NATIONAL TECHNICAL UNIVERSITY OF ATHENS
SCHOOL OF RURAL AND SURVEYING ENGINEERING

DEVELOPMENT OF SPECTRAL UNMIXING METHODS FOR HYPERSPECTRAL DATA EXPLOITATION

Doctoral Thesis

CHAROULA TH. ANDREOU
Rural and Surveying Engineer N.T.U.A.

SUPERVISOR:
Dr. V. Karathanassi
Associate Professor of N.T.U.A.

March, 2014

NATIONAL TECHNICAL UNIVERSITY OF ATHENS

SCHOOL OF RURAL AND SURVEYING ENGINEERING

DEVELOPMENT OF SPECTRAL UNMIXING METHODS FOR HYPERSPECTRAL DATA EXPLOITATION

Doctoral Thesis

CHAROULA TH. ANDREOU
Rural and Surveying Engineer N.T.U.A.

EXAMINATION COMMITTEE:
1. Dr. V. Karathanassi, Associate Professor of N.T.U.A.
2. Dr. D. Argialas, Professor of N.T.U.A.
3. Dr. P. Maragos, Professor N.T.U.A.
4. Dr. K. Koutroumbas, Senior researcher of N.O.A.
5. Dr. N. Doulamis, Assistant Professor of N.T.U.A.
6. Dr. D. Soudris, Assistant Professor of N.T.U.A.
7. Dr. K. Karantzalos, Lecturer of N.T.U.A.

March, 2014

Abstract

Hyperspectral sensors provide high spectral resolution which contributes to material identification. However, due to low spatial resolution, pixels may contain a mixture of more than one distinct material spectrum. In order to exploit the mixed hyperspectral data, spectral unmixing is applied, which involves the decomposition of the mixed spectra into constituent spectra, also called endmembers, weighted by their corresponding fractional abundances. This thesis addresses the spectral unmixing problem so that the full potential of hyperspectral data exploitation to be employed. The main contributions include the exploitation of novel concepts and the development of five new methods dedicated to the steps involved in spectral unmixing; the signal subspace estimation, the dimensionality reduction, the endmember extraction, and the abundance estimation.

The first method, called *outlier detection method* (ODM), is a new automatic non-parametric method for estimating the signal subspace dimension. The number of the signal vectors is much lower compared to the number of the noise vectors. Thus, estimating the population distribution and/or its statistical characteristics could comprise errors. The novelty of ODM lies in the fact that it considers only the existence of noise and treats signals as outliers of noise. It searches for the signals whose radius is by far larger than the one of the noise introducing for the first time in virtual dimension theory, a robust outlier detection method. The ODM achieves the performance of its competitors and outperforms them in case of small image scenes.

The second method is a new band selection (BS) approach for optimizing the performance of the endmember extraction and classification. The few existing BS methods which address spectral unmixing issues set fixed criteria to the spectral information on the whole set of wavelengths. They are based on generalised approaches which disregard the spectral characteristics of a particular material of interest, the image diversity and the endmember variability. The proposed method enables endmember extraction and classification algorithms to act locally in the hyperspectral space. It tries to define subspaces in which spectra of materials vary the most. In case of endmember extraction, the new method accounts for spectrally closed endmembers (SCEs) – a term introduced in the frame of this study- which are likely overshadowed by the prevalent endmembers, by exploiting the original bands. In order to extract the SCEs, the concept of multiple convex hulls is used for the first time in BS. The proposed approach is effective at detecting low contrast materials, which imply different biophysical or chemical properties of a material class. In case of classification, the proposed BS method accounts for subspaces where classes are separable. It achieved the same overall accuracy as if the whole band set is used, by selecting less than 50% of the total bands.

The third method, called *simple endmember extraction* (SEE) is a new convex-based endmember extraction method which searches for the most distinct spectra at the vertices of a simplex defined in the signal subspace. The novelty of the SEE method is that it searches for the extreme values that lie on the end points of the existing transformed axes without further projections that imply iterative procedures. Thus, it bears the advantage of simultaneously extracting endmembers with low computational cost. The fourth method is an enhanced version of SEE, called enhanced-SEE (E-SEE) and it is an empirical method which compensates the tendency that the majority of the convex-based endmember extraction methods encounter to select high contrast endmembers over less contrast endmembers. On this account, it changes the distribution of the initial data sample by increasing the distance between candidate endmembers and the data mean, a technique which has never been explored before. Both methods provided the same and even better performance compared to the state-of-the-arts methods. The computation complexity of the new methods is much lower than those of vertex component analysis (VCA) and N-FINDR, approximately one and three orders of magnitude lower, respectively.

Last but not least, the fifth method is a new multiple endmember spectral signature analysis (MESMA) based on spectral angle distance, called MESMA-SAD. MESMA accounts for within class spectral variability, however, it needs to calculate all the potential endmember combinations of each pixel to find the best-fit one, demanding a time-consuming unmixing technology. The new non-parametric method significantly minimizes the time-processing compared to the existing MESMA algorithms by combining the spectral angle distance values and the mean absolute errors.

Extensive simulated and real image-based experiments indicate the effectiveness of the proposed methods to improve the data exploitation, rendering their implementation very promising in hyperspectral image processing.

Περίληψη

Η πολύ υψηλή φασματική ανάλυση των Υπερφασματικών αισθητήρων συντελεί στο να είναι δυνατή η αναγνώριση υλικών/αντικειμένων τα οποία βρίσκονται στο έδαφος. Ωστόσο, λόγω της χαμηλής χωρικής ανάλυσης τα εικονοστοιχεία της απεικόνισης είναι πιθανόν να περιέχουν φασματική υπογραφή η οποία προέρχεται από μίξη περισσότερων του ενός υλικού. Για να είναι δυνατή η ανάλυση της μικτής υπερφασματικής πληροφορίας, εφαρμόζεται ο φασματικός διαχωρισμός ο οποίος περιλαμβάνει την ανάλυση του μικτού φάσματος σε κύριες φασματικές συνιστώσες ή αλλιώς καθαρές φασματικές υπογραφές, όπου κάθε μία αντιστοιχεί σε ένα μοναδικό υλικό, και στα αντίστοιχα ποσοστά συμμετοχής τους. Η συγκεκριμένη διατριβή πραγματεύεται το φασματικό διαχωρισμό με σκοπό να αξιοποιηθούν πλήρως οι πληροφορίες που εμπεριέχονται στα υπερφασματικά δεδομένα. Τα κύρια σημεία συμβολής της διατριβής περιλαμβάνουν την εισαγωγή καινοτόμων ιδεών στο πεδίο της επεξεργασίας της υπερφασματικής απεικόνισης και την ανάπτυξη πέντε νέων μεθόδων, οι οποίες σχετίζονται με τα στάδια τα οποία απαρτίζουν το φασματικό διαχωρισμό και είναι τα εξής: εκτίμηση της διάστασης του υπόχωρου του σήματος, μείωση της υπερφασματικής διάστασης, εξαγωγή καθαρών φασματικών υπογραφών και εκτίμηση των ποσοστών συμμετοχής τους.

Η πρώτη μέθοδος καλείται *outlier detection method* (ODM). Πρόκειται για μία μη επιβλεπόμενη, μη παραμετρική μέθοδο για την εκτίμηση της διάστασης του υπόχωρου του σήματος. Ο αριθμός των διανυσμάτων που αντιστοιχούν σε σήμα είναι αρκετά μικρότερος από τον αντίστοιχο του θορύβου. Εξαιτίας αυτού, η στατιστική ανάλυση των διανυσμάτων του σήματος μπορεί να επιφέρει σφάλματα. Η καινοτομία του ODM έγκειται στο ότι θεωρεί μόνο την ύπαρξη του θορύβου και μεταχειρίζεται τα σήματα ως ακραίες τιμές του θορύβου. Η μέθοδος εντοπίζει τα σήματα μέσω της μεθόδου ανίχνευσης των ακραίων τιμών, μία προσέγγιση που δεν έχει εφαρμοστεί πρωτύτερα στο συγκεκριμένο θέμα. Η μέθοδος ODM επιτυγχάνει ακρίβεια όμοια με αυτήν των υπαρχουσών μεθόδων, ενώ παρουσιάζει καλύτερα αποτελέσματα στην εφαρμογή της, σε απεικονίσεις μικρού μεγέθους.

Η δεύτερη μέθοδος είναι μία νέα προσέγγιση επιλογής καναλιών για τη βελτιστοποίηση της εξαγωγής καθαρών φασματικών υπογραφών και της διαδικασίας ταξινόμησης. Οι λιγοστές υπάρχουσες μέθοδοι επιλογής καναλιών οι οποίες απευθύνονται αποκλειστικά στο φασματικό διαχωρισμό εφαρμόζονται στο σύνολο της διαθέσιμης φασματικής πληροφορίας χωρίς να λαμβάνουν υπόψη τα φασματικά χαρακτηριστικά κάθε υλικού/αντικειμένου ξεχωριστά και τη φασματική διαφοροποίηση σε κάθε απεικόνιση. Η νέα μέθοδος καθιστά δυνατή την εφαρμογή αλγόριθμων ανίχνευσης καθαρών φασματικών υπογραφών και ταξινόμησης σε τοπικό επίπεδο στον υπερφασματικό χώρο. Προσπαθεί να βρει εκείνους τους φασματικούς υπόχωρους όπου τα υλικά διαφοροποιούνται στο μέγιστο. Όσον αφορά στους αλγορίθμους ανίχνευσης καθαρών υπογραφών, η νέα μέθοδος επικεντρώνεται στον εντοπισμό

των φασματικά κοντινών καθαρών φασματικών υπογραφών, οι οποίες επισκιάζονται από τις φασματικές υπογραφές υψηλής φασματικής αντίθεσης, αξιοποιώντας για πρώτη φορά στην επιλογή καναλιών την έννοια των πολλαπλών κυρτών περιοχών. Η μέθοδος ανιχνεύει αποτελεσματικά τις φασματικές υπογραφές του υπό μελέτη υλικού οι οποίες εμπεριέχουν χρήσιμες πληροφορίες για τη βιοφυσική/χημική ιδιότητα του. Όσον αφορά στην ταξινόμηση, η νέα μέθοδος επικεντρώνεται στην εύρεση εκείνων των φασματικών υπόχωρων όπου οι κατηγορίες είναι διαχωρίσιμες στο μέγιστο. Η προτεινόμενη μέθοδος, μειώνοντας κατά 50% το πλήθος των καναλιών, πέτυχε την ίδια ακρίβεια ταξινόμησης όπως στην περίπτωση που χρησιμοποιήθηκαν όλα τα κανάλια.

Η τρίτη μέθοδος καλείται *simple endmember extraction* (SEE) και πρόκειται για μία νέα μέθοδο εξαγωγής καθαρών φασματικών υπογραφών. Βασίζεται στη γεωμετρική ανάλυση των κυρτών περιοχών. Η καινοτομία της έγκειται στο ότι αναζητά τις κορυφές της χωρικής διάταξης (simplex), οι οποίες βρίσκονται στις άκρες των μετασχηματισμένων αξόνων, χωρίς περαιτέρω προβολές και επαναληπτικές διαδικασίες. Συνεπώς, η μέθοδος έχει το πλεονέκτημα της ταυτόχρονης εξαγωγής καθαρών φασματικών υπογραφών, ενώ είναι απαλλαγμένη από το αυξημένο υπολογιστικό κόστος. Η βελτιωμένη εκδοχή της SEE καλείται enhanced-SEE (E-SEE) και είναι μία εμπειρική μέθοδος, η οποία αντισταθμίζει την τάση των υπαρχουσών μεθόδων εξαγωγής καθαρών φασματικών υπογραφών να ανιχνεύουν τις φασματικές υπογραφές υψηλής αντίθεσης εις βάρος εκείνων με χαμηλή αντίθεση. Αυτό επιτυγχάνεται αυξάνοντας την απόσταση μεταξύ υποψήφιων καθαρών φασματικών υπογραφών και της μέσης τιμής της εικόνας, γεγονός που επιφέρει αλλαγή στην κατανομή των δεδομένων. Οι δύο μέθοδοι παρουσίασαν αποτελέσματα ίδιας και καλύτερης ακρίβειας σε σύγκριση με τις αντίστοιχες μεθόδους, ενώ το υπολογιστικό τους κόστος είναι χαμηλότερο από το αντίστοιχο των μεθόδων VCA και N-FINDR κατά μία και τρεις τάξεις μεγέθους, αντίστοιχα.

Η πέμπτη προτεινόμενη μέθοδος είναι μία νέα εκδοχή της μεθόδου multiple endmember spectral signature analysis (MESMA), η οποία βασίζεται στη φασματική γωνία και καλείται MESMA-SAD. Η πρωτότυπη μέθοδος MESMA αν και επιτρέπει ο αριθμός των καθαρών φασματικών υπογραφών να είναι διαφορετικός για κάθε εικονοστοιχείο, χαρακτηρίζεται υπολογιστικά ακριβή, καθώς εξετάζονται όλοι οι πιθανοί συνδυασμοί καθαρών φασματικών υπογραφών για την εύρεση του βέλτιστου συνόλου. Η νέα μη παραμετρική μέθοδος μειώνει σημαντικά το χρόνο επεξεργασίας αξιοποιώντας τη φασματική γωνία και την τιμή του μέσου απόλυτου σφάλματος.

Η αξιολόγηση των προτεινόμενων μεθόδων ως προς την αποδοτικότερη αξιοποίηση των υπερφασματικών δεδομένων υλοποιήθηκε μέσω εκτεταμένων πειραμάτων σε συνθετικές και πραγματικές απεικονίσεις. Τα αποτελέσματα της αξιολόγησης ήταν πολύ ικανοποιητικά γεγονός που καθιστά τις μεθόδους πολλά υποσχόμενες στο πεδίο της επεξεργασίας των υπερφασματικών απεικονίσεων.

Acknowledgements

To Victor

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Hyperspectral remote sensing

Generally, remote sensing can be defined as the collection of information about an object without making physical contact with it. The sense of sight enables the humans to perceive the surroundings without the need of close contact with the objects. Being visual species, we have evolved the ability to understand the world around us through visual images [1]. Broadly speaking we are performing remote sensing in everyday life. In more restricted sense, remote sensing refers to the technology of collecting information about the Earth's surface and atmosphere using sensors which record electromagnetic radiation. The advantages of this technical phenomenon derive from the fact that remote measurements are related to the familiar realm of what we see on the ground, rendering its use urgent for understanding our changing environment. Large areas of the Earth's surface, among of which locations which are difficult to be approached, or large volume of atmosphere can be recorded in a short pace of time [1], [2]. The applications of remote sensing are enormous; ranging from monitoring environmental changes and evaluating natural resources to military operations. In the 1980's, the convergence of two related but distinct technological fields; spectroscopy and remote imaging led to the development of hyperspectral remote sensing, also called imaging spectroscopy [3], [4]. Since then, the evolution of passive remote sensing has witnessed the collection of measurements with significantly greater spectral

breadth and resolution comparable with the one of laboratory spectrometers. Hyperspectral imaging is produced by complex sensors, called imaging spectrometers, which use hundreds of detectors recording over a wide wavelength range, from visible region through the near-infrared and short-infrared spectrum (typically at least 0.4 to 2.5 μm). In other words, within the sensor the received energy is transformed into an electric impulse which in turn is translated into a data volume. This data volume is a 3-dimensional structure, also called *data cube*, with two spatial dimensions and one spectral dimension. Each plane of the data cube comprises reflected radiance at one spectral band for all pixels, while each pixel column of the data cube is a spectral vector comprising the values of the reflected radiance at the whole wavelength (Fig. 1.1). If we extract a pixel column and plot its spectral values as a function of wavelength, the result is the average spectrum of all the materials in the corresponding ground resolution cell at every wavelength.

Different materials reflect and emit varying amounts of radiance across the electromagnetic spectrum. The reflected light from the surface interacts in a unique way for each material and even slight variations of it are quantified and characterized through hyperspectral imagery [5]. In other words, each material generally has a unique spectral signature which can be treated as its *fingerprint*. Therefore, hyperspectral sensors can be used to identify and distinguish between different materials in a scene [6]. Acquiring the same scene using hundreds of narrow and contiguous spectral bands (on the order of 5nm), hyperspectral images offer an invaluable diversity of information, enabling the accurate physical description and discrimination of the sensed materials. Over the past decades, hyperspectral remote sensing became more and more popular, with a blooming number of available sensors, for airborne or spaceborne platforms. Table 1.1 presents a subset of well known present and future hyperspectral sensors.

The wealth of spectral information available from advanced hyperspectral imaging has opened new perspectives in many application domains, from food safety, precision agriculture and biomedical analysis, to the monitoring of the environment, urban areas and natural hazards, mineral exploration, forestry or defence related issues [5], [7], [8], [9], [10].
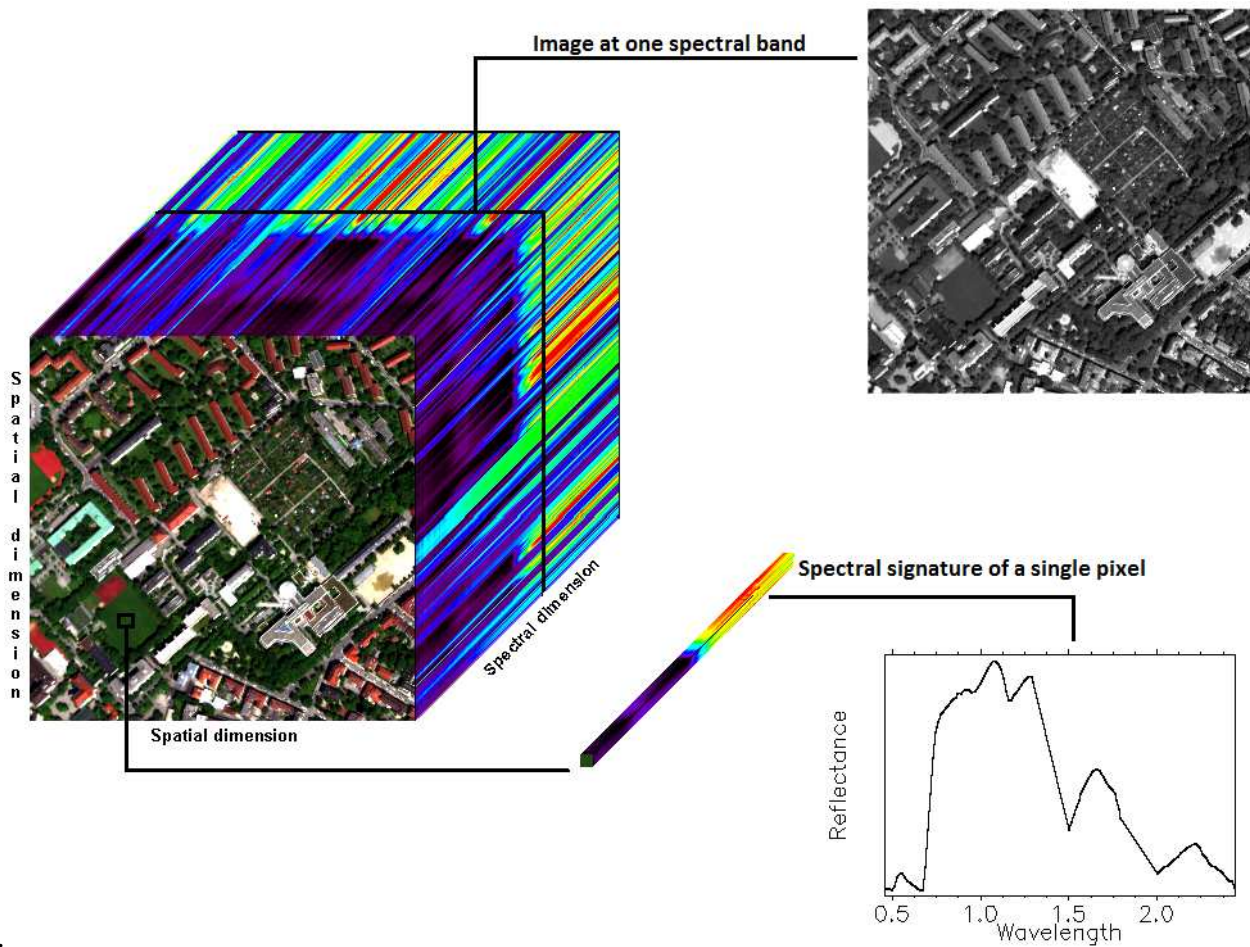
Figure 1.1: Hyperspectral image structure.

Table 1.1: Overview of present and future hyperspectral sensors

Sensor	Platform	Spectral range (μm)	No. of bands	Ground pixel size (m)	Ground swath (km)
AISA[1]	Airborne	0.4-2.5	500	1	0.4
AVIRIS[2]	Airborne	0.4-2.5	224	20	11
CASI[3]	Airborne	0.4-1.1	288	1	1.4
EnMap[4]	Spaceborne	0.42-2.4	232	30	30
HYDICE[5]	Airborne	0.4-2.5	210	3	1
HyMAP[6]	Airborne	0.45-2.5	126	5	2.3
Hyperion[7]	Spaceborne	0.4-2.5	200	30	7.5
HySpex[8]	Airborne	0.4-2.5	316	0.4-7	0.5

---

[1] http://www.spectralcameras.com/aisa

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

[3] http://www.itres.com/products/imagers/casi1500/

[4] http://www.enmap.org/sensor

[5] http://rsd-www.nrl.navy.mil/hydice

[6] http://www.hyvista.com/

[7] http://eo1.usgs.gov/sensors/hyperion

[8] http://www.hyspex.no/

There are many ways to extract information from a spectral image. It would not be wise to anticipate the existence of any single technique that can outperform all others in every practical situation. It depends on the nature of the problem and the available a priori knowledge related to it. Images that fuse spatial, spectral and/or temporal information are needed to be exploited using flexible strategies [1]. As it was mentioned above, hyperspectral images are hyper-dimensional, each pixel has a spectral signature and it is represented as a vector with several hundreds of components. Exploiting these data is very challenging and specific image and signal processing methods have been developed [11]. The existing approaches are related to one of the following processing tasks [6], [5], [12], [13]:

Atmospheric correction:

Atmosphere is responsible for solar radiation's scattering and absorption and affects the radiance recorded by the detectors. Since the radiant energy may not interact with the Earth's surface, it provides no physical meaning in terms of the spectral measurement and the latter should be corrected. Atmospheric correction is a daunting process because the atmosphere's properties vary through space and time. One of the most widely applied approaches is the conversion of radiance to reflectance data. It should be noted that the atmospheric correction process is optional and might not be required depending on the fidelity of the data required.

Dimensionality reduction:

High dimensionality is inherent to hyperspectral data. The trade-off of high spectral resolution which allows for the extraction of material spectra is the volume of data which poses many challenging problems in terms of data storage, computational efficiency, algorithms performance. Therefore, dimensionality reduction techniques aim at defining a new spectral space of much lower dimensionality than the number of the bands of the input hyperspectral data.

Target/anomaly detection:

Target detection aims at identifying a relatively small number of objects with fixed shape or spectrum in a scene. It is possible to detect the targets of interest in an unknown image scene from which the desired target knowledge can be generated directly from it in an unsupervised way using either matched filtering or quadratic forms. In case there is no a priori information to specify the statistics of the target, anomaly detection is applied. Anomaly detector searches for targets which are generally unknown and spectrally distinct from their surroundings, but relatively small with low probabilities of occurrence in an image scene using the statistics derived from the whole image.

Change detection:

Acquiring two or more hyperspectral images of the same geographic location at different times enables the detection of changes of objects of interest which may not be distinguishable at the visible spectrum. Most change detection methods require precise spatial registration between the available images and atmospheric correction in order that pixel to pixel comparison to be effective [13].

Classification:

Classification is a pixel-wise technique which identifies the material class that a pixel spectrum mostly resembles. Classes may or may not be predefined leading to supervised and unsupervised methods, respectively. The output of this process is a classification map where each pixel is assigned to one class label. The classification techniques can be based on various mathematical concepts such as statistical analysis, neural networks, morphology-based approaches, hierarchical segmentation [12], [14].

Spectral unmixing:

Spectral unmixing contributes to the extraction of information at sub-pixel level. It consists of detecting the distinct spectra in the hyperspectral scene which may represent materials and estimating their

apparent quantification in a pixel in terms of fraction. The spectral unmixing is an inversion problem based on the spectral mixture analysis which is addressed in the next session.

1.2 Spectral mixture analysis

Hyperspectral sensors provide high spectral resolution which contributes to material identification. However, due to their low spatial resolution and multiple scattering of light, sensors record mixtures of spectra of materials in the scene [12]. Spatial resolution of the hyperspectral sensors may range from few centimetres to tens of meters depending on the flight altitude. Consequently, pixels may contain a mixture of more than one distinct material spectrum and therefore, they are called *mixed pixels*. Regardless the spatial resolution, mixed pixels may arise due to intimate mixtures existing on the ground [15]. For instance, homogeneous areas of minerals quite often appear in mixtures such as sulphur which may present key absorption features in the spectral signature of another mineral.

In order to be able to exploit the mixed data derived from hyperspectral imaging we need to apply physically based models which connect spectroscopy of materials with remote measurements. Spectral mixture analysis (SMA) [16] is a technique for applying such models to spectral images. According to it, pixel spectrum can be expressed as a mixture of constituent spectra, also called *endmembers*. A large proportion of the relevant literature [17], [18] refers to endmember as a pure pixel, comprising a unique material. Nevertheless, it should be pointed out that definition of endmembers is more complicated since in real applications pure pixels are unlikely to be found, not only due to spatial resolution constrains but also due to the spectral signature contamination of noise, unknown interferes etc. According to authors in [1], endmembers are spectra which may represent materials. Depending on the spatial resolution and the particular application, endmembers for the same scene may be spectra of pure materials or spectra of mixtures of materials at a more fundamental scale [1], [12]. In this thesis, endmember is referred to the dominant distinct spectrum which is constituent part of a spectral mixture.

The process of decomposing the pixel spectrum into endmembers and estimating their proportions, also called *abundances*, is known as *spectral unmixing*. As it was mentioned above, spectral unmixing is an

inversion problem based on spectral mixture analysis. In fact, spectral unmixing is a very active research topic. Its enormous potential has aroused the interest of the relevant scientific community since the beginning of hyperspectral sensor development [16], [19]. This thesis focuses on spectral unmixing and in particular, on developing novel approaches for its optimization.

1.2.1 Linear and nonlinear spectral unmixing

There are two models of spectral mixture analysis; linear and nonlinear [16], [19]. Their difference lies on how the electromagnetic spectrum of a mixed pixel is considered to be recorded. The linear mixture model assumes that incident light interacts only with one material in the field of view and thus, the recorded spectra at the spectrometer can be expressed in the form of a linear combination of endmembers weighted by their corresponding abundances. The nonlinear mixture model assumes that radiance is modified by one material before interacting with another one leading to a random distribution of endmembers [6], [12], [15].

Theoretically, both linear and non-linear models can be effective for interpreting the mixed spectra, depending on the nature of the materials on the ground and the spatial resolution. In practice, spectral unmixing based on nonlinear model is an extremely complex ill-posed problem since it requires complex physical models which infer the dominant spectra and their proportions based on radiative transfer theory [12], [20]. Although there has been an attempt to exploit simpler physically based models for nonlinear unmixing, development of a fully unsupervised method seems a daunting essay. Ideally, the applied physical model should account for all the processes and factors, such as illumination geometry and atmospheric effects, which are associated with the observed spectrum. Since this is not attainable, a good estimate is satisfactory.

In this thesis, focus was given on spectral unmixing based on linear mixture model which is the most widely used model, adequate for unsupervised approaches. The interest of decomposing the observed spectra based on linear mixture model remains unceasing for the last 25 years. Among the advantages of linear unmixing is its effectiveness in a wide range of applications due to its acceptable approximation

of light scattering mechanics in many real scenarios [12]. Indeed, in macroscopic scale which is usually the scale of hyperspectral remote sensing, linear mixing dominates for most scenes because the amount of intimate mixing at patch boundaries is negligible [1], [12]. Another advantage of linear spectral mixture is its clear conceptual meaning which contributes to a computationally simple implementation and an easy interpretation.

Consider that if L is the total number of bands, each observed spectral vector y, $y \in \Re^L$, can be expressed under the linear spectral mixture model, as linear combination of constituent deterministic spectral signatures s_k and their corresponding abundances a_k:

$$y = \sum_{k=1}^{p} a_k s_k + n = S\alpha + n \tag{1.1}$$

where the $L \times p$ matrix $S = [s_1,...,s_P]$ comprises the endmember spectra, the $p \times 1$ vector $\alpha = [a_1,...,a_P]$ their corresponding abundances and $n \in \Re^L$ represents the additive noise.

Two abundance constraints are generally imposed on the model described in eq. (1.1) in order to assure physically meaningful results:

$$\sum_{k=1}^{p} a_k = 1 \qquad \text{abundance non-negativity constraint (ANC)}$$

$$a_k \geq 0 \qquad \text{abundance sum-to-one constraint (ASC)}$$

The geometrical interpretation of the linear mixture model is associated with the mathematical theory of *convex sets* [6], [21], [22]. More precisely, assuming that both constraints are satisfied and noise variance is finite, spectra can be restricted in a simplex lying on a signal subspace of dimension one less than the number of the endmembers.

The only parameter which is given in eq. (1.1) is the observed spectral vector y or, in terms of an image, a set $Y = [y_1,...y_m]$ of m observed L-dimensional spectral vectors. The objective of spectral unmixing is to estimate the matrix S of endmembers and the matrix α of abundances for each pixel.

Since the number of bands is usually larger that the number of the endmembers, a common approach for abundance estimation is the least squares minimization:

$$\min_{a} \|Sa - y\|_2^2 \text{ subject to ANC and ASC} \tag{1.2}$$

1.2.2 Spectral unmixing process

The spectral unmixing process includes several steps, which are crucial for the accuracy of the resulted estimations. Fig. 1.2 shows schematically the outline of spectral unmixing. The first step contains atmospheric correction which leads to the conversion of radiance values to reflectance values. As it has already been noted, this step is optional and linear unmixing may be conducted on radiance values as well. The second step contains the estimation of the number of distinct signals (endmembers) p, which exist in the hyperspectral dataset. An accurate determination of the number of the endmembers infers the signal subspace and significantly contributes to the accuracy of the spectral unmixing. The third step is related to the dimensionality reduction. The number of endmembers is much lower than the number of bands L. Based on the linear spectral mixture analysis the spectral vectors lie on a linear subspace whose dimension is one less than the number of endmembers, i.e. p-1. Thus, dimensionality reduction techniques contribute to challenging problems such as data storage and computational efficiency. The forth step contains the endmember extraction leading to the spectral signatures of the endmembers and finally, last step contains the abundance estimation using a least square method which leads to abundance maps each one assigned to an endmember showing its quantification at each pixel. The two final steps can be implemented sequentially or simultaneously. In the latter case, where no information is provided regarding the endmembers, is more proper to refer to *endmember determination* rather than endmember extraction [13]. As a matter of fact, in such case all the available data sample vectors are considered to be candidate endmembers, and via least square minimization the optimum set of spectral signatures is determined as endmembers.
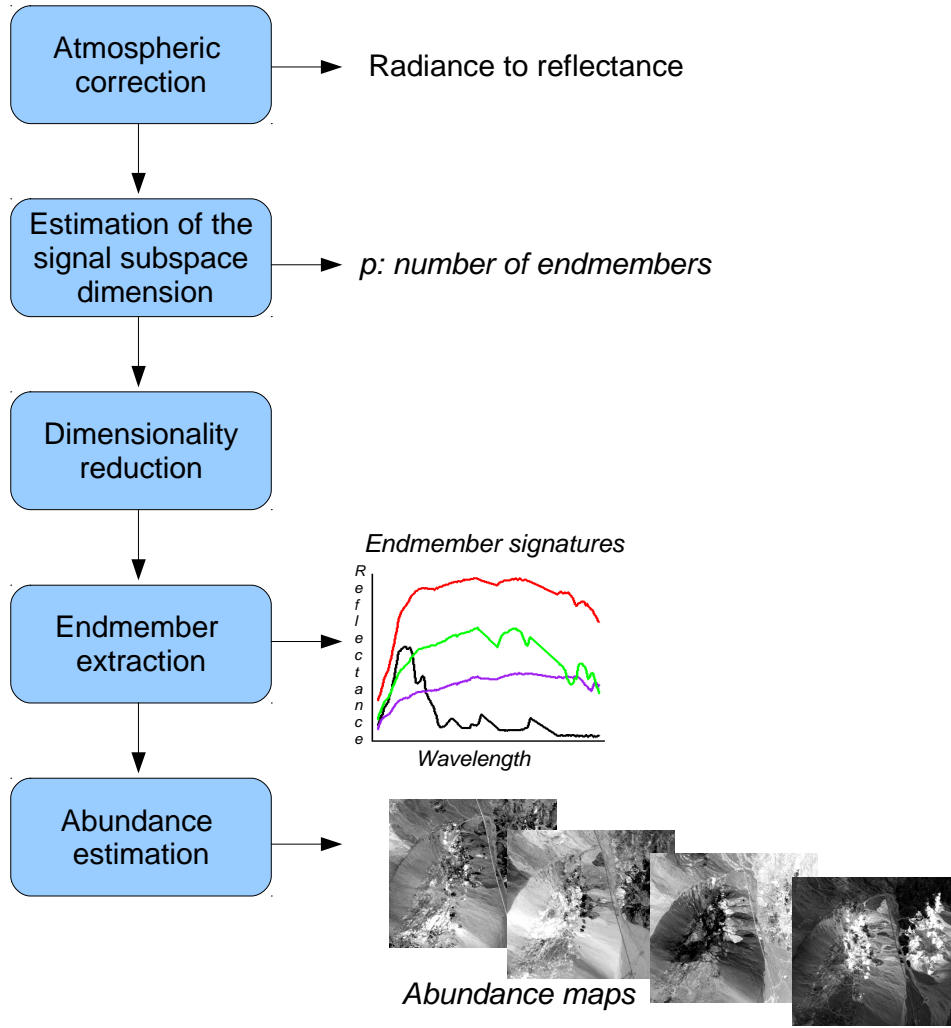
Figure 1.2: Schematic outline of spectral unmixing.

1.3 Motivation

Over the past few years, hyperspectral remote sensing has intrigued scientists from vast disciplinary areas as the potential applications on which it can be implemented are multifarious, ranging from climate change research, natural hazard risk assessment, geological, forestry and urban environment mapping to military services, and there are plenty more to be explored. Hyperspectral sensors provide high spectral resolution which contributes to material identification. However, due to low spatial resolution, multiple scattering of light, and intimate mixtures, sensors record mixtures of spectra of materials in the scene and consequently, pixels may contain a mixture of more than one distinct material spectrum. In order to exploit the mixed hyperspectral data, spectral unmixing should be applied, which provides information contained at sub-pixel level and leads to an efficient overall representation of the

10

materials. On this account, hundreds of new algorithms have been developed from the fast-growing relevant scientific community to optimize spectral unmixing process. Nevertheless, spectral unmixing is still the most fundamental problem in hyperspectral data exploitation since there are several issues which need to be addressed so that its full potential to be employed.

The first issue concerns the estimation of the signal subspace dimension. It has been a long-standing daunting task whose unresolvedness prevents the development of unsupervised methods in terms of dimensionality reduction, endmember extraction, target detection and classification. Estimation of the correct number of endmembers has significant impact on the performance of the endmember extraction algorithms (EEAs) and consequently on the accuracy of the spectral unmixing process. Furthermore, based on linear spectral unmixing, the number of endmembers is associated with the optimal number of dimensions to be retained after dimensionality reduction in order to represent the dataset. Hence, an accurate determination of the number of the endmembers significantly contributes to the accuracy of the spectral unmixing process and enables low-dimensional representation of spectral vectors. The existing methods for estimation of the number of endmembers / signal subspace dimension consider the existence of two different distributions, the one related to noise and the other related to signal, or in geometrical approach they consider two different subspaces one of noise and one of signal. However, in hyperspectral space, signal vectors are very few in order to estimate their population distribution properly or to statistically analyze them and thus, this thesis focuses on designing new approaches to overcome this shortcoming.

Another issue arises in how to reduce the high dimensionality of the hyperspectral data defining the observed spectra into a lower dimension in order to be analyzed more effectively, yielding gains in computational time and complexity, data storage and signal-to-noise ratio. Dimensionality reduction (DR) is widely used as a pre-processing step of great interest as it affects directly the performance of the following processing tasks. Band selection (BS) is an efficient DR approach which searches for a subset of the original spectral channels retaining the physical meaning of the data set, based on a criterion for optimality. There are a few BS methods which have been addressed to spectral unmixing. These

methods set criteria to the spectral information derived by the whole set of wavelengths; a generalised approach which disregards the spectral characteristics of a particular material of interest, the image diversity and the endmember variability. BS approaches existing in the literature do not account for the aforementioned matters. In this thesis a new BS has been designed which is material-based.

The third issue concerns the endmember extraction and the argument over the optimum way to be accomplished. Depending on the spatial resolution and the particular application, endmembers of the same scene may be spectra of pure materials or mixture of spectra at a more fundamental scale [1], [12]. Therefore, it should be wiser to consider the endmember as the dominant distinct spectrum which is constituent part of a spectral mixture. On this account, extracting endmembers from the image can be promising; image endmembers have the advantage of no calibration need, refraining from adding calibration errors in the mixture model, as well as they have the same scale of measurement as the data. Furthermore, according to author in [13], an optimal EEA must select all the endmembers at the same time rather than sequentially. Another concern related to endmember extraction is the tendency of the majority of the convex-geometrical based EEAs to select high contrast endmembers over less contrast endmembers. In this thesis, new simultaneous algorithms for the extraction of the optimum set of reliable endmembers are designed which could contribute to spectral unmixing optimization.

Abundance estimation yields the last output of spectral unmixing which represents the proportion of each endmember in the pixel. Multiple endmember spectral mixture analysis (MESMA) [23] is an effective method which accounts for within class spectral variability, however, it needs to calculate all the potential endmember combinations of each pixel to find the best-fit one, demanding a time-consuming unmixing technology. The growing demand of hyperspectral image processing on real-time applications sets the development of new spectral unmixing methods, comprising EEAs and abundance estimation techniques of low computational complexity, more urgent than ever. Hence, the new methods should be of low computational burden.

1.4 Main Contributions of the Thesis

In wide sense, the particular thesis contribution lies on the optimization of the spectral unmixing process for the hyperspectral imagery. In more restricted sense, the main contributions include the exploitation of novel concepts and the development of five new methods dedicated to the signal subspace estimation, the dimensionality reduction, the endmember extraction and the abundance estimation.

The first method is a new automatic non-parametric method for estimating the signal subspace dimension and, by extension, the number of the endmembers. Its novelty lies in the fact that it considers only the existence of noise and treats signals as outliers of noise. No estimation of statistical distributions is required. The new method, called *outlier detection method* (ODM), explores the geometrical properties of the noise hypersphere. It searches for the signals whose radius is by far larger than that of the noise, introducing for the first time in virtual dimension theory a robust outlier detection method. Parts of the particular research have been published in [24], [25].

The second method is a new band selection (BS) approach for optimizing the performance of the endmember extraction and classification. BS methods often disregard the small differences in spectral signatures which may imply crucial information. In this thesis, it is given an attempt to address the detection of spectrally closed endmembers (SCEs) – a term introduced in the frame of this study- by exploiting the original bands. This approach has been explored before. In order to extract the SCEs, the concept of multiple convex hulls [26] is used for the first time in BS. Contrary to the existing BS methods, the proposed approach accounts for the different distribution of each material's convex hull to the data cloud. For this reason, in terms of endmember extraction, the proposed BS approach allows the final number of endmembers assigned to a material class to vary, resulting in more physically meaningful spectra. In terms of classification, the proposed BS approach results in selected bands of much lower number than the number of the original channels yielding gains in the required training data set quantity. Furthermore, the proposed BS accounts for image diversity and is tied to the available

information derived from the image's spectral characteristics. Thus, different sets of bands are exploited for detecting different materials depicted in the same image scene. Being image-dependent compensates the endmember variability contrary to the existing approaches which use a priori information related to the absorption features at fixed wavelengths. Parts of the particular research have been published in [27], [28].

Two new non sequential, i.e. simultaneous, simplex-based unsupervised endmember extraction methods, the simple endmember extraction (SEE) and the Enhanced-SEE (E-SEE) empirical method, respectively, have been also developed. The proposed methods have light computational burden compared to the-state-of-the-art EEAs and are characterised by their clear conceptual meaning. The novelty of the SEE method is that it searches for the extreme values that lie on the end points of the existing transformed axes without further projections that imply iterative procedures. The E-SEE empirical method compensates the tendency that the majority of the convex-geometrical based endmember extraction methods encounter to select high contrast endmembers over less contrast endmembers. Its novelty lies in the fact that it changes the distribution of the initial data sample increasing the distance between candidate endmembers and the data mean. Parts of the particular research have been published in [9], [29], [30], [31].

Last but not least, the fifth method is a new multiple endmember spectral mixture analysis (MESMA) [23], based on spectral angle distance, called MESMA-SAD, and it is used for the estimation of the fractional abundances for the endmembers in each pixel. The new method significantly minimizes the time-processing compared to the existing MESMA algorithms by combining the spectral angle distance (SAD) values and the mean absolute errors (MAE). The algorithm does not require any threshold. The new method attempts to exploit the advantage of SAD to be insensitive to differences in the albedo of the modeled spectrum. Parts of the particular research have been published in [32].

To conclude this section, simulated and real image-based experiments indicate the effectiveness of the proposed methods to improve the data exploitation, rendering their implementation very promising in

hyperspectral image processing. In addition to this, significant remarks derived from the long-standing study define new research paths to be explored.

1.5 Thesis organisation

This introductory chapter provided a basic background of hyperspectral remote sensing and it covered the main aspects which this thesis is focused on. The remainder of the thesis is organised as follows:

In Chapter 2, the estimation of the signal subspace dimension is described. In particular, the relevant state-of-the-art methods are presented and reviewed. An elaborate description of the proposed method for signal subspace estimation, called *outlier detection method* (ODM) is given along with its evaluation using real and simulated data sets. Remarkable conclusions are presented at the end of the chapter.

In Chapter 3, the dimensionality reduction of hyperspectral data is discussed. A thorough overview of the latest developments related to dimensionality reduction is provided. Then, the new proposed band selection method to optimize the extraction of spectrally close endmembers (SCEs) is described. Its effectiveness is shown in the experiments with real data sets and the results are discussed in the conclusions. In addition to this, the proposed BS method for optimizing classification procedure is precented in second part of this chapter. Results and conclusions of the method are finally given.

In Chapter 4, focus is given on the endmember extraction and several relevant issues of great importance are addressed. A review of the state-of-the-art EEAs is given. Then, two new proposed methods for endmember extraction, called *simple endmember extraction* (SEE) and its empirical enhanced version (E-SEE) are described. Analysis is also provided concerning their computational complexity. For performance evaluation, extensive studies of simulated and real image-based experiments are included. Concluding remarks are discussed at the end of the chapter.

In Chapter 5, abundance estimation is described. An overview of the latest relevant developments is provided. A new proposed method, called *multiple endmember spectral mixture analysis based on*

spectral angle distance (MESMA-SAD) is described. Evaluation is accomplished using real hyperspectral data sets and results are discussed on the conclusions.

Finally, Chapter 6 draws the main conclusions and future research lines derived from the present study.

Chapter 2

Estimation of the signal subspace dimension

2.1 Introduction

Estimation of the number of signals is a fundamental problem in signal processing. In the scientific field of hyperspectral imagery, signals are related to the unique constituent deterministic spectral signatures, called endmembers [6]. A predetermined number of endmembers is required by the majority of the existing endmember extraction methods in order to detect the optimal set of endmembers. Estimation of the correct number of endmembers has significant impact on the performance of the endmember extraction algorithms and consequently on the accuracy of the spectral unmixing process. According to authors in [33], the accuracy of spectral unmixing will be the highest when the exact number of endmembers that are required to account for the spectral variability is utilized in the model. Using fewer endmembers than the actual number would lead to the increase of the root mean square error between the original and the reconstructed image, whilst too many endmembers would make the model sensitive to instrumental noise, atmospheric influences and natural variability in spectra, resulting in abundance estimation error. Furthermore, the number of endmembers is associated with the intrinsic or in a more

wide sense with the virtual dimensionality of a hyperspectral dataset [34], [35], as it determines the optimal number of dimensions to be retained after dimensionality reduction in order to represent the dataset. Hence, an accurate determination of the number of the endmembers significantly contributes to the accuracy of the spectral unmixing processing and enables low-dimensional representation of spectral vectors, yielding gains in computational time and complexity, data storage and signal-to-noise ratio (SNR) [12].

2.2 Review of methods for the estimation of the signal subspace dimension

In recent years, many algorithms have been developed which contribute to the estimation of the number of endmembers. The available methods can be classified into separate categories (Fig.2.1). The first category comprises eigen-based energy methods [12], [36]. Typical examples of this category are the principal component analysis (PCA) and the maximum noise fraction (MNF). These methods involve a dimensionality reduction method and estimate the minimum number of the transformed components for which the total variance of the data is equal to a specified percentage of energy. However, the cut-off threshold should be manually chosen, which is very difficult to determine since the eigenvalues corresponding to signals and noise are sometimes very similar [37]. In the second category, information criteria based on likelihood functions [38], [39] are employed. Two well-known information criteria for model order selection are Akaike information criterion (AIC) [38] and minimum description length (MDL) [39]. Since the criteria require the prior knowledge of the mixture model or likelihood function, the estimation may suffer from model mismatch errors resulting from incorrect prior information. Moreover, it has been shown in [35] that the results of AIC and MDL when applied to hyperspectral data are seriously overestimated due to the invalid Gaussian distribution assumption made on the abundances [37]. The third category consists of eigenvalue-based methods [35], [37]. Harsanyi-Farrand-Chang (HFC) and noise-whitened HFC (NWHFC) [35] methods estimate the virtual dimensionality (VD) based on the fact that the eigenvalues of the correlation matrix and of the covariance matrix

should be equal if noise exists. Thus, eigenvalues of both data correlation and covariance matrices are calculated and if their difference is positive - according to a determined probability false alarm parameter- then a signal source is present. VD methods might overestimate the number of the endmembers because they estimate the spectrally distinct signal sources which could comprise known and unknown image endmembers, background signatures, interferences and anomalies [35]. HFC and NWHFC methods impose limitations to automation since they result in different estimates for different false alarm parameters. Recently, a new empirical method for estimating the number of endmembers presented in [37] modifies the VD concept. The method is called eigenvalue likelihood maximization (ELM) and it is based on the fact that the eigenvalues of the covariance and the correlation matrices which correspond to the noise are identical, whilst eigenvalues of the correlation matrix corresponding to the signal are larger than the ones of the covariance matrix. The eigenvalue-based methods are based only on the eigenvectors of the observed data correlation or covariance matrix. Since signal subspace dimension is unknown in most real applications, it must be inferred from data leading to a model-order problem which may lead to poor results [40]. Authors in [40] presented hyperspectral signal subspace identification by minimum error (HySime) method which selects the subset of eigenvectors that best represents the signal subspace in the minimum mean square error sense.
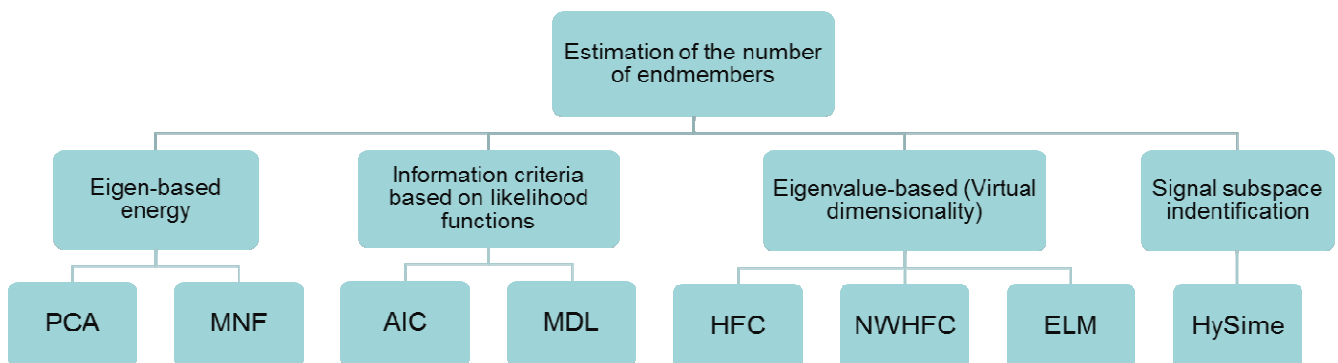
Figure 2.1: Categorization of methods for estimating the signal subspace dimension.

Among the methods for estimating the signal subspace dimension, HFC, NWHFC and HySime are in the front rank of high performance. Therefore, a more detailed description of these methods is given in the next subsections.

Estimation of virtual dimensionality (HFC-NWHFC)

The HFC method [35] uses the eigenvalues to measure signal energies in the detection model. More specifically, let $\{\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq ... \geq \hat{\lambda}_L\}$ the eigenvalues generated by the sample covariance matrix and $\{\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_L\}$ the eigenvalues generated by the sample correlation matrix, where L is the data dimensionality. Since the data dimensionality is equal to the total number of eigenvalues, each eigenvalue specifies a component dimension and provides an indication of the significance of that particular component in terms of energy or variance. Assuming white noise, if there is no signal source contained in a particular component, the corresponding correlation eigenvalue and covariance eigenvalue in this component should reflect only the noise energy and therefore, both eigenvalues should be equal,

$$\hat{\lambda}_l \geq \lambda_l \geq \sigma_{nl}^2 \quad \text{for } l = 1,...,VD$$

$$\hat{\lambda}_l = \lambda_l = \sigma_{nl}^2 \quad \text{for } l = VD+1,...,L \tag{2.1}$$

where σ_{nl}^2 is the noise variance in the lth spectral band and VD is the virtual dimensionality above which noise dominates.

This fact provides us with a base from which we can formulate the difference between the correlation eigenvalue and its corresponding covariance eigenvalue as a binary composite hypothesis testing problem,

$$\begin{aligned} H_0 &: z_l = \hat{\lambda}_l - \lambda_l = 0 \\ H_1 &: z_l = \hat{\lambda}_l - \lambda_l \succ 0 \end{aligned} \quad \text{for } l = 1,2,...,L. \tag{2.2}$$

The null hypothesis represents the case of the zero difference, while the alternative hypothesis indicates the case that the difference is greater than zero. A threshold is needed to be set in order to indicate if the difference between the eigenvalues should be considered important. Thus, the Neyman–Pearson test is applied to each pair of correlation eigenvalue and its corresponding covariance eigenvalue. The detection power P_D is maximized and the probability of the false alarm P_F is set equal to a constant value a. In mathematical terms we have

$$P_F = \int_{\tau_l}^{\infty} N\left(0, \sigma_{z_l}^2\right) dz = a$$

$$P_D = \int_{\tau_l}^{\infty} N\left(\mu_l, \sigma_{z_l}^2\right) dz \tag{2.3}$$

where μ_l is an unknown constant and $\sigma_{z_l}^2$ is the variance of the difference between the eigenvalues $\hat{\lambda}_l - \lambda_l$ and can be approximated as $\sigma_{z_l}^2 = \frac{2\hat{\lambda}_l^2}{N} + \frac{2\lambda_l^2}{N}$ [35]. The threshold τ_l is determined by the fixed value a and it is l-dependent. The number of times the test fails indicates how many signal sources are present in the image. In other words, a failure of the Neyman–Pearson test in a component indicates a truth of the alternative hypothesis, which implies that there is a signal source in this particular component. Using this approach, we can estimate the VD with the receiver operating characteristic (ROC) analysis for evaluating the effectiveness of the decision. This method is sensitive to noise since it does not have a noise-whitening process that de-correlates noise with signal sources to enhance signal detection performance.

An alternative approach includes a noise-whitening process in the HFC method to remove noise effects on eigenvalues [35]. The modified method is referred to as noise-whitened HFC (NWHFC) method. In this case, noise estimation is required for the NWHFC method. Since both the HFC and NWHFC methods model the correlation eigenvalue and the corresponding covariance eigenvalue as random parameters in [40], the sample size must be sufficiently large to ensure that the covariance between these two types of eigenvalues is asymptotically zero.

Hyperspectral signal subspace identification by minimum error (HySime)

HySime [40] estimates the signal subspace dimensionality by searching the subset of eigenvectors of the signal correlation matrix that best represents the signal subspace in the sense of least squares error. Consider that if L is the total number of bands, each observed spectral vector y, $y \in \Re^L$, consists of a signal vector x, $x \in \Re^L$ and an error term n, $n \in \Re^L$ for additive noise which includes sensor noise, endmember variability, and other model inadequacies [6], [19], [37], [40]:

$$y = x + n. \tag{2.4}$$

Assuming that signal and noise matrices are uncorrelated, the correlation matrix of observed spectra y can be expressed by:

$$R_y = R_x + R_n \tag{2.5}$$

The estimate $\hat{R}_n$ of the noise correlation matrix results via multiple regression [41], while the estimate $\hat{R}_x$ of the signal correlation matrix is given by the equation:

$$\hat{R}_x = R_y - \hat{R}_n. \tag{2.6}$$

Let $\{\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_L\}$ the eigenvalues of signal correlation matrix $\hat{R}_x$ and e_i the corresponding eigenvector of λ_i, for $1 \leq i \leq L$. Assume that $E_p = \left\langle \left[e_{i_1}, e_{i_2}, ..., e_{i_p} \right] \right\rangle$ is the subspace inferred by p eigenvectors -not necessary the first p but any possible permutation of p eigenvectors- and the projection of the observed spectra y onto this subspace is expressed by:

$$\hat{x}_p \equiv U_p y, \tag{2.7}$$

where $U_p = E_p E_p^T$ the projection matrix onto $\left\langle E_p \right\rangle$.

The mean square error (MSE) between the signal x from eq. (2.4), (where $x = y - n$) and the estimated $\hat{x}_p$ is computed by $MSE(p/x) = E\left[(x - \hat{x}_p)^T (x - \hat{x}_p)/x \right]$ and $MSE(p) = E[MSE(p/x)]$, where $E(\cdot)$ stands for expectation operator. Noting that $U_p^{\perp} = I - U_p$, the optimization problem is expressed as:

22

$$\hat{p} = \arg\left\{\min_{1 \le p \le L}\left(trace\left[U_p^\perp R_y\right] + 2trace\left[U_p \hat{R}_n\right]\right)\right\} \qquad (2.8)$$

and by further simplification of eq. (2.8) is rewritten:

$$\hat{p} = \arg\left\{\min_{1 \le p \le L}\left[\sum_{j=1}^{p}\left(-v_j + 2\sigma_j^2\right)\right]\right\} \qquad (2.9)$$

where $v_j = e_j^T R_y e_j$, $\sigma_j^2 = e_j^T \hat{R}_n e_j$ for $1 \le j \le L$, and $\{e_1, e_2, ... e_L\}$ are the eigenvectors of $\hat{R}_x$. The first

term, $-v_j$ is related to the projection error and it is a decreasing function of p, while the second term,

$2\sigma_j^2$ is related to the noise power and it is an increasing function of p. The minimization of eq. (2.9) is

accomplished when negative values of the function to be minimized are encountered. The

corresponding number of eigenvectors is the signal subspace dimension.

2.3 Outlier detection method

All the existing methods for estimation of the signal subspace dimension, arguably, consider the

existence of two different distributions, the one related to noise and the other related to signal, or in

geometrical approach they consider two different subspaces one of noise and one of signal. However, in

hyperspectral space, signal vectors are very few in order to estimate their population distribution

properly or to statistically analyze them.

A new automatic non-parametric method for estimating the number of endmembers is introduced. Its

novelty lies in the fact that it considers only the existence of noise and treats signals as outliers of noise.

No estimation of statistical distributions is required. The new method, called outlier detection method

(ODM), explores the geometrical properties of the noise hypersphere. It searches for the signals whose

radius is by far larger than that of the noise introducing for the first time in virtual dimension theory a

robust outlier detection method. In particular, the ODM performs noise estimation and whitening

process. Afterwards, observed data are transformed into a new principal component space, where noise

is expected to lie in a hypersphere of constant radius. Estimation of the number of noise hypersphere

outliers using a robust inter quartile range based outlier detection method [42] results in the estimation of the number of endmembers. In [24] an empirical method for estimating the number of endmembers is presented which implies the approach adopted by ODM.

2.3.1 Data model and problem formulation

Recall the fact that each observed spectral vector y, $y \in \mathfrak{R}^L$, where L is the total number of bands, consists of a signal vector x, $x \in \mathfrak{R}^L$ and an error term n, $n \in \mathfrak{R}^L$ that models the additive noise which includes sensor noise, endmember variability, and other model inadequacies [6], [19], [37], [40], i.e.,

$$y = x + n. \tag{2.10}$$

Furthermore, a signal vector lies in an unknown p-dimensional subspace ($S =< [s_1,...,s_p] >$) of the band space, where $p < L$ and it is described by:

$$x = \sum_{k=1}^{p} a_k s_k = S\alpha. \tag{2.11}$$

Under the subspace model scenario, the $L \times 1$ signal vectors s_k are linearly independent (or otherwise S is a full rank $L \times p$ matrix), serving as a basis for the spectral subspace [6] and α is considered a $p \times 1$ vector containing coefficients a_k. Under the linear spectral mixing concept (see section 1.2.1) [19], matrix $S = [s_1,...,s_p]$ comprises the endmember spectra and $\alpha = [a_1,...,a_p]$ their corresponding abundances. The latter should obey to the sum-to-one and the positivity constraints in order to be physically meaningful. In this work, the subspace model is being studied which specifies the linear vector subspace region of the spectral space in which spectral vectors are allowed to reside regardless the adopted spectral mixing model, linear or non-linear [6], [40].

According to [43], in the case of independent and identically distributed (i.i.d.) zero mean noise with variance $\sigma_n^2 I$, signal subspace can be estimated, even if signal vectors are unknown, by the orthogonal decomposition of the covariance matrix of the observed vectors y, R_y. The estimate of the signal subspace is the span of the eigenvectors of R_y, $\langle M \rangle = \langle [e_1, e_2, .., e_p] \rangle$ whose respective eigenvalues $\lambda_1 > \lambda_2 > ... > \lambda_p$ are larger than σ_n^2 of noise. Of course, in most real applications the dimension p of the signal subspace is unknown and noise is not i.i.d.. Therefore, in many cases noise estimation is a prerequisite for the denoising or whitening process. A plethora of signal subspace estimation methods [35], [36], [37], [38], [39], [44] are based on the eigenvalues of the covariance or correlation matrix of the observed spectral vectors y. The drawbacks of using only the eigenvalues are presented in [40]. In this work, the estimation of the signal subspace dimension is based on the transformation of the observed vectors y using the eigenvectors of R_y. The new transformed space is then statistically analysed based on information theory concepts which are presented in the following section.

2.3.2 Definition of Noise Hypersphere

Multivariate normal distribution

Let $X = [X_1, X_2, ..., X_L]^T$ be a $L \times 1$ random vector. Its mean value is given by $m = E(X)$, $E(\cdot)$ stands for expectation operator, and its covariance matrix by $R_X = E(X - m)(X - m)^T$. Assuming that random X is multivariate normal and R_X is a non-singular matrix, the following quadratic form

$$r^2 = (x - m)^T R_X^{-1} (x - m) \tag{2.12}$$

is a weighted norm which is called the Mahalanobis distance from x to m. The locus of points x for which r^2 is constant is also a locus of points for which the density $f(x)$ is constant. Geometrically, this represents a hyper-ellipse. In case that the locus is a hypersphere, its radius is equal to r [43].

Noise hypersphere

25

Based on information theory [45], [46] the zero mean white Gaussian noise vector $n \sim N(0, \sigma_n^2 I)$ has constant noise spectral density N_0. It is spherically symmetric in all directions in the spectral space and lies on the surface of a hypersphere with radius equal to σ_n. More specifically, in an *L*-dimensional hypersphere, the distance of point $n = [n_1, n_2, ... n_L]^T$ from the origin (zero) point is according to eq. (2.12)

$$r_n = \frac{\sqrt{n_1^2 + n_2^2 + ... + n_L^2}}{\sigma_L}$$ and the distance of the normalized noise vector is $\frac{r_n}{\sqrt{L}}$, which is the σ_n.

The advantages of considering the normalized version of noise vector as well as further details related to the above can be found in [45], [46], [47].

Thereupon, $\sigma_{n1} = \sigma_{n2} = ... = \sigma_{nL}$ are the standard deviations of the normalized noise vector in each dimension of the hypersphere and are equal to its radius as it shown in Fig. 2.2. The signal vector x has evidently $\sigma_x > \sigma_n$ and since σ_x varies in all directions, x lies in a hyperellipsoid. In order to utilize the aforementioned properties of the noise hypersphere, it is requisite that the noise is zero mean i.i.d. or that noise is known and therefore it can be transformed to zero mean i.i.d.. Both requisites do not stand in real applications. However, many approaches have been developed for noise estimation. Two of them are presented in the following section.
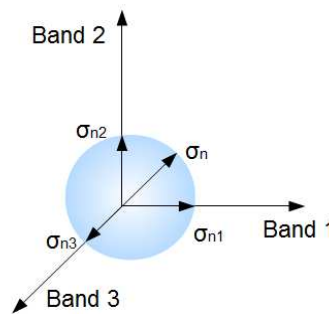
Figure 2.2: Illustration of noise hypersphere in three dimensions.

2.3.3 Noise Estimation

Noise estimation is of great importance not only for hyperspectral imagery but generally for signal processing. In this section, nearest neighbor difference (NND) [48] and multiple regression theory [41] based methods are analyzed since these are widely used by signal subspace estimation algorithms [39], [40]. Both these noise estimation methods are evaluated using simulated data in section 2.4.

Nearest neighbor difference

The nearest neighbour difference (NND) method [48], also called shift difference method, is considered to be the easiest method for noise estimation. The procedure exploits the fact that signal exhibits strong spatial correlation among nearby pixels in an image, whilst the spatial correlation for noise is very weak. Therefore, it is assumed that noise samples are independent and have the same statistics [40]. The shift difference method should be applied on a homogeneous area. More precisely, it is performed on the data by differencing the two adjacent pixels to the right and above each pixel and averaging the two computed differences to obtain the noise value to assign to the pixel being processed. The idea can be illustrated using two adjacent observed vectors, y_1 and y_2, with essentially the same target. Subtracting them yields:

$$y_1 - y_2 = (x_1 + n_1) - (x_2 + n_2) \approx n_1 - n_2 \qquad (2.13)$$

where x_1, x_2 are the signal vectors and n_1, n_2 are noise vectors. Depending on the image, the noise estimation may be performed in a homogeneous subset of pixels, assuming that noise is the same throughout the whole image. Therefore the covariance matrix of noise R_n can be estimated, instead of noise value per observed spectral vector. The drawback of the NND method is that due to its assumption that adjacent pixels have the same signal information, the method is not proper for all the

datasets, because the amount of pixels belonging to homogeneous areas may not be adequate for an accurate calculation of noise statistics.

Multiple regression theory based method

The multiple regression theory based approach [40], [41] is amenable to hyperspectral data since it can accommodate many explanatory variables which may be correlated, such as data in adjacent spectral bands. In particular, let Y be a $L \times N$ data matrix, where N are the $L \times 1$ observed spectral vectors, and L the spectral bands. Defining $Z = Y^T$, a $N \times 1$ vector $z_i = [Z]_{:,i}$ containing the i$^{\text{th}}$ band values of all the pixels and $Z_{\partial i} = [z_1, ..., z_{i-1}, z_{i+1}, ..., z_L]$ is a $N \times (L-1)$ matrix containing the pixel values of all the bands except for band i. Assuming that vector z_i can be expressed as a linear combination of the remaining data of L-1 bands, the following equation can be written:

$$z_i = Z_{\partial i} b_i + n_i \tag{2.14}$$

where $Z_{\partial i}$ is the $N \times (L-1)$ explanatory data matrix, b_i is the $(L-1) \times 1$ regression vector and n_i is the residual error of size $N \times 1$. The linear regression coefficients are determined by:

$$\hat{b}_i = (Z_{\partial i}^T Z_{\partial i})^{-1} Z_{\partial i}^T z_i. \tag{2.15}$$

Noise estimation of band i is accomplished by the following equation:

$$\hat{n}_i = z_i - Z_{\partial i} \hat{b}_i. \tag{2.16}$$

2.3.4 Analysis of the proposed method

In this section the proposed method for estimating the number of endmembers, called outlier detection method (ODM), is introduced and described analytically. The method is fully automatic and non-

parametric. It comprises three steps: 1) noise estimation, 2) MNF transformation and 3) outlier detection. The main key points of the proposed method are summarised as follows:

a) There is a big effort in hyperspectral community to define a threshold between signal and noise [37]. The ODM introduces a new concept which considers only the existence of noise and treats signal as outlier. Consequently, no threshold is needed.

b) Contrary to the existing relevant algorithms which focus on signal subspace, the ODM exploits the properties of noise subspace. It relies on the mathematical description of the noise hypersphere radius which is given by information theory.

c) A new modified version of MNF is introduced which initially performs multiple regression theory based method for noise estimation, instead of NND. Results showed that this modification improves the MNF method.

d) For the first time in virtual dimensionality theory, a robust outlier detection method is used, called inter quartile range (IQR) based method. Its benefit lies in the fact that it can be used when data distribution is unknown and thus, no statistical parameter estimation is needed. The risk of estimating erroneously the signal distribution due to its small population is omitted.

e) The proposed method exhibits high degree of simplicity.

The first step of the proposed method is noise estimation. Experiments with simulated and real data (section 2.4) show that the performance of the proposed method is better when multiple regression based method is applied instead of the NND method.

The second step includes noise whitening and transformation into a new principal component space. More analytically, the noise covariance matrix R_n is estimated. The orthogonal decomposition of R_n results in the matrix $D_n = [d_{n1}, d_{n2}, ..., d_{nL}]$ of size $L \times L$ which consists of noise eigenvectors d_n, each one of size $L \times 1$. Suppose that the observed $L \times N$ data matrix Y, where N are the $L \times 1$ observed spectral vectors and L the spectral bands, is transformed using the noise eigenvectors. The transformed data F is given by

$$F = D_n^T Y. \tag{2.17}$$

The matrix F of size $L \times N$ consists of N transformed spectral vectors f of size $L \times 1$. Defining $W = F^T$,

a $N \times 1$ vector $w_i = [W]_{:,i}$ contains the values of all the transformed pixels in band i. Dividing each data

of band i with the standard deviation of noise $\hat{\sigma}_{ni}$ (symbol ^ stands for the estimated value) of the

corresponding band i

$$W' = [\frac{w_1}{\hat{\sigma}_{n1}}, \frac{w_2}{\hat{\sigma}_{n2}}, ..., \frac{w_L}{\hat{\sigma}_{nL}}] \tag{2.18}$$

results in the $N \times L$ matrix W' which is the transformed data with equal noise variance $\hat{\sigma}_n^2$ in each band,

which means that noise is whitened in the transformed space.

The next step is the orthogonal decomposition of the covariance matrix of W'^T which results in the

$L \times L$ matrix $D_{W'^T}$ containing the $d_{w'^T}$ eigenvectors of size $L \times 1$. The transformation of W'^T using the

eigenvectors of $D_{W'^T}$

$$Y' = D_{W'^T}^T W'^T \tag{2.19}$$

defines a new principal component space in which transformed data of $L \times N$ matrix Y' consist of

uncorrelated noise which increases with the component rank. Thus, the well-known MNF [48] is

modified by applying different noise estimation method.

Assuming that noise is white, rotation of a signal structure, i.e. in eq. (2.19), does not change the noise

distribution [47]. Consequently, noise remains spherically distributed about the mean value and lies in a

hypersphere of radius $\hat{\sigma}_n$. It is reasonable that noise estimation comprises an error, which is justified in

terms of fluctuations. Therefore, it cannot be expected that standard deviations of noise components are

exactly equal to σ_n but it should be expected to be close to zero mean value as the minimum standard

deviation of a component corresponds to the maximum noise fraction [48]. Conversely, standard

deviation $\hat{\sigma}_{xi}$ of signal x_i, where $i = 1,...,p$, is larger than $\hat{\sigma}_n$ and decreases as component's rank increases. Fig. 2.3 shows the histogram of the first three transformed components of a zero mean image with three endmembers and signal-to-noise ratio (SNR) 20:1. As it is observed, the first two transformed components present high variance, while the third transformed component has very low variance which is close to the mean value.
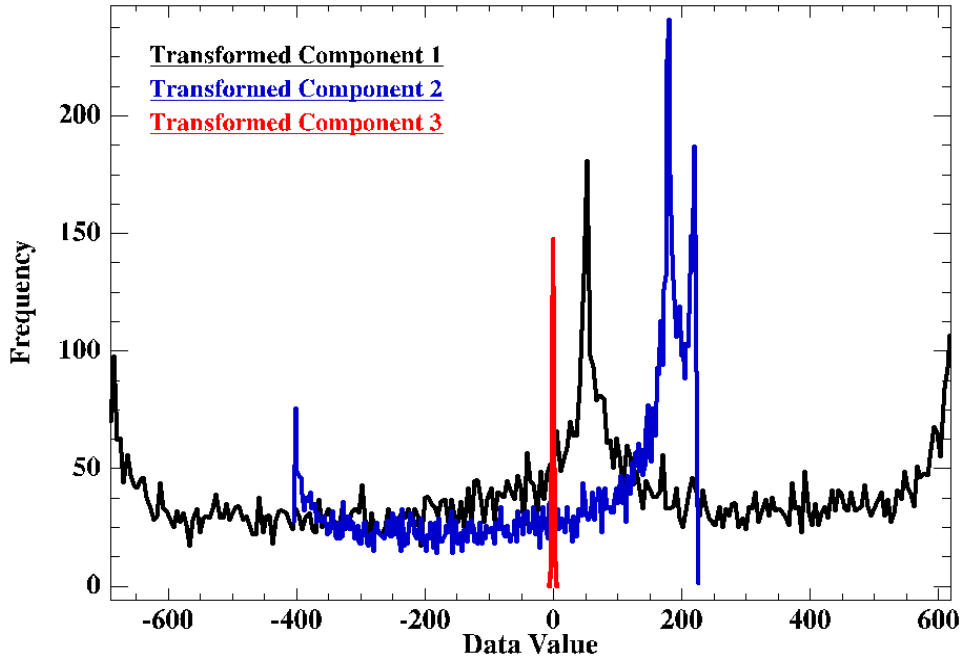
Figure 2.3: Histogram of the three first transformed components of an image with three endmembers.

The third step of ODM includes outlier detection using quartile range. Outlier detection is widely used to detect and/or remove anomalous observations from the data. It is a primary step in many data-mining applications [42]. There are many definitions given for outliers. The one that fits on the particular approach is given by Hawkins [49] who defines an outlier as an observation that deviates so much from other observations as to arouse suspicion that it was generated by a different mechanism.

The sample mean and the sample variance give good estimation for data location and data shape, but they are affected by outliers. Inter quartile range (IQR) based method [42], [50] is one of the most common methods for outlier detection as IQR is a robust statistic compared to total range and standard deviation. The method can be used when data distribution is unknown. Assume that observed values are

placed in ascending order. The lower quartile Q_1 is the observation at the 25th percentile, the second quartile Q_2 is defined the observation at the 50th percentile and the third quartile Q_3 is the observation at the 75th percentile. The quantity Q_3-Q_1 is called the inter quartile range (IQR) and it provides a means to indicate the boundary beyond which the data will be labelled as outliers. More precisely, if an observation is below $Q_1 - 1.5 \times IQR$ or above $Q_3 + 1.5 \times IQR$, it is viewed as being too far from the central values to be reasonable.

In most real applications signal subspace, and consequently signal vectors are unknown and even if they are known they are very few in order to be statistically analysed. Noise subspace consists always of some hundreds of components which are much more than the signal components in the transformed hyperspectral space. Assuming that standard deviations of all the principal components correspond to noise, it is expected that the whole data lie in a hypersphere of radius σ_n. Thus, signal components can be considered as outliers of noise hypersphere.

As it is referred in previous section, the radius of noise hypersphere is much smaller than the radius of the signal hyperellipsoid and since search is focused on detecting noise hypersphere outliers, only the upper bound is of interest in this particular procedure. Let us assume that $\Sigma = [\hat{\sigma}_1, \hat{\sigma}_2, ..., \hat{\sigma}_L]$ is a $L \times 1$ vector which consists of the standard deviations $\hat{\sigma}_i$ of each ith transformed component. The transformed components are ranked according to the SNR, which implies that $\hat{\sigma}_i$ are in descending order and thus, the first p values of Σ correspond to signal vectors. Taking the $\Sigma^R = [\hat{\sigma}_L, \hat{\sigma}_{L-1}, ..., \hat{\sigma}_1]$ as the $L \times 1$ vector which consists of the standard deviations $\hat{\sigma}_i$ in reverse order, meaning in ascending order, the first $L - p$ values of Σ^R correspond to noise vectors. As p is unknown we suppose that all the values of Σ^R correspond to noise. Euclidean distance (ED) of adjacent values of Σ^R, $ED = [ED(\hat{\sigma}_L, \hat{\sigma}_{L-1}), ED(\hat{\sigma}_{L-1}, \hat{\sigma}_{L-2}), ..., ED(\hat{\sigma}_2, \hat{\sigma}_1)]$ reflects possible divergences which are considered reasonable when $\hat{\sigma}_i$ corresponds to noise, but outliers when $\hat{\sigma}_i$ corresponds to signal.

The pseudo-code for ODM is shown in Algorithm 2.1.

Algorithm 2.1: ODM

Data: The LxN matrix Y, where N are the $L \times 1$ observed spectral vectors and L the spectral bands
Result: Estimation of the number of endmembers p

Step 1: *Noise estimation*
$Z = Y^T$
for *i=1* to L, $z_i = [Z]_{:,i}$ all the pixels of band i

$\qquad Z_{\partial i} = [z_1, ..., z_{i-1}, z_{i+1}, ..., z_L]$ the pixels of all the bands except i

$\qquad \hat{b}_i = (Z_{\partial i}^T Z_{\partial i})^{-1} Z_{\partial i}^T z_i$ calculation of coefficient b

$\qquad \hat{n}_i = z_i - Z_{\partial i} \hat{b}_i$ noise estimation for band i

end for
OUTPUT $\hat{n}$

Step 2: *White noise data transformation*
Estimation of the noise covariance matrix R_n

Computational of the matrix associated with the orthogonal decomposition of R_n: $\quad D_n = [d_{n1}, d_{n2}, ..., d_{nL}]$

$$F = D_n^T Y$$
$$W = F^T,$$
$$w_i = [W]_{:,i} \text{ are the transformed pixels in band } i$$

for *i=1* to L, estimation of $\qquad \hat{\sigma}_{ni}$ (standard deviation of noise)

$$w_i' = \frac{w_i}{\hat{\sigma}_{ni}}$$

end for
Estimation of the covariance matrix of $\quad W'^T$
Computational of the matrix associated with the orthogonal decomposition of W'^T: $\quad D_{W'^T}$

Transformation of whitened data $\qquad Y' = D_{W'^T}^T W'^T$
OUTPUT: Transformed whitened data Y'

Step 3: *Outlier detection*
Estimation of standard deviation of Y' $\quad \Sigma = [\hat{\sigma}_1, \hat{\sigma}_2, ..., \hat{\sigma}_L]$
Normalization of Σ
Sorting $\hat{\sigma}_i$ in descending order $\qquad \Sigma^R = [\hat{\sigma}_L, \hat{\sigma}_{L-1}, ..., \hat{\sigma}_1]$
Calculation of Euclidean distance for $\hat{\sigma}_i$ of adjacent bands
$$ED = [ED_{L,L-1}(\hat{\sigma}_L, \hat{\sigma}_{L-1}), ED_{L-1,L-2}(\hat{\sigma}_{L-1}, \hat{\sigma}_{L-2}), ..., ED_{2,1}(\hat{\sigma}_2, \hat{\sigma}_1)]$$
Retrieval of quartiles from EDs:
for i = 1 to 4
$$k = i * (25\%) * (L-1)$$
$$Q_i = ED_{k, k-1}$$
end for
$$IQR = Q_3 - Q_1$$
Definition of the number p of endmembers:
$$p = 0$$
for i = 2 to L
$\qquad$ if $ED_{i,i-1}$ is greater than Q3+1.5*IQR then
$\qquad$ increase p by 1
end for
OUTPUT: Number of endmembers p

Geometrical point of view: An example

For illustration purposes, the following experiment on simulated hyperspectral data is implemented. The simulated data generated according to the linear mixing scattering mechanism using seven random spectral signatures from the U.S. Geological Survey (USGS) digital spectral library and consist of 10^4 pixels and 423 spectral bands. The abundance fractions follow a Dirichlet distribution according to [40] enforcing positivity and full additivity constraints. Gaussian colored noise was added to the data resulting in a signal-to-noise ratio (SNR) of 20dB. The procedure described previously containing noise estimation using multiple regression theory, noise whitening and transformation into a new principal component space and outlier detection is implemented.

Fig. 2.4 shows the distribution of the transformed data through overlapping scattergrams of pairs of adjacent components. Only two axes are used (i, j). Each component is kept on the same axis for the two pairs in which is encountered. The extent of each scattergram in i and j directions implies the magnitude of the standard deviation of Ci and Cj component, respectively. As it is observed, standard deviations of the first six principal components are relatively high and as band rank increases, standard deviations also increase. More precisely, suppose $\hat{\sigma}_i$ denotes the standard deviation of band i. As it is shown in the overlapping scattergram the following relation exists $\hat{\sigma}_1 > \hat{\sigma}_2 > ... > \hat{\sigma}_6 >> \hat{\sigma}_7 \approx \hat{\sigma}_8 \approx ... \approx \hat{\sigma}_{402} \approx \hat{\sigma}_L$. Furthermore, it is remarkable that noise circle (in this case it is not hypersphere since scattergrams are shown in two dimensions) can be detected from the C7-C8 pair (orange circle) and after. This means that in the hyperspectral space, the radius of the noise hypersphere is associated to the standard deviation of the 7th component which is right after the $p-1$ component. It should be noted that since simulated data are generated according to a linear mixing model, the dimension of the signal subspace is $p-1$. The scattergram of the 401st and 402nd components was randomly selected to testify the equality of the noise standard deviation σ_n in all the directions (Fig. 2.4). The difference between the radius of the orange circle (Components 7-8) compared to the radius of

the yellow circle (Components 401-402) can be considered without loss of generality as a result of fluctuations.
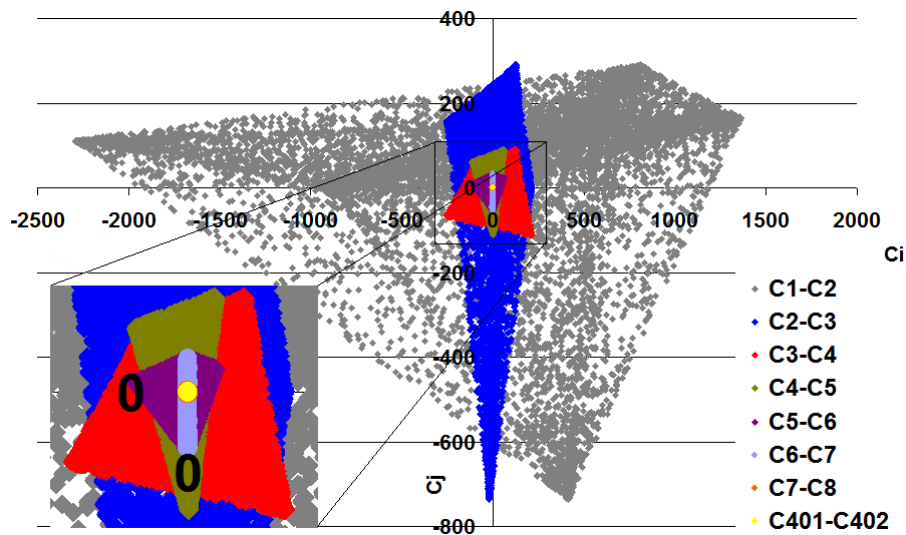
Figure 2.4: Overlapping scattergrams of pairs of adjacent principal components. Only two axes are used (i, j). Each component (B) is kept on the same axis for the pairs in which is encountered.

Another approach for studying the standard deviations of the principal components is by plotting them on a diagram. Fig. 2.5 shows the standard deviations of each principal component resulted from the above experiment. By observing the diagram, it is easy to perceive that standard deviations of the noise (in black color) are almost similar while standard deviations of signal (in blue color) differ a lot from each other. The optimum threshold by which signals are discerned from noise is estimated by using the inter quartile range based method.
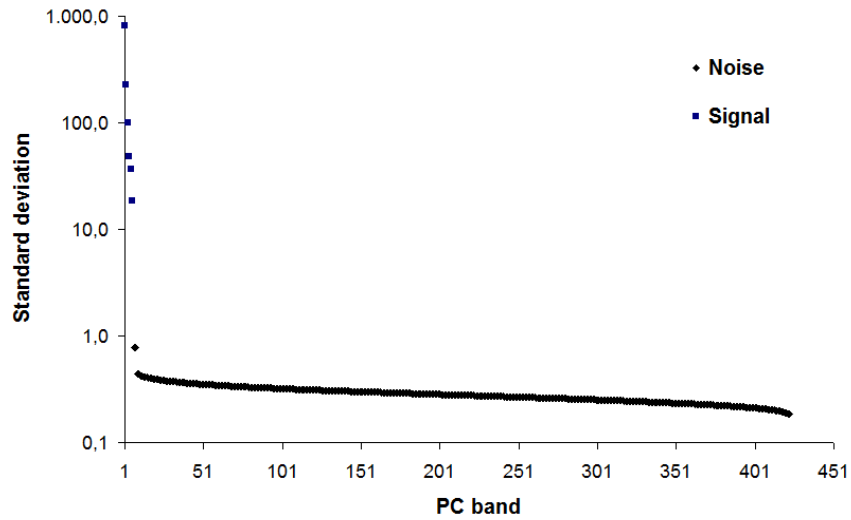
Figure 2.5: Diagram of the standard deviations of each principal component (p=7). A logarithmic scale is used on y-axis.

A graphical display on which outliers can be indicated is a Box plot [42]. Figure 2.6 shows the Box plot created by using the EDs between the standard deviation values. The majority of the EDs are close to one and reside on the left side of the green line, which indicates the upper bound. Black points represent the outliers. As it is observed, the differences in EDs between the three quartiles are negligible compared to the values of the six outliers.
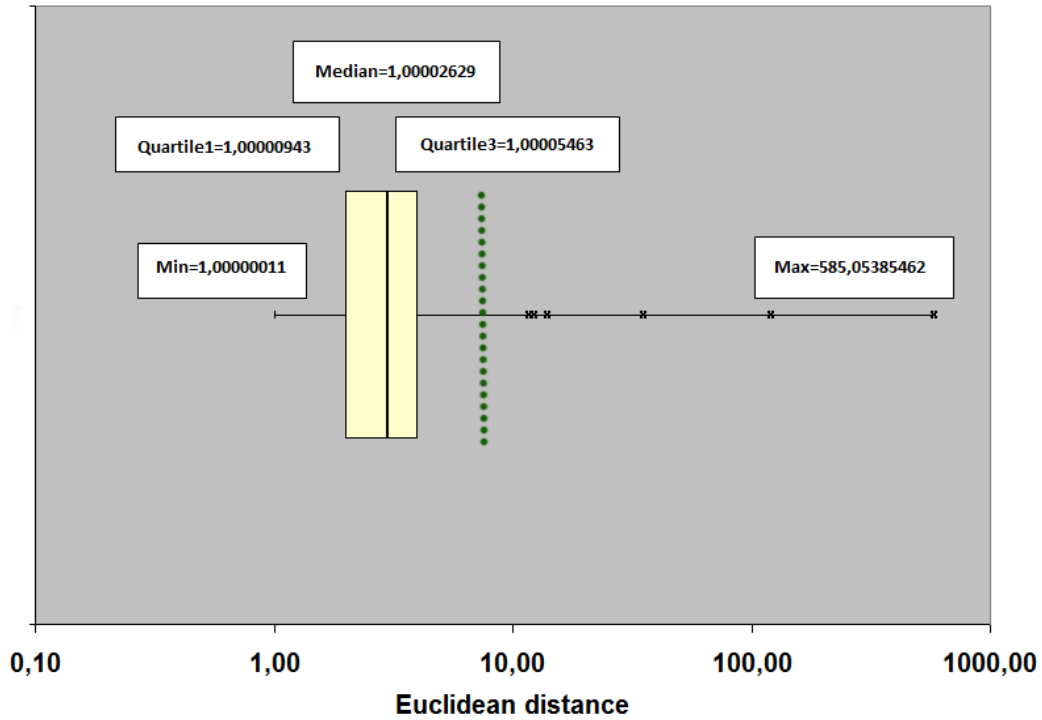
Figure 2.6: Box plot indicating the existence of outliers (on the right side of green line). A logarithmic scale is used on y-axis.

2.4 Experiments

Simulated data experiments

The ODM algorithm was applied on simulated data and compared with the state-of-the-art signal subspace methods, the HySime method and the NWHFC eigen-based Neyman-Pearson detector. The simulated data were generated by a random set of fifteen spectral signatures with 423 spectral bands from the U.S. Geological Survey (USGS) digital spectral library. The abundance fractions follow a Dirichlet distribution according to [40] enforcing positivity and full additivity constraints. Experiments were conducted with respect to 1) the size of the image N, 2) the number of endmembers p, 3) the SNR values, 4) the type of noise (white noise and Gaussian shaped noise)[9] and 5) the existence of outliers.

---

[9] The algorithm which was used for the generation of the simulated data is available at
http://www.lx.it.pt/~bioucas/code.htm.

The reason that different image sizes are introduced is twofold. Firstly, due to sampling error, estimation of the noise covariance matrix R_n and estimation of standard deviation $\hat{\sigma}$ are both affected by the sample size and they should be examined and evaluated using smaller image size, as well. Secondly, recent developed endmember extraction methods tend to integrate spatial information into the endmember extraction process [50]. Towards this direction, these methods search for local endmembers in subsets of image data. Therefore, effectiveness of ODM is examined for such a scenario. Thus, two sets of simulated hyperspectral images were created which differ in size, containing 2500 pixels and 10^4 pixels respectively. Furthermore, evaluation of the proposed method regarding various numbers of endmembers should also be tested. According to [19], the number of endmembers that may be practically identified typically ranges from three to seven, depending on the number of bands and the spectral variability of the scene components. In case of high spectral resolution the hyperspectral datasets may comprise even more, i.e. AVIRIS Cuprite image consists of at least 18 distinct spectral signatures according to USGS. Therefore, the number of endmembers p was determined to be 3, 7, and 15. Two different types of noise were added in the simulated images; white noise and Gaussian shaped noise with variance σ_n^2 equal to 0.02, leading to SNR values of 50dB, 30 dB, 20 dB and 10 dB. Noise estimation step is required in order to transform noise to zero mean i.i.d.. In the case of simulated images with white noise, the last is spherically symmetric in all directions and lies on the surface of a hypersphere with a constant radius Therefore, noise estimation can be omitted.

Figures 2.7 and 2.8 show the standard deviation values for each transformed component of images with $N=10^4$, with p equal to 3, 7 and 15 and white and colored noise respectively. It is observed that standard deviation values minimize and stabilize when the number of the transformed components is equal to the number of the endmembers. For clarity purposes, it was chosen to present a subset of the transformed components of all the simulated images in a stacked plot, and therefore the scale of the values in Figures 2.7 and 2.8 has changed. Table 2.1 shows the results of the applied methods for images with white noise. As it is concluded from the results, regarding the images of 2500 pixels, the ODM yielded quite

satisfactory results outperforming the HySime and NWHFC algorithms when SNR values were very low (30dB-10dB). For the images of 10^4 pixels, all the applied methods yielded the same high performance for the images which contain 3 endmembers, regardless the amount of noise. When the space dimension increased to p=7, the proposed method outperformed the NWHFC method and it had the same high performance with HySime, except for the image with SNR of 10 dB for which the proposed method performed better. For the images with 15 endmembers, the proposed method yielded systematically better results than both HySime and NWHFC.
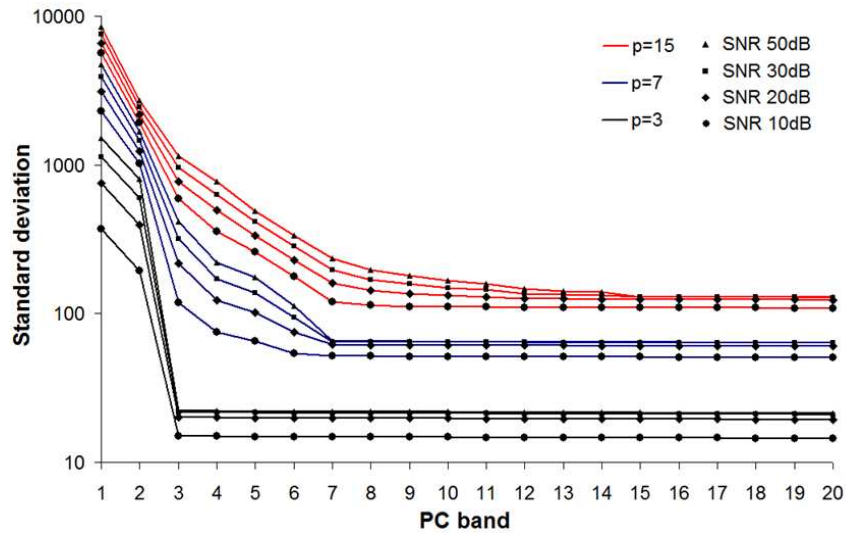
Figure 2.7: Stacked plots of standard deviation values for each PC band for the images with N=10^4 and white noise.
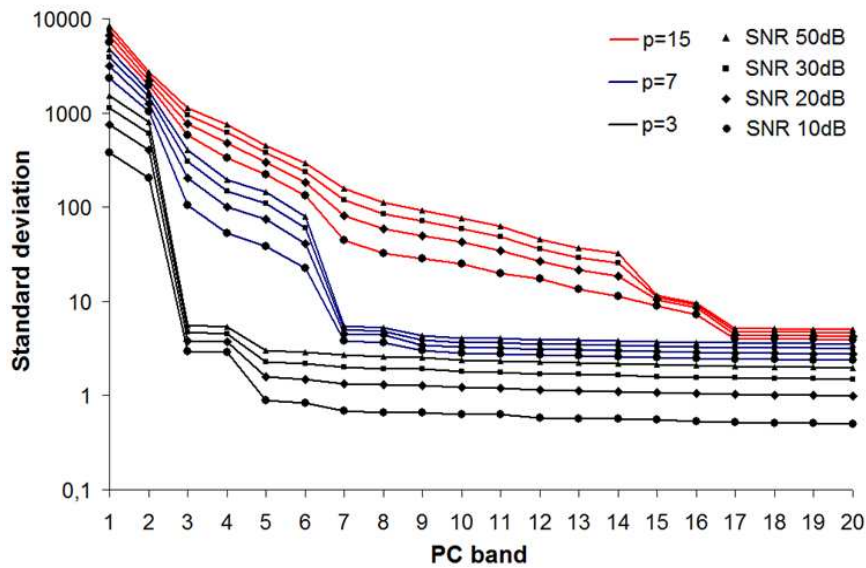
Figure 2.8: Stacked plots of standard deviation values for each PC band for the images with N=10^4 and Gaussian shaped noise.

Table 2.1: Estimated number of endmembers from image with white noise as function of SNR, *p* and *N*.

SNR	Method	White Noise					
		2500 pixels			10^4 pixels		
		p=3	p=7	p=15	p=3	p=7	p=15
50 dB	ODM	3	7	15	3	7	15
	HySime	3	7	15	3	7	15
	NWHFC $P_f=10^{-3}$	3	4	5	3	5	9
	$P_f=10^{-4}$	3	4	4	3	5	7
	$P_f=10^{-5}$	3	4	4	3	5	7
30 dB	ODM	3	7	15	3	7	15
	HySime	21	23	24	3	7	14
	NWHFC $P_f=10^{-3}$	3	5	4	3	6	6
	$P_f=10^{-4}$	3	4	3	3	5	6
	$P_f=10^{-5}$	3	4	3	3	5	5
20 dB	ODM	3	7	13	3	7	13
	HySime	23	26	25	3	7	11
	NWHFC $P_f=10^{-3}$	3	4	1	3	6	5
	$P_f=10^{-4}$	3	4	1	3	5	4
	$P_f=10^{-5}$	3	4	1	3	5	3
10 dB	ODM	3	7	10	3	7	11
	HySime	20	27	24	3	6	7
	NWHFC $P_f=10^{-3}$	3	3	1	3	5	3
	$P_f=10^{-4}$	3	3	1	3	5	3
	$P_f=10^{-5}$	3	3	1	3	4	3

Table 2.2 shows the results of the applied methods for images with Gaussian shaped noise. Two different methods for noise estimation were implemented. As it was expected, NND noise estimation led to the worst results. This is reasonable because NND needs to calculate the shift difference in homogeneous area while pixels in simulated data were created randomly without homogeneous areas. The most satisfactory results were given by ODM for both image sizes when multiple regression theory based method was used for noise estimation. Especially in case of low SNR, results are much more satisfactory compared to the results from HySime, while both methods provided similar results for high SNR. The NWHFC method[10] gave the worst results.

---

[10] Implemented using Open Source MATLAB Hyperspectral Toolbox. 2012. Version 0.06.
http://matlabhyperspec.sourceforge.net/
Endmember Induction Algorithms (EIAs) toolbox. Grupo de Inteligencia Computacional,(UPV/EHU), Spain.
http://www.ehu.es/computationalintelligence/index.php/Endmember_Induction_Algorithms

Table 2.2: Estimated number of endmembers from image with Gaussian shaped noise as function of SNR, p and N. [1] stands for multiple regression, [2] stands for NND.

SNR	Method	Gaussian shaped noise (σ^2=0.02)					
		2500 pixels			10^4 pixels		
		p=3	p=7	p=15	p=3	p=7	p=15
50 dB	ODM [1]	3	7	15	4	7	15
	ODM [2]	49	39	48	32	38	33
	HySime	3	7	15	3	7	15
	NWHFC P_f=10^{-3}	62	13	8	47	23	13
	P_f=10^{-4}	53	12	7	41	22	11
	P_f=10^{-5}	52	12	6	37	20	11
30 dB	ODM [1]	3	8	16	4	7	15
	ODM [2]	39	40	47	34	31	35
	HySime	3	7	13	3	7	14
	NWHFC P_f=10^{-3}	45	58	37	66	48	13
	P_f=10^{-4}	41	54	31	59	43	13
	P_f=10^{-5}	38	47	29	56	36	12
20 dB	ODM [1]	4	8	17	4	8	16
	ODM [2]	48	48	45	30	25	42
	HySime	3	6	6	3	6	8
	NWHFC P_f=10^{-3}	30	23	72	23	56	45
	P_f=10^{-4}	24	20	60	18	48	38
	P_f=10^{-5}	23	17	49	17	43	32
10 dB	ODM [1]	5	9	17	5	8	17
	ODM [2]	47	38	40	37	34	32
	HySime	3	4	5	3	5	5
	NWHFC P_f=10^{-3}	12	16	5	56	67	45
	P_f=10^{-4}	8	14	5	48	60	36
	P_f=10^{-5}	5	14	4	45	50	32

In order to test the method's resistance to outliers in the image pixels, simulated images containing 7 endmembers and Gaussian shaped noise with SNR values of 30dB and 50dB were used. Outliers were added to the images by randomly sampling three outlying points from a uniform distribution, according to [26]. Table 2.3 reports the results. Estimations of the proposed method are satisfactory and verify its resistance to outliers.

Table 2.3: Estimated number of endmembers from image with three outliers.

Method	p=7	
	50dB	30dB
ODM[2]	7	8
HySime	7	6
NWHFC P_f=10^{-3}	7	6
P_f=10^{-4}	6	5
P_f=10^{-5}	6	5

Real data experiments

The proposed algorithm was applied on two real hyperspectral remote sensing images in order to be evaluated in case of unequally distributed noise. The first image was acquired in June, 1992 by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor over an agricultural area of north-western Indiana (Indian Pines) (Fig. 2.9). It consists of 145 x 145 pixels with 220 spectral bands covering a spectral range from 400 to 2500 nm. The number of bands was reduced to 186 after removing 34 bands due to water absorption and low SNR. According to the associate ground based observations[11], 16 land cover classes exist in the image; alfalfa, corn-notill, corn-mintill, corn, grass-pasture, grass-trees, grass-pasture-mowed, hay-windrowed, oats, soybean-notill, soybean-mintill, soybean-clean, wheat, woods, buildings-grass-trees-drives and stone-steel-towers. It should be noted that the aforementioned classes do not represent the entire scene and some of them are not associated with pure materials. Consequently, the number of the endmembers is expected to be higher than 16. Fig. 2.10 shows the standard deviation values for each transformed component of the AVIRIS image and Table 2.4 shows the estimated number of endmembers from the applied methods.

As it is listed in Table 2.4, the ODM using NDD and multiple regression theory based method for noise estimation and the NWHFC resulted in a reasonable number of the distinct classes while the HySime underestimated it. The fact that NWHFC estimates were much higher than its competitors is reasonable since the method searches for signal sources which may include not only endmembers but also unknown interferences, such as clutters, background signatures and anomalies [12].

---

[11] https://engineering.purdue.edu/~biehl/MultiSpec/hyperspectral.html

Figure 2.9: AVIRIS Indian pines hyperspectral dataset.
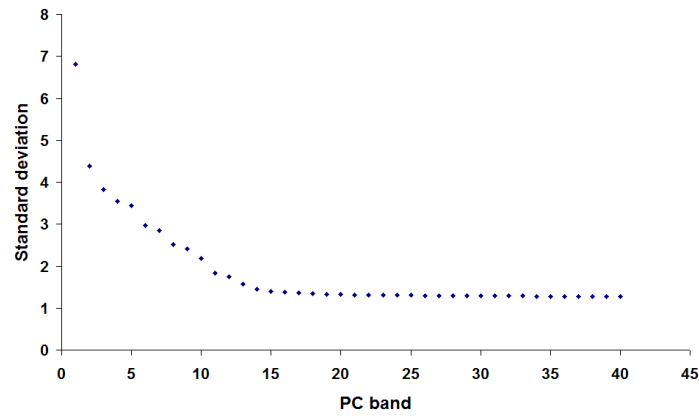
Figure 2.10: Standard deviation values of the transformed bands for AVIRIS Indian pine image.

Table 2.4: Estimated number of endmembers for the AVIRIS Indian Pines image [1] stands for multiple regression, [2] stands for NND.

Method	Estimated number of endmembers (reference number: higher than 16)
ODM [1]	17
ODM [2]	24
HySime	14
NWHFC $P_f=10^{-3}$	27
$P_f=10^{-4}$	23
$P_f=10^{-5}$	22

The second real hyperspectral dataset which has been used for evaluation was collected in 1997 by the

AVIRIS sensor over a well-known mining region of Cuprite in Nevada. The image scene is well

understood mineralogically and the ground truth spectral signatures are available in the USGS digital library. According to the associated ground based observations and the mineral map produced in 1995 by USGS[12], 18 minerals can be identified in the image. Besides minerals, there should be other distinct classes depicted in the image, whose amount is unknown. Thus, the number of the endmembers is expected to be higher than 18. The original image has 220 spectral bands covering a spectral range from 0.4 to 2.5 μm. The number of bands was reduced to 188 after removing bad bands due to water absorption and low SNR. Fig. 2.11 shows the subimage scene of 351 x 350 pixels with reflectance values which was selected for the experiments. Fig. 2.12 shows the standard deviation values for each transformed component of the AVIRIS image. Table 2.5 shows the estimated number of endmembers from the applied methods.

As it is shown in Table 2.5, the ODM using NDD and multiple regression theory based method for noise estimation and the NWHFC resulted in number of the distinct classes higher than 18 while the HySime underestimated it. Particularly, ODM using NDD significantly overestimates the number of endmembers since the Cuprite image does not include adequate number of pixels belonging to homogeneous areas.
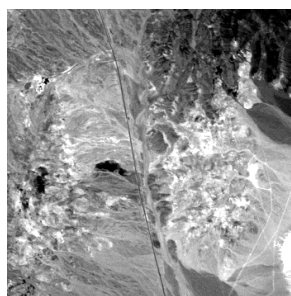
Figure 2.11:   AVIRIS Cuprite hyperspectral data.

---

[12] http://speclab.cr.usgs.gov/cuprite95.tgif.2.2um_map.gif
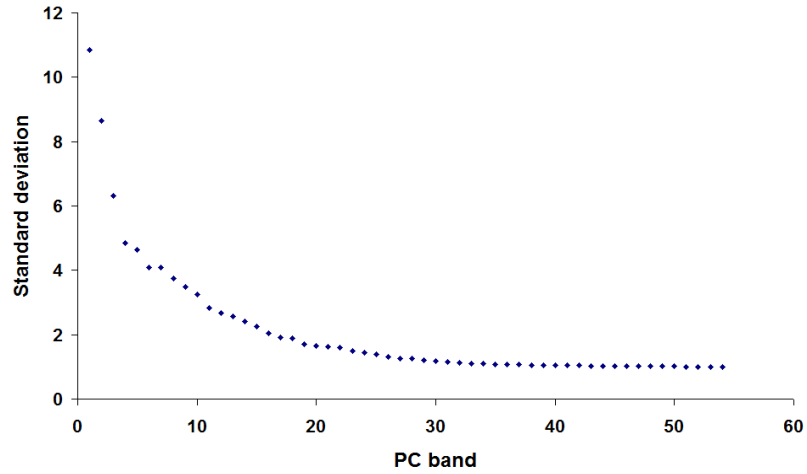
Figure 2.12: Standard deviation values of the transformed bands for AVIRIS Cuprite image.

Table 2.5: Estimated Number of Endmembers for The AVIRIS Cuprite Image
[1] stands for multiple regression, [2] stands for NND.

Method	Estimated number of endmembers (reference number: higher than 18)
ODM [1]	20
ODM [2]	29
HySime	15
NWHFC $P_f=10^{-3}$	22
$P_f=10^{-4}$	21
$P_f=10^{-5}$	19

2.5 Conclusions

In this work, a new automatic and non-parametric method for the estimation of the number of the endmembers in hyperspectral imagery was introduced. The proposed method, called outlier detection method (ODM) develops a novel approach considering signal as outlier of the noise hypersphere. In particular, after noise estimation and whitening process, the transformed data reside in a principal component space where noise presents spherically symmetry towards all the directions, having a constant radius. Conversely, signal radius varies in all the directions and it is much larger than the noise radius in the components which include it. Estimation of the number of noise hypersphere outliers using a robust inter quartile range based outlier detection method results in the estimation of the number of

endmembers. The proposed method is characterized by its simplicity and its significant benefit to refrain from estimation of statistical distributions.

Experiments using simulated data proved the efficiency of the ODM which outperformed compared to its competitors. The performance of the proposed method is quite satisfactory in real data, as well. Through this particular work it is concluded that a successful estimation of the number of endmembers strongly depends on how well signal and noise are discerned. Outlier detection theory could be efficiently used for this goal. Future research should also focus on combining the proposed method with endmember extraction methods which integrate spatial information, taking advantage of its successful estimation of the number of endmembers in small sized images.

Chapter 3

Dimensionality reduction

3.1 Introduction

High dimensionality is inherent to hyperspectral data. Hyperspectral sensors record over hundreds of narrow contiguous spectral bands. The trade-off of high spectral resolution is the huge volume of data which poses many challenging problems in terms of data storage, computational efficiency and algorithms performance [12]. In addition to these issues, high dimensionality arises the curse of dimensionality [51]. According to this, the sparsity increases exponentially with the dimensionality given a fixed amount of data sample and in order to achieve the same accuracy of regression, classification etc. in a higher dimensional space, the data sample needs also to increase which is very difficult, if not impossible, in practice. Dimensionality reduction (DR) aims at mitigating the aforementioned issues. It has been widely used as a preprocessing technique to define the data into a lower dimension in order to be analyzed more effectively. Thus, DR affects directly the performance of the following processing tasks. In this study, focus is given to DR in terms of spectral unmixing and classification purposes.

3.2 Review of the relevant literature

A significantly large proportion of the relevant research is devoted to the development of new approaches of DR since the traditional ones designed for multispectral data are not applicable with high success on hyperspectral data. The large amount of DR methods renders impossible their inclusion in this section, which is also out of the intention of this work. The particular review includes the main and most commonly used approaches for DR of hyperspectral data sets. These DR approaches can be classified into two main groups (Fig. 3.1). The first one includes feature extraction (FE) methods and the second one includes band selection (BS) methods. An FE method uses transformations to represent the data sample in a lower dimensional space exploiting the information given from the entire data set. On the other hand, BS methods search for a subset of the original spectral channels, which comprises the most distinctive and informative ones, based on certain criteria.

Depending on whether there is information in advance related to the materials existing in the scene or not, the FE methods can be further categorised into statistic-based methods and feature-based methods. The former use statistics to de-correlate the data sample and it is an unsupervised approach which requires no a priori knowledge. Such statistic-based FE methods are the second-order statistics transforms, principal component analysis (PCA) which uses the variance [52], maximum noise fraction (MNF) or noise-adjusted principal component (NAPC) transform which use the signal-to-noise ratio (SNR), and the singular value decomposition (SVD) transform [53]. According to [54], [55], second-order statistics transforms cannot retain information related to subtle materials whose proportion compared to the entire data sample is relatively small and thus, their contribution to second-order statistics is very little. To overcome this issue, higher-order statistic-based methods have been developed; third-order and fourth-order statistic-based transforms which use skewness and kurtosis, respectively [13]; and infinite-order statistic-based independent component analysis (ICA) method [56], [55] which is an unsupervised source separation process assuming that sample vectors are linearly mixed by a set of separate independent signal sources and can be used to unmix these signal sources

according to their statistical independency measured by mutual information [57]. The aforementioned high-order statistic-based methods are special cases of projection pursuit (PP) [54], [58], [59], [60]. PP uses a projection index (PI) as a criterion to find directions of interestingness of data to be processed and then represents the data in the data space specified by these new interesting directions [13]. Variance, mutual information, skewness etc. can be considered as PIs.

The feature-based FE methods produce a set of feature vectors by which data sample can be represented. These methods exploit associated a priori information of the data set such as the spectra of the classes or of the endmembers included in the scene and therefore, they are used for representation tasks or classification purposes rather than as preprocessing step for endmember extraction. A representative feature-based FE approach is the spectral mixture analysis (SMA) which besides a technique for applying physical-based model to the observed spectral vectors in order to analyse them, it can also be viewed as DR technique [1]. The SMA results in a reduction in the dimensionality of the data from the initial spectral bands to a few, newly defined axes in the data space. These axes, also called feature vectors, correspond to the p or p-1 endmembers which are quite fewer than the initial L spectral bands and the entire data can then be represented by the p-dimensional endmember space. In section 4.2, there is a description of several developed methods based on SMA. Another well known feature-based FE method is the Fisher's ratio-based linear discriminant analysis (FLDA) [61] which is implemented for supervised classification. The FLDA uses the ratio of among-class variances to within-class variances so as to achieve the best possible class separability. The derived feature vectors define the decision boundaries among the p classes and the data can be represented by them.

Contrary to FE, band selection (BS) is a DR approach which retains the physical meaning of the data set [62]. Based on this approach, a set of spectral channels from the original hyperspectral data set is selected comprising the information needed for subsequent hyperspectral image spectroscopy. Moreover, band selection results in a more general approach than FE as the used features are associated with the original spectral channels of the hyperspectral image and/or with measures that extract information from the original channels and from the spatial context of each single pixel [63]. The

majority of the BS techniques have been developed for the optimization of classification tasks. Assuming that there is information related to the classes of interest, BS methods comprise a search algorithm and a criterion function. The search algorithm generates and compares the solutions resulted from different subsets of bands by applying the criterion function as a measure of the effectiveness of each considered band subset. The criterion functions include class separability based criteria such as Bhatacharyya, Jeffries-Matusita and Mahalanobis distances [64], [65], [66] which according to [67] require enough class samples to examine class statistics; and similarity based criteria such as spectral angle mapper (SAM) [68], euclidean minimum distance (EMD) measures and orthogonal projection divergence (OPD) [11], [69]. The aforementioned criteria measure all the possible pair-wise class distances which has a heavy computational cost. To avoid testing all the possible combinations, subset searching strategies, e.g., sequential forward selection (SFS) and sequential forward floating selection (SFFS), can be used [70]. The minimum estimated abundance covariance (MEAC) method, is a supervised BS method proposed in [67], which selects bands based on class spectral signatures only. BS methods used for classification may also perform genetic algorithm and particle swarm optimization [71], [72] which do not require class knowledge in advance.

A relatively small proportion of the existing band selection DR methods can also be applied for optimizing the spectral unmixing process. Among the state-of-the-art techniques, some may include band ranking using criteria such as correlation [73] and/or information entropy and mutual information [62], [74]. The band selecting sparsity promoting iterated constrained endmember (B-SPICE) algorithm [75] has been developed for simultaneous BS and endmember extraction using the band weighting proposed in [72] and sparsity promoting priors applied to band weights.
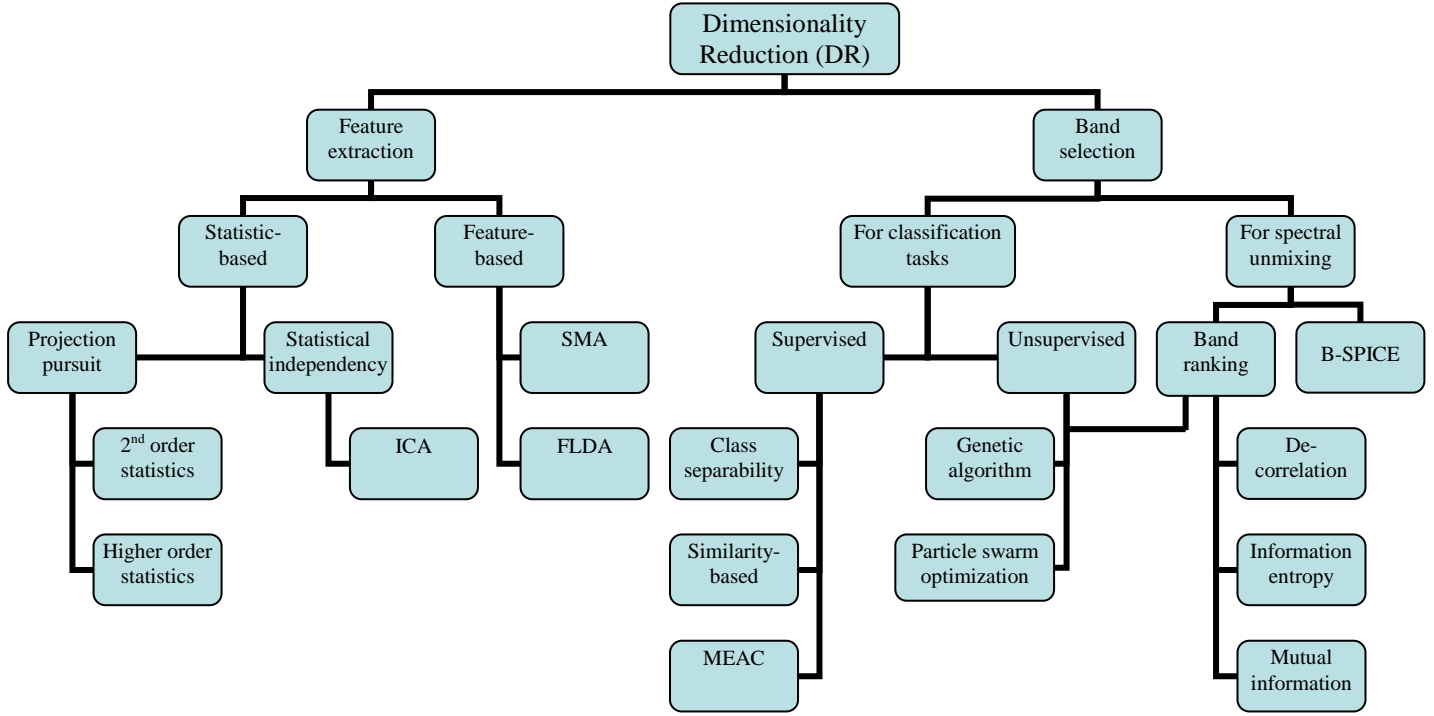
Figure 3.1: Categorization of dimensionality reduction methods.

3.3 Study of new band selection approach

The high spectral resolution of the hyperspectral data offers an invaluable diversity of information, enabling the accurate physical description and discrimination of the sensed materials. Specific spectral characteristics between similar materials are present at certain wavelengths and this crucial and critical information may be compromised and distorted if no original data are used. BS preserves the original data information by selecting a subset of spectral bands based on a criterion for optimality. The majority of the existing BS methods set the fixed criteria to the spectral information on the whole set of wavelengths in order to define the optimum spectral subspace. As a result, the redundant or overlapped data information of adjacent bands is removed. Although, this is an effective way to reduce the vast amount of data information, it is a generalised approach which disregards the spectral characteristics of the particular materials of interest. For instance, in case of an image which depicts water and various types of vegetation, the existing BS techniques would retain the same set of bands whether it is needed

51

to extract a representative spectrum of water and vegetation or it is needed to extract the different types of vegetation. In the latter case, the different types of vegetation present their biophysical variations at specific sequential wavelengths. For instance, vegetation stress can be detected using red edge shift at wavelengths ranging from 650 to 700nm [76]. In such scenario, BS methods would select uncorrelated bands from different wavelengths jeopardizing the loss of important but subtle absorption features which imply such significant biophysical variations.

In this work, a new band selection approach is explored which defines the optimum spectral subspace associated with a specific material. More precisely, each spectral signature of a pure material has spectral characteristics which contribute to distinguish it from other spectral signatures at specific consecutive wavelengths. Band selection should be tied to the desired analysis task and therefore, different sets of bands should be selected for detecting different materials depicted in the same scene, i.e. lentil plants from tare (vicia hirsula), or materials under various biophysical status, i.e. healthy olive trees from those affected by disease. Instead of merging bands or transforming the data, the new approach maintains only those sequential bands that are useful for a specific hyperspectral application. The advantage of maintaining physically meaningful bands has great impact both on the endmember extraction and classification of hyperspectral sensed data. Hence, a new BS method is proposed which can be adopted by the endmember extraction and classification techniques. More precisely, the proposed method enables endmember extraction and classification algorithms to act locally in the hyperspectral space, and instead of a list of selected bands, the outcome results in optimized performance of the endmember extraction and of the classification algorithms.

3.3.1 Exploration of the new band selection method for optimizing spectrally close endmember extraction

Although the majority of the endmember extraction algorithms (EEAs) are preceded by a DR method, as it is concluded from section 3.2., most of the BS methods have been developed for classification purposes and require labeled training data in order to determine the optimum set of bands for discriminating the desired classes. Based on the linear spectral mixture analysis, DR is associated to the number of the endmembers existing in the scene since the dimension of the signal subspace is one less than the number of the endmembers. This is of crucial importance for the implementation of convex-geometry based EEAs (see section 4.3.1). The existing convex-geometry based EEAs usually find the most prevalent endmembers on the scene whilst they fail to extract less ubiquitous materials [77]. Therefore, they implicitly assume that the prevalent endmembers could provide global characterization of the image. Furthermore, by searching for endmembers at the vertices of the simplex, elementary material signatures are imposed to be far from each other. Nevertheless, each elementary material due to biophysical/chemical/environmental factors may present spectral signatures with absorption features that differ at a few wavelengths only. In order to prevent confusion with the prevalent endmembers, the latter are termed in the frame of this work as spectrally close endmembers (SCEs). The SCEs correspond to spectral signatures of a specific material class under different physical and/or chemical status bearing significant detailed information about it.

Fig. 3.2 presents such scenario using a CASI image collected over a littoral Greek region. Fig.3.2(b) shows the clustering using the prevalent endmembers, i.e. seawater, vegetation, soil, sand, roads and buildings whose corresponding spectral classes are plotted in Fig. 3.2(c) using two bands. Each spectral class bears its own participation into the data cloud, defining a particular region. Inside this region, there are spectra which correspond to the same material but with subtle differences due to a) signature variability [78], and b) SCEs whose spectral differences imply different biophysical or chemical properties. The latter are of great importance and two examples of those are shown in Fig. 3.2(c); two correspond to vegetation (grass, tree) and one to seawater (seawater with algae).

Due to inter-band correlation of adjacent bands, the BS methods lack to retain the certain range of wavelengths where these spectral differences exist. As a result, EEAs fail to extract the aforementioned SCEs which present strong spectral similarities along the whole spectrum. Moreover, due to the spectral signature variability [78] and the image diversity, new BS approaches should be devised which will be tied to the spectral characteristics of the materials comprised in the scene. The proposed BS approach aims at overcoming the shortcomings of the existing BS methods by considering the existence of multiple convex hulls [79].

| (a) | (b) | (c) |
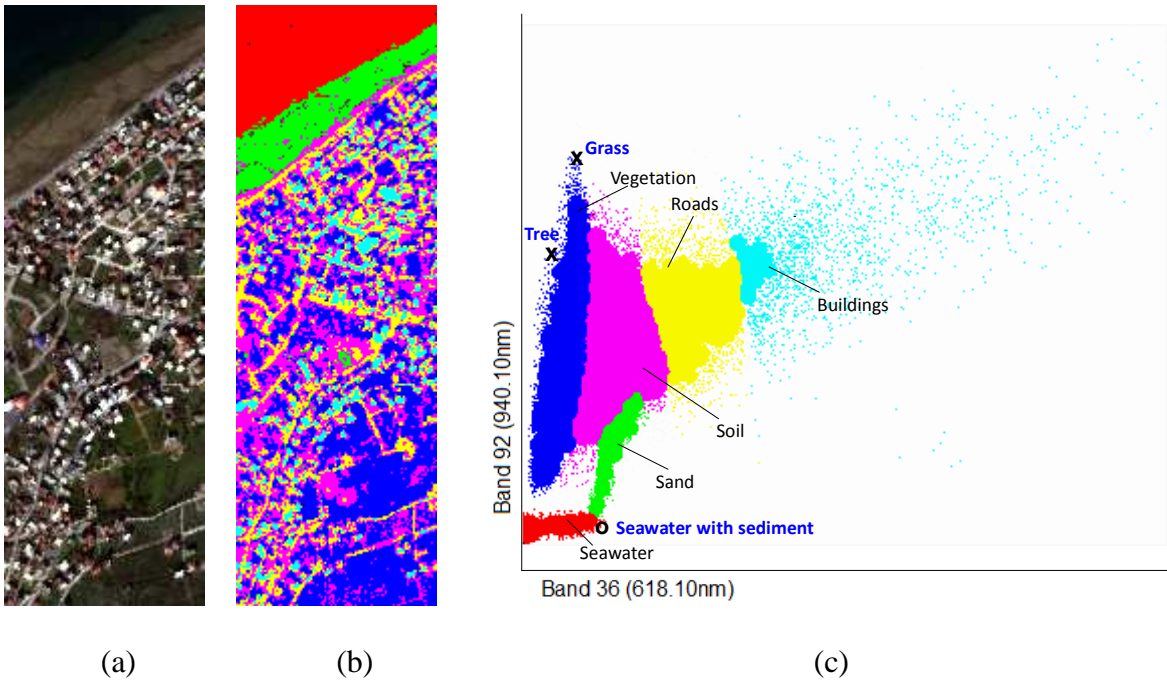
Figure 3.2: (a) CASI RGB (640,8nm, 550,6nm, 461,1nm) image, (b) spectral classes obtained by clustering based on prevalent endmembers, (c) spectral classes illustrated in the scatter plot with two spectral bands. Symbols x and o represent the spectrally close endmembers (SCE).

Analysis of the proposed method

The new band selection approach is based on the theory of convex sets. According to it, all the observed spectral vectors lie in the convex hull defined by the endmember signatures [22], [57], [80]. The new approach considers the existence of more than one convex hulls, each of them defined by the spectral

vectors of a particular material class presented in the scene. Thus, the SCEs are expected to lie at the vertices of the convex hulls which are spanned by spectral vectors of the same material. The concept of considering multiple convex hulls, which may be overlapped, has been introduced recently in [79] for the endmember extraction. The necessity of developing new methods for identifying multiple simplex regions was also pointed out in [12]. In this work, the multiple convex sets are expanded to BS purposes.

More precisely, the new approach explores the information given from the entire range of wavelengths by dividing it into smaller subsets. By this division, it is assured that bands which present the subtle spectral characteristics implying different biophysical or chemical properties of the same material will be retained. For each subset of bands, p candidate SCEs are extracted based on a convex-geometrical based EEA. In particular, the N-FINDR [81] algorithm is employed but instead of using the MNF bands, subsets of original bands are used. N-FINDR is considered one of the state-of-the-art EEAs, characterized by its coherent concept of volume maximization [13]. Then, candidate SCEs are spectrally compared to the material of interest contained in the prevalent endmembers (PE). The PEs can be selected manually from the image or by using an EEA. In our case, the PEs are extracted by the N-FINDR algorithm. Matching between the candidate SCEs and the material of interest is accomplished by using the spectral angle distance (SAD) measure[13]. Based on this measure, the final set of n candidate endmembers related to the material of interest is defined, where $0 \le n \le p$. The procedure is repeated until all subsets of bands are selected and consequently until all the sets of candidates endmembers from each subset of bands are defined. The optimum set of endmembers related to the material of interest is user defined. In case there is no matching, SCEs are considered as new materials and they can be labeled by the user or by using a spectral library. Fig.3.3 presents the flowchart of the proposed method.

---

[13] SAD is computed by $S_{(ce_1)} = \sum^{p-1} \cos^{-1} \left(\dfrac{\langle ce_1, c_{ei} \rangle}{\|ce_1\| \cdot \|c_{ei}\|} \right)$, where e_1 is a prevalent endmember and e_i a candidate SCE.

For each subset, the number of bands is set equal to the signal subspace dimension p-1 , where p is estimated by the ODM [25] because p-1 is much smaller than the initial number of bands which contributes to a desirable division scale of the hyperspectral space, whilst at the same time it is not too small to produce overlapping results. It should be noted that this number serves as an indicator for the maximum possible number of SCEs assigned to each material in one iteration, and it may vary for each material depending on the SAD values. This is of crucial importance as for the first time in BS, multiple convex hulls with various shapes assigned to each material are introduced. Convex hulls of spectrally similar materials may be overlapped and often are not detected. Detecting SCEs compensates this shortcoming and may detect materials which are not extracted by conventional methods. Furthermore, because there is spectral variability among different data sets, detected SCEs are image-based. For these reasons, the proposed BS approach is tied to the available data and the material of interest, and different sets of bands are selected for different images.
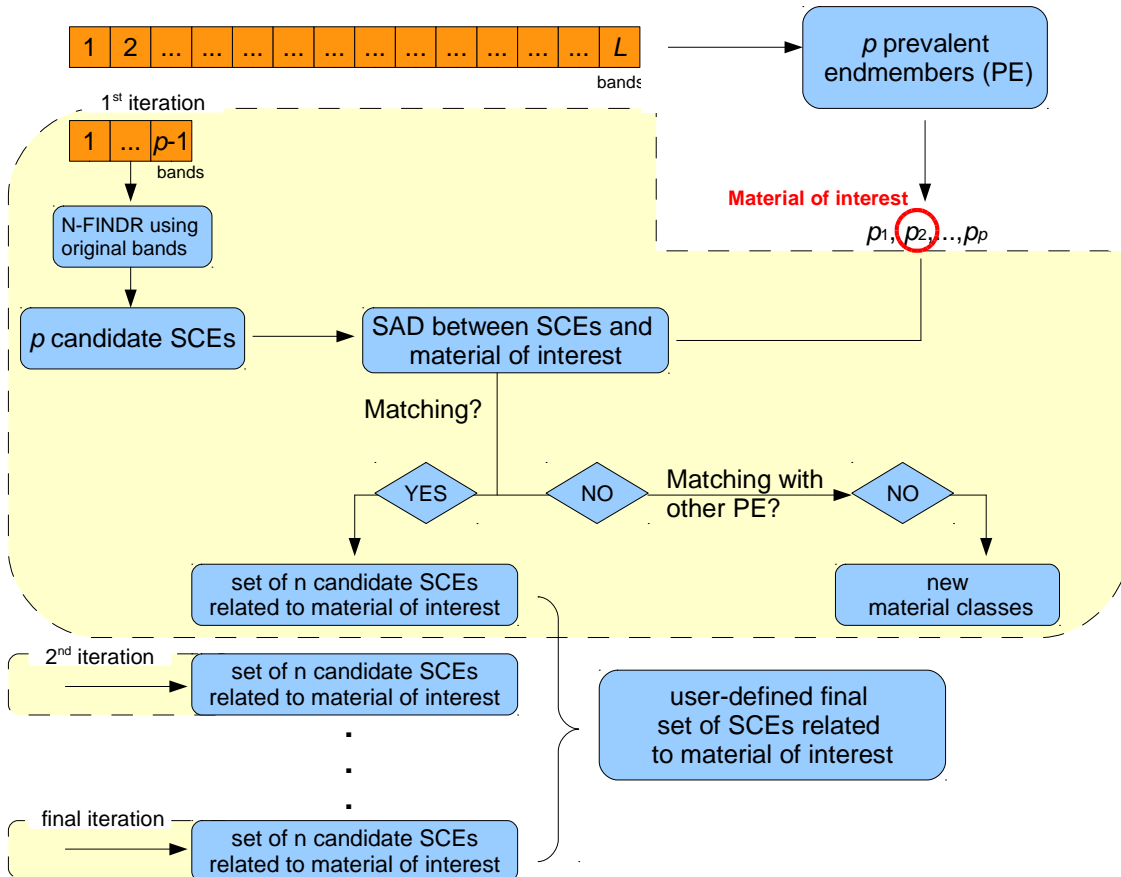
Figure 3.3: Flowchart of the new BS method for endmember extraction.

3.3.1.1 Experiments

A real hyperspectral image collected by CASI-550 over a Greek littoral region in 2007 was used for the experiments (Fig.3.2(a)). It consists of 100 x 300 pixels with 97 spectral bands covering a spectral range from 421nm to 963nm. After bad band removal (1, 61, 95-97) due to low SNR values, 92 spectral bands were retained. The spatial resolution of the image is 4m. The radiance values were converted to top-of-atmosphere reflectance.

In this experiment, the results of the proposed algorithm have been visually interpreted by the user based on the spectral signatures and their location in the area of interest. Extensive reference data certainly could complete the evaluation process but in such case specific laboratory analysis by other scientific groups would be required, which has not been foreseen in the framework of this study.

At first, the conventional N-FINDR was applied for the extraction of the prevalent endmembers. The number of endmembers to be extracted was resulted in 14 using the ODM [25], (see section 2), for the estimation of the signal subspace dimension. Fig. 3.4 contains the spectral signatures of the extracted prevalent endmembers. The majority of the extracted endmembers correspond to roof materials (RM) which can be explained by the fact that N-FINDR implements DR using FE method which has the tendency to select high contrast signals (i.e. urban materials) over low contrast signals [54], [55]. Since urban materials, i.e. tiles, concrete, marbles, metallic roofs etc., are satisfactorily detected and they correspond to different elementary substances, there is no point in searching for SCEs of each RM class. Among the extracted prevalent endmembers, three correspond to vegetation and one to seawater.
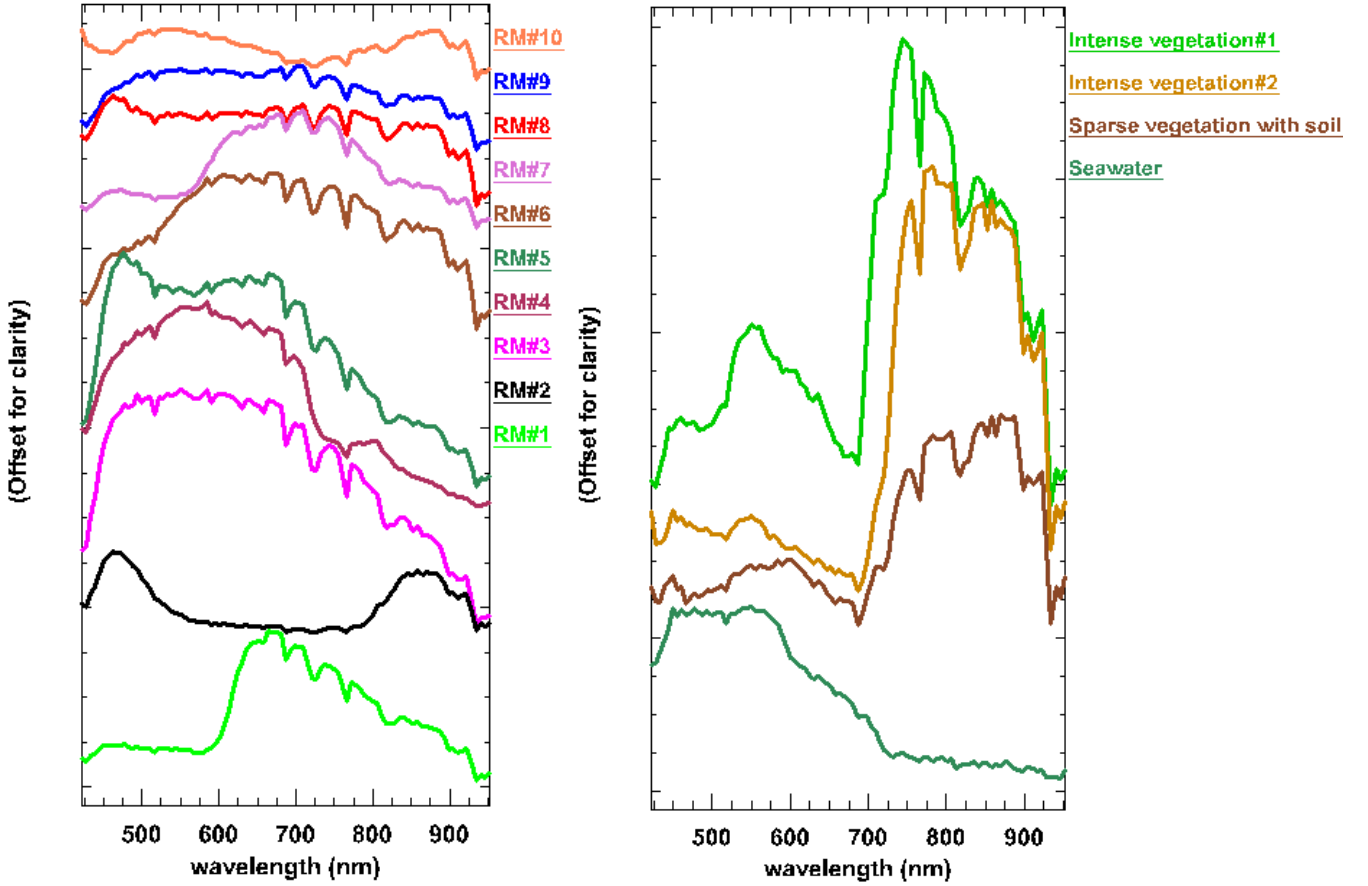
Figure 3.4: Prevalent endmembers extracted by N-FINDR. RM stands for roof materials.

Following the methodology described in Fig. 3.3, several SCEs were extracted. In particular, six SCEs were assigned to vegetation class among which three (endmember#1/#2/#6) belong to the prevalent endmembers, and their spectral signatures are plotted in Fig. 3.5. The spectral signatures present differences regarding the chlorophyll absorption of leaf pigments which is implied by the peak shown at 550nm. Moreover, based on the location of the extracted SCEs in the image and after visual interpretation of the scene, various types of vegetation have been detected, such as grass, trees, and bushes. Parts of the classified images are presented in Fig. 3.5(c) and (d). In Fig.3.5(c), classification uses the prevalent endmembers, while in Fig.3.5(d), classification uses the final set of endmembers related to vegetation. It is worth to mention that the reason for including classification results is only for facilitating the visual interpretation of the SCEs. It is not intended to evaluate the proposed method via the classified images since the extracted endmembers may be present only in few pixels and classification errors may be embedded in the results.
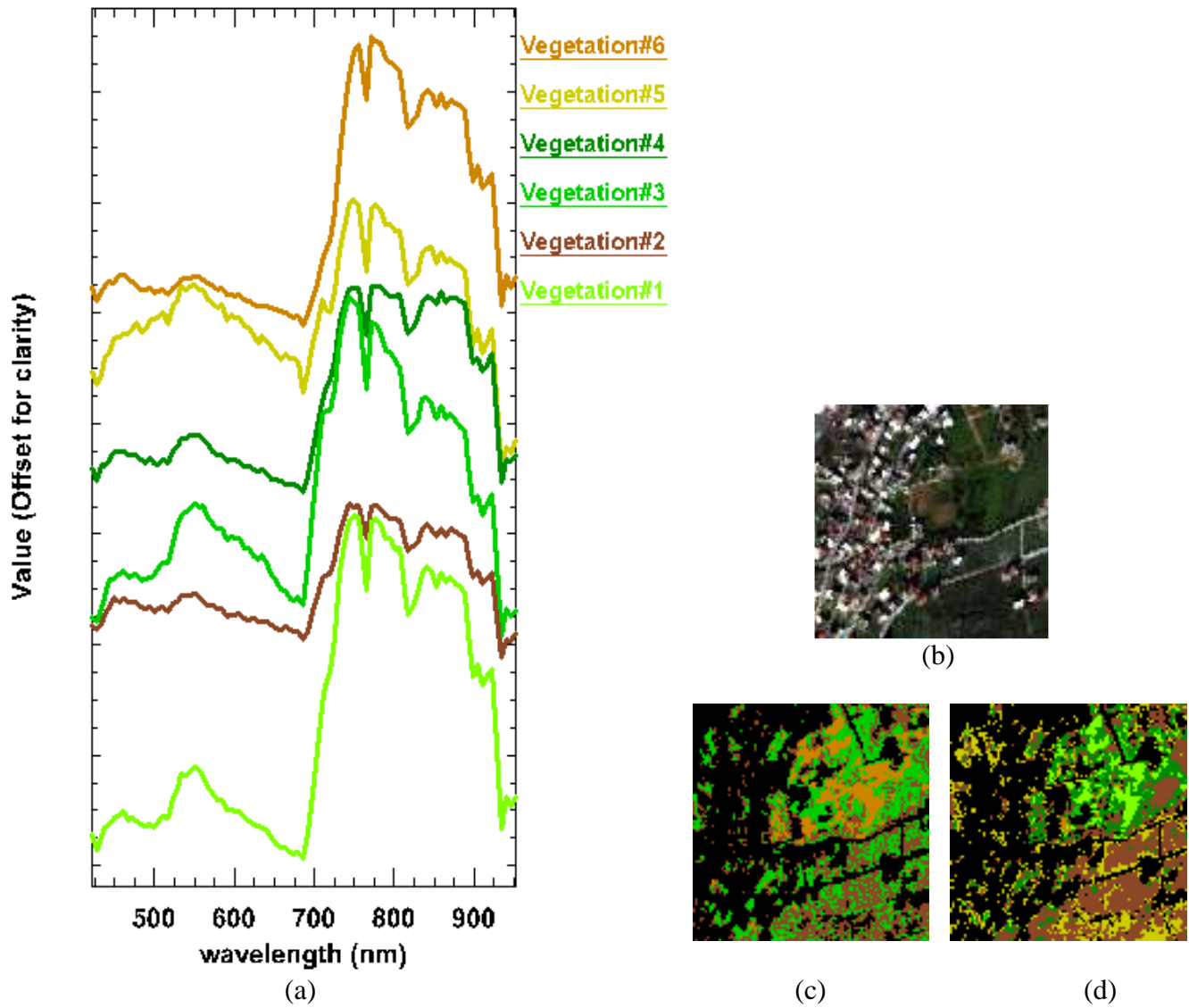
Figure 3.5: (a) Extracted SCEs assigned to vegetation, (b) a part of the original data set, (c) classified image using the prevalent vegetation endmembers, (d) classified image using the final set of endmembers.

A set of seven SCEs were also extracted which correspond to seawater class among which one (seawater#1) belongs to the prevalent endmembers. The effectiveness of using the proposed approach over a conventional EEA is evident in Fig. 3.6.(c) and (d) where in the latter the observed vectors are classified into categories which could be labeled in case that extended reference data were available regarding the bottom depth and cover (algae, sand, etc), as well as the suspended materials.
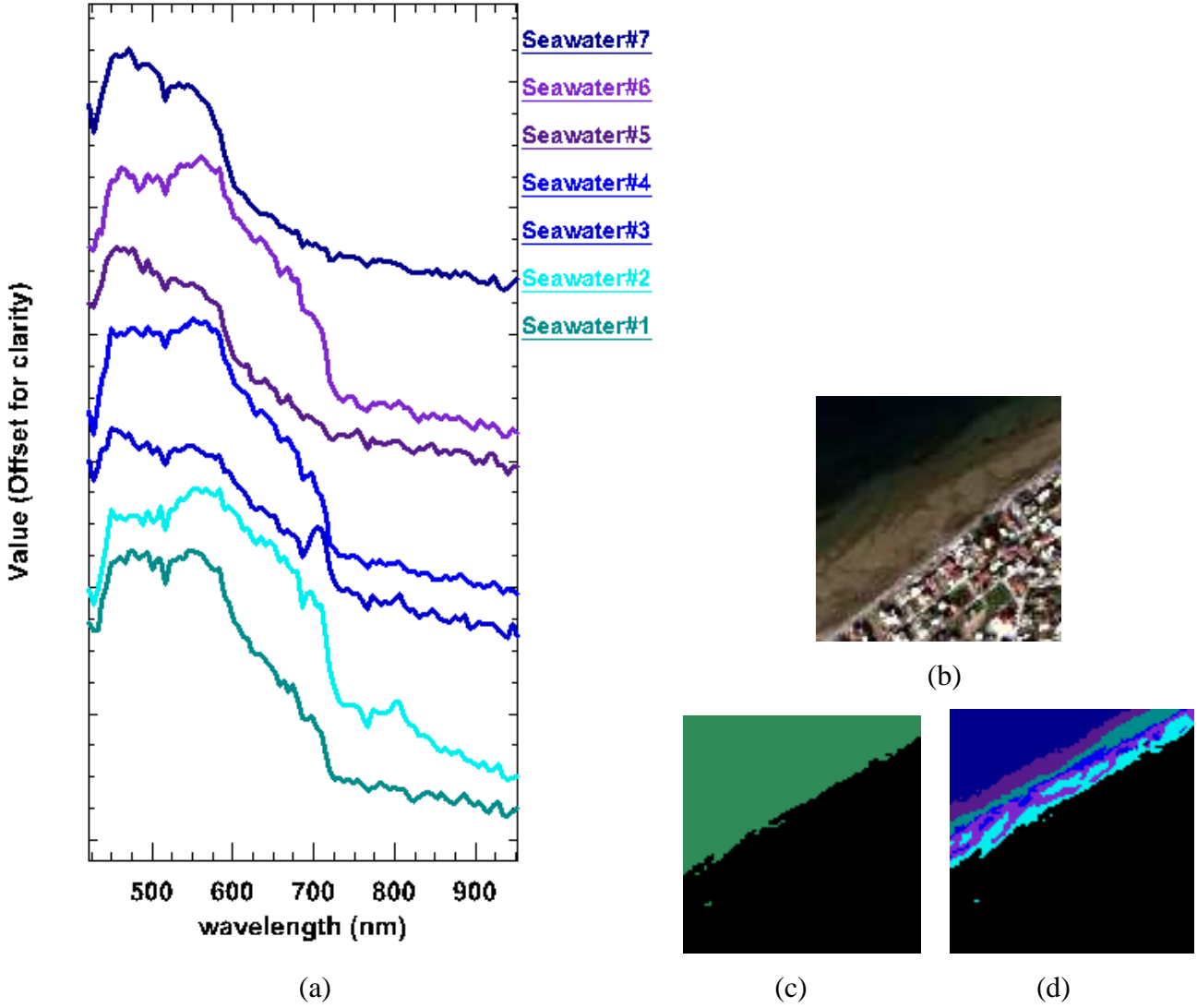
Figure 3.6: (a) Extracted SCEs assigned to seawater, (b) a part of the original data set, (c) classified image using the prevalent seawater endmember, (d) classified image using the final set of endmembers.

Besides the SCEs which match to a class defined from the prevalent endmembers, there are also sets of SCEs which correspond to newly defined classes. More precisely, in our case, extracted SCEs indicate reasonably the existence of three new classes: roads, sand, and shadow.

The SCEs of roads are shown in Fig. 3.7. Based on the location of the extracted SCEs in the image, asphalt road#1 corresponds to the main road network with asphalt, while asphalt road#2 to open-air parking places. Fig. 3.8 shows the SCEs of sand, which mainly differ regarding their concentration in water. Finally, in Fig. 3.9, the two last SCEs that correspond to shadow are shown.
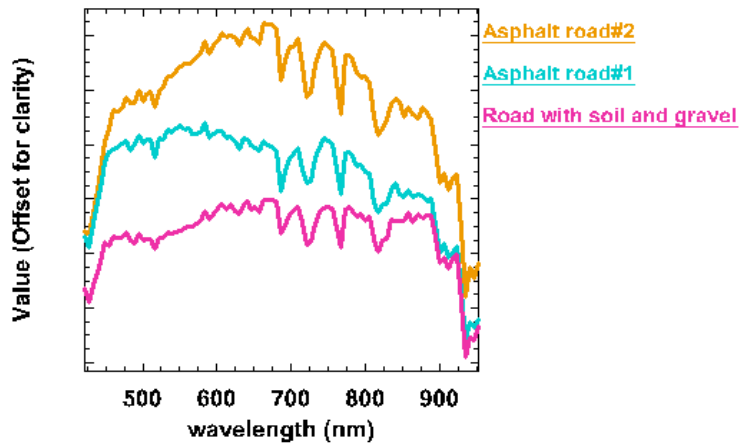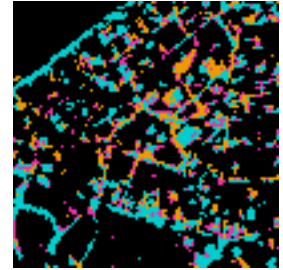
Figure 3.7: (a) Extracted SCEs assigned to roads, (b) a part of the original data set, (c) classified image using the SCEs of roads.
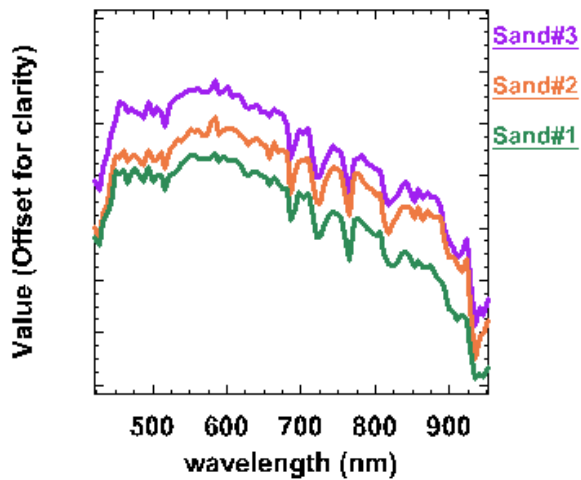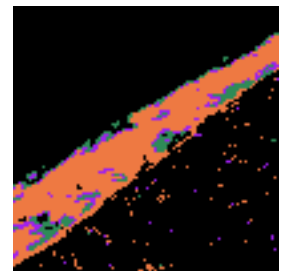
Figure 3.8: (a) Extracted SCEs assigned to sand, (b) a part of the original data set, (c) classified image using the SCEs of sand.
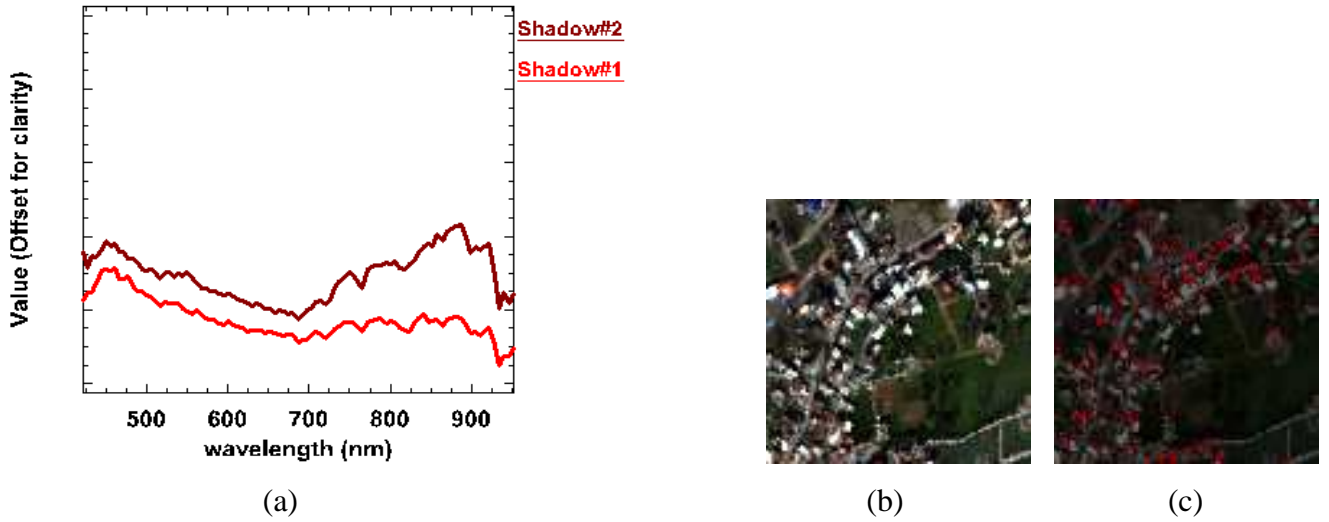
Figure 3.9: (a) Extracted SCEs assigned to shadows, (b) a part of the original data set, (c) a subset of the original data set overlaid with the classified image using the SCEs of shadows.

3.3.1.2 Conclusions

In this work, a new band selection approach has been introduced for optimizing the performance of the endmember extraction. The proposed method accounts for the existence of elementary material which due to biophysical/chemical/environmental factors may present spectral signatures with absorption features that differ at a few wavelengths only. These spectra are termed in the frame of this work as *spectrally close endmembers* (SCEs). The SCEs correspond to spectra of the same material class which present subtle spectral differences, crucial for their distinction. In order to extract the SCEs, the concept of multiple convex hulls [79] is exploited for the first time in BS. Each convex hull is defined by the spectral vectors of a particular material class presented in the scene and the SCEs which lie at its vertices. Contrary to the existing BS methods, the proposed approach accounts for the different distribution of each material's convex hull to the data cloud. For this reason, the proposed BS approach allows the number of SCEs for each material class to vary resulting in more physically meaningful spectra. Furthermore, the proposed BS accounts for image diversity and is tied to the available information derived from the image's spectral characteristics. Thus, different sets of bands are exploited for detecting different materials depicted in the same image. Being the method image-dependent

compensates the endmember variability, contrary to the existing approaches which use a priori information related to the absorption features at fixed wavelengths.

Experiments were implemented using a CASI real hyperspectral image leading to the following concluding remarks. Firstly, the proposed approach contributed to the detection of low contrast materials, disregarded by the N-FINDR algorithm which –as the majority of the convex-based EEAs- has the tendency to detect the prevalent endmembers (i.e. urban materials) over less ubiquitous endmembers. Secondly, the extracted SCEs provided wealth and elaborate information for material classes defined by the prevalent endmembers. In addition, the effectiveness of the proposed approach is also shown on classification results. The inclusion of the optimum set of spectral signatures, which represent a certain material class in the classification process, enables studying it without need of masking the background. Last but not least, it should be noted that the aim of the proposed approach is to develop a BS method able to extract spectral signatures of SCEs which contain subtle information for a material.

3.3.2 Exploration of the new band selection method for optimizing classification

High dimensionality demands special attention in the classification process. The main problem caused by the increased dimensionality is the reduction on the efficiency of the classifiers. This problem is known as the Hughes phenomenon. According to this, as the dimensionality increases and the number of the available training samples is kept fixed, the predictive power of the classification methods reduces. The number of pixels included in training samples is affected by the number of bands [82]. In case of hyperspectral imagery, training data sets are never enough to achieve classification requirements. Thus, dimensionality reduction besides yielding gains in computational time and complexity, and data storage, it also contributes to optimizing the classification performance when the available training data sets are limited.

The advantage of maintaining physically meaningful bands is associated to the identification of the useful wavelengths for a particular classification task [27]. Furthermore, it is critical to find useful and effective features associated to the defined classes of interest for image classification tasks of hyperspectral sensed data. The spectrum of each class of interest has unique spectral characteristics in a narrow region of the spectrum, which contribute to distinguish it from other classes at specific sequential wavelengths. Thus, investigation on the band selection technique can be performed using the original data, instead of merging bands or transforming the data. The main goals of the present study are:

1) to examine how robust are the classification results when different bands are selected to be used,

2) to explore whether there is an optimum set of sequential bands which contribute more for a specific hyperspectral application compared to other sets of bands, and

3) to explore whether the classification of different classes can be improved by using different set of bands for each class.

In order to accomplish this, the spectral space is divided into subsets of equal number of sequential bands. Each subset is used as input to a classification algorithm and results are examined separately. The dimension of the band subsets is equal to p-1, where p is the number of signals. The outlier detection method (ODM) [24], [25] is implemented for the estimation of the number of signals. SVM is implemented for all the possible subsets of p-1 sequential bands and classification accuracy is examined per class. SVM has been chosen to be implemented as it has proven to be much more effective than other parametric classifiers in terms of classification accuracies, computational complexity and robustness to parameter setting [83]. Those subsets of bands which contribute to the highest classification accuracy of each class are kept as the optima. Fig.3.10 shows the flowchart of the proposed method.

For evaluation purposes, two well-known unsupervised band selection methods based on information entropy and correlation were also applied prior to classification. Evaluation of the classification accuracy leads to remarkable conclusions.
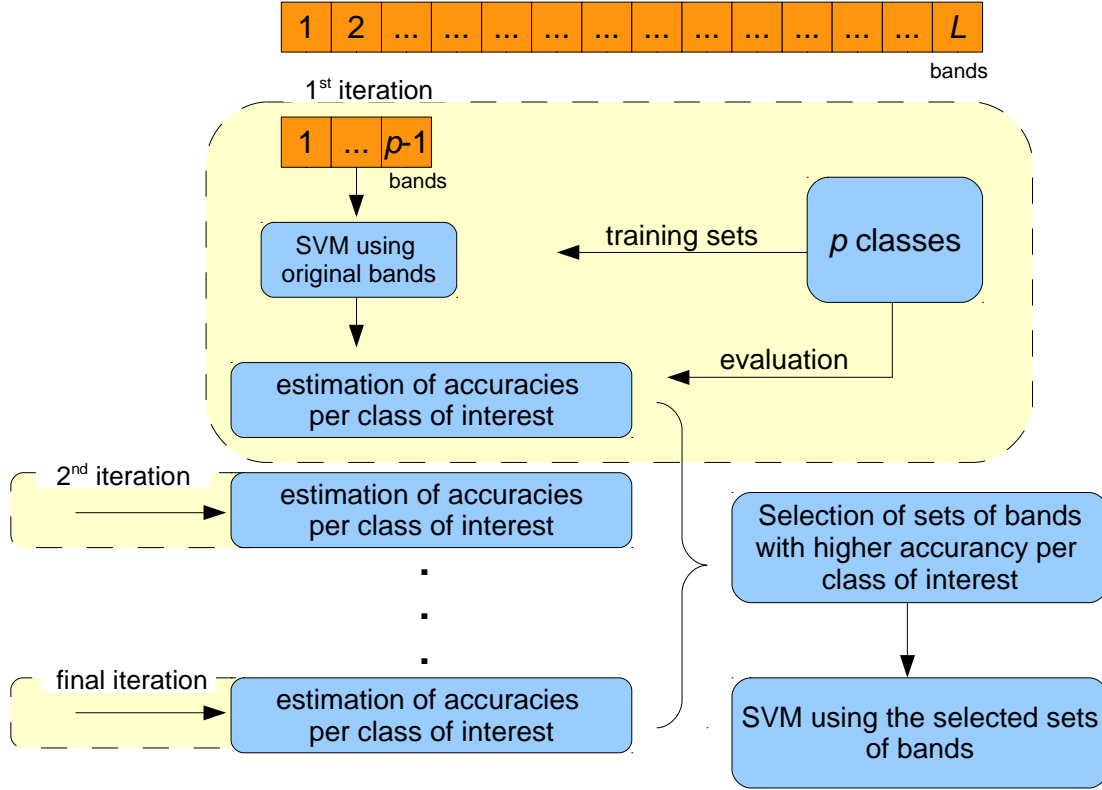
Figure 3.10: Flowchart of the new BS method for classification.

Support vector machine

Support vector machines (SVM) are widely known classifiers based on statistical learning theory. The use of kernels leads to data transformation into a new feature space where the data is hoped to be nearly linearly separable [84]. The SVM method attempts to separate training samples belonging to different classes by tracing maximum margin hyperplanes in the space where the samples are mapped.

Assume a binary classification problem in a B-dimensional space $\Re^B$. Let N be the training samples $x_i \in \Re^B, i = 1,...,N$ and let $y_i = \pm 1$: $\{(\mathbf{x}_i, y_i) / i \in [1,N]\}$ be their corresponding labels. The SVM method aims at finding a hyperplane H_p that has the largest distance from the closest training data points in both classes. Denoting $\mathbf{w} \in \Re^B$ as the vector normal to the hyperplane H and $b \in \Re$ as its bias, H_p is given by

$$\mathbf{w} \cdot \mathbf{x} + b = 0, \ \forall \mathbf{x} \in H_p. \tag{3.1}$$

In case that $\mathbf{x} \notin H_p$, the distance of $\mathbf{x}$ to H_p is given by

$$f(x) = \frac{|\mathbf{x} \cdot \mathbf{w} + b|}{\|\mathbf{w}\|} \tag{3.2}$$

Maximizing the distance of samples from the optimal decision hyperplane is equivalent to minimizing the norm of $\mathbf{w}$, leading to the following quadratic optimization problem:

$$\min\left\{\frac{\|\mathbf{w}\|^2}{2}\right\} \tag{3.3}$$

subject to $\quad y_i(\mathbf{w} \cdot \mathbf{x}_i + b) > 1, \ \forall I \in [1,N]$ in case of linearly separable data

$$\text{or } \min\left\{\frac{\|\mathbf{w}\|^2}{2} + C\sum_i \xi_i\right\} \tag{3.4}$$

subject to $y_i(\mathbf{w} \cdot \mathbf{x}_i + b) > 1 - \xi i, \ \xi i \geq 0 \ \forall I \in [1,N]$ in case of non-linearly separable data,

where $C \ (0 < C < \infty)$ is a regularization parameter that controls the penalty value and ξ are the so-called slack variables which are introduced to deal with misclassified samples. Based on Lagrangian formulation of the problem, $\mathbf{w}$ is given by

$$\mathbf{w} = \sum_{i=1}^{N_s} a_i y_i x_i \ , \tag{3.5}$$

where N_s is the number of the support vectors, i.e. the closest vectors to the hyperplane, and a_i are the positive Lagrange multipliers. For an unseen sample x_u, label y_u is derived by the following equation:

$$y_u = sgn(\mathbf{w} \cdot \mathbf{x}_u + b) = \text{sgn}\left(\left(\sum_{i=1}^{N_s} a_i y_i x_i\right) x_u + b\right) = \text{sgn}\left(\sum_{i=1}^{N_s} a_i y_i \left(x_i^T x_u\right) + b\right) \tag{3.6}$$

where $(\mathbf{w}, b)$ are the hyperplane parameters found after the completion of the training process. Eq. (3.6) uses dot products of x_u with $\mathbf{w}$, which can be replaced by kernel functions, in case of nonlinear SVM solutions in the data.

Two kernel functions are widely used for hyperspectral image classification, the polynomial function and the Gaussian radial basis function (RBF). More detailed analysis of the SVM can be found in [84], [85].

3.3.2.1 Experiments

Experiments were implemented using the AVIRIS Indian pines real hyperspectral remotely sensed image which was described in section 2.4. Classification of this dataset is very challenging since a) a significant amount of pixels has no ground truth information, b) the crops in the image are in early growth stages and thus have only about a 5% crop cover, c) moderate spatial resolution is 20m.
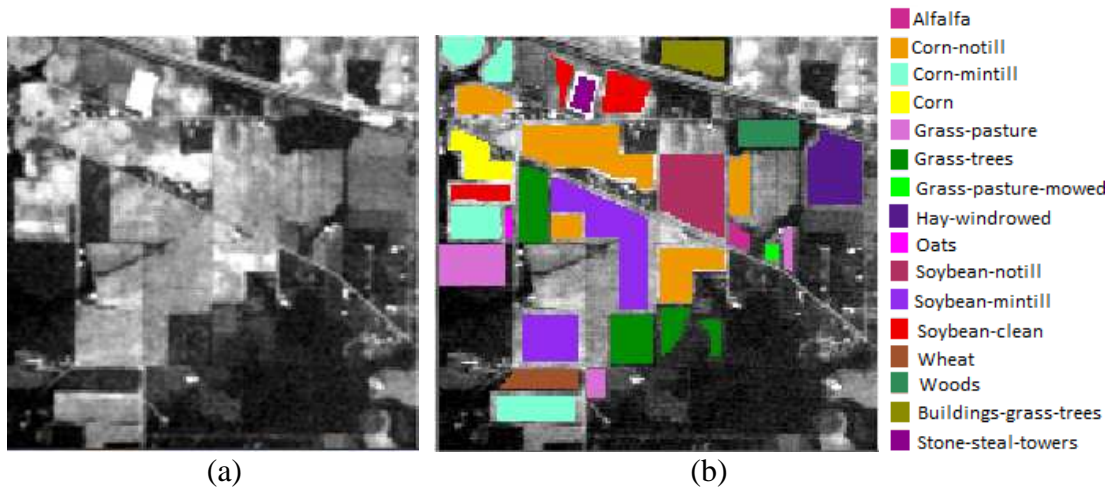
Figure 3.11: (a) Sample band of Indian Pine dataset, (b) ground-truth dataset.

At first, ODM was applied and resulted in the number of signals equal to 17. Therefore, the original data was divided into subsets of 16 sequential bands. The 16 land-cover classes were used to generate a set of 3046 training samples (used for learning the classifiers) and a set of 3058 test samples (exploited for assessing their accuracies) (see Fig.3.11 and Table 3.1).

Table 3.1: Number of training and test samples used in the experiments.

Category	Training data	Testing data
Soybean-notill	254	260
Soybean-mintill	468	466
Corn-notill	627	624
Grass/Pasture	207	206
Corn-mintill	311	315
Grass/Trees	326	327
Soybean-clean	170	169
Buildings/Grass/Trees	102	102
Stone-Steel-Towers	35	36
Alfalfa	27	26
Corn	97	105
Grass-pasture-moved	12	12
Hay-windrowed	213	216
Oats	10	10
Wheat	87	84
Woods	100	100
Total	3046	3058

In total, the SVM method was run 172 times, (186 original bands – 14 dimensions of the signals). A nonlinear SVM based on Gaussian radial basis kernel functions has been considered for the experiments. The nonlinear SVM requires the determination of the width parameter of the Gaussian radial basis kernels, which tunes the smoothing of the discriminant function. For the considered dataset, the best value of the parameter C was 5000. The optimal kernel width parameter was found equal to 0.005. These values were estimated empirically on the basis of the available training samples. Table 3.2 reports the highest accuracy per class using different subsets of original bands, which are equal to 86 non-overlapped bands.

Table 3.2: The highest accuracy per class using different subsets of original bands.

Bands	Prod. Acc/Category															
	Soybean-notill	Soybean-mintill	Corn-notill	Grass-pasture	Corn-mintill	Grass-trees	Soybean-clean	Buildings-Grass	Stone-Steel-Towers	Alfalfa	Corn	Grass-pasture moved	Hay-windrowed	Oats	Wheat	Woods
2-17	38.08	81.55	57.73	88.83	21.27	87.77	7.10	26.47	94.44	92.31	4.76	0.00	89.81	0.00	88.10	49.00
12-27	63.08	77.25	40.71	89.81	6.03	93.88	2.37	39.22	100.00	11.54	14.29	58.33	97.69	10.00	86.90	71.00
13-28	52.69	77.90	38.46	90.29	7.30	93.58	1.78	38.24	100.00	11.54	10.48	41.67	98.15	70.00	86.90	69.00
16-31	60.38	86.70	24.84	93.69	11.75	95.72	4.14	37.25	100.00	11.54	23.81	25.00	100.00	50.00	97.62	66.00
18-33	70.38	86.48	22.60	97.57	27.30	93.58	12.43	37.25	100.00	7.69	48.57	50.00	98.61	30.00	96.43	60.00
23-38	58.08	89.48	40.87	92.23	34.92	95.41	11.83	42.16	100.00	19.23	48.57	25.00	97.69	50.00	95.24	48.00
25-40	51.54	85.84	35.58	90.78	42.54	96.94	9.47	45.10	100.00	11.54	42.86	41.67	97.69	50.00	96.43	39.00
28-43	55.38	83.26	47.92	87.38	29.21	97.25	8.88	48.04	100.00	3.85	34.29	25.00	99.54	50.00	95.24	32.00
47-62	36.92	46.14	41.83	86.41	19.05	85.63	4.73	69.61	77.78	61.54	5.71	8.33	96.76	0.00	100.00	35.00
132-147	65.00	76.18	79.49	69.42	16.19	83.79	1.78	19.61	75.00	3.85	6.67	0.00	98.61	0.00	92.86	55.00
154-169	0.00	56.87	65.87	65.05	16.83	74.62	23.08	16.67	100.00	0.00	9.52	0.00	98.61	0.00	71.43	67.00

In order to investigate the effectiveness of the reported 86 bands (Table 3.2) on the overall accuracy, SVM was applied using them (Table 3.3). Table 3.4 shows the SVM results using the whole set of 186 bands.

Table 3.3: Overall and class-by-class accuracies using subset of 86 bands.

Overall Accuracy = (2388/3058) 78.0903% Kappa Coefficient = 0.7505				
Class	Prod. Acc. (Percent)	User Acc. (Percent)	Prod. Acc. (Pixels)	User Acc. (Pixels)
Soybean-notill	91.15	99.58	237/260	237/238
Soybean-mintill	97.42	64.58	454/466	454/703
Corn-notill	80.93	68.24	505/624	505/740
Grass-pasture	100.00	78.03	206/206	206/264
Corn-mintill	22.54	88.75	71/315	71/80
Grass-trees	96.94	91.62	317/327	317/346
Soybean-clean	50.30	85.00	85/169	85/100
Buildings-Grass-Trees	46.08	75.81	47/102	47/62
Stone-Steel-Towers	100.00	100.00	36/36	36/36
Alfalfa	92.31	100.00	24/26	24/24
Corn	37.14	46.99	39/105	39/83
Grass-pasture-moved	75.00	64.29	9/12	9/14
Hay-windrowed	100.00	98.63	216/216	216/219
Oats	90.00	100.00	9/10	9/9
Wheat	100.00	97.67	84/84	84/86
Woods	49.00	90.74	49/100	49/54

Table 3.4: Overall and class-by-class accuracies using whole set of bands.

Overall Accuracy = (2417/3058) 79.0386% Kappa Coefficient = 0.7616				
Class	Prod. Acc. (Percent)	User Acc. (Percent)	Prod. Acc. (Pixels)	User Acc. (Pixels)
Soybean-notill	92.31	99.59	240/260	240/241
Soybean-mintill	97.21	62.31	453/466	453/727
Corn-notill	81.41	71.75	508/624	508/708
Grass-pasture	100.00	81.10	206/206	206/254
Corn-mintill	20.95	90.41	66/315	66/73
Grass-trees	96.02	88.95	314/327	314/353
Soybean-clean	59.76	90.99	101/169	101/111
Buildings-Grass-Trees	43.14	80.00	44/102	44/55
Stone-Steel-Towers	100.00	100.00	36/36	36/36
Alfalfa	96.31	100.00	25/26	25/25
Corn	45.71	52.75	48/105	48/91
Grass-pasture-moved	58.33	100.00	7/12	7/7
Hay-windrowed	100.00	99.08	216/216	216/218
Oats	90.00	100.00	9/10	9/9
Wheat	100.00	100.00	84/84	84/84
Woods	60.00	90.91	60/100	60/66

For comparison reasons, SVM was also applied after band selection using information entropy and correlation criteria. Using the entropy criterion, 59 bands were selected. These bands had entropy value higher that 1. Using the correlation criterion 19 bands were used, which had correlation value below 0.2. The results are shown on Tables 3.5 and 3.6.

Table 3.5: Overall and class-by-class accuracies using information entropy criterion.

Overall Accuracy = (2311/3058) 75.5723%
Kappa Coefficient = 0.7221

Class	Prod. Acc. (Percent)	User Acc. (Percent)	Prod. Acc. (Pixels)	User Acc. (Pixels)
Soybean-notill	91.92	100.00	239/260	239/239
Soybean-mintill	96.14	60.46	448/466	448/741
Corn-notill	72.12	64.94	450/624	450/693
Grass-pasture	100.00	78.63	206/206	206/262
Corn-mintill	17.78	61.54	56/315	56/91
Grass-trees	97.25	89.08	318/327	318/357
Soybean-clean	42.01	86.59	71/169	71/82
Buildings-Grass-Trees	44.12	77.59	45/102	45/58
Stone-Steel-Towers	100.00	100.00	36/36	36/36
Alfalfa	76.92	100.00	20/26	20/20
Corn	52.38	56.12	55/105	55/98
Grass-pasture-moved	50.00	100.00	6/12	6/6
Hay-windrowed	100.00	98.63	216/216	216/219
Oats	50.00	71.43	5/10	5/7
Wheat	100.00	100.00	84/84	84/84
Woods	56.00	86.15	56/100	56/65

Table 3.6: Overall and class-by-class accuracies using correlation criterion.

Overall Accuracy = (2066/3058) 67.5605%
Kappa Coefficient = 0.6306

Class	Prod. Acc. (Percent)	User Acc. (Percent)	Prod. Acc. (Pixels)	User Acc. (Pixels)
Soybean-notill	75.77	98.01	197/260	197/201
Soybean-mintill	86.91	54.00	405/466	405/750
Corn-notill	61.22	52.33	382/624	382/730
Grass-pasture	91.26	82.46	188/206	188/228
Corn-mintill	19.05	71.43	60/315	60/84
Grass-trees	95.72	87.92	313/327	313/356
Soybean-clean	7.10	38.71	12/169	12/31
Buildings-Grass-Trees	44.12	84.91	45/102	45/53
Stone-Steel-Towers	100.00	100.00	36/36	36/36
Alfalfa	73.08	86.36	19/26	19/22
Corn	28.57	19.61	30/105	30/153
Grass-pasture-moved	50.00	66.67	6/12	6/9
Hay-windrowed	98.61	94.67	213/216	213/225
Oats	60.00	75.00	6/10	6/8
Wheat	100.00	96.55	84/84	84/87
Woods	70.00	82.35	70/100	70/85

Finally, Fig. 3.12 shows the SVM classification maps using the whole set of bands and subsets of bands resulted from the aforementioned band selection techniques.
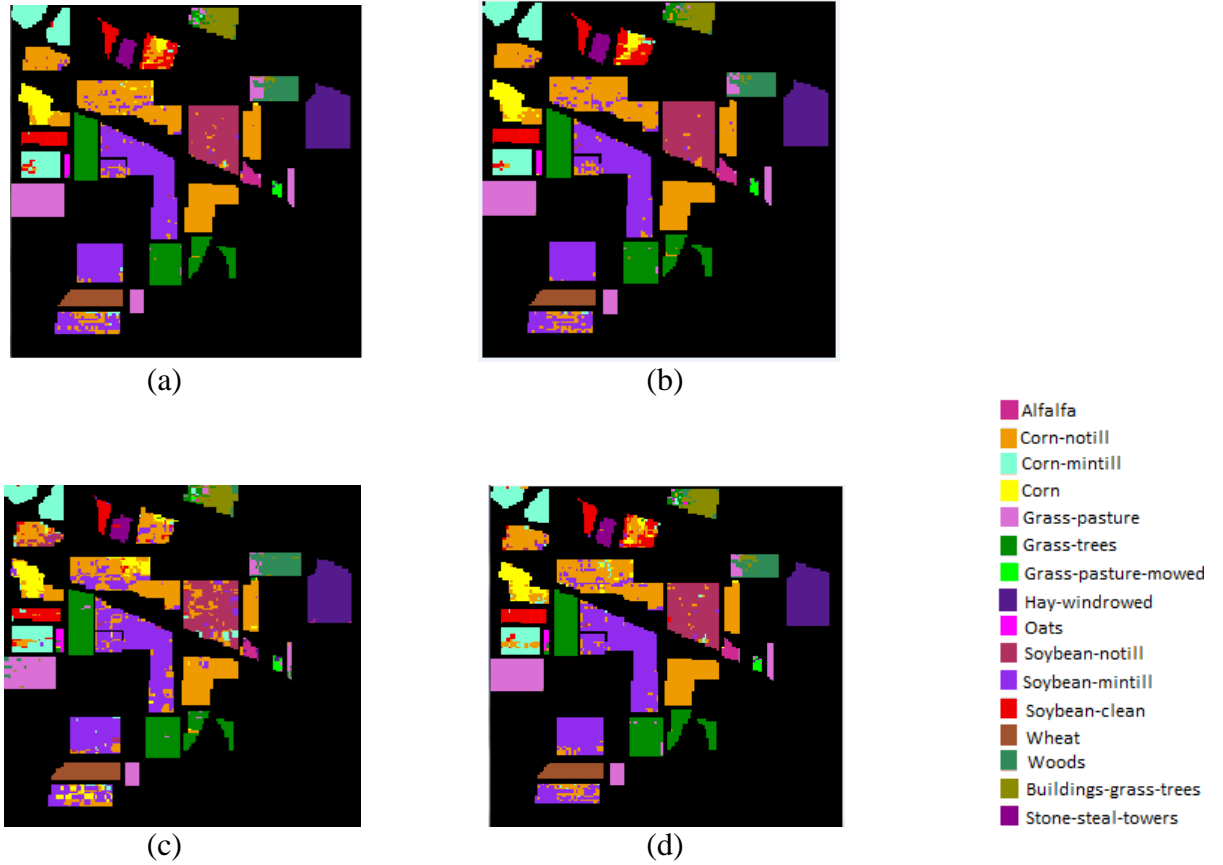
Figure 3.12: Classification maps using (a) whole set of bands , (b) subset of 86 bands selected from the new approach, (c) entropy criterion, (d) correlation criterion.

3.3.2.2 Conclusions

Implementation of the SVM using all the possible subsets of 16 sequential bands led to remarkable conclusions. The majority of the selected bands belongs to the visible light and few to the mid-infrared. Compared to the classification accuracies when no dimensionality reduction method was applied (Table 3.4), classification accuracy per class using different set of bands was improved for 7 out of 16 classes and it was deteriorated for the same amount (7 out of 16). Analysis on the experimental results showed that there was no optimum subset of 16 sequential bands which could lead to the improvement of the classification accuracy for all the classes simultaneously. Nevertheless, it should be noted that setting the number of bands to be equal to the number of signals is rather arbitrary. Although, in the frame of the particular study it seemed a good indicator, other criteria for separating the spectral space should also be considered. The overall accuracy when the 86 selected bands were used (Table 3.3) was almost

1% less than the overall accuracy when no dimensionality reduction method was applied. Taking into account that the number of bands was significantly reduced from 186 to 86, the results are very satisfactory.

For comparison reasons, two well-known criteria for band selection were used; entropy and correlation. It should be referred that these two criteria are not very practical to be used as there is no defined threshold which would lead to a certain amount of bands. Consequently, several trials should be made which lead to a time-consuming process. The proposed method outperformed compared to the two criterion-based methods. Analysis of the results contributes to remarkable conclusions; SVM presents higher accuracy when the number of bands is increased. However, there is a threshold on the number of selected bands above which the rate of the improvement of the classification accuracy is very low. The selected bands in this case are the optimal ones, because they facilitate the SVM regarding the hyperplanes searching. The proposed BS method achieved overall accuracy very close to the corresponding one (only 1% less) when all available bands are used, by reducing the number of selected bands more than 50%. This is of great importance since besides yielding gains in terms of computational time and data storage, it significantly reduces requirements in terms of the number of training data sets.

Chapter 4

Endmember extraction

4.1 Introduction

As it is previously noted in section 1.2.2, analysis of hyperspectral images requires spectral unmixing, which consists of identifying the unique constituent deterministic spectral signatures, called endmembers [6] and determining their apparent quantification, also called abundance fraction, at pixel level [19]. The accuracy of the quantification depends strongly on the accuracy with which endmembers are identified [86]. In the past, due to low spatial and spectral resolution of multispectral sensors, it was unlikely to detect endmembers and therefore, endmember extraction has presented low activity. Recent advances in hyperspectral imaging sensors have contributed to the exposure of many indistinct material substances which are unknown *a priori* and can be analyzed only by high spectral resolution [13]. These substances can be considered as endmembers.

There are three main ways to find endmembers; from the image (image endmembers), from a spectral library (reference endmembers) and by creating virtual endmembers using a trial-and-error approach. Working with reference endmembers premises that spectral library comprises the corresponding spectral signatures of the substances existing in the scene, which is not always the case. Assuming that for a particular study the desired reference spectral signatures exist in a spectral library, we have to take

into consideration that these signatures have been acquired in totally different conditions compared to those the remotely sensed image has been acquired (atmospheric conditions, illumination effects). Consequently, data sample vectors need to be calibrated and corrected for radiometric effects in order to be properly compared with the reference endmembers. However, the calibration process is not always likely to be accomplished nor to be evaluated. Thus, calibration errors are embedded in the spectral unmixing error. Another issue that is encountered when using reference endmembers is the difficulty to interpret the scene of remote-sensing scale using spectra measured at centimetre or millimeter scales [1], [87]. On the other hand, creating virtual endmembers contributes to minimizing the spectral unmixing error but mathematically correct endmembers do not always represent physically meaningful spectra able to interpret the imagery. This issue is overcome by extracting endmembers from the image. Image endmembers have the advantage of no calibration need, refraining from adding calibration errors in the mixture model. Another advantage of image endmembers is that they have the same scale of measurement as the data. In this work, new endmember extraction methods have been developed for extracting image endmembers.

Depending on the captured area, endmembers might be represented by a few sample vectors or even by one sample vector which in fact could mislead to treating endmembers as anomalies. Because of the aforementioned characteristics, detecting endmembers is very challenging.

Over the last decades, researchers have been focused on devising several models which result in more stable and efficient endmember extraction algorithms (EEAs). However, the majority of these models are characterized by their high computational complexity which imposes limits to endmember extraction on real-time application demands.

4.2 Review of endmember extraction methods

In recent years, a high level of activity has been observed regarding the development of new methods for endmember extraction. Categorization of the existing endmember extraction methods is a daunting task as a method may fit to more than one category. As it is mentioned in section 1.2, endmember extraction and abundance estimation may be implemented simultaneously instead of sequentially. In such case it should be more accurate to refer to endmember determination rather than endmember extraction since the entire sample vectors serve as potential endmembers and their selection is accomplished based on the minimization of the fraction error. Nevertheless, in this section an attempt was given to include the main aspects of finding the endmember whether this includes simultaneously abundance estimation or not. With this clarification, the majority of the existing methods are based on one of the four following approaches (see Fig. 4.1): convex-geometrical approach, statistical approach, sparse regression approach and spatial-contextual based approach.

Convex geometry-based EEAs assume that the measured spectra can be expressed as a linear combination of the endmembers presented in the mixed pixel and thus, mixed vectors lie in a p-vertex simplex, whose p vertices correspond to the endmembers. Two principal criteria to materialise the concept of convex-geometry are the orthogonal projection and the simplex volume [13]. The convex geometry-based EEAs could be categorized into two main categories. The first category contains EEAs which assume the presence of pure pixels in the image and seek for them at the vertices of the simplex defined by the data. Some of the most widely used pure pixel based algorithms are a) maximum volume-based algorithms: the N-finder algorithm (N-FINDR) [81], the simplex growing algorithm (SGA) [88] and the alternative volume maximization [89](AVMAX), b) orthogonal subspace-based algorithms: the automated target generation process (ATGP) [90], the vertex component analysis (VCA) [91], the pixel purity index (PPI) [92] and the successive volume maximization (SVMAX) [89], c) other algorithms: the sequential maximum angle convex cone (SMACC) [93], the iterative error analysis (IEA) [94] and lattice associative memories (LAM) [95] algorithms. The second category of

geometrical-based EEAs comprise methods which do not assume the presence of pure spectra in the image such as the convex cone analysis (CCA) [96], the iterative constrained endmembers (ICE) [77] and the sparsity-promoting ICE (SPICE) [97] which is an extension of the ICE algorithm that estimates the number of endmembers incorporating sparsity-promoting priors, or minimum volume based algorithms [98] such as the minimum volume constrained nonnegative matrix factorization method (MVC-NMF) [99], the minimum volume simplex analysis (MVSA) algorithm [100], the convex analysis-based minimum volume enclosing simplex algorithm (MVES) [101], [89] and the simplex identification via variable splitting and augmented Langrangian algorithm (SISAL) [102]. Recently, a new concept of geometric endmember extraction methods has been examined which assumes the existence of more than one convex region in hyperspectral space. These methods are called piecewise convex methods [79] and account for endmember variability. For each convex region, an individual set of endmember distributions and proportion values are determined using either fuzzy or probabilistic clustering.

EEAs based on a statistical approach assume that the endmember components are randomly distributed in the image. The main idea is that the endmembers yield the most uncorrelated data sample among the same number of other sample vectors [13]. In [12] authors present a comprehensive review of parametric EEAs such as Bayesian self-organizing maps (BSOM), and non parametric algorithms such as independent component analysis ICA [103] and dependent component analysis (DECA) [104]. The majority of the statistical based methods impose constraints and regularizations in order to give physically meaningful solutions [12]. Compared with the geometrical based approaches, the statistical methods have higher computational complexity.

The sparse regression based methods are basically semi-supervised and assume that the observed data can be expressed by linear combinations of a set of a priori known pure spectral signatures [105], [106]. In this case, inversion problem is decomposed with altering direction method of multipliers (ADMM) [107] into a sequence of simpler ones. ADMM can be derived as a variable splitting procedure, which is followed by the adoption of an augmented Lagrangian method to solve the inversion problem. The

sparse unmixing via variable splitting and augmented Lagrangian (SUnSAL) [108] computes abundances through the minimization of an objective function which measures the lack of fitness to the observed data vectors and the lack of sparsity of a candidate solution. By imposing ACS and ANC constraints the constrained version of SUnSAL (C-SUnSAL) is obtained [105].

Linear sparse regression is a very active research area but there are some issues to be noted. The pure spectral signatures are presumed to be available in a spectral library which is not always the case, and even if there are available they are not necessarily acquired under the same conditions as airborne or satellite image data and may not be good representations of the image component [109]. Searching for the optimum set of signatures which model each pixel is a combinatorial problem which requires efficient linear regression techniques since the ever-growing dimensionality and availability of the spectral libraries is much higher than the numbers of endmembers participated in a mixed pixel [19]. A step forward, termed *sparse coding* [110], consists of learning the dictionary from the data and, thus, avoiding not only the need of libraries but also calibration issues related to different conditions under which the libraries and the data were acquired [12].

Spatial-spectral contextual based methods use spatial statistics to improve the geometrical selection of endmembers, but with additional computational cost. Well known EEAs which use such an approach are the automated morphological endmember extraction (AMEE) [111], the spatial-spectral endmember extraction algorithm (SSEE) [112], and the spatial preprocessing (SPP) algorithm [113]. The SPP can also be used in order to boost the endmember extraction performance of other EEAs.
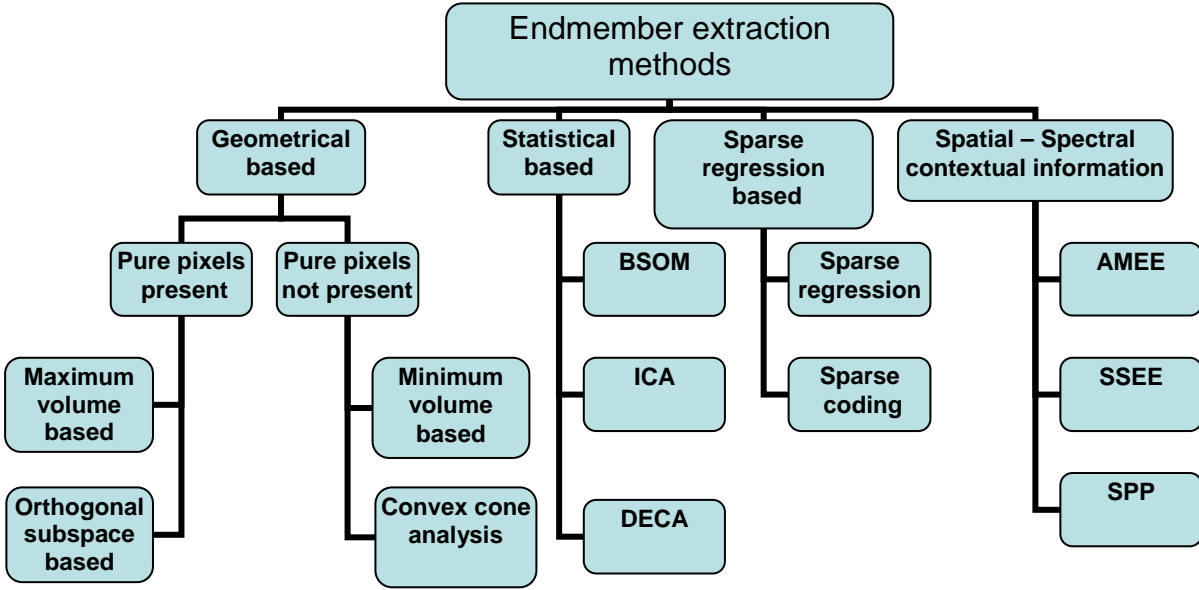
Figure 4.1: Categorization of EEAs.

4.2.1 Convex geometry-based endmember extraction methods

Endmember extraction methods based on convexity geometry use either the criterion of orthogonal projection or the criterion of simplex volume [13]. According to the first, endmember candidates are considered those data samples whose orthogonal projections on selected vectors lie at their end points. Three well known methods using orthogonal projections are PPI, ATGP and VCA. Based on the simplex volume criterion, endmembers are those sample vectors which are found at the vertices of the simplex yielding the maximum volume among all simplexes formed by the same number of sample vectors as vertices. N-FINDR is a representative method of such approach.

In this sub-section, a more detailed description of the state-of-the-art convex geometry-based endmember extraction methods is given.

Pixel purity index

The pixel purity index (PPI) [92] was the first method to exploit the principle of orthogonal projection in order to detect endmembers. Due to convexity, orthogonal projections of potential endmembers onto randomly generated unit vectors, called *skewers*, should present minimal or maximal at their end points

compared to the projections of the data sample vectors. PPI is a simultaneous endmember extraction algorithm, extracting all the desired endmembers at the same time which according to [13] is the optimal approach technically speaking.

In particular, dimensionality reduction is accomplished by applying MNF transform [48] and thus, a signal subspace is defined where convexity exists. Then, each data sample vector is orthogonally projected onto a skewer and those who appear at extreme points of the skewer are counted. The procedure is repeated for several skewers and those data sample vectors, whose frequency of having extreme values as they are projected on the random skewers is above a certain cut-off threshold, are considered candidate endmembers.

There are several issues related to the PPI method. Firstly, no criteria are provided for how to determine the number of endmembers and therefore the dimensions to be retained after MNF transform. Secondly, results are sensitive to tuning parameters, the number of skewers, K and the threshold factor τ which effects the extreme values selection. Thirdly, the randomness of the skewers creation results in different subsets of candidate endmembers each time the method is applied. Consequently, setting number of $K+1$ skewers leads to implementation of all the steps from the beginning without exploiting the results from a previous implementation using K skewers. The last issue of PPI is that it needs user interference in order to define the optimum set of endmembers among the candidate ones, using a visualization tool.

Automated target generation process

The automated target generation process (ATGP) [90] extends the orthogonal subspace projection (OSP) [114], which was developed for abundance estimation in order to perform mixed pixel classification, to an unsupervised version of OSP [13], [11]. ATGP is a sequential endmember extraction method which finds one endmember at a time exploiting the endmembers already extracted using a sequence of orthogonal subspaces with reduced dimensionality.

Let the initial endmember s_1 be the pixel vector with a maximum length. The ATGP begins by applying an orthogonal subspace projector:

$$P_U^{\perp} = I - U(U^TU)^{-1}U^T \tag{4.1}$$

to every pixel vector y in the data, specified by $U = \begin{bmatrix} s_1 \end{bmatrix}$. It then finds the second endmember, which is denoted by s_2, with the maximum absolute projection in the space orthogonal to the space linearly spanned by s_1. The third endmember s_3 can be found by applying another orthogonal subspace projector $P_U^{\perp}$ with $U = \begin{bmatrix} s_1 & s_2 \end{bmatrix}$ to the original image, and taking the sample vector with the maximum orthogonal projection in the space that is orthogonal to the one spanned by the first two extracted endmembers. The preceding procedure is repeated until a set of a pre-defined number of endmembers is extracted.

Using sequence orthogonal projections the computational complexity of ATGP is high. Furthermore, ATGP searches for the spectra which are most distant from the data sample and thus, it has the tendency to extract high contrast endmembers (e.g. urban materials) over less contrast endmembers (e.g. different types of vegetation or soil).

Vertex component analysis

Vertex component analysis (VCA) [91] works similar to ATGP but instead of selecting the vector with the largest magnitude as the next orthogonal projection, it chooses a random one. Additionally, it uses the maximum orthogonal projection as a criterion as PPI does to grow its convex hulls but instead of using random skewers it exploits the subspace spanned by the endmembers already extracted. For that reason, VCA can be considered as a random version of ATGP and a sequential EEA version of PPI [13].

VCA initially performs dimensionality reduction and a scaling procedure. The dimensionality reduction depending on the signal-to-noise (SNR) estimate is carried out via either singular value decomposition (SVD) or via PCA. The scaling procedure is to mitigate topographic modulation. It projects the original data onto a direction orthogonal to the subspace spanned by the endmembers already extracted and selects as new endmember the vector with the maximum projection.

Because of random initialization, a final set of endmembers produced by VCA is not repeatable. Each implementation leads to a different set of endmembers which should be carefully examined by the user.

N-finder

N-FINDR [81] is the most popular simplex-based algorithm and has been served as a base to develop new simplex-based algorithms [88], [115], [116]. It is a simultaneous algorithm which finds the set of p data sample vectors that define the p-vertex simplex with the maximum volume among all possible p-vertex simplexes formed by any set of p data sample vectors. First, a dimensionality reduction of the original image is accomplished by using the MNF transform [48]. Next, randomly selected data sample vectors qualify as endmembers, and a trial volume is calculated as follows. Let **E** be defined as

$$\mathbf{E} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \vec{s}_1 & \vec{s}_2 & \dots & \vec{s}_p \end{bmatrix} \tag{4.2}$$

where $\vec{s}_p$ are 1x(p-1) endmember column vectors, and p is the number of endmembers used to calculate the simplex volume. The volume of the simplex formed by the endmembers is proportional to the determinant of **E**

$$V(\mathbf{E}) = \frac{1}{(p-1)!} abs(|\mathbf{E}|) \tag{4.3}$$

In order to refine the initial volume estimate, a trial volume is calculated for every pixel in each endmember position by replacing that endmember and recalculating the volume. If the replacement results in a volume increase, the pixel replaces the endmember. This procedure, which does not require any input parameters, is repeated until there are no replacements of endmembers left.

In order to find the optimum set of endmembers which forms the simplex with the maximum volume, all the possible permutations of p data sample vectors are needed to be checked. This is an exhausting search which requires huge computing time. Several new versions of N-FINDR have been developed which aim at decreasing the computational burden. Further description of these new methods can be found at [13].

Sequential maximum angle convex cone

The sequential maximum angle convex cone (SMACC), [93] method solves linear unmixing using a convex factorization technique that simultaneously generates the set of endmembers and their abundances. When the abundance coefficients are constrained to sum to one, the convex cone reduces to a convex hull and the extreme vectors form a simplex. The endmembers and abundance coefficients are determined sequentially. To obtain the endmembers, SMACC first selects a group of pixels that are extreme vectors in the data. They become a basis and form a convex cone within their subspace. The data that is outside of the cone is called residual. The determination of the next endmember is based on the spectral angle that it makes with the existing convex cone. The data sample vector which has the maximum spectral angle with the cone is selected as the next endmember. After an endmember has been identified, its contribution to the residual is removed by oblique projections. The sequence is repeated until the desired number of endmembers is obtained.

4.3 Proposed endmember extraction methods

As the research activity is increased more complicated approaches are adopted in order to develop more efficient endmember extraction algorithms. The possible approaches and their combinations can be endless, but it should be taken into consideration the computational burden and the reliability of each method. For instance, if the performance of an endmember extraction algorithm is slightly better than its competitors but its computational complexity is few orders of magnitude higher, this should be carefully evaluated. Moreover, the reliability of the algorithms which do not extract endmembers directly from the image can be questioned if the estimated pure signatures lead to a satisfying unmixing result but they have no physical meaning.

This work introduces new simultaneous simplex-based unsupervised endmember extraction methods, the so-called simple endmember extraction (SEE) method and the Enhanced-SEE (E-SEE) empirical

method which have light computational burden and are characterised by their clear conceptual meaning. The SEE method exploits the dimensionality reduction ability of the eigen-based projection techniques for seeking the vertices of the simplex which is defined by the new lower dimensional space. The main concept of the SEE algorithm is that a subset of the extreme values of the projected pixels onto the first p-1 transformed components, where p is the number of the endmembers to be defined, corresponds to the vertices of the simplex. Analysis of the defined simplexes using images of various proportions of endmembers in the mixed pixels concluded that there is a weakness in extracting those endmembers which are comprised in many mixed pixels and therefore, the data mean is closer to them. This weakness was overcome by the development of the empirical E-SEE method. The E-SEE exploits the fact that the maximum projected value (extreme) onto the first transformed component is always the endmember which is far away from the data mean. Adding pixels in the image, whose spectral signatures correspond to the endmember which is the most distant from the data mean leads to a new data space where the distance between candidate endmembers and data mean is increased. The proposed methods were evaluated using simulated and real hyperspectral data and they were also compared with well-known simplex-based endmember extraction methods. The proposed methods have lower computational complexity compared with their competitors and despite their simplicity, they can be promising in the field of endmember extraction.

4.3.1 Theoretical Background

Simplex approach

Recall the fact that each observed spectral vector y, $y \in \Re^L$, can be expressed, under the linear spectral mixing concept [19], as linear combination of pure pixels s_k, each one multiplied by its corresponding abundances a_k, plus a noise vector n, $n \in \Re^L$:

$$y = \sum_{k=1}^{p} a_k s_k + n = S\alpha + n \qquad (4.4)$$

where L is the total number of bands, the $L \times p$ matrix $S = [s_1,...,s_P]$ comprises the endmember spectra and the $p \times 1$ vector $\alpha = [a_1,...,a_P]$ their corresponding abundances. The solution of the linear spectral mixture problem described by eq. (4.4) relies on a successful estimation of the number of endmembers, p, presented in the input hyperspectral scene Y, and also on the correct determination of a set of endmembers and their corresponding fractional abundance [117]. Two physical constrains are generally imposed on the fractional abundances: the abundance non-negativity constraint (ANC) and the abundance sum-to-one constraint (ASC). The geometrical interpretation of the linear mixture model is associated with the mathematical theory of *convex sets* [6], [21], [22]. More precisely, assuming that both constraints are satisfied and noise is limited, spectra can be restricted in a simplex lying on a signal subspace of dimension one less than the number of the endmembers.

The following points are the basis for the proposed methods:

1. Abundances are unit-sum and mixed spectra are all-positive linear combinations of the pure endmember spectra [22].

2. Mixed spectra are interior to the convex hull defined by the endmember vertices [22].

3. The endmembers can be found at the vertices of the simplex whose existence is guaranteed when the dimensionality of the mixed data is one less than the number of linearly independent endmembers [22].

4. In case of uncorrelated noise with equal variance in all bands, the noise is spherically distributed about the data mean [48].

5. A data sample vector with maximum Euclidean norm (magnitude) must be located at one of the vertices of the simplex [115], [118].

According to the first three points, inference of matrix S is equivalent to identifying the vertices of the defined simplex. Based on points (3) and (4) an eigen-based projection technique such as principal component analysis (PCA) [52] or minimum noise fraction (MNF) [48] could determine the lowest

dimensional subspace that spans all the information but not the noise, if the number p is known. This p-1 dimensional subspace is defined by the mean of the data and the eigenvectors of the significant dimensions. Point (5) has been implied in [90], where authors claim that the first generated endmember from SGA is always a pixel which has either a maximum or a minimum value in the first principal component. Nevertheless, there was no further analysis on this observation.

Eigen-based projection techniques

Principal component analysis (PCA) [52] is a well known mathematical procedure which is used for dimensionality reduction, feature extraction, visualization of high-dimensional data, etc. As its concept is well known, analysis will be limited to the basics. PCA is the simplest of the true eigenvector-based multivariate analyses. It uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of uncorrelated variables called principal components (PCs). The number of PCs is equal to the number of original variables. An orthogonal linear transformation U transforms the data $\mathbf{Y}$ to a new coordinate system $\mathbf{X}$ such that the greatest variance among all the variances resulting by the projection of the data set at any direction is associated with the first coordinate, called the first PC, the second greatest variance on the second coordinate, and so on [52]. Basically, the covariance matrix of the data $\mathbf{Y}$ in the new coordinate system is diagonal. The original covariance matrix, $\Sigma_{\mathbf{Y}}$, becomes the diagonalized covariance matrix, $\Sigma_{\mathbf{X}}$. The solution to this problem becomes a generalized eigenvalue problem of the form:

$$\Sigma_{\mathbf{Y}}U = U\Sigma_{\mathbf{X}} \tag{4.5}$$

where the eigenvalues are the diagonal elements of $\Sigma_{\mathbf{X}}$, and the eigenvectors form the columns of U. Each eigenvalue is proportional to the portion of the variance, and more accurately to the sum of the squared distances of the points from their multidimensional mean that is correlated with each eigenvector. The original data $\mathbf{Y}$ is multiplied by the eigenvectors of the original data covariance $\Sigma_{\mathbf{Y}}$. The PCA transformation is then computed by:

$$\mathbf{X} = U^T\mathbf{Y} \tag{4.6}$$

PCA rotates the set of points around their mean in order to align with the PCs. This moves as much of the variance as possible into the first few dimensions. Given a set of points in Euclidean space, the first PC corresponds to a line that passes through the multidimensional mean and minimizes the sum of squares of the distances of the points from the line. The next PC is orthogonal to the first PC and passes through the data mean.

Another eigen-based method, the MNF is effective at creating a set of images that is ordered according to image quality. MNF consists of two separate PCA rotations and a noise whitening process. More precisely, decomposition of the noise covariance matrix is performed using PCA and then noise is rescaled resulting in transformed data in which noise is uncorrelated and has unit variance. Afterwards, PCA is performed on the noise-whitened data.

Noise estimation

As it was mentioned previously, the proposed methods exploit the fact that in case of uncorrelated noise with equal variance in all bands, the noise is spherically distributed about the data mean. In case of independent and identically distributed (i.i.d.) noise, by applying either PCA or MNF on zero mean data, the same set of eigenvectors will be produced. If the noise is not known it cannot be transformed to zero mean i.i.d., which stands in real applications. Thus, noise estimation should be performed. Many approaches have been developed for noise estimation. Usually, the nearest neighbour difference (NND) [48] method is initially performed on real datasets for noise estimation and then noise-whitening is applied using MNF.

4.3.2 Simple endmember extraction (SEE) method

The SEE method is a simultaneous convex-geometrical based endmember extraction method which seeks for the unique constituent deterministic spectral signatures, image endmembers, contained in hyperspectral images. Principles of the SEE method have been initially introduced in [29]. The SEE method exploits the lower dimensional space defined by the eigenvectors computed by the minimum

noise fraction (MNF) [48]. More precisely, the first p-1 eigenvectors, where p is the number of the linearly independent endmembers, define a space of much lower dimension than that of the initial hyperspectral space. Given that endmembers are found at the vertices of the simplex that exists when the dimensionality of the mixed data is one less than the number of linearly independent endmembers [22], the subset of the minimum and maximum values of the projected pixels onto the first p-1 transformed components correspond to the vertices of the simplex which is created by the data. To find the optimum subset, the method computes the spectral angle distances (SAD) between each candidate endmember with the others. Consequently, the proposed method only projects data onto the first p-1 MNF axes and uses the two sample vectors with the maximum and minimum projected value per axis. SEE results in a stable solution compared to other EEAs such as PPI and VCA which use projections onto random vectors presuming different set of extracted endmembers for each implementation. Moreover, unlike other orthogonal-based algorithms, SEE is a simultaneously endmember extraction algorithm exploiting the total data sample at the same time and omitting the computational burden of iterative projections or projective projections on the space defined by the already determined endmembers.

The pseudo-code of SEE is shown in Algorithm 4.1.

Algorithm 4.1: SEE

INPUT: The LxN matrix Y, where N are the $L \times 1$ observed spectral vectors and L the spectral bands

Step 1: Estimation of number p using ODM
Step 2: MNF transformation
Step 3: Projection of data sample vectors onto the first p-1 eigenvectors
Step 4: Selection of the maximum and minimum projected value for each eigenvector $\rightarrow$ 2p-2 candidate endmembers
Step 5: Definition the set of unique endmembers using SAD

OUTPUT: Image endmembers

The novelty of the method is that it searches for the extreme values that lie on the end points of the existing transformed axes without further projections that imply iterative procedures. It uses only the most extreme projected vectors of the p-1 transformed components unlike the existing EEAs which set thresholds for defining a set of extremes pixels at each random vector, estimating their frequencies [92]. The SEE merely uses the projected sample vectors of the transformed subspace without proceeding to

high computational approaches such as orthogonal projections onto the space defined by the already estimated endmembers or iterative volume computational of all the possible combinations of endmembers [81], [91].

Assuming that the number of endmembers p is known using a signal subspace identification technique [25], SEE implements MNF in order to eliminate noise components and to ensure that the greatest variances of signals are estimated by the projected data. Then, SEE seeks for the endmembers at the vertices of the simplex which is defined by the mean of the noise-whitened data and the p-1 eigenvectors. For illustration purposes, simulated data were generated by three spectral signatures from U.S. Geological Survey (USGS) digital spectral library which are presented in Fig. 4.2; (A) almandine, (B) soil, (C) glauconite. The 50x50-pixel simulated hyperspectral image with 423 spectral bands was created in such a way that the abundance fractions follow a Dirichlet distribution enforcing positivity and full additivity constraints. White Gaussian noise was also added.

Since the simulated image has 3 endmembers, data sample vectors must lie in a triangle with the endmembers as its three vertices. Fig. 4.3 presents the concept of SEE on the scattergram of the data using the first two eigenvectors. As it is observed, projection of the data sample onto the first PC, meaning multiplying the data with the first eigenvector (ev_1), leads to two candidate endmembers; maximum projected value corresponds to the endmember (A) and minimum projected value corresponds to the endmember (C). Projected data with the greatest variance lie onto the first PC. The maximum projected pixel value on the first PC will always be the endmember whose spectral signature is the highest concerning radiance/reflectance values and the minimum projected pixel will be the endmember with the lowest spectral signature. This is why the first generated endmember from SGA [88] is always a pixel which has either a maximum or a minimum value onto the first component of dimensionality reduction. Proceeding to the projection of data onto the second PC, two more candidate endmembers are defined; maximum projected value corresponds to the endmember (C) and minimum projected value corresponds to the endmember (B). Endmember (B) which has the least distance from the data mean, compared to the other endmembers, is found at the last projection. Selecting the maximum and

minimum projected values to the p-1 PCs leads to $2p$-2 candidate endmembers. In order to select the optimum subset of p endmembers, sum S of the SAD between each candidate endmember (ce_1) with the others (ce_i) is calculated for each candidate endmember:

$$S_{(ce_1)} = \sum_{i=0}^{p-1} \cos^{-1}\left(\frac{\langle ce_1, ce_i \rangle}{\|ce_1\| \cdot \|ce_i\|}\right) \qquad (4.7)$$

Spectral signatures which are extracted more than once have the same sum of SAD and the duplicates can be omitted. In this case, the endmember (C) has been extracted twice.
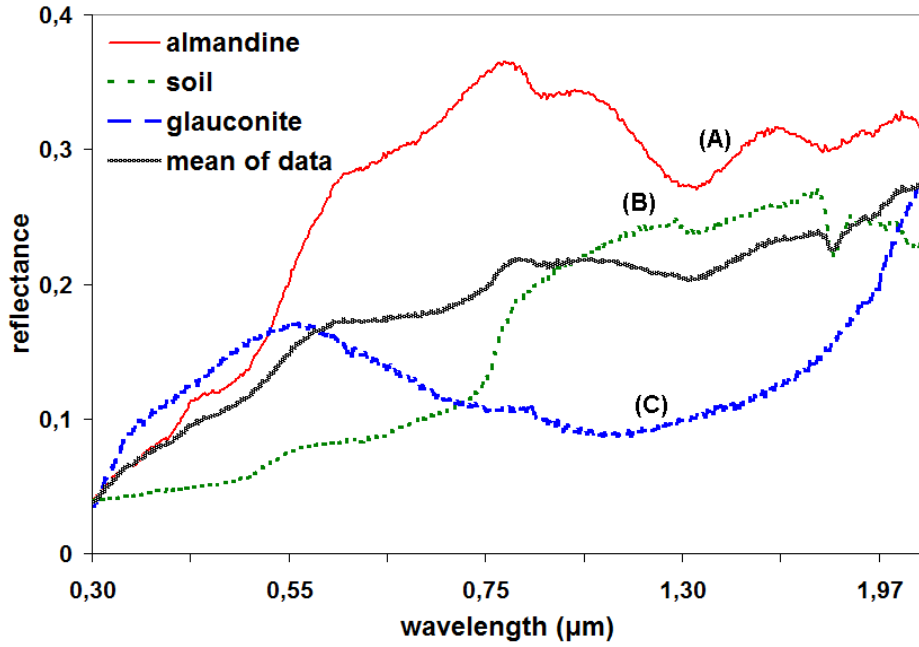
Figure 4.2: Three USGS spectral signatures used for simulated data.
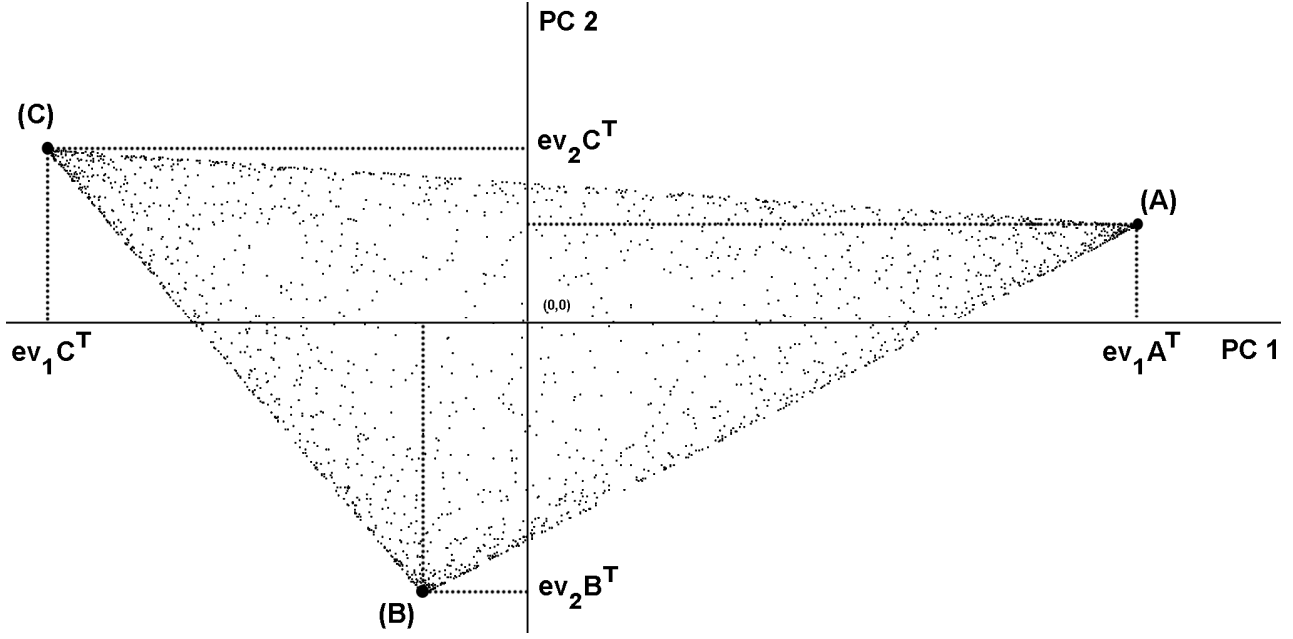
Figure 4.3: Illustration of SEE.

4.3.3 Enhanced SEE method

An important issue that the majority of the convex-geometrical based endmember extraction methods encounter is the tendency to select high contrast endmembers over less contrast endmembers. Imagine an image depicting mainly sea and a relatively small area of the neighboring coastal zone with urban materials and vegetation. In this case, data mean is closer to the spectral signatures of seawater and endmembers which represent different types of seawater could be in the interior of the convex region and therefore they would fail to be identified as extreme sample vectors. On the other side, high contrast endmembers are distant from the mean data and yield as extreme end points during several projections. In order to overcome this issue, a new empirical method has been developed, called enhanced SEE (E-SEE) whose novelty lies in the fact that it changes the distribution of the initial data sample increasing the distance between candidate endmembers and the data mean. E-SEE exploits the fact that the maximum projected value (extreme) onto the first transformed component is always the endmember which is the most distant from the data mean. Thus, adding pixels in the original image whose spectral

signatures represent the first extracted endmember lead to the less contrast endmembers shifting away from the data mean.

Simulated data used in the previous section consist of almost equal proportions of pure and mixed pixels, which are indicated by the location of the zero mean data quite close to the centroid of the data. In real hyperspectral data, proportions of mixed pixels may be uneven. In order to evaluate the proposed method in such a scenario, a new simulated hyperspectral image was generated by the same spectral signatures showed in Fig. 4.3, but frequency of pixels containing spectral signature of (B) is higher. The reason for changing the frequency of (B) spectral signature is because it is closer to the mean of the data than the other endmembers as it is showed in Fig. 4.2 and it is the last extracted endmember. Generally, it is difficult to extract endmembers which are close to each other as well as to the mean. Extreme endmembers are found easier. Fig. 4.4 presents the new scattergram of the data for the first two PCs. It is clear that the zero mean is close to the endmember (B). Repeating the processing of SEE, projection of the data onto the first PC, leads to two candidate endmembers; maximum projected value corresponds to the endmember (A) and minimum projected value corresponds to the endmember (C). Projection of the data onto the second PC leads to the same candidate endmembers; maximum projected value corresponds to the endmember (C) and minimum projected value corresponds to the endmember (A). Despite being at the vertex of the simplex, the endmember (B) is not found when its distance from the mean of the data is decreased. It is clear that if PCs rotate a bit, then the projection of (B) on them will be one of the extreme values. The question which is raised is how to effectively rotate PCs without changing the variance of the data. A simple way to accomplish it is by adding pixels with the spectral signature of the maximum projected pixel onto the first PC. As it is mentioned before, this pixel corresponds to the endmember which is far away from the mean of the data and is the brightest. In this case is endmember (A). By adding these pixels, the mean of the data moves away from the endmember (B) and goes towards the endmember (A). The variance would be almost the same as only the frequency of the maximum value is increased. This increase leads to a slight rotation of the first PC towards the

endmember (A). Fig. 4.5 shows the scattergram after the addition of 17% more pixels than the initial image. The slight rotation of the PCs led to the extraction of the optimum set of extracted endmembers.
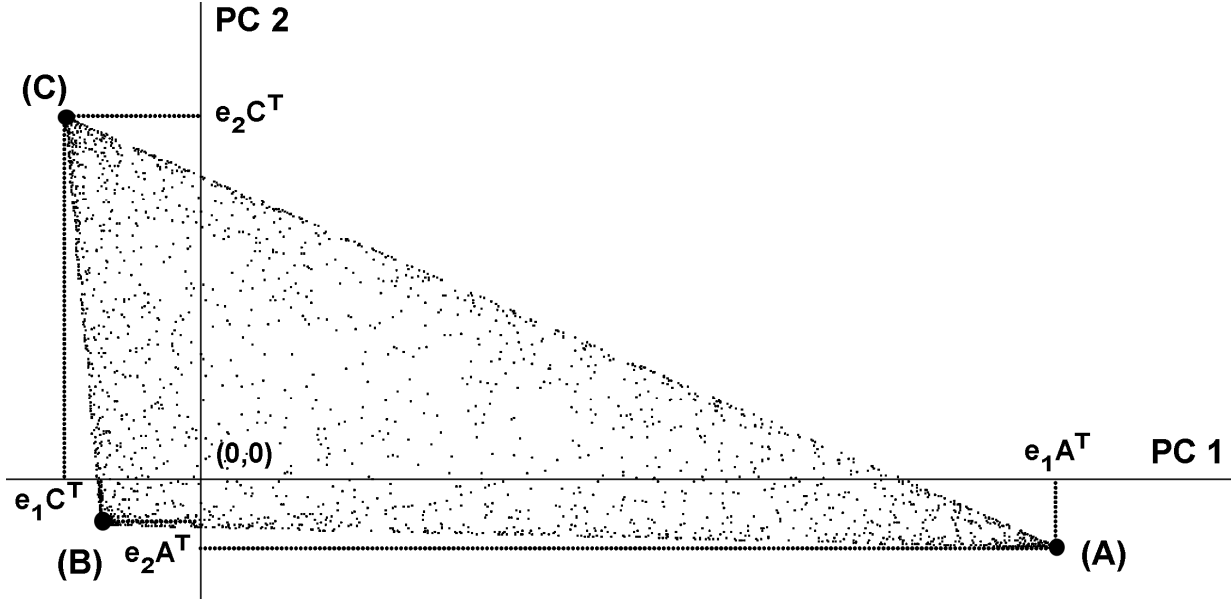
Figure 4.4: Illustration of SEE when the frequency of (B) is higher than the others.
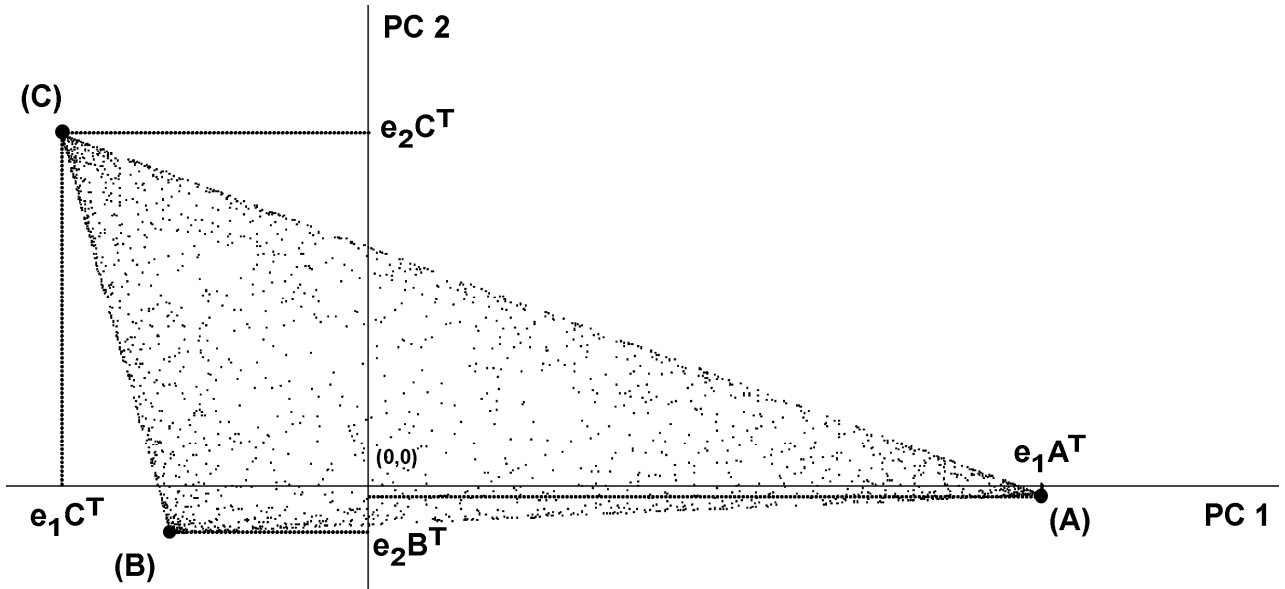
Figure 4.5: Illustration of Enhanced SEE.

In order to define the optimum number of additional pixels, k, several experiments took place. Results showed that the addition of pixels with the maximum projected value on the first PC does not change significantly the direction of the defined PCs using the new shifted data. As shown in Fig. 4.6, the

angle differences between the first eigenvector of the initial image and the first eigenvector of the image with the additional pixels are bellow 0,01 degrees and as the number of the extra pixels increases the angle tends to stabilize. This is reasonable because eventually the first eigenvector will pass through the additional pixels.
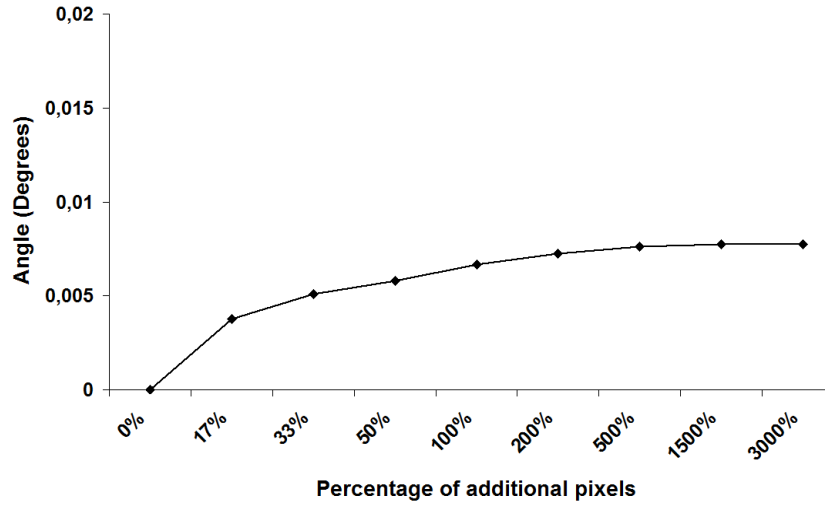
Figure 4.6: Angle rotation of eigenvectors regarding different amount of additional pixels.

In addition, experiments on images with noise showed that as the signal to noise ratio (SNR) increases, fewer pixels are needed to be added. Moreover, experimental results were satisfying when adding at least 300% more pixels than the pixels of the initial image.

The pseudo-code of E-SEE is shown in Algorithm 4.2.

Algorithm 4.2: E-SEE

INPUT: The LxN matrix Y, where N are the $L \times 1$ observed spectral vectors and L the spectral bands

Step 1: Estimation of number p using ODM
Step 2: MNF transformation
Step 3: Projection of original data onto the first 1 eigenvector→select the vector with maximum projected value
      as 1st endmember
Step 4: Addition of pixels having the spectral signature of the 1st endmember (at least 300%) to the original data
      →creation of shifted data
Step 5: MNF transformation to the shifted data
Step 6: Selection of the maximum and minimum projected value for each eigenvector → 2p-2 candidate endmembers
Step 7: Definition the set of unique endmembers using SAD

OUTPUT: Image endmembers

4.3.4 Computational Complexity of SEE and E-SEE

Computational complexity of SEE and E-SEE algorithms has been evaluated by means of the number of the floating point operations (flop). SEE performs dimensionality reduction (DR) which includes computation of covariance matrix of complexity $O(2NL^2)$ and eigen-decomposition of complexity $O(L^3)$, where N is the number of pixels and L is the number of bands. According to [88], [91], computational cost of performing dimensionality reduction is relatively small and can be negligible since: 1) for the sample covariance computation there is no need to use all the N pixels to infer the appropriate signal subspace, but a subset of N pixels, called N', with $N'<<N$; 2) concerning the eigen-decomposition, only partial decomposition can be performed for the computation of the p-1 eigenvectors. Consequently, SEE has computational complexity of $O(2N'L+(p-1)^3)$. It is reasonable that computation of spectral angles has a negligible complexity and it is omitted. E-SEE firstly computes the sample covariance, it uses the first eigenvector and then operations are identical with SEE except for the amount of pixels which is kN' instead of N', where k represents the times of additional pixels by means of the initials.

At this stage, computational complexity of VCA and N-FINDR has also been estimated in order to be compared with the proposed methods. These two algorithms were selected because a) they are widely used for endmember extraction presenting satisfying results, 2) they belong to two different approaches; orthogonal subspace approach and volume maximization, respectively and 3) they share the same concept with the proposed method according to which the endmembers locate at the vertices of the simplex defined by the data of lower dimension. (Please refer to section 4.2.1 for further elaboration).

VCA projects all the data vectors onto the p-dimensional subspace and onto the direction which is orthogonal to the determined endmembers. N-FINDR computes the determinant of a $p \times p$ matrix Np times and the computational complexity of each time is p^δ with $2.3 < \delta < 2.9$ [119]. In order to present proper results, despite being negligible compared to the remaining operations, computational complexity of dimensionality reduction has been added to the overall complexity of N-FINDR and

VCA algorithms as they performed it as well. Table 4.1 reports the approximated expressions for the estimation of the number of flop for each EEA.

Table 4.1: Computational complexity of applied algorithms

Algorithm	Number of flop
Enhanced SEE	$2N'L^2+2kN'L^2+(p-1)^3$
SEE	$2N'L^2+(p-1)^3$
VCA	$2p^2N+2N'L^2+(p-1)^3$
N-FINDR	$p^{\delta+1}N+2N'L^2+(p-1)^3$

Fig. 4.7(a) plots the number of flop versus p for each of the considered EEAs, with $N=10^6$ and $N'=10^3$ [91], L=50 and k=4. As shown, SEE and E-SEE are invariant with regard to the number of endmembers and require less number of flop compared to the other EEAs. Compared to SEE, E-SEE has more flop but their difference is less than one order of magnitude, while for p=15, E-SEE complexity is three orders of magnitude lower that N-FINDR and one order lower than VCA. Fig. 4.7(b) plots the number of flop versus the number of pixels, with p=5. As it is observed, SEE presents the lower computational complexity and along with E-SEE present lower increment of the number of flop as a function of the number of pixels compared to the N-FINDR and VCA.
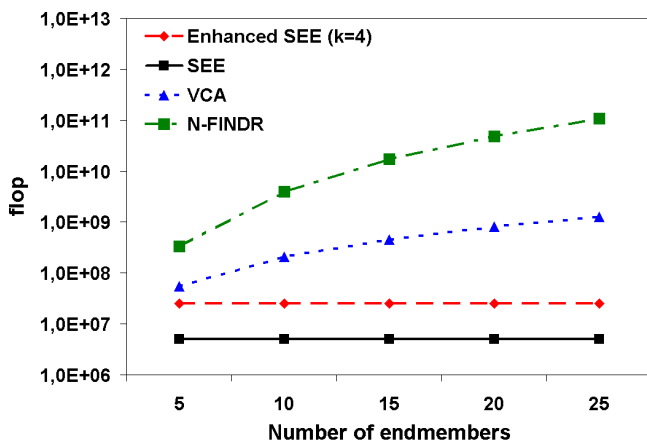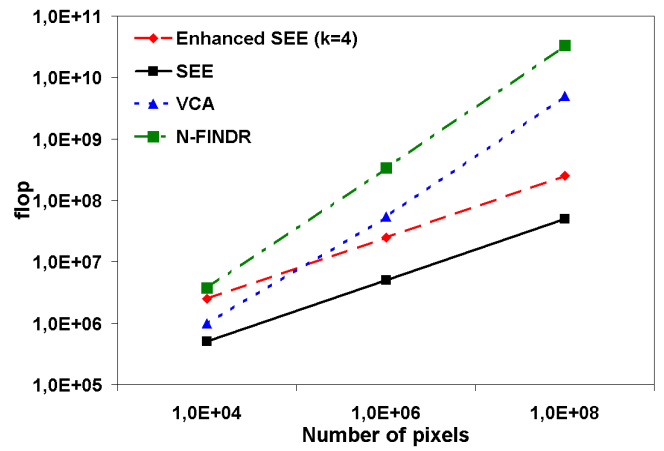
(a)   (b)

Figure 4.7: Computational complexity measured in flop (a) in terms of the number of endmembers, (b) in terms of the number of pixels.

Fig. 4.8 plots the number of flop in terms of the percentage of additional pixels required for the implementation of the E-SEE. It is shown that even after adding 500% more pixels, computational complexity was increased by less than one order magnitude. Since this addition does not make significant difference to the computational complexity it is proposed to add at least 300% more pixels (k=4) than the pixels of the initial image.
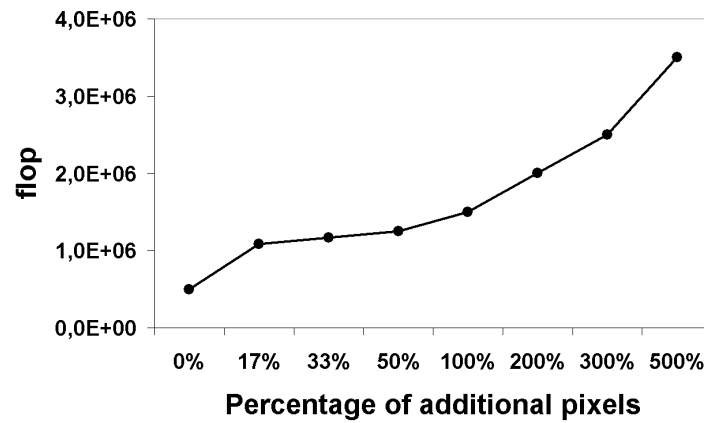
Figure 4.8: Number of flop regarding different amount of additional pixels.

4.4 Experiments

Simulated data experiments

The proposed methods have been tested in several simulated data. Two sets of simulated data are selected to be presented among those which have been generated for testing and evaluating the proposed methods. In both cases, the same set of 7 spectral signatures from USGS digital spectral library was used which correspond to sea, blackbrush, soil, stonewall, andradite, almandine, and glauconite (Fig. 4.9).
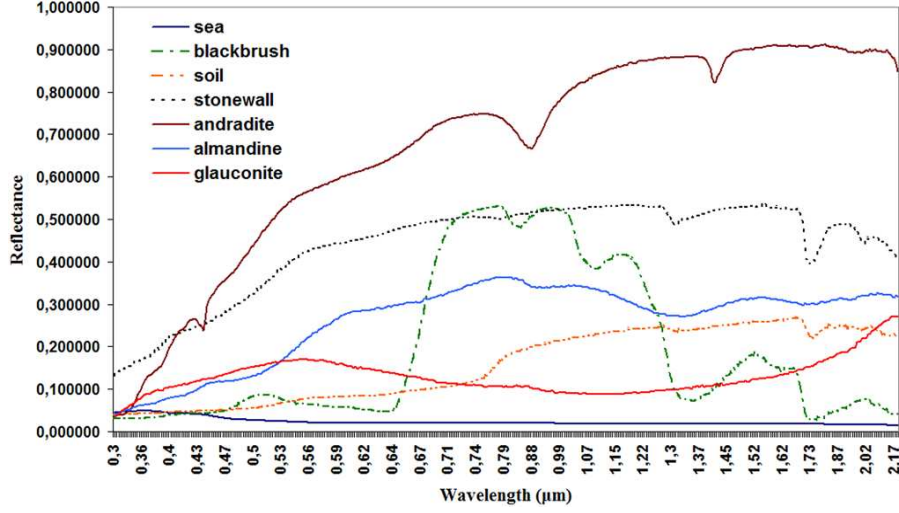
Figure 4.9: Spectral signatures from USGS library.

Scenario 1: The first simulated data set was created in such a way so that the fractional abundances of each endmember were known for every pixel. In particular, the 21x45-pixel simulated data was created in such a way that the abundances are assigned to the pixels in progressive linear mixture for all endmember combinations. Sum-to-one constraint has been applied to the abundances. Pure pixels are in the first and last row and randomly in the middle of the image. Different levels of noise have been added with signal to noise ratio (SNR) values 10dB, 20dB and 40dB, respectively. These SNR values are representative of what is typically found in real sensors [96]. The SNR for each band is defined as the ratio of the 50% reflectance signal level to the standard deviation of the noise [11]. The additive random noise was simulated by using numbers with a standard normal distribution obtained from a pseudorandom number generator. The simulated data were obtained by:

$$y_i = \left(\frac{SNR}{2} + n_i \right)(Ma_i) \tag{4.8}$$

where y_i is a vector containing the simulated spectrum, n_i is the noise factor, $\boldsymbol{M}$ is the matrix of endmember spectra and α_i is the fractional abundance of endmembers at pixel y_i. The values of the pure pixels were computed by setting one component of the $\boldsymbol{\alpha}$ vector to one and the rest to zero to indicate class membership. Scaling the signal by 50% of the SNR is equivalent to reducing the noise standard

97

deviation by the inverse factor (2 / SNR), so that the simulated data meets the SNR definition. The vector terms in the parentheses are multiplied element by element.

Besides SEE, N-FINDR, VCA, PPI, ICA-EEA, SMACC and ATGP were also applied. Independent component analysis endmember extraction algorithm (ICA-EEA) [55] is a nonparametric statistical approach. It finds pure endmembers by decomposing the data using the ICA algorithm into independent components. The components are scored and ordered; only the p highest components are kept. Pure endmembers are derived from the maximum absolute value of each component. ICA algorithm which was used for our experiments was FastICA as it was chosen by [55].

The endmember extraction methods have been repeatedly applied to the simulated dataset and the best results are presented on Table 4.2. As it is observed, the proposed method, the SEE, had the best performance for images with noise, compared to the other EEAs. Almost all the extracted endmembers apart from one, are pure pixels regardless the amount of noise in the image. All the other algorithms extract mixed pixels as endmembers and/or fail to extract all the endmembers. The ICA-EEA presents the worst performance, while SMACC failed to extract "almandine" in all images.

Table 4.2: Fractional abundances of extracted endmembers from the applied EEAs using images with different noise level. Symbol 'o' means that endmember was not extracted.

EEA \ Endmember		Sea	Blackbrush	Soil	Stonewall	Andradite	Almandine	Glauconite
N-FINDR	SNR 10	o	100%	100%	100%	100%	100%	100%
	SNR 20	o	100%	100%	100%	100%	100%	100%
	SNR 40	100%	100%	100%	100%	100%	100%	100%
	no noise	100%	100%	100%	100%	100%	100%	100%
VCA	SNR 10	95%	100%	o	100%	100%	o	100%
	SNR 20	95%	100%	o	100%	100%	100%	100%
	SNR 40	95%	100%	100%	100%	100%	100%	100%
	no noise	100%	100%	100%	100%	100%	100%	100%
PPI	SNR 10	100%	95%	95%	o	o	o	80%
	SNR 20	100%	65%	90%	o	70%	o	80%
	SNR 40	100%	95%	95%	80%	90%	70%	90%
	no noise	100%	100%	100%	100%	100%	100%	100%
ICA-EEA	SNR 10	o	60%	o	o	o	90%	90%
	SNR 20	o	95%	o	o	o	90%	55%
	SNR 40	80%	o	o	50%	o	75%	80%
	no noise	100%	100%	100%	100%	o	100%	100%
SMACC	SNR 10	100%	100%	o	o	100%	o	o
	SNR 20	100%	100%	o	o	100%	o	100%
	SNR 40	100%	100%	o	100%	100%	o	100%
	no noise	100%	100%	100%	100%	100%	o	100%
ATGP	SNR 10	o	100%	o	o	100%	o	o
	SNR 20	o	100%	o	95%	100%	o	100%
	SNR 40	o	100%	o	100%	100%	o	100%
	no noise	100%	100%	100%	100%	100%	100%	100%
SEE	SNR 10	100%	100%	100%	100%	100%	95%	100%
	SNR 20	100%	100%	100%	100%	100%	100%	100%
	SNR 40	100%	100%	100%	100%	100%	100%	100%
	no noise	100%	100%	100%	100%	100%	100%	100%

Fig. 4.10 presents the number of the extracted endmembers per algorithm, as well as, their mixture in the simulated data. For the images without noise, SMACC and ICA-EEA missed to extract one endmember (Figure 4.10(a)). For images with SNR 10 and 20, N-FINDR presents the second best performance after the proposed method, extracting correctly all the endmembers, but failing to extract sea.

(a) Image without noise

(b) Image with SNR 10dB

(c) Image with SNR 20dB

(d) Image with SNR 40dB
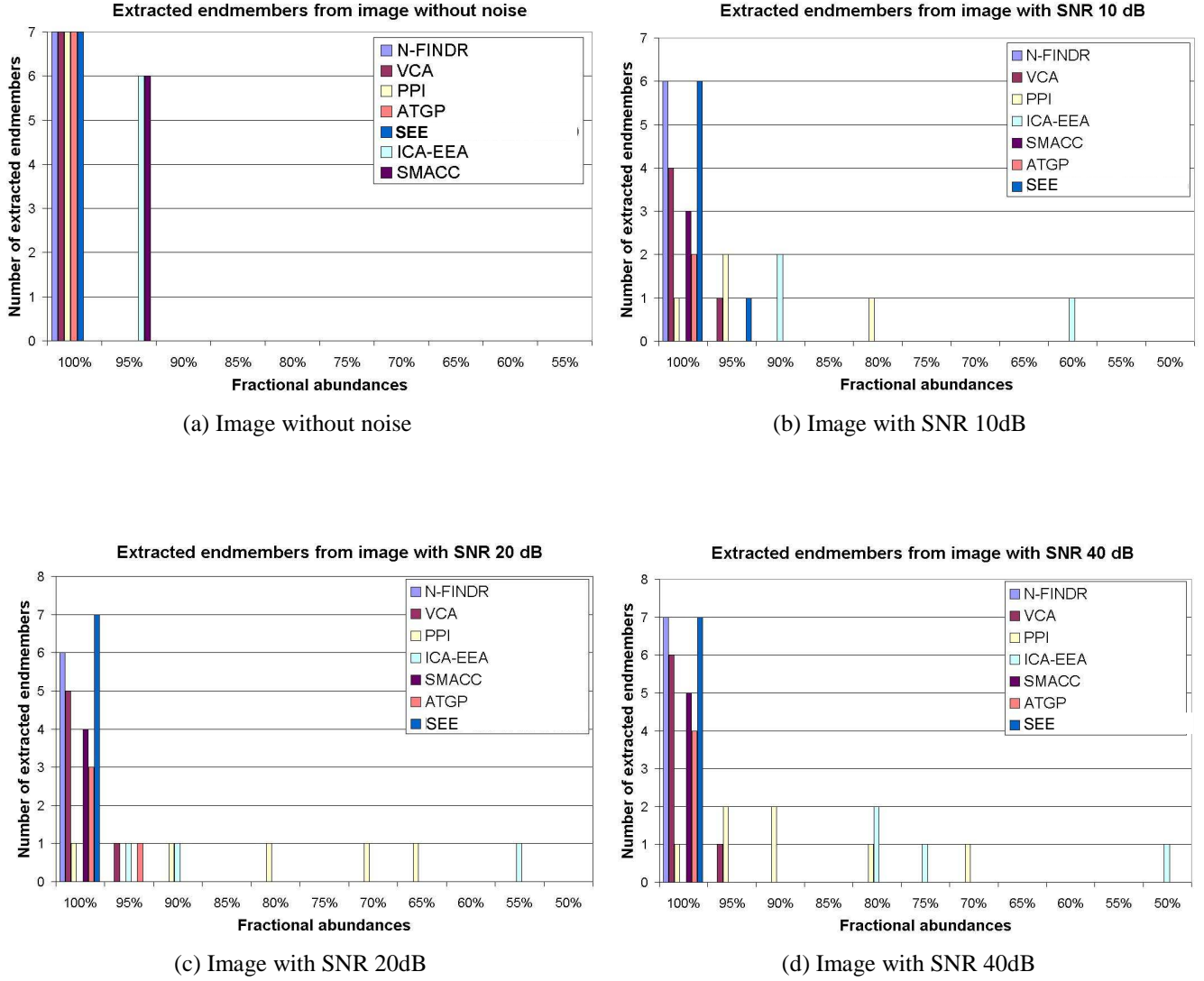
Figure 4.10: Number and abundance of the extracted endmembers per algorithm.

Scenario 2: The 50x50-pixel simulated hyperspectral images[14] with 423 spectral bands were created in such a way that a) the abundance fractions follow a Dirichlet distribution enforcing positivity and full additivity constraints and b) the noise is zero-mean white Gaussian, leading to SNR values of 50dB, 30 dB, 20 dB and 10 dB. In this section N-FINDR, VCA, SEE and E-SEE are tested and compared.

E-SEE was run after adding 300% more pixels of the initial image having the spectral value of the maximum projected pixel onto the first transformed component derived by PCA. PCA was applied since noise is zero-mean white Gaussian. Spectral angle distance was used as the spectral similarity

---

[14] The algorithm for the generation of the simulated data is available at http://www.lx.it.pt/~bioucas/code.htm.

measure between the estimated endmembers from the four EEAs and their corresponding reference spectral signatures from USGS library.

Compared to all the remaining applied EEAs, E-SEE performed better when it was applied on image with SNR 10 and SNR 20. As it is reported in Tables 4.3 and 4.4, E-SEE presented the lowest spectral angles at five out of seven endmembers, whereas SEE yields the second best performance presenting the lowest spectral angles at three out of seven endmembers when applied on the image with SNR 10. For images with higher SNR, all the applied EEAs performed quite similar.

Table 4.7 reports the average spectral angles. The performance is considered to be better if the average spectral angle is smaller. E-SEE presents better performance compared to the other EEAs concerning the average SAD whereas SEE and VCA yield the second best performance.

Table 4.3: Spectral angles (in rads) between extracted endmembers from each EEA (SNR 10) and their corresponding reference spectral signature from the USGS library.

| | SNR 10 | | | | | | |
	sea	blackbrush	soil	stonewall	andradite	almandine	Glauconite
N-FINDR	0.989	0.341	**0.460**	**0.222**	0.152	0.353	0.576
VCA	0.941	0.341	0.467	0.243	0.15	**0.334**	0.576
SEE	**0.924**	**0.316**	**0.460**	0.232	**0.139**	0.350	0.510
E- SEE	**0.924**	**0.298**	**0.460**	0.232	**0.139**	0.350	**0.422**

Table 4.4: Spectral angles (in rads) between extracted endmembers from each EEA (SNR 20) and their corresponding reference spectral signature from the USGS library.

| | SNR 20 | | | | | | |
	sea	blackbrush	soil	stonewall	andradite	almandine	Glauconite
N-FINDR	**0.726**	0.120	0.183	0.085	**0.073**	**0.123**	0.259
VCA	**0.726**	0.120	**0.181**	**0.083**	0.075	0.125	0.259
SEE	0.730	**0.118**	0.185	0.085	**0.073**	0.125	0.258
E-SEE	**0.726**	**0.118**	0.185	**0.083**	**0.073**	0.125	**0.248**

Table 4.5: Spectral angles (in rads) between extracted endmembers from each EEA (SNR 30) and their corresponding reference spectral signature from the USGS library.

| | SNR 30 | | | | | | |
	sea	blackbrush	soil	stonewall	Andradite	almandine	Glauconite
N-FINDR	0.371	0.039	**0.082**	0.049	0.059	0.067	0.122
VCA	0.371	0.039	0.085	0.049	0.059	0.067	0.122
SEE	**0.367**	0.039	0.085	0.049	0.059	0.067	0.122
E- SEE	**0.367**	0.039	0.085	0.049	0.059	0.067	0.122

Table 4.6: Spectral angles (in rads) between extracted endmembers from each EEA (SNR 50) and their corresponding reference spectral signature from the USGS library.

	sea	blackbrush	soil	stonewall	andradite	almandine	Glauconite
SNR 50							
N-FINDR	0.062	0.010	0.062	0.043	0.058	0.056	0.098
VCA	0.062	**0.009**	0.062	0.043	0.058	0.056	0.098
SEE	0.062	0.010	0.062	0.043	0.058	0.056	0.098
E- SEE	0.062	0.010	0.062	0.043	0.058	0.056	0.098

Table 4.7: Average spectral angles over the endmembers (in rads).

	SNR 10	**SNR 20**	**SNR 30**	**SNR 50**
N-FINDR	0.44186	0.22414	0.11271	0.05557
VCA	0.43600	0.22414	0.11314	**0.05543**
SEE	0.41871	0.22486	**0.11257**	0.05557
E-SEE	**0.40357**	**0.22257**	**0.11257**	0.05557

Real data experiments

The proposed algorithms, SEE and E-SEE were applied on a real hyperspectral remote sensing image which was collected in 1997 by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor over a well-known mining region of Cuprite in Nevada. The image scene, which is available on line in reflectance values after atmospheric correction[15], is well understood mineralogically and it is widely used to validate the performance of endmember extraction algorithms. Several of the exposed minerals are included in the USGS digital library[16]. According to the associated ground based observations a wide range of minerals can be identified in the image [120], [121].

The original image has 220 spectral bands covering a spectral range from 0.4 to 2.5 μm. A sub-image of 301 x 300 pixels of AVIRIS scene was selected for the experiments (Fig. 2.11). This dataset contained 50 contiguous spectral bands in the range of 2-2.5 μm. This range present key absorption features associated with minerals [76].

Outlier detection method [25] resulted in the number of 20 endmembers. According to [88], [116], [122], [123] five highly representative minerals (alunite, buddingtonite, calcite, kaolinite, muscovite)

---

[15] http://aviris.jpl.nasa.gov/html/aviris.freedata.html
[16] http://speclab.cr.usgs.gov/spectral-lib.htm

depicted in the sub-image are extracted with success by both N-FINDR and VCA algorithms. In the current study, an experiment-based cross examination has been performed in order to evaluate whether the proposed methods can extract efficiently the aforementioned minerals. Analysis of the endmembers found by the applied EEAs showed that two more minerals (chabazite and sillimanite) - i.e. totally seven minerals- have also been extracted by all the four methods. Consequently for a more complete comparison, each method is evaluated in terms of its accuracy to extract the seven aforementioned minerals. The USGS spectral signatures of these minerals, shown in Fig. 4.11, are selected as ground truth references for the evaluation of the applied EEAs. The extracted endmembers were compared with the ground truth signatures using spectral angles after continuum removal function [124].

The spectral angles between the reference minerals spectral signatures and the corresponding extracted signatures by each EEA are reported in Table 4.8. The average spectral angle of each EEA is also reported. As listed in Table 4.8, E-SEE provided better performance compared to the SEE and it presents the lowest spectral angles for four out of seven minerals. N-FINDR performed quite well with three out of seven lowest spectral angles, and SEE follows with two out of seven lowest spectral angles; VCA performed the worst. The performance of the EEAs is also reflected to the average spectral angles; E-SEE presented the lowest average spectral angle and SEE yielded the second best performance.

Table 4.8: Spectral angles (in rads) between extracted endmembers from each EEA and their corresponding reference spectral signature from the USGS library.

Algorithm	USGS	Alunite	Buddingtonite	Calcite	Chabazite	Kaolinite	Muscovite	Sillimanite	Average
N-FINDR		0.059	0.050	**0.045**	**0.022**	0.060	**0.052**	0.041	0.0470
VCA		0.075	0.054	0.050	0.037	**0.033**	0.055	0.038	0.0489
SEE		0.058	**0.043**	0.051	0.029	0.038	**0.052**	0.031	0.0431
E-SEE		**0.056**	**0.043**	**0.045**	0.029	0.038	0.063	**0.024**	**0.0426**

Fig. 4.11 shows one-to-one comparison of continuum removed spectra between the estimated endmembers by the applied EEAs and the USGS spectral library spectra. Topological changes of data resulted from the E-SEE was implemented by adding 300% more pixels. Differences between the

reference spectra and the extracted endmembers spectra in the absolute value of reflectance are likely due to atmospheric transmission effects.

(a)

Alunite

USGS Spectral library spectrum
Endmember found by E-SEE
Endmember found by VCA
Endmember found by N-FINDR

wavelength (µm)

(b)

Buddingtonite

USGS Spectral library spectrum
Endmember found by E-SEE
Endmember found by VCA
Endmember found by N-FINDR

wavelength (µm)

(c)

Calcite

USGS Spectral library spectrum
Endmember found by E-SEE
Endmember found by VCA
Endmember found by N-FINDR

wavelength (μm)

(d)

Chabazite

USGS Spectral library spectrum
Endmember found by E-SEE
Endmember found by VCA
Endmember found by N-FINDR

wavelength (μm)

(e)

Kaolinite

USGS Spectral library spectrum

Endmember found by E-SEE

Endmember found by VCA

Endmember found by N-FINDR

(f)

Muscovite

USGS Spectral library spectrum

Endmember found by E-SEE

Endmember found by VCA

Endmember found by N-FINDR

(g)

Figure 4.11: Comparison of continuum removed spectra of endmembers found by several EEAs with USGS spectral library spectra.

Effectiveness of E-SEE was also evident when it was implemented for the detection of oil spills. In particular, in the frame of ARGOMARINE[17] airborne hyperspectral imagery was acquired using the CASI-550 hyperspectral sensor[18]. The scene was depicting the seawater area of Laganas bay of Zakynthos island. In this bay a natural non-continuant submarine oil outflow exists, resulting in the appearance of natural oil-spills on the sea surface. Fig. 4.12 shows the CASI image, the location of the detected endmember of oil spill and the classified image using spectral angle mapper (classifier) (SAM) [68].
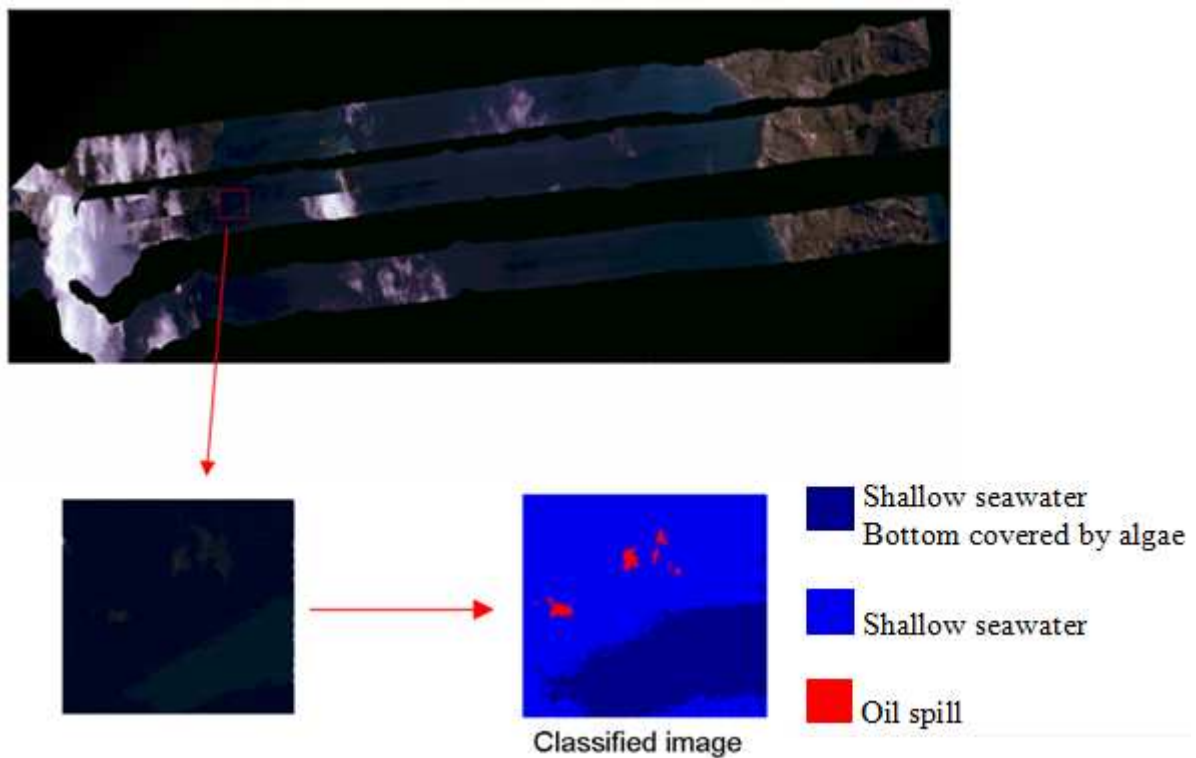
Figure 4.12: Oil spill detection using E-SEE.

4.5 Conclusions

In this work, new simultaneous simplex-based unsupervised endmember extraction methods have been introduced; the simple endmember extraction (SEE) and the Enhanced-SEE (E-SEE) empirical method

---

[17] FP7 European program
[18] The image was acquired by the Laboratory of Remote Sensing of NTUA

which have light computational burden and are characterised by their clear conceptual meaning. The SEE method seeks for endmembers at the vertices of the simplex which is spanned by the zero mean data and a few eigenvectors. The method is simple with very low computational cost. E-SEE has been developed in order to overcome the tendency that the majority of the convex-geometrical based endmember extraction methods encounter to select high contrast endmembers over less contrasted endmembers. This is accomplished by changing the data distribution. More precisely, E-SEE moves the mean of the data by adding pixels which have the spectral signature of the maximum projected pixel onto the first transformed component. In comparison to SEE, E-SEE presented higher performance regardless the distance of the extracted endmembers from the data mean and the amount of noise.

For evaluation purposes, VCA and N-FINDR were also performed. Experiments on synthetic data demonstrated the effectiveness of the proposed methods in comparison with VCA and N-FINDR for various proportions of pure and mixed pixels. E-SEE presented the best performance, whereas SEE and VCA yielded the second best performance in terms of average spectral angles. In case of the AVIRIS Cuprite real hyperspectral data, results showed that E-SEE method outperformed the other methods. It should be noted that despite its simplicity, the SEE algorithm yielded a better performance than VCA and N-FINDR in terms of average spectral angle. Finally, computation complexity of the proposed methods is much lower than those of VCA and N-FINDR, approximately one and three orders of magnitude lower, respectively. The very low computation complexity of SEE as well as its performance makes the method a very fast and effective tool for endmember extraction, whereas low computation complexity and high performance of E-SEE make the E-SEE a reliable and fast tool for endmember extraction.

Chapter 5

Abundance estimation

5.1 Introduction

In the previous section we have seen the main aspects of endmember extraction and the approaches which address it. Abundance estimation yields the last output of spectral unmixing which represents the proportion of each endmember in the pixel. Abundance estimation can be accomplished having previously defined the endmembers or at the same time with endmember determination. It is an inversion problem based on spectral mixture analysis (see section 1.2). Given the observed spectral vectors y and the matrix containing the endmembers S in eq. (1.1), the inversion step can be considered as an optimization problem which minimizes the residual between the observed vectors and the linear space spanned by the inferred spectral signatures [12] using least squares estimation described in eq. (1.2). Methods which address abundance estimation as a least squares problem can be found in [11], [13], [125], [126]. Another group of inversion methods address the abundance estimation as a multiple signal detection problem [114]. In this group, implementation is based on orthogonal subspace projection (OSP) and, instead of estimating the abundances of all the endmembers at the same time, it regards a single endmember as the desired signal source to be detected, and all the others as the undesired sources. Then, all the data sample vectors are orthogonally projected onto the space spanned by the undesired sources leading to the suppression of the latter. More elaborate description of OSP-

based inversion methods can be found in [11], [13], [90], [127], [128], [129], [130]. Besides optimization-based and OSP-based algorithms, there have been developed approaches which model the underlying physics using either neural networks [131], [132], networks of endmembers [133], or more complicate expressions of multiple photon interactions [15].

The majority of the existing inversion methods use a unique set of endmembers for spectral unmixing of the entire image, failing to account that each pixel may comprise a different combination of endmembers. For instance, imagine a hyperspectral scene depicting 1) a built-up area with metal roofs, and 2) sea. It is reasonable that pixels with seawater have no participation into metal roof spectra. Nevertheless, the standard application of spectral unmixing would use the spectral signature of the metal roof as one of the input endmembers for the unmixing of seawater pixels. Consequently, the accuracy of the unmixing method would be low since utilizing more endmembers than the actual set makes the model sensitive to instrumental noise, atmospheric contamination and natural variability in spectra, resulting in fraction error [33].

5.2 Multiple endmember spectral mixture analysis

Multiple endmember spectral mixture analysis (MESMA) [23] extends the linear spectral mixture analysis (LSMA) [16] and allows the number and types of endmembers to vary on a per-pixel basis. LSMA assigns each endmember to one spectral signature, neglecting the spectral variability of the same material under different scene conditions. On the other hand, MESMA can be used to account for within class spectral variability because it creates a large endmember pool in which multiple endmembers belong to the same class. MESMA is a spectral matching technique which premises the set of candidate endmembers contained in the endmember pool -either derived from the entire data sample or reference endmembers- to be known a priori and seeks for the optimum set of endmembers that composes the spectral signature of each pixel. Three techniques have been used for determining the most appropriate set of endmembers for each pixel; count-based endmember selection (CoB) [134], endmember average

root mean squared error (EAR) [135] and minimum average spectral angle (MASA) [136]. The CoB selects the endmembers which model the greatest number of endmembers within their class whilst minimizing endmember overlapping with other classes. The EAR selects the endmembers which produce the lowest RMSE within a class and MASA selects those with the lowest average spectral angle.

According to [23], the selection of the optimum set of endmembers using MESMA is achieved by setting a threshold to the RMSE between the original and the reconstructed pixel spectra and by keeping combinations that produce acceptable abundances (e.g. values between -0,01 and 1,01). In [137], authors proposed two new algorithms - the second is a fast version of the first - in order to select the optimum set of endmembers for each pixel, without any threshold requirement. The first algorithm performs unconstrained least squares (UCLS) method for all the possible combinations of endmembers extracted by an endmember extraction algorithm (EEA), and for each pixel retains as final set, the one with the minimum RMSE and nonnegative abundances. The aforementioned algorithm is very time-consuming because it runs for all the possible combinations of endmembers. The second algorithm, proposed in [137], is mainly based on the RMSE and aims to fasten the searching process. It relies on the concept that if an endmember contained in a given pixel is removed the RMSE will become larger. Results showed that nonnegativity and sum-to-one constrains on abundances can be automatically satisfied but compared to UCLS method, the two algorithms proposed by authors in [137] presented very high computational complexity and a bit lower accuracy.

5.3 MESMA-SAD

MESMA needs to calculate all the potential endmember combinations of each pixel to find the best-fit one. Therefore, MESMA is a time-consuming and low-efficiency unmixing technology, especially for hyperspectral images. In this study, a novel MESMA is introduced, called *multiple endmember spectral mixture analysis based on spectral angle distance* (MESMA-SAD), [32] which aims to minimize the

time-processing by combining the Spectral Angle Distance (SAD) values and the mean absolute errors (MAE). The algorithm does not require any threshold. The new method attempts to exploit the advantage of SAD to be insensitive to differences in the albedo of the modeled spectrum and computational inexpensive [136]. More specifically, estimation of SAD between a set of extracted endmembers and a pixel spectrum provides the reduction of the number of endmember combinations from 2^p -1 to p, where p is the number of endmembers, by retaining the endmember combinations with the minimum SAD values. The fully constraint least squares (FCLS) method runs for the p combinations and the endmember combination with the minimum MAE between the original and reconstructed pixel spectrum is selected as the optimum. The MESMA-SAD requires the endmembers presented in the image to be known previously to its application. In this study, E-SEE algorithm was firstly applied for endmember extraction.

5.3.1 Analysis of the proposed method

Assuming that there are p endmembers in the image, for each pixel, the MESMA-SAD process can be summarized by the following steps:

1. Estimation of the SAD values between the pixel under processing and the endmembers using 2^p -1 endmember combinations. 2^p -1 SAD values will be produced. For combination sets which contain more that one endmember, a SAD is initially estimated between each endmember and the pixel, and the sum of the estimated SAD values is calculated. In case of a single endmember, there is a unique SAD value. Hereinafter, both the summed SAD and the unique SAD values will be referred as summed spectral angle distance (SSAD).

2. The estimated SSAD values are categorized into p classes in such a way that each class comprises SSAD values which are estimated from the same number of endmembers.

3. Selection of one endmember combination with the minimum SSAD from each class.

4. Overall, p endmember combinations are selected for which FCLS method is performed.

5. The endmember combination with the minimum MAE between the original and reconstructed pixel spectrum is selected as the optimum.

For illustration purposes, Fig. 5.1 presents four imaginary spectra plotted as vectors in two spectral bands which represent 3 endmembers, A, B, and C and a random mixed pixel M. Assuming that the spectral signature of pixel M is composed by endmembers A and B only, its vector in the spectral space will be close to these endmembers, as shown in Fig. 5.1. Since there are three endmembers, there will be 7 endmember combinations, and 3 SSAD classes, shown in Table 5.1. For each class, the combination with the minimum SSAD is selected. As it is shown in Fig. 5.1, from the first class the $SSAD_{MB}$, from the second class the $SSAD_{MA + MB}$, and from the third class, the $SSAD_{MA+MB+MC}$ combination is selected. FCLS is performed for these three endmember combinations. The endmember combination with the minimum MAE between the original and reconstructed pixel spectrum is selected as the optimum one. Table 5.2 reports the number of iterations during the unmixing process. The MESMA-SAD algorithm significantly minimizes the computational time compared to other MESMA algorithms as the 2^p-1 iterations during the unmixing process become p.
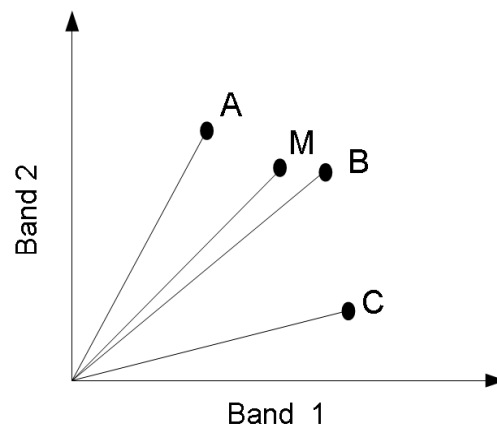
Figure 5.1: Imaginary spectra in two dimensional diagram.

Table 5.1: The three SSAD classes and the combinations included in each class for the given example.

CLASS	COMBINATIONS	SSAD
	A	$SSAD_{MA}$
1st	B	$SSAD_{MB}$
	C	$SSAD_{MC}$
	AB	$SSAD_{MA + MB}$
2nd	AC	$SSAD_{MA + MC}$
	BC	$SSAD_{MB + MC}$
3rd	ABC	$SSAD_{MA+MB+MC}$

Table 5.2: Number of iterations during the unmixing process.

Method	Number of iterations
MESMA	2^p-1
MESMA-Algorithm 2[19]	$p\times(p+1)/2$
MESMA-SAD	p

5.4 Experiments

The proposed algorithm was applied on the same AVIRIS Cuprite dataset used in section 2.4.

For comparison purposes, UCLS and FCLS algorithms were also applied on the AVIRIS data. As it is reported in Table 5.3, the MAE between the original and the reconstructed image is a bit lower when MESMA-SAD was applied compared to FCLS. UCLS presented the lowest MAE but it should be noted that the estimated fractional abundances do not comply with sum-to-one and non-negativity constrains and results may be unrealistic.

Table 5.3: MAE between the original and the reconstructed image by each applied algorithm.

ALGORITHM	MAE between the original and reconstructed image
MESMA-SAD	0.00264468
UCLS	0.00210867
FCLS	0.00264565

---

[19] Algorithm 2 is referred to [137].

This fact is depicted in Fig. 5.2-5.4, where the estimated abundances of each applied method are presented for three minerals. Results can be evaluated by visually comparing them with the classification maps [105] produced by USGS Tricorder algorithm [120] (Fig.5.5). As it is observed, MESMA-SAD appeared the best performance regarding the estimation of buddingtonite fractional abundances. Concerning the muscovite, MESMA-SAD presented satisfactory results since it presented higher abundances to the pixels indicated by Tricorder map but it presented lower performance for kaolinite compared to FCLS algorithm. The fractional abundances resulted from UCLS differ significantly from those given in the Tricorder map.
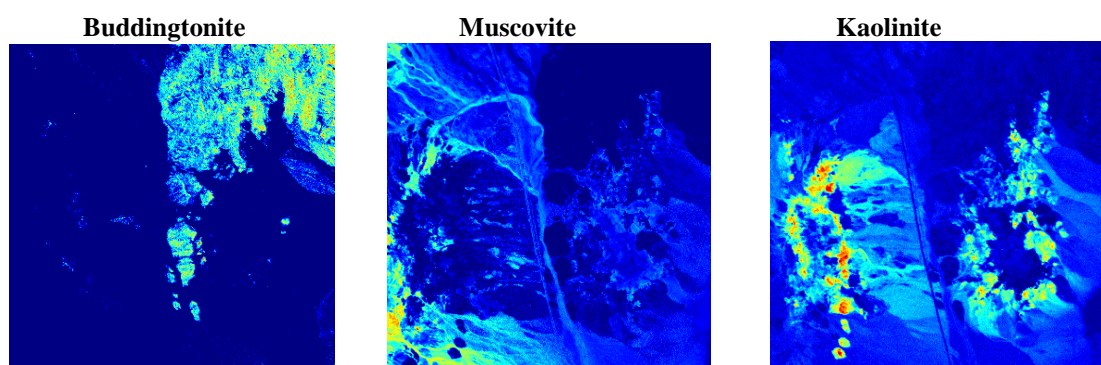
Figure 5.2: Fractional abundances resulted from UCLS.
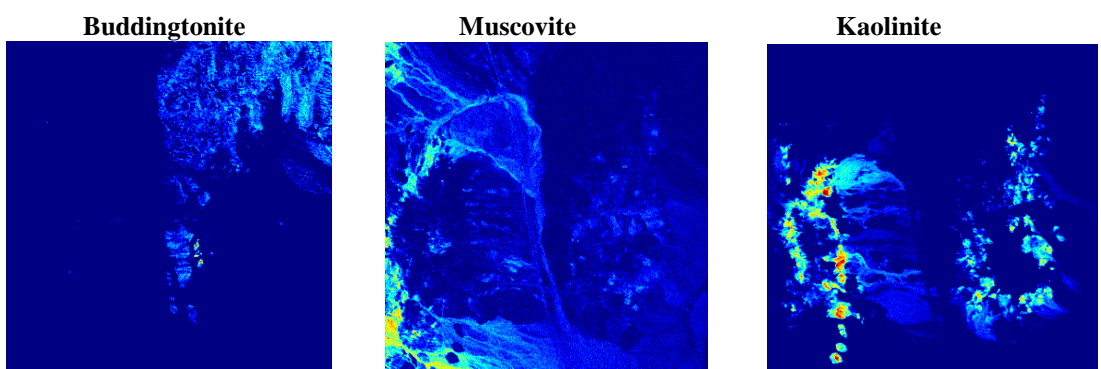
Figure 5.3: Fractional abundances resulted from FCLS.
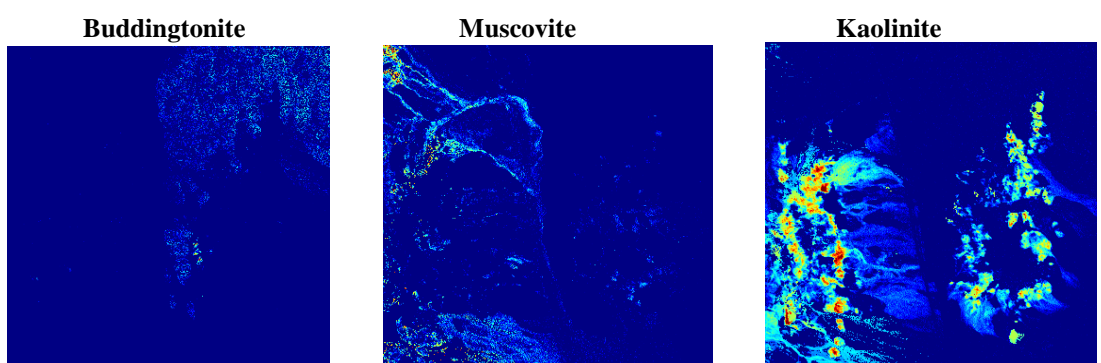
Figure 5.4: Fractional abundances resulted from MESMA-SAD.
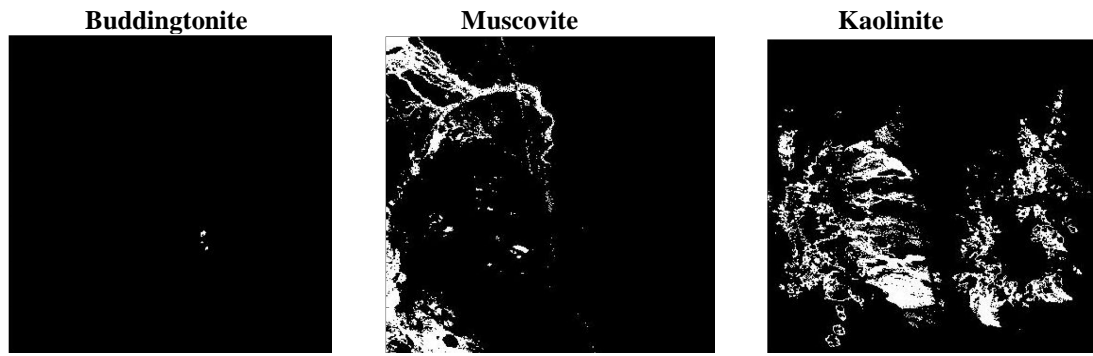
Figure 5.5: Classification maps produced by Tricorder.

5.5 Conclusions

MESMA concept considers all the possible different sets of endmembers for each pixel for performing unmixing. The main constraint of MESMA is the computational complexity required for the retrieval of the optimum endmember combination which composes the signature of each pixel. This work presents a new MESMA method, the MESMA-SAD, which uses the SAD values in order to reduce the number of endmember combinations, and MAEs between the original and the reconstructed pixel spectrum for the selection of the optimum endmember combination. The MESMA-SAD does not require tuning parameters and reduces significantly the time-processing as the 2^p-1 iterations required during the unmixing process in the MESMA become p. In order to evaluate the effectiveness of the proposed method, experiments on AVIRIS CUPRITE hyperspectral data was performed. Results showed that the proposed algorithm has much potential in spectral unmixing field but further work is needed to be done in order to have a more thorough view of the stong and the weak points of the method.

Chapter 6

Conclusions and future work

Spectral unmixing is the process of decomposing the observed pixel spectrum into endmembers and estimating their proportions, also called abundances. It is a very active research topic as its enormous potential has aroused the interest of researchers from many different disciplinary areas. This thesis focused on several issues related to spectral unmixing which need to be addressed so that the full potential of hyperspectral data exploitation to be employed. These issues regard the steps which are involved in spectral unmixing process, including signal subspace estimation, dimensionality reduction, endmember extraction and abundance estimation.

The performance of each of the aforementioned steps may affect directly the performance of the following ones. The following facts are adduced to show the importance of developing effective methods for improving the spectral unmixing process:

- An accurate determination of the number of the endmembers significantly 1) contributes to the development of unsupervised methods in terms of dimensionality reduction, endmember extraction and target detection, 2) contributes to the accuracy of the spectral unmixing process and 3) enables low-dimensional representation of spectral vectors.

- Dimensionality reduction methods define the hyperspectral data into a lower dimension in order to be analyzed more effectively, yielding gains in computational time and complexity, data

storage, algorithm performance and signal-to-noise ratio. It is widely used as a pre-processing step of great interest as it affects directly the performance of the following processing tasks.

> ➢ Band selection is a dimensionality reduction technique which selects the original spectral channels, retaining the spectral characteristics of the materials. Selecting the optimum set of bands which comprise the most informative ones is of crucial importance for the data representation in lower dimension and for the identification of desired materials.

- Development of new endmember extraction methods by introducing new concepts for optimizing their performance presents a high level of research activity as the image scenes become more complex and spectral resolution of hyperspectral sensors offer an invaluable diversity of information, enabling the accurate physical description and discrimination of the sensed materials.

- Designing spectral unmixing algorithms of low complexity and/or improving the computational time of existing algorithms contributes to real-time applications.

In the frame of this thesis, existing relevant methods have been elaborately studied and extensive analysis of the hyperspectral data sets have been thoroughly conducted seeking to discover the main issues related to the spectral unmixing process that needs to be explored. The main issues that are uncovered and investigated are as follows:

- A successful estimation of the number of endmembers strongly depends on how well signal and noise are discerned.

- The existing methods for estimation of the signal subspace dimension consider the existence of two different distributions, the one related to noise and the other related to signal, or in geometrical approach they consider two different subspaces one of noise and one of signal. However, in hyperspectral space, signal vectors are very few compared to noise vectors in order to estimate their population distribution properly or to statistically analyze them.

- Specific spectral characteristics between similar materials are present at certain wavelengths and this crucial and critical information may be compromised and distorted if no original data are used.

- There are a few band selection (BS) methods which have been addressed to spectral unmixing and these methods set criteria to the spectral information derived by the whole set of wavelengths; a generalised approach which disregards the spectral characteristics of a particular material of interest, the image diversity and the endmember variability.

- Working with reference endmembers implies several matters: firstly, it is based on the premise that the data sample vectors are calibrated and corrected for radiometric effects in order to be properly compared with them, a process which is not always likely to be accomplished nor to be evaluated; secondly, when using reference endmembers there exists a difficulty in interpreting the scene of remote-sensing scale using spectra measured at centimetre or millimeter scales; thirdly, the computational complexity of such approaches using the available spectral libraries can be quite high.

- Creating virtual endmembers contributes to minimizing the spectral unmixing error but mathematically correct endmembers do not always represent physically meaningful spectra able to interpret the imagery.

- An optimal EEA must select all the endmembers at the same time using the whole spectral data (simultaneous extraction), rather than sequentially extracting them. However, the former is characterised by its high computational burden.

- The majority of the convex-geometrical based endmember extraction methods have the tendency to select high contrast endmembers over less contrast endmembers.

- The spectral vectors of endmembers which are closer to data mean are more likely to be disregarded by the endmember extraction methods.

- Multiple endmember spectral mixture analysis (MESMA) [23] is an effective spectral matching technique which accounts for within class spectral variability. However, it needs to calculate all

the potential endmember combinations of each pixel to find the best-fit one, demanding a time-consuming unmixing technology.

The main contributions of this thesis include the exploitation of novel concepts and the development of five new methods dedicated to the spectral unmixing process, aiming at overcoming the shortcomings described above.

- A new method for estimation of the signal subspace dimension, called *outlier detection method* (ODM) has been introduced. ODM is a new automatic non-parametric method whose novelty lies in the fact that it considers only the existence of noise and treats signals as outliers of noise. It searches for the signals whose radius is by far larger than that of the noise and introduces for the first time in virtual dimension theory a robust outlier detection method. Briefly mentioned, the main key points of ODM are as follows:

  1. ODM exploits the geometrical properties of noise hypersphere which are given by information theory.

  2. Noise vectors lie in a hypersphere of radius equal to their standard deviation value while the signal vectors have evidently larger standard deviations values which vary in all directions and consequently, they lie in a hyperellipsoid.

  3. No threshold is needed between signal and noise vectors.

  4. A robust outlier detection method is used, called inter quartile range (IQR) based method. Its benefit lies in the fact that it can be used when data distribution is unknown and thus, no statistical parameter estimation is needed. The risk of estimating erroneously the signal distribution due to its small population is omitted.

  5. The proposed method is characterized by its simplicity.

Experiments using simulated and real images indicated the efficiency of the ODM, even when small sets of sample vectors were used, as it outperformed compared to its competitors.

- A new band selection (BS) approach has been developed which can be adopted by endmember extraction and classification methods leading the latter to enhanced performance. The method is tied to the materials/classes of interest. It is the first attempt to address the issue of disregarding spectrally closed endmembers (SCEs) – a term introduced in the frame of this study- exploiting the original bands. In order to extract the SCEs, the concept of multiple convex hulls [79] is used for the first time in BS. The main key points of the proposed BS approach are as follows:

  1. It accounts for the different distribution of each material's convex hull to the data cloud. For this reason, the proposed BS approach allows the final number of endmembers assigned to a material class to vary resulting in more physically meaningful spectra.

  2. It accounts for image diversity and is tied to the available information derived from the image's spectral characteristics. Thus, different sets of bands are exploited for detecting different materials depicted in the same image scene.

  3. Being image-dependent compensates the endmember variability, contrary to the existing approaches which use a priori information related to the absorption features at fixed wavelengths.

  4. The proposed BS contributes to optimizing the classification performance when the available training data sets are limited and facilitates the SVM classification method.

Experiments were implemented using two real hyperspectral images. In terms of endmember extraction, the proposed approach is effective at detecting low contrast materials, which were disregarded by the N-FINDR algorithm. The extracted SCEs provided wealth and elaborate information for material classes defined by the prevalent endmembers. The potential of the proposed approach can contribute to the extraction of spectral signatures which imply different biophysical or chemical properties of a material class. In terms of classification, the proposed method achieved overall accuracy very close to the one where all available bands are used (only 1% less) while, at the same time, it reduces the number of selected bands more than 50%.

- Two new simultaneous simplex-based unsupervised endmember extraction methods, the simple endmember extraction (SEE) and the Enhanced-SEE (E-SEE) empirical method, respectively, have been also developed. The proposed methods are based on the linear spectral mixture model which is the most widely used model, adequate for unsupervised approaches, due to its acceptable approximation of light scattering mechanics in many real scenarios. Its clear conceptual meaning contributes to a computationally simple implementation and an easy interpretation. The geometrical interpretation of the linear mixture model is associated with the mathematical theory of *convex sets* based on which, spectra can be restricted in a simplex lying on a signal subspace of dimension one less than the number of the endmembers. The novelty of the SEE method is that it searches for the extreme values that lie on the end points of the existing transformed axes without further projections that imply iterative procedures. The E-SEE empirical method compensates the tendency that the majority of the convex-geometrical based endmember extraction methods encounter to select high contrast endmembers over less contrast endmembers. Its novelty lies in the fact that it changes the distribution of the initial data sample increasing the distance between candidate endmembers and the data mean. The main key points of the proposed EEAs are as follows:

  1. Both SEE and E-SEE are simultaneous EEAs exploiting the information derived from the whole data in contrast to the existing sequential EEAs.

  2. The proposed methods have light computational burden compared to the-state-of-the-art EEAs and are characterised by their clear conceptual meaning.

  3. Endmembers are considered as the dominant distinct spectra which are constituent parts of a spectral mixture. On this account, the proposed methods extract image endmembers which have the advantage of no calibration need, refraining from adding calibration errors in the mixture model, as well as they have the same scale of measurement as the data.

Extensive experiments with simulated and real hyperspectral data sets showed that the proposed methods can be promising in the endmember extraction process. The very low computation complexity of SEE as well as its performance makes the method a very fast and effective tool for endmember extraction, whereas the low computational complexity and the enhanced performance of E-SEE make the E-SEE a reliable and fast tool for endmember extraction.

- Last but not least, the fifth method is a new multiple endmember spectral mixture analysis (MESMA) based on spectral angle distance, called MESMA-SAD, and it is used for the estimation of the fractional abundances for the endmembers in each pixel. The new method applies fully constraint least squares (FCLS) and combines the spectral angle distance (SAD) values and the mean absolute errors (MAE). The algorithm does not require any threshold. The new method attempts to exploit the advantage of SAD to be insensitive to differences in the albedo of the modelled spectrum. Results using a real hyperspectral data set showed that MESMA-SAD significantly minimizes the time-processing compared to the existing MESMA algorithms leading to satisfactory results.

Through the development and the testing of the proposed methods, several directions for future research have been revealed. Taking advantage of ODM to successfully estimate the number of endmembers in small sized images, future research could focus on combining it with endmember extraction methods which integrate spatial information. Regarding the proposed method for band selection, the number of p is selected to be used for defining the sets of bands to be exploited. Experimental research could be conducted to determine the optimum number of bands to be selected using data sets with ground reference data. In addition to this, based on the empirical E-SEE an effort is given to develop a robust criterion according to which the data distribution will be changed.

List of Publications

Publications in scientific journals:

-C. Andreou and V. Karathanassi (2014)."Estimation of the number of endmembers using robust outlier detection method ", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote,* vol.7, no.1, pp.:247–256.

- C. Andreou, V. Karathanassi and P. Kolokousis, (2011). "Investigation of hyperspectral remote sensing for mapping asphalt road conditions", *International Journal of Remote Sensing,* Taylor & Francis, vol. 32, no. 21, pp. 6315-6333.

- V. Karathanassi and C. Andreou. "A new band selection method for optimizing spectrally close endmembers extraction", *GIScience and Remote Sensing Journal* (under review).

- C. Andreou and V. Karathanassi, "New simple endmember extraction method using transformed components for hyperspectral images, *IEEE Trans. on Geoscience and Remote Sensing* (under review).

Publications in proceedings of international conferences:

-C. Andreou and V. Karathanassi (2013). "Effects of band selection on the hyperspectral classification", proc. *SPIE Optics and Photonics*, 25-29 August, San Diego U.S.A.

- C. Andreou and V. Karathanassi (2012). "A novel multiple endmember spectral mixture analysis using spectral angle distance*" proc. IEEE International Geoscience and Remote Sensing Symposium*, 22-27 July 2012, Munich, Germany.

- C. Andreou and V. Karathanassi (2012). "New automated method for estimating the number of endmembers in hyperspectral images", *proc. IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS),* 4-7 June, 2012, Shanghai, China.

- D. Sykas, V. Karathanassi, C. Andreou and P. Kolokoussis (2011). "Oil Spill Mapping Using Hyperspectral Methods and Techniques", *Proceedings of the Tenth International Conference on the Mediterranean Coastal Environment, (MEDCOAST),* October 2011, Rhodes, Greece.

- C.Andreou and V. Karathanassi (2011). "Endmember detection in marine environment with oil spill event", *proc. SPIE Remote Sensing*, 19-22 September 2011, Prague, Czech Republic.

- C.Andreou and V. Karathanassi (2011). "Exploitation of hyperspectral imagery for the detection of oil spill characteristics", proc. CEST 8-10 September, 2011, Rhodes, Greece.

- C. Andreou, V. Karathanassi (2011). "Using principal component analysis for endmember extraction", proc. *IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 5-9 June, 2011, Lisbon, Portugal.

- D. Sykas, V. Karathanassi, C. Andreou and P. Kolokoussis (2011). "Oil spill thickness estimation using unmixing methods", *proc. IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS),* 5-9 June, 2011, Lisbon, Portugal.

- *C.* Andreou, V. Karathanassi, P. Kolokoussis (2011). "Spectral Library for Oil Types", *proc. 34th International Symposium on Remote Sensing of the Environment (ISRSE).* April 10-15, 2011, Sydney, Australia.

References

[1]   J. B. Adams and A. R. Gillespie, *Remote sensing of landscapes with spectral images : A physical modeling approach*. Cambridge, New York, Cambridge University Press, 2006.

[2]   G. Rees, *Physical principles of remote sensing*, 2nd ed. Cambridge, U.K., New York, Cambridge University Press, 2001.

[3]   A. F. Goetz, G. Vane, J. E. Solomon, and B. N. Rock, "Imaging spectrometry for earth remote sensing," *Science*, vol. 228, no. 4704, pp. 1147–53, 1985.

[4]   R. O. Green, M. L. Eastwood, C. M. Sarture, T. G. Chrien, M. Aronsson, B. J. Chippendale, J. A. Faust, B. E. Pavri, C. J. Chovit, M. Solis, M. R. Olah, and O. Williams, "Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS)," *Remote Sensing of Environment*, vol. 65, no. 3, pp. 227 – 248, 1998.

[5]   M. Borengasser, W. S. Hungate, and R. L. Watkins, *Hyperspectral remote sensing : Principles and applications*. London, CRC Press, 2008.

[6]   D. Manolakis, D. Marden, and G. A. Shaw, "Hyperspectral image processing for automatic target detection applications," *Lincoln Laboratory Journal*, vol. 14, no. 1, pp. 79 – 116, 2003.

[7]   A. Gowen, C. O'Donnell, P. Cullen, G. Downey, and J. Frias, "Hyperspectral imaging : An emerging process analytical tool for food quality and safety control," *Trends in Food Science & Technology*, vol. 18, no. 12, pp. 590 – 598, 2007.

[8]   F. Dell'Acqua, P. Gamba, A. Ferrari, J. Palmason, J. Benediktsson, and K. Arnason, "Exploiting spectral and spatial information in hyperspectral urban data with high resolution,", *IEEE Geoscience and Remote Sensing Letters,* vol. 1, no. 4, pp. 322–326, Oct 2004.

[9]   C. Andreou and V. Karathanassi, "Endmember detection in marine environment with oil spill event," vol. 8180, pp. 81800P–81800P–8, 2011.

[10]   C. Andreou, V. Karathanassi, and P. Kolokoussis, "Investigation of hyperspectral remote sensing for mapping asphalt road conditions," *International Journal of Remote Sensing*, vol. 32, no. 21, pp. 6315–6333, 2011.

[11]   C.-I. Chang, *Hyperspectral imaging: techniques for spectral detection and classification*. Dordrecht, London, Kluwer Academic/Plenum, 2003.

[12]   J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Apr. 24 2012.

[13]   C.-I. Chang, *Hyperspectral data processing : Algorithm design and analysis*. Hoboken, NJ: Wiley-Interscience, 2013.

[14]   A. Plaza, J. A. Benediktsson, J. W. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni, "Recent advances in techniques for hyperspectral image processing," *Remote Sensing of Environment*, vol. 113, no. 1, pp. 110 – 122, 2009.

[15]   N. Keshava, "A survey of spectral unmixing algorithms," *Lincoln Laboratory Journal*, vol. 14, no. 1, pp. 55 – 78, 2003.

[16]   J. B. Adams, M. O. Smith, and P. E. Johnson, "Spectral mixture modeling: A new analysis of rock and soil types at the viking lander 1 site," *Journal of Geophysical Research: Solid Earth*, vol. 91, no. B8, pp. 8098–8112, 1986.

[17]   J. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geoscience and Remote Sensing Magazine,* vol. 1, no. 2, pp. 6–36, June 2013.

[18]   R. A. Schowengerdt, R. A. T. f. i. p. Schowengerdt, and classification in remote sensing, *Remote sensing, models, and methods for image processing*, 2nd ed. San Diego, Calif., Academic Press, 1997.

[19]   N. Keshava and J. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine,* vol. 19, no. 1, pp. 44–57, Jan 2002.

[20]   S. Chandrasekhar, *Radiative transfer*. New York, Dover Publications, 1960.

[21]   J. Boardman, "Automating spectral unmixing of AVIRIS data using convex geometry concepts," *Summaries of the Fourth Annual JPL Airborne Geoscience Workshop*, 1993, vol. 1.

[22]   J. W. Boardman, "Analysis, understanding, and visualization of hyperspectral data as convex sets in n space," *Proc. SPIE*, 1995, vol. 2480, pp. 14–22.

[23]   D. Roberts, M. Gardner, R. Church, S. Ustin, G. Scheer, and R. Green, "Mapping chaparral in the Santa Monica mountains using multiple endmember spectral mixture models - ii. environmental influences on regional abundance," *Remote Sensing of Environment*, vol. 65, no. 3, pp. 267–279, 1998.

[24]   C. Andreou and V. Karathanassi, "New automated method for estimating the number of endmembers in hyperspectral images," in *4th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS),* June 2012, pp. 1–4.

[25]   C. Andreou and V. Karathanassi, "Estimation of the number of endmembers using robust outlier detection method," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,*, vol. 7, no. 1, pp. 247–256, 2014.

[26]   A. Zare and P. Gader, "L1-endmembers: a robust endmember detection and spectral unmixing algorithm," *Proc. SPIE*, vol. 7695, 2010, pp. 76951L–76951L–10.

[27]   C. Andreou, V. Karathanassi, and G. Diamantopoulou, "Effects of band selection on the hyperspectral classification," *Proc. SPIE*, 1993, vol. 8870, pp. 88700T–88700T–9.

[28]   V. Karathanassi and C. Andreou, "A new band selection method for optimizing spectrally close endmembers extraction," *GIScience and Remote Sensing Journal* (under review).

[29]   C. Andreou and V. Karathanassi, "Using principal component analysis for endmember extraction," in *3rd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS),* June 2011, pp. 1–4.

[30]   C. Andreou and V. Karathanassi, "Exploitation of hyperspectral imagery for the detection of oil spill characteristics," *Proc. SPIE*, 2011, vol. B, pp. 51–58.

[31]   C. Andreou and V. Karathanassi, "New simple endmember extraction method using transformed components for hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing* (under review).

[32]   C. Andreou and V. Karathanassi, "A novel multiple endmember spectral mixture analysis using spectral angle distance," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS),* July 2012, pp. 4110–4113.

[33]   D. E. Sabol, J. B. Adams, and M. O. Smith, "Quantitative subpixel spectral detection of targets in multispectral images," *Journal of Geophysical Research: Planets*, vol. 97, no. E2, pp. 2659–2672, 1992.

[34]   K. Fukunaga, "Intrinsic dimensionality extraction," in *Classification Pattern Recognition and Reduction of Dimensionality*, ser. Handbook of Statistics, P. Krishnaiah and L. Kanal, Eds. Elsevier, 1982, vol. 2, pp. 347 – 360.

[35]   C.-I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 42, no. 3, pp. 608–619, March 2004.

[36]   S. Moussaoui, H. Hauksdottir, F. Schmidt, C. Jutten, J. Chanussot, D. Brie, S. Doute, and J. Benediktsson, "On the decomposition of mars hyperspectral data by ICA and Bayesian positive source separation," *Neurocomputing*, vol. 71, no. 10-12, pp. 2194 – 2208, 2008.

[37]   B. Luo, J. Chanussot, S. Doute, and L. Zhang, "Empirical automatic estimation of the number of endmembers in hyperspectral images," *IEEE Geoscience and Remote Sensing Letters,* vol. 10, no. 1, pp. 24–28, Jan 2013.

[38]   H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control,* vol. 19, no. 6, pp. 716–723, Dec 1974.

[39]   J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465 – 471, 1978.

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

[41]   R. E. Roger and J. F. Arnold, "Reliably estimating the noise in AVIRIS hyperspectral images," *International Journal of Remote Sensing*, vol. 17, no. 10, pp. 1951–1962, 1996.

[42]   J. W. Tukey, *Exploratory Data Analysis*. Addison Wesley, 1977.

[43]   L. L. Scharf, *Statistical signal processing : Detection, estimation, and time series analysis*. Reading, Mass.: Addison-Wesley Pub. Co, 1991.

[44]   J. A. Richards, *Remote sensing digital image analysis : An introduction*, 2nd ed. Springer-Verlag, 1993.

[45]   C. Shannon, "Communication in the presence of noise," *Proceedings of the IEEE*, vol. 86, no. 2, pp. 447–457, Feb 1998.

[46]   J. R. Pierce, *An introduction to information theory : Symbols, signals and noise*, 2nd ed. New York: Dover Publications, 1980.

[47]   F. M. Willems, "*Information and communication theory: communication theory,*" Eindhoven University of Technology, 2010.

[48]   A. Green, M. Berman, P. Switzer, and M. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 26, no. 1, pp. 65–74, Jan 1988.

[49]   D. M. Hawkins, *Identification of outliers*. London: Chapman & Hall, 1980.

[50]   I. Ben-Gal, "Outlier detection," in *Data Mining and Knowledge Discovery Handbook*, O. Maimon and L. Rokach, Eds. Springer US, 2005, pp. 131–146.

[51]   R. Bellman, *Dynamic programming*. Princeton University Press, 1957.

[52]   I. T. Jolliffe, *Principal Component Analysis*. Berlin; New York: Springer-Verlag, 1986.

[53]   G. Golub and C. Reinsch, "Singular value decomposition and least squares solutions," *EnglishNumerische Mathematik*, vol. 14, no. 5, pp. 403–420, 1970.

[54]   H. Ren, Q. Du, J. Wang, C.-I. Chang, J. Jensen, and J. Jensen, "Automatic target recognition for hyperspectral imagery using high-order statistics," *IEEE Transactions on Aerospace and Electronic Systems,* vol. 42, no. 4, pp. 1372–1385, October 2006.

[55]   J. Wang and C.-I. Chang, "Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 44, no. 6, pp. 1586–1600, June 2006.

[56]   P. Comon, "Independent component analysis, a new concept?" *Signal Processing*, vol. 36, no. 3, pp. 287 – 314, 1994.

[57]   J. Nascimento and J. Bioucas Dias, "Does independent component analysis play a role in unmixing hyperspectral data?" *IEEE Transactions on Geoscience and Remote Sensing,* vol. 43, no. 1, pp. 175–187, Jan 2005.

[58]   A. Ifarraguerri and C.-I. Chang, "Unsupervised hyperspectral image analysis with projection pursuit," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 38, no. 6, pp. 2529–2538, Nov 2000.

[59]   L. Jimenez and D. Landgrebe, "Hyperspectral data analysis and supervised feature reduction via projection pursuit," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 37, no. 6, pp. 2653–2667, Nov 1999.

[60]   J. Friedman and J. Tukey, "A projection pursuit algorithm for exploratory data analysis," *IEEE Transactions on Computers,* vol. C-23, no. 9, pp. 881–890, Sept 1974.

[61]   R. O. Duda and P. E. Hart, *Pattern classification and scene analysis*. New York, London, Wiley-Interscience, 1973.

[62]   B. Guo, S. Gunn, R. I. Damper, and J. D. B. Nelson, "Band selection for hyperspectral image classification using mutual information," *IEEE Geoscience and Remote Sensing Letters,* vol. 3, no. 4, pp. 522–526, Oct 2006.

[63]   L. Bruzzone and C. Persello, "A novel approach to the selection of spatially invariant features for the classification of hyperspectral images with improved generalization capability," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 47, no. 9, pp. 3180–3191, Sept 2009.

[64]   A. Ifarraguerri and M. Prairie, "Visual method for spectral band selection," *IEEE Geoscience and Remote Sensing Letters,* vol. 1, no. 2, pp. 101–106, April 2004.

[65]   R. Huang and M. He, "Band selection based on feature weighting for classification of hyperspectral data," *IEEE Geoscience and Remote Sensing Letters,* vol. 2, no. 2, pp. 156–159, April 2005.

[66]   S. De Backer, P. Kempeneers, W. Debruyn, and P. Scheunders, "A band selection technique for spectral classification," *IEEE Geoscience and Remote Sensing Letters,* vol. 2, no. 3, pp. 319–323, July 2005.

[67]   H. Yang, Q. Du, H. Su, and Y. Sheng, "An efficient method for supervised hyperspectral band selection," *IEEE Geoscience and Remote Sensing Letters,* vol. 8, no. 1, pp. 138–142, Jan 2011.

[68]   F. Kruse, A. Lefkoff, J. Boardman, K. Heidebrecht, A. Shapiro, P. Barloon, and A. Goetz, "The spectral image processing system (SIPS) - interactive visualization and analysis of imaging spectrometer data," *Remote Sensing of Environment*, vol. 44, no. 2–3, pp. 145 – 163, 1993.

[69]   N. Keshava, "Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 42, no. 7, pp. 1552–1565, July 2004.

[70]   P. Pudil, J. Novovicova, and J. Kittler, "Floating search methods in feature selection," *Pattern Recognition Letters*, vol. 15, no. 11, pp. 1119 – 1125, 1994.

[71]   R. Huang and X. Li, "Band selection based on evolution algorithm and sequential search for hyperspectral classification," in *International Conference on Audio, Language and Image Processing, (ICALIP),* July 2008, pp. 1270–1273.

[72]   Y.-L. Chang, J.-P. Fang, J. Benediktsson, L. Chang, H. Ren, and K.-S. Chen, "Band selection for hyperspectral images based on parallel particle swarm optimization schemes," in *IEEE International Geoscience and Remote Sensing Symposium, (IGARSS),* July 2009, vol. 5, pp. V–84–V–87.

[73]   C.-I. Chang and S. Wang, "Constrained band selection for hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 44, no. 6, pp. 1575–1585, June 2006.

[74]   W. Baijie, W. Xin, and C. Zhangxin, "Spatial entropy based mutual information in hyperspectral band selection for supervised classification," *International Journal of Numerical Analysis & Modeling*, vol. 9, no. 2, pp. 181–192, May 2012.

[75]   A. Zare and P. Gader, "Hyperspectral band selection and endmember detection using sparsity promoting priors," *IEEE Geoscience and Remote Sensing Letters,* vol. 5, no. 2, pp. 256–260, April 2008.

[76]   J. R. Jensen, *Remote sensing of the environment : An earth resource perspective*, 2nd ed. Upper Saddle River, N.J.: Pearson/Prentice Hall Pearson Prentice Hall, 2007.

[77]   M. Berman, H. Kiiveri, R. Lagerstrom, A. Ernst, R. Dunne, and J. Huntington, "Ice: a statistical approach to identifying endmembers in hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 42, no. 10, pp. 2085–2095, Oct 2004.

[78]   B. Somers, G. P. Asner, L. Tits, and P. Coppin, "Endmember variability in spectral mixture analysis: A review," *Remote Sensing of Environment*, vol. 115, no. 7, pp. 1603 – 1616, 2011.

[79]   A. Zare and P. Gader, "Pce: Piecewise convex endmember detection," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 48, no. 6, pp. 2620–2632, June 2010.

[80]   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 Transactions on Geoscience and Remote Sensing,* vol. 51, no. 5, pp. 2753–2769, May 2013.

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

[82]   T. G. V. Niel, T. R. McVicar, and B. Datt, "On the relationship between training sample size and data dimensionality: Monte carlo analysis of broadband multi-temporal classification," *Remote Sensing of Environment*, vol. 98, no. 4, pp. 468 – 480, 2005.

[83]   G. Licciardi, F. Pacifici, D. Tuia, S. Prasad, T. West, F. Giacco, C. Thiel, J. Inglada, E. Christophe, J. Chanussot, and P. Gamba, "Decision fusion for the classification of hyperspectral data: Outcome of the 2008 GRSS data fusion contest," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 47, no. 11, pp. 3857–3865, Nov 2009.

[84]   J. A. Gualtieri and R. F. Cromp, "Support vector machines for hyperspectral remote sensing classification," *Proc. SPIE*, 1999, vol. 3584, pp. 221–232.

[85]   F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 42, no. 8, pp. 1778–1790, Aug 2004.

[86]   A. Plaza, P. Martinez, R. Perez, and J. Plaza, "Spatial/spectral endmember extraction by multidimensional morphological operations," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 40, no. 9, pp. 2025–2041, Sep 2002.

[87]   C. Quintano, A. Fernandez-Manso, and D. A. Roberts, "Multiple endmember spectral mixture analysis (MESMA) to map burn severity levels from Landsat images in mediterranean countries," *Remote Sensing of Environment*, vol. 136, no. 1, pp. 76 – 88, 2013.

[88]   C.-I. Chang, C.-C. Wu, W. min Liu, and Y.-C. Ouyang, "A new growing method for simplex-based endmember extraction algorithm," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 44, no. 10, pp. 2804–2819, Oct 2006.

[89]   T.-H. Chan, W.-K. Ma, A. Ambikapathi, and C.-Y. Chi, "A simplex volume maximization framework for hyperspectral endmember extraction," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 49, no. 11, pp. 4177–4193, Nov 2011.

[90]   H. Ren and C.-I. Chang, "Automatic spectral target recognition in hyperspectral imagery," *IEEE Transactions on Aerospace and Electronic Systems,* vol. 39, no. 4, pp. 1232–1249, Oct 2003.

[91]   J. Nascimento and J. Bioucas Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 43, no. 4, pp. 898–910, April 2005.

[92]   J. Boardman, K. Kruse, and R. Green, "Mapping target signatures via partial unmixing of AVIRIS data," in *Proc. Summaries JPL Airborne Earth Sci. Workshop*, 1995, pp. 23–26.

[93]   J. H. Gruninger, A. J. Ratkowski, and M. L. Hoke, "The sequential maximum angle convex cone (SMACC) endmember model," *Proc. SPIE*, 2004, vol. 5425, pp. 1–14.

[94]   R. A. Neville, K. Staenz, T. Szeredi, J. Lefebvre, and P. Hauff, "Automatic endmember extraction from hyperspectral data for mineral exploration," in *Proc. of the Fourth International Airborne Remote Sensing Conference*, 1999, pp. 891–897.

[95]   G. X. Ritter and G. Urcid, "A lattice matrix method for hyperspectral image unmixing," *Information Sciences*, vol. 181, no. 10, pp. 1787 – 1803, 2011.

[96]   A. Ifarraguerri and C.-I. Chang, "Multispectral and hyperspectral image analysis with convex cones," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 37, no. 2, pp. 756–770, Mar 1999.

[97]   A. Zare and P. Gader, "Sparsity promoting iterated constrained endmember detection in hyperspectral imagery," *IEEE Geoscience and Remote Sensing Letters,* vol. 4, no. 3, pp. 446–450, July 2007.

[98]   M. Craig, "Minimum-volume transforms for remotely sensed data," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 32, no. 3, pp. 542–552, May 1994.

[99]   L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 45, no. 3, pp. 765–777, March 2007.

[100]   J. Li and J. Bioucas-Dias, "Minimum volume simplex analysis: A fast algorithm to unmix hyperspectral data," in *IEEE International Geoscience and Remote Sensing Symposium, (IGARSS)*, July 2008, vol. 3, pp. 250–253.

[101]   A. Ambikapathi, T.-H. Chan, W.-K. Ma, and C.-Y. Chi, "Chance-constrained robust minimum-volume enclosing simplex algorithm for hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 49, no. 11, pp. 4194–4209, Nov 2011.

[102]   J. Bioucas-Dias, "A variable splitting augmented lagrangian approach to linear spectral unmixing," in *First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, (WHISPERS),* Aug 2009, pp. 1–4.

[103]   T.-M. Tu, "Unsupervised signature extraction and separation in hyperspectral images: a noise-adjusted fast independent component analysis approach," *Optical Engineering*, vol. 39, no. 4, pp. 897–906, 2000.

[104]   J. Nascimento and J. Bioucas-Dias, "Hyperspectral unmixing based on mixtures of dirichlet components," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 50, no. 3, pp. 863–878, March 2012.

[105]   M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 49, no. 6, pp. 2014–2039, June 2011.

[106]   D. Rogge, B. Rivard, J. Zhang, and J. Feng, "Iterative spectral unmixing for optimizing per-pixel endmember sets," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 44, no. 12, pp. 3725–3736, Dec 2006.

[107]   J. Eckstein and D. Bertsekas, "On the Douglas - Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming*, vol. 55, no. 1-3, pp. 293–318, 1992.

[108]   J. Bioucas-Dias and M. A. T. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, June 2010, pp. 1–4.

[109]   J. Zhang, B. Rivard, and D. M. Rogge, "The successive projection algorithm (spa), an algorithm with a spatial constraint for the automatic search of endmembers in hyperspectral data," *Sensors*, vol. 8, no. 2, pp. 1321–1342, 2008.

[110]   B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, no. 6583, pp. 607–9, 1996.

[111]   A. Plaza, P. Martinez, R. Perez, and J. Plaza, "A new approach to mixed pixel classification of hyperspectral imagery based on extended morphological profiles," *Pattern Recognition*, vol. 37, no. 6, pp. 1097 – 1116, 2004.

[112]   D. Rogge, B. Rivard, J. Zhang, A. Sanchez, J. Harris, and J. Feng, "Integration of spatial-spectral information for the improved extraction of endmembers," *Remote Sensing of Environment*, vol. 110, no. 3, pp. 287 – 303, 2007.

[113]   M. Zortea and A. Plaza, "Spatial preprocessing for endmember extraction," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 47, no. 8, pp. 2679–2693, Aug 2009.

[114]   J. Harsanyi and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 32, no. 4, pp. 779–785, Jul 1994.

[115]   M. E. Winter, "A proof of the n-findr algorithm for the automated detection of endmembers in a hyperspectral image," *Proc. SPIE,* vol. 5425, 2004, pp. 31–41.

[116]   C.-I. Chang, C.-C. Wu, and C.-T. Tsai, "Random n-finder (N-FINDR) endmember extraction algorithms for hyperspectral imagery," *IEEE Transactions on Image Processing,* vol. 20, no. 3, pp. 641–656, March 2011.

[117]   M. Zortea and A. Plaza, "Improved spectral unmixing of hyperspectral images using spatially homogeneous endmembers," in *IEEE International Symposium on Signal Processing and Information Technology,(ISSPIT),* Dec 2008, pp. 258–263.

[118]   P. Bajorski, "Stepwise simplex projection method for selection of endmembers in hyperspectral images," vol. 5806, 2005, pp. 318–329.

[119]   E. Kaltofen and G. Villard, "On the complexity of computing determinants," *Comput. Complex.*, vol. 13, no. 3-4, pp. 91–130, Feb. 2005.

[120] R. N. Clark, G. A. Swayze, K. E. Livo, R. F. Kokaly, S. J. Sutley, J. B. Dalton, R. R. McDougal, and C. A. Gent, "Imaging spectroscopy: Earth and planetary remote sensing with the USGS tetracorder and expert systems," *Journal of Geophysical Research: Planets*, vol. 108, no. E12, 2003.

[121] G. Swayze, R. N. Clark, F. Kruse, S. Sutley, and A. Gallagher, "Ground-truthing AVIRIS mineral mapping at Cuprite, Nevada," in *Summaries of the Third Annual JPL Airborne Geosciences Workshop*, vol. 1, pp. 47–49, 1992.

[122] Q. Du, N. Raksuntorn, N. H. Younan, and R. L. King, "End-member extraction for hyperspectral image analysis," *Applied. Optics*, vol. 47, no. 28, pp. F77–F84, Oct 2008.

[123] A. Plaza, G. Martin, J. Plaza, M. Zortea, and S. Sanchez, "Recent developments in endmember extraction and spectral unmixing," in *Optical Remote Sensing*, ser. Augmented Vision and Reality, S. Prasad, L. M. Bruce, and J. Chanussot, Eds. Springer Berlin Heidelberg, 2011, vol. 3, pp. 235–267.

[124] R. N. Clark and T. L. Roush, "Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications," *Journal of Geophysical Research: Solid Earth*, vol. 89, no. B7, pp. 6329–6340, 1984.

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

[126] C.-I. Chang and B. Ji, "Weighted abundance-constrained linear spectral mixture analysis," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 44, no. 2, pp. 378–388, Feb 2006.

[127] K. Arai and H. Chen, "Unmixing method for hyperspectral data based on sub-space method with learning process," *Advances in Space Research*, vol. 44, no. 4, pp. 517 – 523, 2009.

[128] C.-I. Chang, "Target signature-constrained mixed pixel classification for hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 40, no. 5, pp. 1065–1081, May 2002.

[129] T.-M. Tu, C.-H. Chen, and C.-I. Chang, "A posteriori least squares orthogonal subspace projection approach to desired signature extraction and detection," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 35, no. 1, pp. 127–139, Jan 1997.

[130] J. C. Harsanyi, "Detection and classification of subpixel spectral signatures in hyperspectral image sequences," Ph.D. dissertation, 1993.

[131] P. Demartines and J. Herault, "Curvilinear component analysis: a self-organizing neural network for nonlinear mapping of data sets," *IEEE Transactions on Neural Networks,* vol. 8, no. 1, pp. 148–154, Jan 1997.

[132] G. Licciardi and F. Del Frate, "A neural network approach for pixel unmixing in hyperspectral data," in *2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS),* June 2010, pp. 1–4.

[133] V. Karathanassi, D. Sykas, and K. Topouzelis, "Development of a network-based method for unmixing of hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 50, no. 3, pp. 839–849, March 2012.

[134] D. Roberts, P. Dennison, M. Gardner, Y. Hetzel, S. Ustin, and C. Lee, "Evaluation of the potential of hyperion for fire danger assessment by comparison to the airborne visible/infrared imaging spectrometer," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 41, no. 6, pp. 1297–1310, June 2003.

[135] P. E. Dennison and D. A. Roberts, "Endmember selection for multiple endmember spectral mixture analysis using endmember average {RMSE}," *Remote Sensing of Environment*, vol. 87, no. 2-3, pp. 123 – 135, 2003.

[136] P. E. Dennison, K. Q. Halligan, and D. A. Roberts, "A comparison of error metrics and constraints for multiple endmember spectral mixture analysis and spectral angle mapper," *Remote Sensing of Environment*, vol. 93, no. 3, pp. 359 – 367, 2004.

[137] N. Raksuntorn and Q. Du, "A new linear mixture model for hyperspectral image analysis," in *IEEE International Geoscience and Remote Sensing Symposium, (IGARSS),* July 2008, vol. 3, pp. 258–261.

Brief Biography

Charoula Andreou received the Dipl.Ing. degree in Rural and Surveying Engineering from National Technical University of Athens, Athens, Greece, in 2008. Since 2009, she is a Research Associate with the Laboratory of Remote Sensing, National Technical University of Athens, Athens, Greece, participating in several European and national research projects. She has been a Visiting Researcher with the Remote Sensing Technology Institute, German Aerospace Center (DLR), Munich, Germany. Her main research interests are focused on hyperspectral remote sensing, spectral unmixing, feature extraction, and target detection.