Let $\mathbf{Y}_e$ and $\hat{\mathbf{Y}}$ denote the true and estimated endmember signatures, respectively. Let $\mathbf{A}_e$ and $\hat{\mathbf{A}}$ denote the true and estimated endmember abundance, respectively. Moreover, $\mathbf{p}_m = (\mathbf{y}_m / \sum_{l=1}^{L} y_{jm})$ is the probability distribution vector of the mth true endmember signature. The vector $\mathbf{p}_m$ can capture the variableness of spectral signatures. Let $\mathbf{P}_e \in \mathbb{R}^{L \times M}$ and $\hat{\mathbf{P}} \in \mathbb{R}^{L \times M}$ be the probability distribution matrix for $\mathbf{Y}_e$ and $\hat{\mathbf{Y}}$, respectively. p_{lm} is the (i, j)th element of matrix $\mathbf{P}_e$, and $\hat{p}_{lm}$ is the (i, j)th element of matrix $\hat{\mathbf{P}}$. We have

$$
D(\mathbf{y}_m | \hat{\mathbf{y}}_m) = \sum_{l=1}^{L} p_{lm} \log \frac{p_{lm}}{\hat{p}_{lm}}.
\tag{7}
$$

The relative entropy between $\mathbf{Y}_e$ and $\hat{\mathbf{Y}}$ is $D(\mathbf{Y}_e | \hat{\mathbf{Y}}) = [D(\mathbf{y}_1 | \hat{\mathbf{y}}_1), \ldots, D(\mathbf{y}_M | \hat{\mathbf{y}}_M)] \in \mathbb{R}^{1 \times M}$. SID, AID, and RMSE are given as follows:

$$
\begin{aligned}
\text{SID} &= D(\mathbf{Y}_e | \hat{\mathbf{Y}}) + D(\hat{\mathbf{Y}} | \mathbf{Y}_e) \\
\text{AID} &= D(\mathbf{A}_e | \hat{\mathbf{A}}) + D(\hat{\mathbf{A}} | \mathbf{A}_e) \\
\text{RMSE} &= \sqrt{\frac{\sum_{i=1}^{N} \sum_{j=1}^{M} R_{ij}^4}{NM}}
\end{aligned}
\tag{8}
$$

where N is the number of pixels, M is the number of endmembers, and R_{ij} is the (i, j)th element of the matrix $\mathbf{A}_e - \mathbf{A}$. We note that EEAs obtain the best results when the values of SID, AID, and RMSE are zero.

A. Synthetic Data

We use the United States Geological Survey (USGS) mineral spectral library[1] to create synthetic data. We utilize the linear mixture method [7] to generate two different types of synthetic data: 1) synthetic data with pure pixels and 2) synthetic data with different levels of noise. We employ SID, AID, and RMSE to assess the performance of our method in the task of extracting endmembers from two types of synthetic data, that is, with pure pixels and with different levels of noise and ten endmembers. Fig. 3(a)–(c) illustrates the accuracy of different endmember extraction methods on synthetic data with pure pixels. From Fig. 3(a)–(c), we observe that the values of SID, AID, and RMSE are near to zero for all methods, and this fact confirms that all of them exhibit high endmember extraction accuracy. To test the robustness of our method, we add 10-, 15-, 20-, and 25-dB white Gaussian noise to the synthetic data and report the results of different methods in Fig. 3(d)–(f). From Fig. 3(d)–(f), we see that the values of SID and RMSE are in general close to zero, and the value of AID is a bit high. This fact reflects that our method exhibits

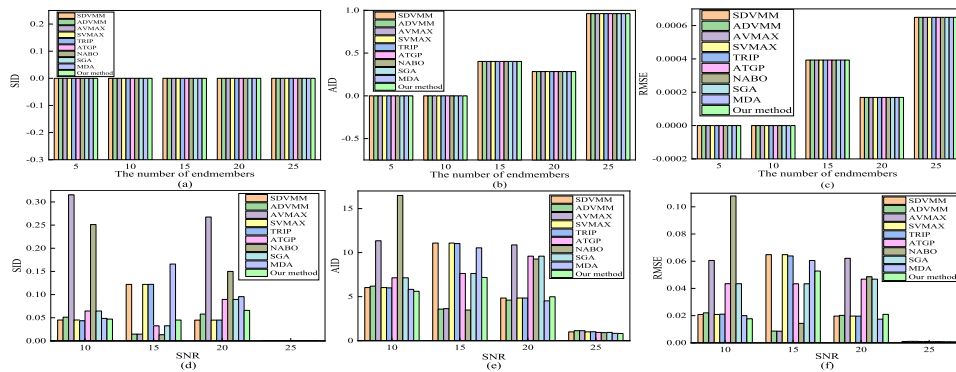[1] https://speclab.cr.usgs.gov/spectral-lib.html

Fig. 3. Performance of EEAs. (a)–(c) Synthetic data with pure pixels. (d)–(f) Synthetic data with different levels of noise and five endmembers.

TABLE I
NUMBER OF ENDMEMBERS AND THE MAXIMUM DISTANCE OBTAINED BY OUR PROPOSED NMDA METHOD
ON SYNTHETIC DATA WITH PURE PIXELS AND DIFFERENT LEVELS OF NOISE

Actual number of endmembers	Estimated number of endmembers + the maximum distance									
	With pure pixels		SNR = 10dB		SNR = 15dB		SNR = 20dB		SNR = 25dB	
5	5 + 0.0329	6 + 0	5 + 0.0481	6 + 0	5 + 0.0305	6 + 0	5 + 0.025	6 + 0	5 + 0.0246	6 + 0
10	10 + 0.027	11 + 0	10 + 0.0035	11 + 0	10 + 0.0216	11 + 0	10 + 0.0081	11 + 0	10 + 0.0347	11 + 0
15	15 + 0.0006	16 + 0	15 + 0.0699	16 + 0	15 + 0.0229	16 + 0	15 + 0.0193	16 + 0	15 + 0.0196	16 + 0
20	20 + 0.0003	21 + 0	20 + 0.075	21 + 0	20 + 0.0289	21 + 0	20 + 0.0084	21 + 0	20 + 0.0258	21 + 0
25	25 + 0.0015	26 + 0	25 + 0.039	26 + 0	25 + 0.0542	26 + 0	25 + 0.0253	26 + 0	25 + 0.0046	26 + 0

TABLE II
AVERAGE PROCESSING TIME (SECONDS) FOR ESTIMATING THE NUMBER AND SIGNATURES
OF ENDMEMBERS ON DIFFERENT TYPES OF SYNTHETIC DATA

The types of synthetic data	Methods												
	SDVMM	ADVMM	AVMAX	SVMAX	TRIP	ATGP	NABO	SGA	Hysime	HFC	GENE-AH	MDA	Our method
With pure pixels	1.2179	1.1892	1.1921	1.1296	1.0602	0.9673	0.8356	1.1021	0.1629	0.0928	5.3545	1.0535	**0.9903**
SNR = 10	1.5179	1.5384	1.4853	1.4039	1.4253	1.2495	1.4616	1.1283	0.1481	0.067	5.9397	1.1702	**1.0485**
SNR = 15	1.512	1.5747	1.4695	1.4348	1.4236	1.3104	1.4792	1.1463	0.1539	0.068	6.0469	1.1797	**1.0784**
SNR = 20	1.6468	1.7343	1.6237	1.5999	1.591	1.3606	1.6055	1.294	0.1648	0.0719	6.3767	1.2392	**1.1576**
SNR = 25	1.5132	1.6649	1.5289	1.4345	1.4711	1.2931	1.5003	1.2091	0.1477	0.0681	6.0306	1.2159	**1.0972**

good and robust behavior with different noise levels and achieves accurate results in terms of extracting endmembers. It is worth mentioning that only MDA and our method can complete the overall task of endmember estimation, and other methods can only count or extract endmembers. Moreover, our method has a slight advantage over MDA.

Table I shows the relationship between the estimated and actual number of endmembers and the maximum distance on synthetic data with pure pixels and different levels of noise. From Table I, we can see that the maximum distance is zero when the assessed number of endmembers goes beyond the actual number of endmembers. This fact indicates that the proposed NMDA method obtains accurate estimations of the number of endmembers on all synthetic datasets. Simultaneously, we apply Hysime, HFC, and GENE-AH to compare their performance with that of the proposed NMDA method in counting endmembers on different synthetic datasets. Fig. 4 displays the number of endmembers determined by different methods on various synthetic datasets. From Fig. 4, we can see
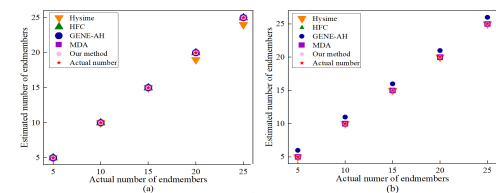

Fig. 4. Performance of different methods in estimating the number of endmembers. (a) Synthetic data with pure pixels. (b) Synthetic data with noise.

that HFC, MDA, and our NMDA provide accurate estimates on the number of endmembers, and Hysime and GENE-AH obtain some erroneous results on synthetic data with pure pixels and synthetic data with noise, respectively. Similarly, only MDA and our NMDA method can complete the overall task of endmember estimation, and other methods can only determine the number of endmembers or extract endmembers.

Table II shows the average processing time of different methods on different synthetic datasets. From Table II, we see

TABLE III

EFFECTIVENESS AND EFFICIENCY FOR DIFFERENT EEAs APPLIED TO THE SAMSON SUBIMAGE

Metrics	Methods									
	SDVMM	ADVMM	AVMAX	SVMAX	TRIP	ATGP	NABO	SGA	MDA	Our method
SID	0.0274	0.0270	0.0279	0.0277	0.0201	1.0830	0.0778	1.1329	0.0194	**0.0061**
AID	7.8269	7.7682	7.7619	7.8258	7.8180	20.1406	7.9293	19.9537	7.8344	**7.4205**
RMSE	0.1726	0.1715	0.1714	0.1727	0.1729	0.3766	0.1815	0.3732	0.1729	**0.1620**
Time (seconds)	1.8305	1.8130	1.7754	1.6125	1.6474	1.7134	**1.4292**	1.7346	1.4917	1.4743

TABLE IV

SID AND RUNNING TIME FOR DIFFERENT EEAs APPLIED TO THE CUPRITE IMAGE

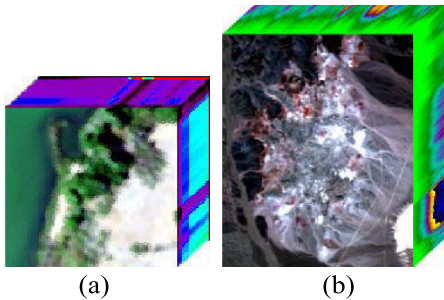| Metrics	Methods									
	SDVMM	ADVMM	AVMAX	SVMAX	TRIP	ATGP	NABO	SGA	MDA	Our method
SID	0.0655	0.0571	0.0616	0.0655	0.0655	0.0656	0.0580	0.0654	0.0844	0.0444
Time (seconds)	12.3193	11.7942	11.2322	11.3414	12.1164	13.1332	11.8322	15.4361	13.4752	10.1035

Fig. 5. False color composition of the considered. (a) Samson subimage. (b) Cuprite image.

that our proposed method can not only simultaneously estimate the number and signatures of endmembers, but also has high efficiency compared with other methods that can only extract endmembers or determine the number of endmembers. Moreover, although MDA can also determine the number of endmembers and extract their spectral signatures, our proposed method has higher accuracy and efficiency than MDA.

B. Real Data

We use the Samson subimage and Cuprite image[2] to assess the performance of our proposed NMDA method in the task of extracting endmembers. The false color composition of the Samson subimage and Cuprite image are given in Fig. 5.

The unmixing results of different methods on Samson subimage are shown in Table III. From Table III, we see that our method acquires the most accurate results in terms of SID, AID, and RMSE, and NABO has the highest efficiency. However, our method can simultaneously complete the two tasks, that is, counting and extracting endmembers, and NABO can only extract endmembers and the time gap between it and our method is very small. Moreover, our method is more accurate and efficient than MDA. The unmixing results of different methods on Cuprite image are shown in Table IV.

[2]http://lesun.weebly.com/hyperspectral-dataset.html

Especially, since Cuprite only has real endmember signatures and does not have their corresponding real abundance maps, we only present the results of SID for evaluating estimated endmembers. From Table IV, we see that our proposed method not only acquires more accurate results but also has higher efficiency than other methods. Therefore, our method is also very effective and efficient on real scenes.

IV. CONCLUSION AND FUTURE WORK

We have proposed a new endmember estimation method, called NMDA, for simultaneously counting and extracting endmembers. We have evaluated our NMDA method using both synthetic and real hyperspectral datasets, and the experimental results reveal that our method estimates the number of endmembers and extracts their spectral signatures effectively and efficiently compared with other state-of-the-art methods. Further studies will focus on how to apply our method to nonlinear unmixing problems.

REFERENCES

[1] X. Tao, T. Cui, and P. Ren, "Cofactor-based efficient endmember extraction for green algae area estimation," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 6, pp. 849–853, Jun. 2019.

[2] A. M. Ambikapathi, T. H. Chan, C. Y. Chi, and K. Keizer, "Hyperspectral data geometry-based estimation of number of endmembers using p-norm-based pure pixel identification algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2753–2769, May 2013.

[3] T.-H. Chan, J.-Y. Liou, A. Ambikapathi, W.-K. Ma, and C.-Y. Chi, "Fast algorithms for robust hyperspectral endmember extraction based on worst-case simplex volume maximization," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2012, pp. 1237–1240.

[4] T. Chan, W. Ma, A. Ambikapathi, and C. Chi, "A simplex volume maximization framework for hyperspectral endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4177–4193, Nov. 2011.

[5] R. Marrero *et al.*, "A novel negative abundance-oriented hyperspectral unmixing algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3772–3790, Jul. 2015.

[6] X. Tao, M. E. Paoletti, J. M. Haut, P. Ren, J. Plaza, and A. Plaza, "Endmember estimation with maximum distance analysis," *Remote Sens.*, vol. 13, no. 4, p. 713, Feb. 2021.

[7] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 765–777, Mar. 2007.