Analytica Chimica Acta 762 (2013) 25–38

Contents lists available at SciVerse ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

Application of Multivariate Curve Resolution Alternating Least Squares (MCR-ALS) to remote sensing hyperspectral imaging☆

Xin Zhang, Romà Tauler*

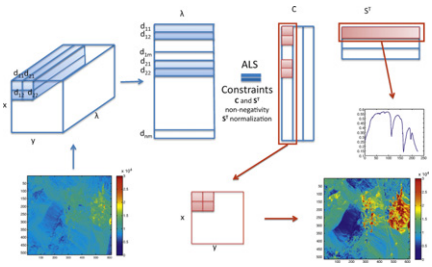
IDAEA-CSIC, Jordi Girona 18-26, Barcelona 08034, Spain

HIGHLIGHTS

- MCR-ALS is successfully applied to remote sensing hyperspectral images.
- Pure spectra and relative concentrations of image constituents were obtained.
- MCA-ALS results are favorably compared to results obtained by MVSA and VCA methods.
- Physical constraints were implemented to decrease the rotational ambiguities.
- MCR-BANDS is used to evaluate the presence and extent of rotational ambiguities.

GRAPHICAL ABSTRACT

This article summarizes the use of the MCR-ALS method as a powerful tool for the resolution of hyperspectral images on their constituents. Non-negativity, spectral normalization and local rank constraints are used to get physical meaningful resolution results and decrease rotation ambiguities. MCR-BANDS method is used to evaluate the presence and extend of rotational ambiguities.



ARTICLE INFO

Article history:
Received 17 September 2012
Received in revised form
14 November 2012
Accepted 26 November 2012
Available online 3 December 2012

Keywords:
Multivariate Curve Resolution
Remote sensing hyperspectral imaging
MCR-ALS
MVSA
VCA
Endmember

ABSTRACT

The application of the MCR-ALS method is demonstrated on two simulated remote sensing spectroscopic images and on one experimental reference remote sensing spectroscopic image obtained by the Airborn Visible/Infrared Imaging Spectrometer (AVIRIS). By application of MCR-ALS, the spectra signatures of the pure constituents present in the image and their concentration distribution at a pixel level are estimated. Results obtained by MCR-ALS are compared to those obtained by other methods frequently used in the remote sensing spectroscopic imaging field like VCA and MVSA. In the case of the analysis of the experimental data set, the resolved pure spectra signatures were compared to reference spectra from USGS library for their identification. In all cases, results were also evaluated for the presence of rotational ambiguities using the MCR-BANDS method. The obtained results confirmed that the MCR-ALS method can be successfully used for remote sensing hyperspectral image resolution purposes. However, the amount of rotation ambiguity still present in the solutions obtained by this and other resolution methods (like VCA or MVSA) can still be large and it should be evaluated with care, trying to reduce its effects by selecting the more appropriate constraints. Only in this way it is possible to increase the reliability of the solutions provided by these methods and decrease the uncertainties associated to their use.

1. Introduction

Hyperspectral imaging is a breakthrough in remote sensing technology [1], which can be obtained by Raman, infrared and fluorescence spectroscopies [2], and it is a useful methodology that can be applied for analytical purposes to agriculture [3], biology [4], environmental [5] and other earth science fields [6]. It has emerged as a very important field in recent years [7]. Remote sensing is the

☆ Paper presented at the XIII Conference on Chemometrics in Analytical Chemistry (CAC 2012), Budapest, Hungary, 25–29 June, 2012.

* Corresponding author. Tel.: +34 93 4006140; fax: +34 93 2045904.
E-mail address: Roma.Tauler@idaea.csic.es (R. Tauler).

0003-2670/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.aca.2012.11.043

acquisition of information about an object or phenomenon, without making physical contact with the object. In modern usage, the term generally refers to the use of aerial sensor technologies to detect and classify objects on Earth. Remote sensing hyperspectral imaging is a combination of both technologies, hyperspectral imaging and remote sensing, which provides spectral and spatial analytical information about a particular geographical area. Based on the resolved pure spectra of the constituents of the image and on their constituents distributions, additional information can be derived, such as the possible identification of these constituents from available libraries and their relative quantitation within the image. One of the implications of remote sensing hyperspectral imaging is that at every pixel of the measured image, a mixture of several spectral signatures of different materials is generally present. Resolution of the hyperspectral image is required to extract these signatures (pure spectra) and to figure out how is the distribution of the corresponding image constituents.

Hyperspectral remote sensing includes large data sets generally composed of about 100–200 spectral bands of relatively narrow bandwidths for about 5–10 nm. Hyperspectral images are represented in the form of data cubes. The spatial information is collected in the *X*–*Y* plane, and the spectral information is represented in the *Z*-direction. The analysis of these multidimensional datasets requires sensitive detectors, fast computers, and large data storage capacities, potentially exceeding hundreds of megabytes. PCA (principal component analysis) [8] allows processing spectroscopic image datasets by reducing their dimensions without a significantive loss of relevant information. PCA is a useful tool to estimate the number of more significantive components (constituents) of the image and to remove noise and non-informative parts of it. However, because of the maximum variance criterion and of applied orthogonal constraints, PCA do not provide the true signatures (spectra) and relative concentration profiles of the image constituents directly.

The concept of endmember has been proposed, to refer to the pure constituent spectra present in the image pixels. Diverse mathematical methods based on convex geometry and subspace projection like the PPI (Pixel Purity Index) [9], *N*-FINDR (*N*-finder) [10], and VCA (Vertex Component Analysis) [11] methods have been proposed and applied. PPI projects the vector of pixels on unit vectors selected at random directions, and counts for the number of times the value of each projected pixel reaches an extreme value (this extreme value either can be a maximum or a minimum projected value). The endmembers are identified as those pixels with the highest scores. The *N*-FINDR method is based on the fact that in *N* spectral dimensions, the *N*-volume contained by a simplex formed by the purest pixels is larger than any other volume formed from any other combination of pixels. It begins with an initial simplex composed by a set of random pixels, then iteratively the simplex volume is increased inside the data set until the simplex with the smallest volume containing all data pixels is found. VCA projects the data to the identified orthogonal subspace in an interactive way, and finds the endmembers by repeated iteration. The endmember are at the vertices of the simplex. In all cases, these methods find the most likely pure pixels by an approximate method, but they do not estimate the smallest simplex (or the convex hull) directly. All these methods are based on the assumption that pure pixels do exist in the measured dataset, which of course may not be the general case for most of the natural systems and situations. More recently MVSA (minimum volume simplex analysis) [12] which is based on geometrical image analysis [13] and AMEE (Automated Morphological Endmember Extraction) [14] which is based on morphological image analysis have been reported too. The MVSA method has been proposed to cope with the situation where no pure pixel exists in the measured image. It is also based on the concept of the minimum simplex volume estimation, unmixing the image by fitting a minimum volume simplex to the data, constraining the abundance fractions to be positive and belong to the most probable simplex. It is a fast method, but, in its present implementation, it does not guarantee non-negative spectra, which would not have physical nor chemical sense.

Multivariate Curve Resolution Alternating Least Squares (MCR-ALS) has been proposed and extensively used to resolve multiple pure responses and concentrations of the components present in unknown mixtures [15]. It has been applied to analyze multicomponent chemical systems like chemical reactions [16], industrial processes [17], chromatographic coelution problems [18], spectroscopic mixtures [19], environmental monitoring data [20], and it can be applied to many other type of mixture analysis cases. MCR-ALS has been reported also to be a useful method for the resolution and segmentation of hyperspectral biomedical and other type of hyperspectral imaging [21–25]. In MCR-ALS, the measured analytical signals are assumed to follow a generalized bilinear additive model (like the extension of Beer's law in absorption spectroscopy [26]). The contribution of each component to the measured signal depends on its concentration and on its own spectral sensitivity response (pure spectrum). MCR-ALS can also be applied to obtain quantitative information and it provides physical and chemical meaningful solutions. This is accomplished because in MCR-ALS, any type of constraints can be easily applied to the sought solutions, like non-negativity [27,28], unimodality [29,30], local rank [31,32], and trilinearity [29,33].

In this work, three different spectrsocopic images have been analyzed to test the MCR-ALS method and to compare its results with those obtained by VCA and MVSA methods. Two of these images are simulated data sets and the third one is an experimental remote sensing hyperspectral airborne image from the Cuprite area in Nevada (USA). Finally, the MCR-BANDS method [34] is applied to evaluate the amount of rotation ambiguity associated to the image resolved pure spectra and concentration profiles under the constraints imposed during their estimation.

2. Methods

2.1. MCR-ALS method

The MCR-ALS method is used to decompose the hyperspectral image into the signatures or pure spectra of the image constituents and into their concentration (relative amounts) on the image (distribution map). MCR-ALS is based on a bilinear model which assumes that the observed spectra are a linear combination of the spectra of the pure components in the system [21,26]. This model can be written in matrix form as:

$$\mathbf{D} = \mathbf{C}\mathbf{S}^{\mathrm{T}} + \mathbf{E} \quad (1)$$

where **D** is the reshaped image data matrix. **C** is the matrix of the relative amounts or concentrations, $\mathbf{S}^{\mathrm{T}}$ is the pure spectra. **E** is the matrix associated to noise or experimental error.

In hyperspectral imaging using visible/infrared spectroscopies, the intensity or absorption of the radiation in every pixel should not be negative and neither should be negative the concentrations of the different constituents of the image (physical constraints). Thus, during the alternating least squares procedure previously described, non-negative constraints are applied on both, on the pure spectra and on the image concentrations. Moreover, to avoid scale indeterminacies and stabilize the ALS iterative optimization, spectra matrix $\mathbf{S}^{\mathbf{T}}$ is normalized. Instead of spectra normalization, an alternative possible constraint to be used to scale adequately the obtained solutions is a mass balance or closure condition on C matrix, i.e., that concentrations or relative amounts of the constituents on the pixels sum equal to a constant value. This is for