Contributions to hyperspectral image processing from Lattice Computing and Computational Intelligence

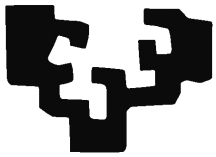
By

Miguel Angel Veganzones Bodón

http://www.ehu.es/ccwintco

Dissertation presented to the Department of Computer Science and Artificial Intelligence in partial fulfillment of the requirements for the degree of

Doctor of Philosophy



PhD Advisor:

Prof. Manuel Graña Romay

At

The University of the Basque Country
Donostia - San Sebastián
2012

Acknowledgments

To my huge family, which born long ago with my closest parents and now spreads out all over the world with friends and colleagues.

Special thanks to my advisor, Prof. Manuel Graña, for his support and patience.

Special thanks to Prof. Marine Campedel from the Institut Telécom ParisTech, for the confidence she put in me giving me the opportunity to participate of the CNES-DLR-TELECOM ParisTech Competence Centre on Information Extraction and Image Understanding for Earth Observation.

Special thanks to Prof. Mihai Datcu from the German Aerospace Center (DLR) to allow me be one more of his team, I really enjoyed it.

Miguel A. Veganzones

"Learn is fun. Learn is multiplayer."

Some gamified children.

Contributions to hyperspectral image processing from Lattice Computing and Computational Intelligence *by*

Miguel Angel Veganzones Bodón

Submitted to the Department of Computer Science and Artificial Intelligence on May 15th, 2012,
in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Abstract

The main material on which this Thesis works are remote sensing hyperspectral images. It is expected of the deployment of hyperspectral imaging sensors to improve our ability for remote recognition of material components in a scene. They already have applications in a variety of fields, ranging from geology, mineralogy to forestry sciences. However, these new kind of images pose new kinds of challenges. This Thesis has grown along two main lines of work. First, the idea of developing Content Based Image Retrieval (CBIR) systems for hyperspectral images. Second, the idea of applying Lattice Computing to the diverse computational tasks on hyperspectral images.

The CBIR system provides answers to image database queries computed on the basis of the contents of image, and not on the metadata associated to the image. The growing mass of remote sensing hyperspectral image data is the main motivation for this line of research. The definition of CBIR systems for hyperspectral images requires the definition of a image feature extraction process that benefits as much as possible from all the information in the image, the definition of a suitable similarity measure, which could be a distance in the feature space, allowing to search in the database for the answer to a given query; and, the definition of a search strategy, which can be direct or having a relevance feedback, including the definition of the query information. In CBIR systems, query by example is the typical option. we have defined three hyperspectral CBIR systems. The first two are defined using spectral and spectral-spatial features defined as the endmembers and fractional abundancies computed over the images. the third one follows a completely different philosophy, using the dictionaries that a lossless compression algorithm generates from the images. In all cases, specific similarity measures have been proposed. The Thesis includes some experiences on the use of a relevance feedback strategy based on dissimilarity spaces and a classifier trained on the images labeled by the user in previous iterations of the search algorithm.

Lattice Computing has been defined as the construction of algorithms on the ring defined by the real numbers, the lattice operators infimum and supremum

as the additive operator, and the real number addition as the multiplicative operator. This definition encompasses several research lines in computational intelligence, including morphological associative memories and some forms of fuzzy systems. In this thesis, Lattice Computing has been applied to the task of hyperspectral image analysis in the formulation of endmember induction algorithms and the formulation of Multivariant Mathematical Morphology for hyperspectral image segmentation. The research group previously defined some endmember induction algorithms following Lattice Computing approaches. This Thesis contributes a new algorithm combining the endmember extraction from Lattice AutoAssociative Memories (LAAM) with a Multi-Objective Genetic Algorithm (MOGA) for the optimal selection of endmembers from the pool of endmembers provided by the LAAMs. The definition of Mathematical Morphology on vector spaces is hindered by the difficulty of defining appropriate order relations building complete lattices. A relevant approach is the definition of such ordering on the basis of the classification of the image pixels using specific training data and appropriate classifiers whose output provides the required output. This Thesis builds such an ordering using the distance to the LAAM recall as the classification output, obtaining a Lattice Computing Multivariant Mathematical Morphology. This order is used for image segmentation, including the definition of morphological operators, erosion, dilation, gradient, and a subsequent watershed segmentation.

Contents

List of Algorithms

List of Figures

List of Tables

Chapter 1

Introduction

This introductory chapter is aimed to provide a quick overlook of the thesis that may allow the casual reader to assess its contents and main contributions. Its structure is as follows: Section 1.1 gives some general guidelines and motivation of the Thesis contents. Section 1.2 gives a general description of the Thesis contents. Section 1.3 summarizes the main contributions of the Thesis. Section 1.4 enumerates the publications obtained along the works. Section 1.5 describes the actual contents of the chapters in the Thesis.

1.1 Introduction and motivation

The main material on which this Thesis works are remote sensing hyperspectral images. It is expected of the deployment of hyperspectral imaging sensors to improve our ability for remote recognition of material components in a scene. They already have applications in a variety of fields, ranging from geology, mineralogy to forestry sciences. However, these new kind of images pose new kinds of challenges.

This Thesis has grown along two main lines of work. First, the idea of developing Content Based Image Retrieval (CBIR) systems for hyperspectral images. Second, the idea of applying Lattice Computing to the diverse computational tasks on hyperspectral images.

1.1.1 CBIR systems

The CBIR system provides answers to image database queries computed on the basis of the contents of image, and not on the metadata associated to the image. The growing mass of remote sensing hyperspectral image data is the main motivation for this line of research. The definition of CBIR systems for hyperspectral images requires:

- The definition of a image feature extraction process that benefits as much as possible from all the information in the image. Originally we were

interested in exploiting the spectral content of the image.

- The definition of a suitable similarity measure, which could be a distance in the feature space, allowing to search in the database for the answer to a given query.

- The definition of a search strategy, which can be direct or having a relevance feedback, including the definition of the query information. In CBIR systems, query by example is the typical option.

Feature extraction. This Thesis has followed several approaches in the definition of a feature extraction process. First, aiming to take into account the spectral information of the image as much as possible, we have defined the image endmembers as the image features. The feature extraction process then consists of the endmember induction algorithm. In a subsequent work, we include spatial information in the feature vector computed over the abundance images corresponding to the image endmembers. Abundance images are assumed as a dimensionality reduced version of the original image so that their correlation among hyperspectral images gives a good measure of the spatial similarity of the original images.

Following a completely different philosophy, a feature extraction process based on general image compression algorithms has been defined and tested. Specifically, the dictionaries that a lossless compression algorithm generates from the images for the actual compression computation have been used as image features. This approach is blind in the sense that all image information is used regardless of its spatial spectral meaning. As such it is the ultimate semantic-free feature extraction process.

Distances in feature spaces. This Thesis proposes some specific similarity measures on the feature spaces defined for hyperspectral images. For the spectral features a specific distance is proposed that combines the similarity among individual image endmembers. It is compared with other conventional similarity measures on synthetic and real hyperspectral image data. The extension of the feature space to spatial-spectral features lead us to consider a conventional earth-moving distance combining the aforementioned spectral distance with the spatial information. Finally, the dictionary based features use the dictionary intersection to measure their similarity.

Search strategy. In most of the Thesis works, the search strategy is straightforward: all the database images are compared to the query image, their similarities computed, and the images returned are the closest ones within the query span. The Thesis includes some experiences on the use of a relevance feedback strategy based on a classifier trained on the images labeled by the user in previous iterations of the search algorithm. This approach effectively introduces an adaptation of the topology of the feature space according to the user preferences.

1.1.2 Lattice Computing

Lattice Computing has been defined as the construction of algorithms on the ring defined by the real numbers, the lattice operators infimum and supremum as the additive operator, and the real number addition as the multiplicative operator. This definition encompasses several research lines in computational intelligence, including morphological associative memories and some forms of fuzzy systems. In this thesis, Lattice Computing has been applied to the task of hyperspectral image analysis in two specific forms:

- The formulation of endmember induction algorithms.
- The formulation of Multivariant Mathematical Morphology for hyperspectral image segmentation.

Endmember induction. Spectral unmixing provides sub-pixel resolution analysis results of the hyperspectral images. It needs the knowledge of the image endmembers, which can be provided a priori or induced from the image data by so called Endmember Induction Algorithms (EIA). The research group previously defined some EIAs following Lattice Computing approaches. This Thesis contributes a new algorithm combining the endmember extraction from Lattice AutoAssociative Memories (LAAM) with a Multi-Objective Genetic Algorithm (MOGA) for the optimal selection of endmembers from the pool of endmembers provided by the LAAMs.

Multivariant Mathematical Morphology. The definition of Mathematical Morphology on vector spaces is hindered by the difficulty of defining appropriate order relations building complete lattices. A relevant approach is the definition of such ordering on the basis of the classification of the image pixels using specific training data and appropriate classifiers whose output provides the required ordering. This Thesis builds such an ordering using the distance to the LAAM recall as the classification output, obtaining a Lattice Computing Multivariant Mathematical Morphology. This order is used for image segmentation, including the definition of morphological operators, erosion, dilation, gradient, and a subsequent watershed segmentation.

1.2 General description of the contents of the Thesis

Dimensionality reduction, spectral unmixing and endmember induction algorithms We provide a state of the art on dimensionality reduction (DR), spectral unmixing (SU) and endmember induction algorithms (EIAs). Together with machine learning techniques, these are the main computational techniques in hyperspectral image analysis. We show that hyperspectral image analysis can be often described as a chain of DR, EIA and SU techniques, and sometimes as an hybrid process.

We propose an endmember induction algorithm using Lattice Auto-Associative Memories (LAAMs) and Multi-Objective Genetic Algorithms (MOGA), called WM-MOGA. The WM-MOGA algorithm first produces a large set of candidate endmembers from LAAM constructed with the image data, applying the WM Algorithm proposed by Ritter at al. Then, we apply an specific MOGA minimizing the unmixing residual error and the number of endmembers. WM-MOGA last step selects an appropriate set of endmembers tailored to the data. To that effect, an Occam razor selection over the Pareto front obtained from the MOGA is applied. The WM-MOGA compares well to a recent state-of-the-art endmember induction heuristic in terms of the correlation of the induced abundance images with the given ground truth class spatial distribution. Furthermore, we propose an approximation to the MOGA which does not need to compute the linear unmixing based on the individual chromosomes at each generation. This faster process compares well with the reference heuristic regarding the identification of the ground truth classes. However, it overestimates the set of endmembers, including some redundant or irrelevant endmembers.

Hyperspectral CBIR systems The main body of the Thesis work addresses the definition and validation of an hyperspectral CBIR system. We give a general introduction to CBIR systems with some emphasis in Remote Sensing CBIR systems, including the performance measures and validation methodologies used to assess the goodness of the proposed hyperspectral CBIR systems.

The validation of remote sensing techniques is a hard problem by itself due to the scarcity of groundtruth data. Collecting remote sensing groundtruth is an expensive, tedious and error prone process which is the reason of its scarcity. This issue is worsened by the difficulties that users (even expert users) have to visually evaluate the data. To alleviate this problem, we describe a remote sensing CBIR validation methodology that does not require any a-priori knowledge about the data ground-truth. The proposed validation methodology is inspired in the DAMA strategy proposed by Baraldi et al. to address the same issue in remote sensing image data classification. As in the DAMA strategy, we make use of unsupervised clustering processes to discover data inherent structures that are later used in the validation process. In remote sensing CBIR, we propose to use the discovered inherent structures to simulate potential user queries.

The first approach to define an hyperspectral CBIR system described in the Thesis is focused on exploiting the rich spectral information of hyperspectral data. We define an hyperspectral CBIR system based on a spectral characterization of the hyperspectral images by means of the endmembers induced by some endmember induction algorithm. The Spectral CBIR system works by comparing the endmembers of two images by means of a spectral dissimilarity function. We proof that the proposed spectral dissimilarity function is a distance, giving exhaustive experimental results on synthetic and real hyperspectral datasets. The natural way to extend the Spectral CBIR system is to include spatial information from the hyperspectral data. The Spectral-Spatial CBIR system includes spatial information regarding the abundances of the end-

members induced from each image. In order to build the Spectral-Spatial CBIR system, we define a dissimilarity function on the basis of the spectral and spatial features of hyperspectral images. This Spectral-Spatial dissimilarity function s designed according to the IRM dissimilarity function. In this case the proposed Spectral-Spatial dissimilarity function is not a distance, as the IRM itself is not a distance either. Again, experimental results on synthetic and real hyperspectral datasets validate the use of the proposed Spectral-Spatial CBIR system.

Next, the Thesis explores the use of distance functions coming from the Information Theory field. The Normalized Information Distance (NID) is an universal distance formulated over the notion of Kolmogorov Complexity,, which, Unfortunately, is not computable in a Turing sense. The Normalized Compression Distance (NCD) is a computable approximation to NID defined over the compression factor calculated by some lossless compressor. We show that NCD allows the comparison of hyperspectral images by means of classification experiments. However, NCD is not feasible as a similarity measure for hyperspectral CBIR due to its unaffordable high computational cost. To overcome this problem, define a Dictionary CBIR system that makes use of a Dictionary distance that approximates NCD using dictionaries of words extracted from the hyperspectral images by dictionary-based lossless compressors. We give performance results of the proposed Dictionary CBIR over a real hyperspectral dataset, comparing the use of dictionary distances to NCD.

The final approach to designing CBIR system for hyperspectral images treated in this Thesis follows a Relevance Feedback (RF) methodology to enhance the different defined hyperspectral CBIR systems. Conventional RF methods are defined as a two-class (or one-class) classification process, often solved by an SVM's family classifier. However, the proposed hyperspectral CBIR system do not fit easily in a common classification scenario. The spectral or spectral-spatial features representing the hyperspectral images information cannot be represented as a point in a vector space, but as a set of points which cardinality depends on each image. Moreover, most of the defined dissimilarity functions do not comply with Mercer's conditions to be valid kernel functions. To overcome this issues, we propose a RF method that makes use of dissimilarity spaces to represent the user's queries. Dissimilarity spaces are defined as relative spaces, compared to the common machine learning's feature spaces, which are absolute spaces. A dissimilarity space spans over the axes defined by the dissimilarity of each sample data to some given prototypes. Each prototype's dissimilarity values define a dimension of the dissimilarity space. In the proposed RF method for hyperspectral CBIR systems, the prototypes are given by the images returned by the system as response to an user's query. The prototypes are evaluated by the user as relevant or irrelevant for the query and then, are compared among them building a prototype dissimilarity matrix with which a classifier is trained. Each image in the database is compared to the prototypes giving place to a dissimilarity feature vector which is the input to the trained classifier. A ranking over the database images is then defined using the classifier outputs. The RF method is validated showing the improvement over the Dictionary CBIR using real hyperspectral datasets.

Multivariant mathematical morphology Lattice Theory allows the most general formulation of Mathematical Morphology (MM). Any MM operator can be defined as a mapping between complete total-ordered lattices. We were interested in the application of MM to the analysis of hyperspectral data. However, hyperspectral data is a case of multivariate data and MM is not well defined for multivariate data. The reason is that multivariate data do not posses a natural total order. Existing techniques for multivariant MM make not use of all the spectral information or incur in some undesirable side-effects, like the false color problem of the component-wise ordering. Recently, a new way to define a multivariate ordering taking advantage of all the spectral information and without side effects has been proposed, the supervised orderings. A supervised ordering defines an order among multivariate data according to the distance to two training sets, named the foreground and background sets. We propose a supervised ordering using LAAMs. Actually, we provide three different supervised orderings based on LAAMs, a one-side LAAM supervised ordering and a relative and absolute background/foreground LAAM-supervised orderings. We also provide an unsupervised methodology based on endmember induction algorithms to automatically select the training sets from an hyperspectral image. We apply the proposed LAAM-supervised orderings to the watershed segmentation of hyperspectral data by means of the morphological Beucher gradient. We validate the LAAM-based morphological segmentations by comparing them to the segmentation obtained by usual component-wise MM on the contextual classification of real hyperspectral images.

1.3 Contributions

The contributions of the Thesis from a methodological point of view are the following ones:

1. Provides a review of the state of the art in two research areas: endmember induction algorithms and content based image retrieval.

2. The experimental methodology employed to obtain the experimental results is exhaustive and free of circular analysis. In the case of CBIR experiments, the recall performance measures were computed on independent features obtained for each image. In the case of classification experiments, all care was taken to ensure that the test data was completely independent of the train data.

3. Provides a methodology for the generation of experimental synthetic images. A free implementation is available.

4. Provides a validation methodology in the absence of ground truth.

5. Provides formal proofs of some assertions, such as the nature of the dissimilarity distances.

6. Results on real hyperspectral images are provided, some on well known benchmark images allowing for qualitative and quantitative comparison with the literature.

From a computational point of view the contributions of the Thesis are the following ones:

1. An innovative endmember induction algorithm (WM-MOGA) based on Lattice Computing.

2. The extension of the CBIR paradigm to hyperspectral images.

3. The definition of a distance on hyperspectral images based on its spectral characterization by a set of induced endmembers.

4. The definition of a distance on hyperspectral images based on its spectral and spatial characterization by a set of induced endmembers and its abundances.

5. The definition of dictionary based distances between hyperspectral images.

6. The formulation of a relevance feedback strategy based on dissimilarity spaces for hyperspectral images.

7. The definition of a Multivariate Mathematical Morphology based on the recall of LAAMs, including the definition of a morphological gradient and a watershed segmentation algorithm.

1.4 Publications obtained

1. M.A. Veganzones and M. Graña, «A Spectral/Spatial CBIR System for Hyperspectral Images», Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of, (online pre-publication), 2012.

2. M. Graña and M.A. Veganzones, «An endmember-based distance for content based hyperspectral image retrieval», Pattern Recognition, (online pre-publication), 2012.

3. M. Graña and M.A. Veganzones, «Endmember induction by lattice associative memories and multi-objective genetic algorithms», EURASIP Journal on Advances in Signal Processing, (online pre-publication), mar. 2012.

4. M.A. Veganzones and M. Graña, «On hyperspectral morphology by lattice auto-associative memories supervised orderings», in Proceedings of the 16th international conference on Knowledge-Based Intelligent Information and Engineering Systems, San Sebastián, Spain, 2012.

5. M.A. Veganzones and M. Graña, «Hybrid Computational Methods for Hyperspectral Image Analysis», in Hybrid Artificial Intelligent Systems, vol. 7209, E. Corchado, V. Snášel, A. Abraham, M. Wozniak, M. Graña, y S.-B. Cho, Eds. Springer Berlin / Heidelberg, 2012, pp. 424–435.

6. M.A. Veganzones and M. Graña, «Relevance feedback by dissimilarity spaces for hyperspectral CBIR», in ESA-EUSC-JRC 2012 Conference Proceedings, Oberpfaffenhofen, Germany, 2012.

7. M.A. Veganzones and M. Graña, «Dictionary based hyperspectral image retrieval», in Proceedings of the 1st International Conference on Pattern Recognition Applications and Methods, Vilamoura, Portugal, 2012.

8. M.A. Veganzones and M. Graña, «Lattice Auto-Associative Memories induced Supervised Ordering defining a Multivariate Morphology on Hyperspectral data», in Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2012 4rd Workshop on, Shangai, China, 2012.

9. M. A. Veganzones and M. Graña, «On the validation of a spectral/spatial CBIR system for hyperspectral images», in Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2011 3rd Workshop on, Lisbon, Portugal, 2011, pp. 1 –4.

10. M.A. Veganzones and M. Graña, «Validation of a Hyperspectral Content-Based Information Retrieval (RS-CBIR) System Upon Scarce Data», in Soft Computing Models in Industrial and Environmental Applications, 6th International Conference SOCO 2011, vol. 87, E. Corchado, V. Snášel, J. Sedano, A. Hassanien, J. Calvo, y D. Sle˛zak, Eds. Springer Berlin / Heidelberg, 2011, pp. 47–56.

11. M.A. Veganzones and M. Graña, «Validation of remote sensing content-based information retrieval (RS-CBIR) systems upon scarce data», in ESA-EUSC-JRC 2011 Conference Proceedings, Ispra, Italy, 2011.

12. A. Roman-González, M.A. Veganzones, M. Graña and M. Datcu, «A novel data compression technique for remote sensing data mining», in ESA-EUSC-JRC 2011 Conference Proceedings, Ispra, Italy, 2011.

13. M.A. Veganzones and C. Hernández, «On the use of a hybrid approach to contrast endmember induction algorithms», in Proceedings of the 5th international conference on Hybrid Artificial Intelligence Systems - Volume Part II, Berlin, Heidelberg, 2010, pp. 69–76.

14. M.A. Veganzones and M. Graña, «Endmember Extraction Methods: A Short Review», in Proceedings of the 12th international conference on Knowledge-Based Intelligent Information and Engineering Systems, Part III, Berlin, Heidelberg, 2008, pp. 400–407.

15. M.A. Veganzones, J. Maldonado and M. Graña, «On Content-Based Image Retrieval Systems for Hyperspectral Remote Sensing Images», in Computational Intelligence for Remote Sensing, vol. 133, M. Graña y R. Duro, Eds. Springer Berlin / Heidelberg, 2008, pp. 125–144.

16. J.O. Maldonado, D. Vicente, M. Graña, y M.A. Veganzones, «Spectral indexing for Hyperspectral image CBIR», in ESA-EUSC 2006 Conference Proceedings, Torrejon air base, Madrid (Spain), 2006.

1.5 Structure of the Thesis

The thesis starts with the contributions related to endmember induction, including a new approach based on Lattice Computing. The bulk of the Thesis contents are referred to Content-Based Image Retrieval (CBIR) systems for hyperspectral datasets, where some new proposals are discussed along with some methodological questions. Finally, the Thesis last part is devoted to Mathematical Morphology over multivariate data with application to hyperspectral segmentation. Appendices contain the detailed description of additional materials. The chapters that contain some specific contribution, have a specific conclusions section. Therefore, the concluding chapter is strictly devoted to enunciate ideas for future research.

The contents of the chapters is as follows:

- Chapter 1 contains the general description of the works done, its motivation, the specification of the Thesis objectives, contributions and publications obtained as a result of the works achieved.

- Chapter 2 introduces the linear mixing model of image generation and the spectral unmixing analysis. The chapter provides a review of the state-of-the-art of related techniques such as dimensionality reduction methods and endmember induction algorithms.

- Chapter 3 contains the proposal of an endmember induction algorithm using Lattice Auto-Associative Memories (LAAMs) and Multi-Objective Genetic Algorithms (MOGA), called WM-MOGA, which is an improvement over Ritter's WM endmember induction algorithm.

- Chapter 4 is a general introduction to CBIR systems with some emphasis in Remote Sensing CBIR systems. This chapter discusses the performance measures and validation methodologies used to assess the goodness of CBIR systems for hyperspectral images. A validation methodology that does not require any *a-priori* knowledge about the data ground-truth is presented.

- Chapter 5 contains the definition of a hyperspectral CBIR system based on a spectral characterization of the hyperspectral images by means of the induced endmembers. We propose a spectral dissimilarity function

allowing to compare the endmember characterization of two images. We proof that the proposed spectral dissimilarity function is a distance, giving exhaustive experimental results on synthetic and real hyperspectral datasets.

- Chapter 6 contains the definition of a Spectral-Spatial CBIR system which extends the spectral characterization of hyperspectral images by including spatial information related to the induced endmembers. In order to build it, we define a dissimilarity function that compares the spectral and spatial features of two images. We check that this Spectral-Spatial dissimilarity function is not a distance. Experimental results on synthetic and real hyperspectral datasets validate the use of the proposed Spectral-Spatial CBIR system.

- Chapter 7 contains the definition of a Dictionary based CBIR system for hyperspectral datasets, exploring the use of distance functions between hyperspectral images grounded in Information Theory using Kolmogorov Complexity and compression factors. The Normalized Compression Distance (NCD) and several dictionary distances are studied by means of computational experiments for hyperspectral CBIR. We give performance results over a real hyperspectral dataset, comparing the use of dictionary distances to NCD.

- Chapter 8 contains the definition of a Relevance Feedback (RF) methodology to enhance the hyperspectral CBIR systems, making use of dissimilarity spaces to represent the user's query, which are computed from the prototypes labeled by the user as relevant or irrelevant by means of a supervised classifier. Then, a ranking of the images is defined by the response of the trained classifier to the database dissimilarity vectors given as inputs. We validate the proposed RF method on a real hyperspectral dataset.

- Chapter 9 contains a proposal for multivariate Mathematical Morphology (MM) using LAAMs to define a supervised ordering using of a background and foreground training set. We provide three supervised orderings based on LAAMs, and an unsupervised methodology to select the training sets from an hyperspectral image. The chapter contains the definition of a watershed algorithm based on the morphological gradient computed according to the LAAMs based MM. The chapter contains validation results on classification of real hyperspectral images.

- Chapter 10 contains a description of future research lines.

- Appendix A provides a theoretical background on Lattice Theory and Lattice Computing by means of LAAMs.

- Appendix B provides a brief review of the endmember induction algorithms used on the experiments.

- Appendix C explains the methodology used to build synthetic hyperspectral images and gives details on the toolbox implemented to that effect.
- Appendix D contains information about the real hyperspectral scenes used on the experiments.

Chapter 2

Dimensionality reduction, spectral unmixing and endmember induction: a state of the art

This chapter provides an state of the art on basic techniques, setting the grounds for some of our contributions. The structure of the chapter is as follows: Section 2.1 gives introductory definitions and remarks. Section 2.2 gives an state of the art on dimensionality reduction. Section 2.3 defines the spectral unmixing process and current trends. Section 2.4 reviews endmember induction algorithms. There is no conclusions section in this chapter.

2.1 Introduction

An hyperspectral image analysis process can often be described as a chain of different computational techniques. Figure 2.1 shows a typical flow chart of an hyperspectral image analysis process. Before the actual hyperspectral image analysis, a pre-processing step is usually required, including hyperspectral sensor calibration and image acquisition, radiometric correction, data geo-registration and atmospheric correction if the sensor is mounted in a platform for Earth Observation. The acquired data must be indexed and stored together with metadata containing information about the sensed scene. The analysis of the hyperspectral image can be performed over the images resulting from any of the pre-processing levels.

Dimensionality Reduction (DR) is in many cases the first step in the analysis because of the high dimensionality of hyperspectral data. DR tries to find a low-dimensional representation of the hyperspectral data which is optimal in some sense, improving either the time requirements for the data or the subsequent

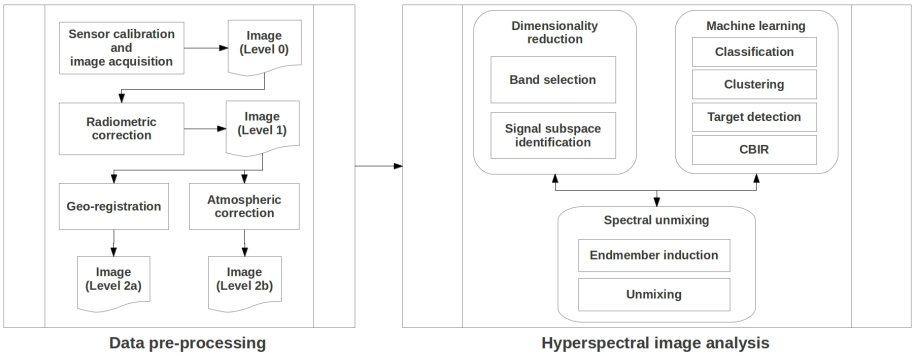


Figure 2.1: Flow Diagram of a conventional hyperspectral image analysis.

analysis results. *Spectral Unmixing* (SU) [98] are a collection of subpixel resolution analysis techniques of hyperspectral data. In SU the hyperspectral image is assumed to be a (non-)linear mixture of the spectral signatures of the materials in the image (endmembers), according to their fractional spatial distributions (abundances) plus additive noise. Given the endmembers presented in an hyperspectral image, the unmixing process looks for the fractional abundances of such endmembers at each pixel in the image. The abundances obtained by the SU can be the goal of the hyperspectral analysis or an intermediate step to further subpixel analysis. The endmembers of the materials in the image are seldom known *a-priori*. Thus, SU often requires the estimation of the endmembers in the image by means of manual or automatic methods to make the unmixing process possible. *Machine Learning* (ML) techniques, such as classification, clustering, target detection and CBIR, can be applied directly to the hyperspectral data or to the results of DR and/or SU methods. Sometimes, DR, SU and ML methods are combined not in a chain process but hybridized. For instance, a dimensionality reduction technique can be modeled as function of a classifier performance; or, an spectral unmixing technique can iteratively perform an endmember induction and a (non-)linear unmixing so the later obtains abundance images given the induced endmembers, and the former induce the endmembers depending on the error obtained in the unmixing.

The problem of the endmember induction algorithms initialization is discussed in [138], proposing an Endmember Initialization Algorithm (EIA). The number of endmember spectral signatures that form an hyperspectral image is usually unknown. Recently, a new concept denoted Virtual Dimensionality (VD) [29, 138, 30] has been used for automated search of the optimal number of endmembers in an image. [39, 60] are novel techniques to calculate the VD of an hyperspectral image, although VD has been questioned as a good interpretation for the estimate of the number of materials in the image [8]. Most endmember induction algorithms are quite computationally expensive, so a mention is due to the efforts to obtain distributed implementations [142] that may help them to be a feasible approach for real-life applications. The use of off-the-shelf

Graphical Processing Units (GPU) [165] are a low cost way to obtain substantial speed-ups.

2.2 Dimensionality reduction

Hyperspectral imaging measures a large number of spectral bands, usually hundreds or even thousands. Dimensionality Reduction (DR) is often the first step in hyperspectral analysis. DR allows to have a better and more compact representation of the data, alleviating the computational cost and the memory allocation requirements of posterior analysis methods. Hyperspectral high-dimensional data incorporate redundancies due to the high correlation between bands. Thus, it can be assumed that the discriminant spectral information lies in a lower dimensional subspace. *Signal Subspace Identification* (SSI) [3] is the appropriate way to perform dimensionality reduction on hyperspectral data. However, finding the correct subspace dimensionality where the hyperspectral data lies is a very challenging problem.

SSI can be performed in two ways: space transformation and band selection. The former SSI consists of some analytical transformation of the high-dimensional data space onto a different space estimated according to some criteria, which can be statistical or other. The dimensionality reduction is achieved by selecting the first N_D axes and discarding the remaining ones. Axes are ordered according to some value assigned to the dimensions related to the criteria used for the transformation estimation. The later SSI consists in a selection of a reduced number of bands, N_B, from the original set of bands. The selected spectral bands must contain most of the relevant information according to some criteria. In both cases, the identified subspace has a dimensionality, N_D or N_B, lower than the dimensionality of the original spectral space.

2.2.1 Space transformation SSI

The most popular SSI by space transformation method is Principal Component Analysis (PCA). PCA calculates the eigenvectors and eigenvalues of the sample data covariance matrix. DR is achieved by selecting the eigenvectors with largest eigenvalues, which is equivalent to maximize the variance of the projected data . Maximum Noise Fractions (MNF) and Singular Value Decomposition (SVD) are techniques similar to PCA, that also estimate second-order statistics. MNF maximizes Signal-to-Noise Ratio (SNR) and, SVD obtains a transformation with maximum power. Application of PCA, MNF and SVD to hyperspectral data has been criticized in [143, 9]. Hyperspectral images contain many subtle materials with subpixel sizes that are missed by second-order statistics. Recent techniques [103, 2, 3] try to address this problem by preserving rare materials dividing the subspace in two, the signal subspace and the rare vector subspace.

Independent Component Analysis (ICA) has been proposed for hyperspectral DR in [189]. ICA looks for statistical independent sources using high-order statistics or mutual information-based criteria. However, statistical indepen-

dence assumption in hyperspectral data has been criticized in [130]. Progressive dimensionality reduction by transform (PRDT) [33] performs DR in terms of progressive information preservation. Each spectral transformed component is ranked respect to its information preservation capability. Authors in [33] propose two procedures to perform DR. One starts by a reduced number of spectral transformed components and expands it progressively. The other starts by a large number of components and progressively removes them.

Some recent papers introduce manifold and tensor based techniques in the Remote Sensing field. Works in [6, 7] propose improvements to the ISOMAP method [180] to find the non-linear structure of hyperspectral data, consisting in a manifold coordinate system that preserves geodesic distances in the high-dimensional hyperspectral data-space. Additionally, [146] proposes the use of tensors to jointly process spatial and spectral information for denoising and DR.

Other recent techniques make use of *a-priori* information about the data, in the form of partial labeling, to perform SSI. In [40] SSI is achieved by looking for the subspace that minimizes two terms: a discriminative term that assesses the pairwise class separability of the labeled data samples and, a regularization term that characterizes some property of the original data space. The paper proposes two approaches for the regularization term, maximizing data global variance and minimizing reconstruction errors. In [201], authors pursue both DR and classification in a row by pruning a neural network input layer. [127] proposes to project labeled samples onto a new 'prototype space' defined by the bands, and then, cluster projected data by grouping bands containing similar information in order to perform DR. [89] proposes improvements to two extensions of Fisher's linear discriminant analysis (LDA), the non-parametric discriminant analysis and the non-parametric weighted feature extraction, to reduce the dimensionality and increase the classification accuracy. The use of generalized relevance learning vector quantization (GRLVQ) to extract those features interesting to classification purposes is proposed in [124].

2.2.2 Band selection SSI

The conventional feature selection sub-optimal search strategies, i.e. sequential forward selection, steepest ascent and fast constrained search, present the same problems as subspace identification algorithms. The work in [162] models the hyperspectral DR as a band selection problem where subsets of the original bands are selected and averaged to form a spectral band, proposing modifications of the common feature selection strategies to find the optimal solution minimizing the probability of classification error in the spectral space.

Some methods take advantage of labeled data samples, such as [34] making use of a well-know algorithm for target detection, the constrained energy minimization (CEM) algorithm, to perform linear constraint on an isolated band image while minimizing band correlation or dependence provided by other band images. In [23] authors evaluate labeled sample data statistical dependence by the so-called Hilbert-Schmidt independence criterion (HSIC). The labeled samples are used to calculate the Hilbert-Schmidt norm of the cross-covariance

kernel operator. Then, the proposed method looks for those bands that minimize the associated HSIC p-value.

Other recent works make use of patterns of known spectral signatures that are of interest for the posterior analysis. For instance, [133] views the hyperspectral sensing process as a projection of the scene space, the scene of all spectral patterns of interest, onto the spectral bands, called the sensor space. Authors provide a method based on the canonical correlation feature selection (CCFS) algorithm to find optimal superpositions of the spectral bands representing the most informative directions in the sensor space for specific patterns in presence of noise. [191] uses the spectral band-to-band correlation within a single spectral signature and proposes a method based on orthogonal subspace projections (OSP) to select a variable number of different bands for each of the spectral signatures of interest.

The work in [199] follows a different approach, modifying the sparsity promoting iterated constrained endmember (SPICE) algorithm to perform band selection, endmember induction and spectral unmixing simultaneously. Band selection is achieved by incorporating band weights and a band sparsity promoting term to the SPICE objective function.

2.3 Spectral unmixing

In the field of hyperspectral image processing, Spectral Unmixing [98] is the computation of the fractional contribution of elementary spectra, called endmembers because they constitute the vertexes of a convex polytope covering (most of) the image data points in high dimensional space. The underlying image model is usually a linear mixture of the endmembers, with positive coefficients that sum up to one. Let $E = [\mathbf{e}_1, \ldots, \mathbf{e}_p]$ be a set of spectral signatures where each $\mathbf{e}_i \in \mathbb{R}^q$ is a q-dimensional vector. Then, the hyperspectral signature $\mathbf{h}(x, y)$ at the pixel with coordinates (x, y) is defined by the expression:

$$\mathbf{h}(x, y) = \sum_{i=1}^{p} \mathbf{e}_i a_i(x, y) + \mathbf{n} \tag{2.1}$$

where $\mathbf{a}(x, y)$ is the q-dimensional vector of fractional abundances at pixel (x, y) and $\mathbf{n}$ is an independent additive noise component. There are two common constraints to equation 2.1: the abundance non-negative constraint (ANC) and the abundance sum-to-one constraint (ASC), respectively defined as $a_i(x, y) \geq 0$, for all $1 \leq i \leq p$, and $\sum_{i=1}^{p} a_i = 1$. The unmixing process looks for the fractional abundances of such endmembers. In order to perform spectral unmixing, the spectral signatures (endmembers) of the materials in the image must be known. This is not the common scenario and often some method must be used to select endmembers from an spectral library or automatically induce them from the image itself. Algorithms estimating the endmembers from the image itself are named endmember induction algorithms (EIAs).

2.3.1 Linear and non-linear spectral unmixing

Given the image endmembers, enforcing the ANC and ASC conditions on Spectral Unmixing involves performing constrained non-negative least squares estimation, the so-called full-constrained linear spectral unmixing (FCLSU), can be a very computationally expensive process by itself. When the convex polytope defined by the provided endmembers does not cover all the data points, it is not possible to enforce these conditions. Least squares and orthogonal subspace projections are the widely used linear unmixing methods because their computational requirements are lesser than FCLSU.

Sometimes, the unmixing problem is relaxed by dropping one of the constraints or both, called partially-constrained LSU (PCLSU) and unconstrained LSU (ULSU) respectively. These methods have some limitations derived from the linearity of the assumed model and deficiencies in the selection of endmembers. The work in [14] proposes a new approach called spectral-angle measure-based spectral unmixing (SAMSU) that uses spectral angle measures to reduce the spectral unmixing error due to spectral within-class confusion derived from the variability on the spectral signatures amplitude. Same variability problem is addressed in [96] by a method based on the Fisher's discriminant null space (FDNS). [24] proposes dividing the hyperspectral scene in small tiles, inducing a set of endmembers for each local tile. Then, all locally induced endmembers are clustered together to find groups before unmixing the image globally. [61] adopts a Bayesian formulation to exploit spatial correlations between pixels. In [86] authors propose an analytical solution to FCLSU by projecting image pixels onto the simplex formed by the endmembers and [167] proposes the use of fuzzy membership functions that are equivalent to the least square solution of the FCLSU problem.

Non-linear spectral unmixing (NLSU) has attracted increasing attention in the last years. Kernelization of LSU methods [113] allows to estimate non-linear abundances. The generalized bilinear model (GBM) [62] is a model for non-linear unmixing of hyperspectral data due to multipath effects. [81] studies a Bayesian algorithm to estimate the abundance coefficients and the noise variance of the GBM. [145] presents a simple but effective process to non-linear unmixing, where the multiplication of each pair of endmembers results in a virtual endmember representing multiple scattering effects during pixel construction process. Virtual endmembers resulting of non-linear relationships between actual endmembers can yield poor unmixing results as it is shown in [41]. [97] proposes a non-linear unmixing method based on relative distances through networks. The assumption is that for a pixel, the abundance fraction of an individual endmember is inversely proportional to its distance to the endmember and proportional to the sum of distances of the other endmembers. In [136] labeled data samples are used to estimate the non-linear abundances of given endmembers through Gaussian synapse artificial neural networks. [87] presents an algorithm capable of extracting endmembers and determining their abundances in hyperspectral imagery under nonlinear mixing assumptions. The algorithm is based upon simplex volume maximization, and uses shortest-path

distances in a nearest-neighbor graph in spectral space, hereby respecting the nontrivial geometry of the data manifold in the case of nonlinearly mixed pixels.

2.4 Endmember induction algorithms

Current approaches try to induce the endmembers from the image data. Either, trying to select some image pixel spectra as the best approximation to the endmembers in the image [139, 69, 73], or computing estimations of the endmembers on the basis of the transformations of the image data (i.e. [93, 31]). The latter is the predominant class of techniques in the literature. The review given in [140] makes some emphasis on the degree of automation to classify the algorithms. In [184], we make emphasis on the computational foundations, assuming that user interaction must be minimal or null. We distinguish three fundamental approaches:

- Geometric approaches, that try to find a simplex that covers the image data.
- Lattice computing approaches, that use some kind of lattice theoretic formalism or mathematical morphology approach.
- Heuristic approaches, that are not very rigorously formalized under a theoretical framework.

2.4.1 Geometric endmember induction methods

Geometric methods follow the formal definition of the endmembers, they search for the vertexes of a convex set that covers the image data. Because the distribution of the data in the hyperspace is usually tear-shaped, they look for the minimum simplex that covers all the data. Unless said otherwise, the algorithms search for a prefixed number of endmembers, defined by the user.

The first such methods is the Minimum Volume Transform, proposed by [44], that introduces two non-orthonormal transforms, the dark-point-fixed (DPF) transform and the fixed-point-free (FPF) transform that map the data onto the minimal simplex that contain all the data points. One of the earliest approaches is the N-FINDR algorithm proposed in [194]. The N-FINDR algorithm is a selection algorithm. It starts with a random collection of image pixel spectra, corresponding to the initial set of endmembers. Then, each of the remaining image pixels is considered as a candidate to replace each endmember. If doing so the volume of the simplex increases, then it is accepted as the new endmember. The process ends when no more replacements are possible. The N-FINDR algorithm requires a dimension reduction step, originally an orthogonal subspace projection (OSP) to an space of dimension N-1, where N is the number of endmembers. The set of endmembers found by the N-FINDR would not allow the nonnegative unmixing of the pixel spectra in general.

The Convex Cone Analysis (CCA) [93] is based on the fact that the vectors formed by discrete radiance spectra are linear combinations of nonnegative components, and they lie inside a nonnegative convex region. The object of CCA is to find the boundaries of this convex region, which can be used as endmember spectra. The algorithm performs a Principal Component Analysis (PCA) dimension reduction based on the sample spectral correlation matrix of the image. In this reduced space, the endmembers must define a convex cone on the positive hyperquadrant of the space, whose apex is in the space origin. Endmembers are points with exactly $c-1$ zero coefficients in the PCA decomposition, c being the number of eigenvectors selected.

The approach followed in [13] searches for the optimal simplex using a simulated annealing algorithm (SA) whose state configuration is given by the partition of the faces of the convex hull of the image pixel spectra. Before, a reduction to N-1 dimensions by the Minimum Noise Fraction (MNF) algorithm is performed. The partition in the configuration space defines a simplex covering the image data whose vertexes are the candidate endmembers. The minimized objective function is the simplex volume. This approach is followed by the generation of endmember bundles that allow the computation of bounds on the abundance images.

The Iterated Constrained Endmembers (ICE) [16] algorithm performs the minimization of a regularized residual sum of squares (RSS). The regularization term is the volume of the simplex. The name of the algorithm comes from the applied minimization schema. Given that the free parameters are both, the endmembers and the proportions (abundances) for each pixel, the algorithm iterates the solution of the tow interleaved and interdependent minimization problems (much like in an Expectation Maximization process): first the proportions are computed by quadratic programming problem solving assuming that the endmembers are known, then the endmembers are computed as the direct minimization of the RSS functional. The addition of an sparsity promoting term in the RSS functional gives way to SPICE algorithm [198]. This sparsity promoting term is derived as the substitution of a Gaussian prior by a Laplacian prior in a Bayesian formulation of the RSS functional. The SPICE algorithm allows the selection of the appropriate number of endmembers based on the sparsity measure. Both, ICE and SPICE algorithms do need a dimension reduction step, performed by the MNF algorithm.

The Vertex Component Analysis algorithm (VCA) is presented in [131]. The algorithm is unsupervised and exploits that the affine transformation of a simplex is also a simplex. It works with projected and unprojected data. The algorithm iteratively projects data onto a direction orthogonal to the subspace spanned by the endmembers already determined. The new endmember signature corresponds to the extreme of the projection. The algorithm iterates until all endmembers are exhausted.

In [36] a simplex-based endmember extraction algorithm, called Simplex Growing Algorithm (SGA), is presented. It is a sequential algorithm that finds a simplex with maximum volume by every time a new vertex is added. Virtual Dimensionality (VD) is applied as stopping rule to determine the number of

vertexes required. SGA improves N-FINDR by including a process of growing simplexes one vertex at a time until the desired number of vertexes is reached, which results in a high computational complexity reduction; and by selecting an appropriate initial vector to avoid the use of random vectors as initial condition, which produces different sets of final endmembers if different sets of randomly generated initial endmembers are used.

In [126], a method for endmember extraction for highly mixed data, when there are not pure pixels in the hyperspectral image, is presented. The proposed method, called Minimum Volume Constrained Nonnegative Matrix Factorization (MVC-NMF) takes advantage of the fast convergence of NMF schemes and, at the same time, eliminates the pure-pixel assumption. It consists in the reformulation of an NMF cost function introducing an volume regularization term, much like the ICE, substituting the RSS by the NMF criteria.

Recent works present new geometrical simplex-based methods or improvements to the existing ones: [28] proposes a minimum volume enclosing simplex (MVES) algorithm for highly mixed data and [4] extends it by a robust MVES (RMVES) that deals with uniform/nonuniform additive Gaussian noise, [37] proposes a simplex growing algorithm (SGA) that works in real time and [112] a SGA based on the Householder transformation, [64] employs support vector machines to cluster the data and build simplexes enclosing each cluster, [123] proposes a simplex-based algorithm that improves endmember induction in the presence of anomalous materials, [178, 195] are improved versions of the N-FINDR algorithm [194].

2.4.2 Lattice computing endmember induction methods

Lattice computing [75, 155, 72, 68, 153] is an alternative to geometrical simplex formulation, where a connection between linear mixing model algebraic properties and lattice independence is established.

Lattice computing can be defined as the collection of computational methods that either are defined on the algebra of lattice operators inf and sup, with the addition, or employs lattice theory to generalize previous approaches. Mathematical Morphology is a very successful case of this paradigm, but it also encompasses some fuzzy systems approaches and neural networks. The Automated Morphological Endmember Extraction (AMEE) method [139] is a mathematical morphology inspired algorithm for the extraction of the endmembers from the data. It is based on the definition of multispectral erosion and dilation operators, which are then used to compute, for all the pixels in the image, the Morphological Eccentricity Index (MEI) over kernels of increasing size. The result is a MEI image whose maxima correspond to the endmember pixels. The method does not need a dimension reduction step.

The concept of morphological independence, later reformulated as lattice independence, was the basic tool in the approach proposed in [69, 73, 71, 70]. The set of endmembers was formulated a set of morphologically independent vectors, either in a dilative or erosive sense, or both. There the Associative Morphological Memories, later renamed Lattice Associative Memories, are proposed as

detectors of morphologically independent vectors. The algorithm works in a single pass over the sample data.

This approach has been followed by the one proposed in [155, 153]. The relationship between strong lattice independence and affine independence was proven. Then it was found that most vectors in the erosive and dilative lattice memories are strong lattice independent. Therefore, the mere construction of the lattice memories provide a way to obtain the convex hull of the data. Provided an endmember selection mechanism, the algorithm can obtain in a single pass over the image a set of endmembers.

2.4.3 Heuristic endmember extraction methods

The heuristic methods collects a set of heterogeneous endmember extraction methods that uses different approaches not grouped under a strict theoretical background for endmember induction. The most famous and widely used method, due to its inclusion in the ENVI software package, it is the Pixel Purity Index (PPI) algorithm introduced in [94]. The algorithm reduces the data dimensionality and makes a noise whitened process by MNF method. Then, it determines the pixel purity by repeatedly projecting data onto random unit vectors. The extreme pixel in each projection is counted, identifying the purest pixels in scene. PPI requires the human intervention to select those extreme pixels that best satisfy the target spectrum.

Although PPI has been intensively used, its implementation aspects are kept unknown due to the limited published results. In [31], PPI is investigated and a fast iterative algorithm to implement PPI is proposed. The Fast Iterative PPI algorithm (FIPPI) improves PPI in several aspects. FIPPI produces an appropriate initial set of endmembers to speed up the process. Additionally, it estimates the number of endmembers to be generated by Virtual Dimensionality (VD). FIPPI is also an unsupervised and iterative algorithm, where an iterative rule is developed to improve each of the iterations until it reaches a final set of endmembers.

In [188], the well known Independent Component Analysis (ICA) method is the base of the proposed approach for endmember extraction and abundance quantification. The algorithm, called ICA-based Abundance Quantification Algorithm (ICA-AQA), is a high-order statistics-based technique, that can accomplish endmember extraction and abundance quantification simultaneously in one-shot operation. [130] analyzes the use of ICA and Independent Factor Analysis (IFA) for unmixing tasks, showing that the statistically independent of the sources, assumed by ICA and IFA, is violated in the hyperspectral unmixing, compromising the performance of ICA/IFA algorithms for this purpose. It concludes that the accuracy of this ICA/IFA-based methods tends to improve with the increase of the signature variability and the signal-to-noise ratio.

The Spatial-Spectral Endmember Extraction algorithm (SSEE) proposed in [156] is another projection based method that works by analyzing a scene in parts (subsets), such that it increases the spectral contrast of low contrast endmembers, thus improving the potential for these endmembers to be selected.

The SSEE method uses a singular value decomposition (SVD) to determine a set of basis vectors that describe most of the spectral variance for subsets of the image. Then the full image dataset is projected onto the locally defined basis vectors to determine a set of candidate endmember pixels from where the final endmembers are selected. For that, it searches for spectrally similar but spatially independent endmembers. This is realized by imposing spatial constraints for averaging spectrally similar endmembers.

Non-negative matrix factorization (NMF) is another alternative technique that have attracted recently a lot of attention [95, 91, 196, 197] and that exploits the positive matrix representation of hyperspectral linear mixing model. Spatial information is exploited in [206, 122, 121] to improve endmember induction.

2.5 Hybrid approaches

Some recent methods follow a hybrid approach where endmember induction and spectral unmixing are performed simultaneously. For instance, [54, 53, 59] propose a Bayesian formulation to solve endmember induction and spectral unmixing at the same time . Other works [35, 38] combine well-know EIAs, the Pixel Purity Index (PPI) and the N-FINDR algorithms, with linear unmixing to optimize the number of induced endmembers. Similar approaches [147, 74] propose genetic algorithms to select the optimal number of the set of induced endmembers. Others [55] deal with the induction of endmembers preserving rare materials. In [22], the spectral unmixing is modeled as a dependent component analysis, proposing a method based on maximum non-Gaussianity and Parzen windows technique to separate the dependent sources and estimate their fractional abundances. Another approach, [114, 115] combines a Bayesian self-organizing map (BSOM), a supervised method to induce the endmembers, with a Gaussian mixture model (GMM) in the former and a fuzzy membership (FM) in the later to perform the unmixing. Additionally, [203, 202] combine endmember induction and linear unmixing into a combinatorial optimization problem solved by either particle swarm optimization (PSO) or ant colony optimization (ACO), where unmixing error is the objective function and the particles/ants represent different sets of pixels for endmember determination. Finally, [200] allows spectral unmixing to be performed directly on compressed data without any need to reconstruct hyperspectral imagery prior to analysis.

Chapter 3

Endmember induction by lattice associative memories and multi-objective genetic algorithms

In the Linear Mixing Model (LMM), endmembers are defined as the vertexes of a convex polytope covering the data. Affine Independence is a sufficient condition for a set of vectors to be the vertexes of a convex polytope, and thus to be considered as endmembers. This chapter uses recent theoretical results showing that a set of Affine Independent vectors can be extracted from the rows and columns of Lattice Autoassociative Memories (LAAM). Therefore it is possible to extract the image endmembers from the dual LAAMs built to store the spectra of the hyperspectral image pixels. The WM algorithm proposed by Ritter et al. uses these results to find a convex polytope covering the hyperspectral image data. However, the number of induced endmembers obtained by this procedure is too high for practical purposes, besides they are highly correlated. We apply a Multi-Objective Genetic Algorithm (MOGA) to the optimal selection of the image endmembers. Two fitness functions are used in the works of this chapter, the residual error of the unmixing process and the size of the set of endmembers. The final set of endmembers is selected on the MOGA's Pareto front by examining the decrease in residual error obtained by increasing the number of endmembers. Additionally, this chapter proposes a faster MOGA where the error fitness function is replaced by a fitness function based on the correlation between endmembers, comparing the proposed process with a state-of-the-art EIA on well known benchmark images.

The chapter is organized as follows. Section 3.1 gives an introduction to the chapter. Sections 3.2 reviews the WM algorithm. Section 3.3 reviews the specific MOGA applied to the problem. Section 3.4 introduces the proposed

WM-MOGA approach for endmember induction. In section 3.5 we define the experimental research, and in section 3.6 we analyze the results. Finally, we give some conclusions in section 3.7.

3.1 Introduction

The set of image endmember spectra define a convex polytope[1] in the high-dimensional space defined by the image pixel spectra. The fractional abundance of the endmembers at each pixel correspond to the convex coordinates relative to the convex polytope vertexes. The set of endmembers can be defined on the basis of *a priori* knowledge about the imaged scene. A library of known pure ground signatures or laboratory samples could be used. Alternatively, the set of endmembers must be induced from the hyperspectral image data by means of Endmember Induction Algorithms (EIA) [140, 184].

Lattice based EIA (L-EIA) are based on lattice computing techniques [67]. For instance, the works in [73, 181, 75, 155, 182, 68, 153] are based on the notion of Strong Lattice Independence (SLI), following the conjecture in [150] that SLI vectors are Affine Independent vectors and thus its convex hull defines a simplex. Using Lattice Auto-Associative Memories (LAAM) [151, 154] built from the hyperspectral image data, sets of SLI vectors where induced and used as endmembers. Recent works [153] have shown how to obtain sets of Affine Independent vectors from the rows and columns of the LAAM constructed using the hyperspectral data.

Specifically, the WM algorithm [153] computes the erosive and dilative LAAM, the hyperbox enclosing the data, defined by the minimum and maximum values of the data at each band, transforming the columns or rows of the erosive and dilative LAAMs to become the vertexes of a convex polytope covering the image data. This algorithm has the following advantages:

- it is very fast,
- performs only addition, subtraction and max/min operations,
- induced endmembers are directly related to the actual data in the image.

However, the WM algorithm always returns the same number of endmembers $p = 2 * (L + 1)$, where L is the number of spectral bands in the image. This number is too high for the actual distinct constituent materials in hyperspectral images. Besides, these endmembers are highly correlated and identification of useful endmembers with some physical interpretation is tricky [152].

Genetic Algorithms (GA) are random optimization algorithms inspired on natural evolution, a population of individuals evolves through mutation and crossover to maximize a fitness function. Multi-objective GA (MOGA) are specific GAs dealing with the optimization of several objective functions. This

[1] The convex polytope is a simplex when it is defined by $d + 1$ vertexes, where d is the dimension of the space.

chapter proposes the use of specific MOGA for the selection of the endmembers from the large set of endmember candidates provided by the WM algorithm. A three-step process for endmember induction, called WM-MOGA, is proposed and tested.

1. First, compute the set of candidate endmembers applying the WM algorithm.

2. Secondly, apply a MOGA looking for an optimal set of endmembers minimizing both the residual error (RMSE) from the unmixing process and its cardinality, which amounts to minimizing the complexity of the solution. This WM-MOGA process returns a set of solutions that form a Pareto front on the solutions space formed by the two objective functions.

3. Thirdly, apply an Occam razor threshold criterion [132] to select the optimal set of induced endmembers as in [38].

To speed up the MOGA phase an alternative definition of the objective functions, minimizing the maximum correlation between endmembers, is used. This MOGA does not need to compute the unmixing process for each individual and, therefore, it is several orders of magnitude faster.

The WM-MOGA process is tested on real hyperspectral scenes, comparing it against the random search approach [38] based on the N-FINDR algorithm [193]. For validation, we calculate the correlation between the fractional abundances of the optimal endmembers induced by both, the WM-MOGA and the N-FINDR, and the available ground truth class spatial distribution. As WM-MOGA is an unsupervised process, the assignment of an abundance image to a ground truth class implies computing all possible combinations and selecting the best match.

3.2 WM Algorithm

The WM algorithm [153] is a Lattice Computing based EIA that builds a convex polytope containing the data from the rows and columns of the Lattice Auto-Associative Memories (LAAMs) constructed to store the image data. Algorithm 3.1 shows a pseudo-code specification for the WM algorithm. Given an hyperspectral image H, it is reshaped to form a matrix X of dimension $N \times L$, where N is the number of image pixels, and L is the number of spectral bands. The algorithm starts by computing the least hyperbox covering the data, $\mathcal{B}(\mathbf{v}, \mathbf{u})$, where $\mathbf{v}$ and $\mathbf{u}$ are the *minimal* and *maximal corners,* respectively, whose components are computed as follows:

$$v_k = \min_{\xi} x_k^{\xi} \text{ and } u_k = \max_{\xi} x_k^{\xi};\ k = 1, \ldots, L;\ \xi = 1, \ldots, N. \tag{3.1}$$

Next, the WM algorithm computes the dual erosive and dilative LAAMs $\mathbf{W}_{XX}$ and $\mathbf{M}_{XX}$. Next, the columns of $\mathbf{W}_{XX}$ and $\mathbf{M}_{XX}$ are scaled by $\mathbf{v}$ and $\mathbf{u}$, forming the additive scaled sets $W = \left\{\mathbf{w}^k\right\}_{k=1}^{L}$ and $M = \left\{\mathbf{m}^k\right\}_{k=1}^{L}$:

Algorithm 3.1 Pseudo-code specification of the WM algorithm.

1. L is the number of the spectral bands and N is the number of data samples.
2. Compute $\mathbf{v} = [v_1, \ldots, v_L]$ and $\mathbf{u} = [u_1, \ldots, u_L]$,
$$v_k = \min_{\xi} x_k^{\xi}; u_k = \max_{\xi} x_k^{\xi}$$
for all $k = 1, \ldots, L$ and $\xi = 1, \ldots, N$,
3. Compute the LAAMs
$$\mathbf{W}_{XX} = \bigwedge_{\xi=1}^{N} \left[\mathbf{x}^{\xi} \times \left(-\mathbf{x}^{\xi}\right)'\right]; \mathbf{M}_{XX} = \bigvee_{\xi=1}^{N} \left[\mathbf{x}^{\xi} \times \left(-\mathbf{x}^{\xi}\right)'\right]$$
where $\times$ is any of the $\boxtimes$ or $\boxtriangle$ operators.
4. Build $W = \left\{\mathbf{w}^1, \ldots, \mathbf{w}^L\right\}$ and $M = \left\{\mathbf{m}^1, \ldots, \mathbf{m}^L\right\}$ such that
$$\mathbf{w}^k = u_k + \mathbf{W}^k; \mathbf{m}^k = v_k + \mathbf{M}^k; k = 1, \ldots, L.$$
5. Return the set $V = W \cup M \cup \{\mathbf{v}, \mathbf{u}\}$.

$$\mathbf{w}^k = u_k + \mathbf{W}^k;\ \mathbf{m}^k = v_k + \mathbf{M}^k,\ \forall k = 1, \ldots, L, \tag{3.2}$$

where $\mathbf{W}^k$ and $\mathbf{M}^k$ denote the k-th column of $\mathbf{W}_{XX}$ and $\mathbf{M}_{XX}$, respectively.

Finally, the set $V = W \cup M \cup \{\mathbf{v}, \mathbf{u}\}$ contains the vertexes of the convex polytope $F(X) \cap \mathcal{B}(\mathbf{v}, \mathbf{u})$ which covers the convex hull of the data, $C(X)$, as a subset:

$$X \subset C(X) \subset F(X) \cap \mathcal{B}(\mathbf{v}, \mathbf{u}), \tag{3.3}$$

where $F(X)$ denotes the set of fixed points for the LAAMS. The WM algorithm returns the set V as the set of induced endmembers. The algorithm is simple and fast, but the number of induced endmembers in V can be too large for practical purposes. Furthermore, some of the endmembers can show high correlation even if they are affine independent. To obtain a meaningful set of endmembers, we search for an optimal subset of V in the sense of minimizing the unmixing residual error and the number of endmembers.

3.3 NSGA-II

MOGA's fitness function is vector valued [43, 51, 52, 101]. A minimization multi-objective problem is stated as follows: Given an n-dimensional variable

vector $\mathbf{x} = \{x_1, \ldots, x_n\}$ in the solution space $\mathbf{X}$, find a vector $\mathbf{x}^*$ that minimizes the K objective functions $\mathbf{z}(\mathbf{x}^*) = \min_{\mathbf{x}} \{z_1(\mathbf{x}), \ldots, z_K(\mathbf{x})\}$. The region of feasible solutions in the solution space $\mathbf{X}$ is often specified by a collection of constraints, such as $g_j(\mathbf{x}^*) = b_j$ for $j = 1, \ldots, m$. An ideal solution that simultaneously optimizes each objective function is impossible to find in most cases because of the mutual conflicts between objectives. Multi-objective optimization algorithms search for balanced solutions, providing not a single solution but a collection of them, an approximation to the *Pareto set*. A solution $\mathbf{x}$ *dominates* another solution $\mathbf{x}'$ if it improves it for all objective functions, i.e. $z_1(\mathbf{x}') < z_1(\mathbf{x}), \ldots, z_K(\mathbf{x}') < z_K(\mathbf{x})$. The Pareto set is the set of solutions that are not dominated by any other solution

$$\mathbf{P} = \{\mathbf{x} \,|\neg\exists\mathbf{x}'; z_1(\mathbf{x}') < z_1(\mathbf{x}), \ldots, z_K(\mathbf{x}') < z_K(\mathbf{x})\}.$$

In the function domain, the *Pareto front* is constituted by the function values of the solutions in the Pareto set. If a single solution is sought, then an additional selection must be performed on the Pareto set of non-dominated solutions.

The Non-dominated Sorting Genetic Algorithm II (NSGA-II) [52] is a fast and elitist multi-objective genetic algorithm. The NSGA-II algorithm starts by creating a random initial parent population P_0, which is sorted based on non-domination such that a rank is assigned to each solution according to its level of non-domination (rank 1 corresponds to non-dominated solutions in the Pareto front). Conventional tournament selection, recombination and mutation operators for binary chromosomes are used to create an offspring population Q_0 of size N.

Algorithm 3.2 gives the pseudo-code for a single NSGA-II generation. First, a combined population $R_t = P_t \bigcup Q_t$ of size $2N$ is formed. Elitism is ensured because the best individuals from the parents and offsprings are always retained. Then, R_t is sorted according to non-domination level using a fast sorting algorithm. The non-dominated set of solutions in $\mathcal{F}_1$ are included in new population P_{t+1}. Once $\mathcal{F}_1$ is removed from R_t, then the solutions in $\mathcal{F}_2$ are the new set of non-dominated solutions in $R_t - \mathcal{F}_1$, and are thus included in the new population P_{t+1}. This procedure is repeated adding subsequent non-dominated fronts in the order of their ranking until reaching the required number of solutions N. Often, not all the solutions in the last considered front F_l are included. The solutions of the last front are sorted in descending order using the crowded-comparison operator $\preceq_n$ which favors solutions with lower (better) non-domination rank and, if both solutions belong to the same front, favors the solution located in a lesser crowded region. Best solutions are chosen up to fill P_{t+1}, which is now used for crossover and mutation to create a new offspring population Q_{t+1} of size N. The overall complexity of the algorithm is $O\left(MN^2\right)$ which is governed by the non-dominated sorting part of the algorithm. The diversity among non-dominated solutions is introduced by using the crowding-comparison procedure, which makes unnecessary any niching parameter.

Algorithm 3.2 NSGA-II algorithm iteration

1. Combine parent and offspring population: $R_t = P_t \bigcup Q_t$
2. Calculate all the non-dominated fronts of R_t: $\mathcal{F} = (\mathcal{F}_1, \mathcal{F}_2, \ldots) =$ fast-non-dominated-sort (R_t)
3. Do until filling the parent population: $|P_{t+1}| + |\mathcal{F}_i| \leq N$
   (a) Calculate crowding-distance in $\mathcal{F}_i$: crowding-distance-assignment $(\mathcal{F}_i)$
   (b) Include i-th non-dominated front in the parent population: $P_{t+1} = P_{t+1} \bigcup \mathcal{F}_i$
   (c) Check the next front for inclusion: $i = i + 1$
4. Sort in descending order using the crowding-comparison operator $\preceq_n$: Sort $(\mathcal{F}_i, \preceq_n)$
5. Choose the first $(N - |P_{t+1}|)$ elements of $\mathcal{F}_i$: $P_{t+1} = P_{t+1} \bigcup \mathcal{F}_i [1 : (N - |P_{t+1}|)]$
6. Use crossover and mutation to create a new offspring population Q_{t+1}
7. Increment the generation counter: $t = t + 1$

3.4 WM-MOGA

The WM-MOGA is a three step process specified in Algorithm 3.3. WM-MOGA starts by computing a set of candidate endmembers using the WM algorithm. Given an hyperspectral image H, it is reshaped to a matrix X of size $N \times L$, where N is the number of pixels in the image, and L is the number of spectral bands. Applying WM algorithm to X obtains a set of candidate endmembers, denoted $E_{\mathrm{WM}} = \left\{\mathbf{e}^1, \ldots, \mathbf{e}^p\right\}$. The second step of WM-MOGA finds the optimal subset of endmembers in terms of unmixing residual error and complexity, by using a MOGA to calculate the *Pareto front* of non dominated solutions, as defined in the introduction section, $\mathbf{P} = \left\{E^1, \ldots, E^q\right\} \subseteq \mathcal{P}(E_{\mathrm{WM}})$, where $\mathcal{P}(E_{\mathrm{WM}})$ is the power-set of E_{WM}.

We define two fitness functions. One is the unmixing residual error of equation (3.4) denoted by $f_{\mathrm{RMSE}}(E)$,

$$f_{\mathrm{RMSE}}(E) = \mathrm{RMSE}(E, X) = \frac{1}{N} \sum_{i=1}^{N} (\mathbf{x}_i - E\boldsymbol{\alpha}_i)^2, \tag{3.4}$$

where $\mathbf{x}_i$ is the i-th pixel in the hyperspectral image X, and $\boldsymbol{\alpha}_i$ is the vector of fractional abundances for the ith pixel calculated by Full Constrained Least Squares Unmixing (FCLSU) [32]. The second fitness function is a measure of the

Algorithm 3.3 Pseudo-code for the WM-MOGA process

1. Apply WM (X) to obtain $E_{\mathrm{WM}} = \left\{\mathbf{e}^1, \ldots, \mathbf{e}^p\right\}$
2. Apply MOGA (E_{WM}) to obtain the Pareto set of solutions $\mathbf{P} = \left\{E^i,\, i = 1, \ldots, q\right\}$
3. Apply the Occam razor selecting $E^*(\varepsilon) = \arg\min_{\mathbf{P}} \left\{ \left| \frac{f_{\mathrm{RMSE}}\left(E^{i+1}\right)}{f_{\mathrm{RMSE}}\left(E^i\right)} - \frac{f_{\mathrm{RMSE}}\left(E^i\right)}{f_{\mathrm{RMSE}}\left(E^{i-1}\right)} \right| < \epsilon \right\}$
4. Return $E^*(\varepsilon)$

solution complexity given the relative size of the set of endmembers as specified in equation (3.5), denoted by $f_{|\cdot|}(E)$,

$$f_{|\cdot|}(E) = \frac{|E|}{|E_{\mathrm{WM}}|}, \tag{3.5}$$

where $|\cdot|$ denotes the cardinality of a set. The MOGA requires to encode the problem so each individual in the search population represents a solution. A k-th individual *chromosome* is defined as a binary vector $\mathbf{b}_k = \{b_1, \ldots, b_p\}$; $b_i \in \{0, 1\}$; $\forall i = 1, \ldots, p$; being p the number of candidate endmembers returned by WM algorithm. If $b_i = 1$, the i-th candidate endmember $\mathbf{e}^i \in E_{WM}$ belongs to the set of induced endmembers, E_k, corresponding to $\mathbf{b}_k$.

The final third step applies the Occam razor [132] to select the size set of endmembers where the reduction in the unmixing residual error obtained accepting another endmember is below a given threshold. Sorting the solutions $E^i \in \mathbf{P}$ by their cardinality so that $i = \left|E^i\right|$, the Occam razor condition is specified in equation (3.6) for a given selection threshold ε on the difference of relative errors between consecutive solutions according to the number of endmembers:

$$E^*(\varepsilon) = \arg\min_{\mathbf{P}} \left\{ \left| \frac{f_{\mathrm{RMSE}}\left(E^{i+1}\right)}{f_{\mathrm{RMSE}}\left(E^i\right)} - \frac{f_{\mathrm{RMSE}}\left(E^i\right)}{f_{\mathrm{RMSE}}\left(E^{i-1}\right)} \right| < \epsilon \right\}. \tag{3.6}$$

The algorithm returns $E^*(\varepsilon)$ as the final set of induced endmembers from the hyperspectral image X.

3.4.1 WM-MOGA-CORR

Much of the computational cost of the MOGA is due to the computation of the fractional abundances $\boldsymbol{\alpha}$ for each population individual at each generation. We propose a faster approximation in which MOGA looks for subsets of endmembers that minimize the maximum of the between endmembers correlation while keeping as much endmembers as possible. This is done by substituting the unmixing error fitness function in equation (3.4) by a fitness function based on the maximum correlation between the endmembers

$$f_{\text{CORR}}(E) = \max\{c_{ij}; \mathbf{e}_i, \mathbf{e}_j \in E\}, \tag{3.7}$$

where c_{ij} is the Pearson's correlation between endmembers. Thus, we try to minimize the maximum correlation between endmembers. This fitness function is far less computationally expensive than the fitness function f_{RMSE} of equation (3.4). The solution complexity related fitness function of equation (3.5) can not be applied in conjunction with the correlation based fitness function of equation (3.7) because it leads to the trivial result of a single endmember. Instead, we use its inverse:

$$f_{|\cdot|^{-1}}(E) = \frac{|E_{\text{WM}}|}{|E|}. \tag{3.8}$$

We call this approximation to the WM-MOGA using fitness functions $\left\{f_{\text{CORR}}, f_{|\cdot|^{-1}}\right\}$ WM-MOGA$_{\text{CORR}}$. Note that the optimality criteria for the set of endmembers sought is the minimization of the unmixing residual error and number of endmembers. The approximation WM-MOGA$_{\text{CORR}}$ does not attempt to minimize these criteria directly, but nevertheless the quality of its achieved solution will be evaluated on the basis of the unmixing residual error.

3.5 Experimental methodology

We have perform experiments over two well known hyperspectral images, the Indian Pines and the Salinas scenes. The experiments have been run using the MATLAB[2] implementation of NSGA-II [52] multiobjective genetic algorithm. The number of individuals in the population was set to 100 for the WM-MOGA_{RMSE} and 1000 for the WM-MOGA_{CORR}. To perform the comparison we calculate the correlation between the fractional abundance images corresponding to the optimal sets of endmembers induced by the different approaches and the ground truth classes from the different hyperspectral scenes. All the hyperspectral scenes and code of the algorithms and methods implemented are freely accessible from the Computational Intelligence group (Basque Country university, UPV/EHU) website[3][4].

The explored research questions are the following ones:

- (a) is it possible to obtain a reduced set of endmembers from the WM algorithm using a search process based on the quality and size of the set of endmembers?,
- (b) is it possible to speed up the search process using indirect information such as the correlation between endmembers?, and

[2] http://www.mathworks.com/products/matlab/
[3] http://www.ehu.es/ccwintco/index.php/GIC-source-code-free-libre
[4] http://www.ehu.es/ccwintco/index.php/GIC-experimental-databases

- (c) how those endmember induction processes compare with a state-of-the-art algorithm?.

We have defined the WM-MOGA and WM-MOGA$_{\text{CORR}}$ algorithms to answer the first two questions. The experimental results provide answers to the last question. We compare the WM-MOGA (denoted WM-MOGA_{RMSE} in the figures) and WM-MOGA$_{\text{CORR}}$ processes with a recent random search approach based on N-FINDR [38] which runs the N-FINDR algorithm several times for increasing values of the number of desired endmembers, p, and then applies the Occam razor specified by equation (3.6) to determine the optimal set of endmembers $E^*_{\text{N-FINDR}}(\varepsilon)$.

Note The endmember induction processes are unsupervised, therefore the meaning of the endmembers found and their relation to the ground-truth classes is unknown. However, we want to support our work on the knowledge of a given ground-truth for the benchmark images. The evaluation process looks for the best match between the abundance images produced by the unmixing and the image regions identified with each class in the ground-truth[5]. We compute all the possible spatial correlations between them, obtaining a matrix of correlation indexes. The examination of this matrix gives information about the correspondence between endmembers and ground-truth classes, and the uncertainty of this association.

3.6 Experimental results

Quantitative results We first provide the plots of the unmixing residual error (RMSE) versus the number of endmembers of the solutions found by the WM-MOGA, WM-MOGA$_{\text{CORR}}$ and N-FINDR based approach of [38]. The Pareto front of WM-MOGA$_{\text{CORR}}$ refers to the correlation between endmembers, not to the unmixing residual error, however, the selection of the solution is based on the same criteria for all algorithms. Figures 3.1 and 3.2 show the RMSE plots for the Indian Pines and Salinas images, respectively. The curve corresponding to the WM-MOGA is smooth because the MOGA searches for the Pareto front based on these criteria. The curves corresponding to WM-MOGA$_{CORR}$ and N-FINDR are more irregular, with several local minima. The Occam razor tries to determine the optimal solution based on the relative error decrease according to equation (3.6).

We plot in figures 3.3 and 3.4 the relative error

$$\frac{f_{\text{RMSE}}\left(E^i\right)}{f_{\text{RMSE}}\left(E^{i-1}\right)}$$

evolution for the algorithms. It can be appreciated that the WM-MOGA provides a smooth relative error curve that allows an easy setting of the threshold

[5] We do not have knowledge of the ground truth endmembers.

parameter of the Occam razor and gives sensible results in the solution selection applying $\varepsilon = 10^{-2}$.

For the WM-MOGA$_{CORR}$ selection of the definitive threshold required inspection of the relative error curve, a threshold $\varepsilon = 10^{-2}$ gives sensible results for the Indian Pines image, but $\varepsilon = 10^{-1}$ is required for the Salinas image. The N-FINDR approach gives very irregular relative error curves, however the standard threshold $\varepsilon = 10^{-2}$ gives sensible results for both images. The endmembers selected as a result of these decisions are plot in figures 3.5 and 3.6 for the Indian Pines and Salinas images, respectively.

Endmembers found by the N-FINDR are also plot for comparison. It can be appreciated that WM-MOGA$_{\mathrm{CORR}}$ provides more uncorrelated endmembers than the other approaches. The WM-MOGA$_{\mathrm{CORR}}$ is the most relaxed approach, finding the highest number of endmembers, however, a strong correlation can be appreciated in figures 3.5 and 3.6 among the endmembers found by all three approaches, though the WM-based endmembers are not pixel spectra from the image, while N-FINDR endmembers are pixel signatures selected from the image.

Qualitative evaluation of the results. Figures 3.7 and 3.8 provide the thematic maps and the images containing the maximum abundance value per pixel for each of the tested approaches. The thematic maps are computed as follows:

- Compute the correlation coefficient between each abundance image and each binary image corresponding to a ground-truth class spatial distribution.
- Assign to each endmember the set of ground truth regions with positive correlation coefficients.
- For each pixel select the endmember with the maximal abundance value and we assign to the pixel the linear combination of the colors of ground-truth positively correlated with the endmember abundance, i.e. the orange color corresponds to the mixture of red and yellow.
- Remove the background class in these computations.

It can be appreciated examining the thematic maps in figures 3.7(b,d,f) and 3.8(b,d,f) that the WM-MOGA$_{\mathrm{CORR}}$ provides more clean recognition of the ground-truth class areas in both Indian Pines and Salinas, maybe due to its emphasis in uncorrelated endmembers. Besides, there is little correspondence between the classes and the endmembers in all cases: most of the ground truth regions are not recognized in their original spatial localization. Attending to the abundance coefficients shown in figures 3.7(c,e,g) and 3.8(c,e,g) there are few pixels with a pure endmember matching, most abundance values are moderate implying some degree of mixture of the real classes.

In average, the WM-MOGA$_{\mathrm{CORR}}$ provides the greater values of the abundance coefficients, improving over WM-MOGA and N-FINDR. Finally, to asses

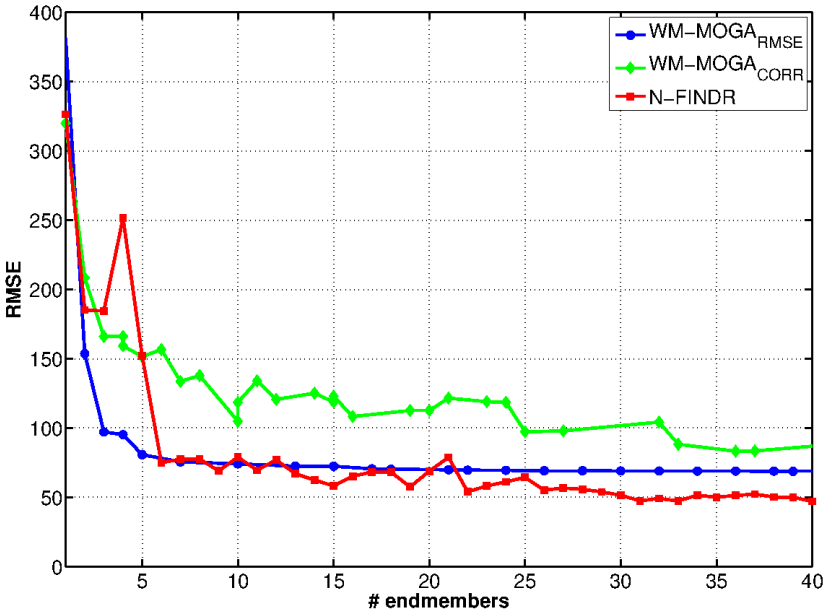


Figure 3.1: Indian Pines image. Plot of the unmixing residual error versus the number of endmember for the Pareto set of non-dominated solutions found by WM-MOGA and WM-MOGA$_{CORR}$, and the solutions found by the N-FINDR based approach of [38].

the degree of ground-truth class discovery by the endmembers we plot the maximum correlation coefficient *per* abundance and *per* ground truth class in figures 3.9 and 3.10 for the Indian Pines and Salinas images, respectively. Examining the maximum correlation *per* ground truth class (figures 3.9(a) and 3.10(a)) the results are not exactly equal in both images, however the trends are similar.

The WM-MOGA provides the best identification of the class almost for all of them. For some classes WM-MOGA$_{\mathrm{CORR}}$ performs better, for a couple of classes N-FINDR is the best detector. We can say that WM-MOGA compares well or improves the ground-truth class detection over N-FINDR. Also the correlation approximation of WM-MOGA$_{\mathrm{CORR}}$ gives surprising good detections of some ground truth classes, and in general is a good approximation to the detection obtained by WM-MOGA. Examining the maximum correlation per abundance image (figures 3.9(b) and 3.10(b)), we find the same kind of results. For most of the induced abundances, the WM-MOGA provides the best correspondence to some ground-truth class. The approximation provided by the WM-MOGA$_{\mathrm{CORR}}$, which is several orders of magnitude faster, does a good job of finding meaningful endmembers, but it finds many endmembers so that there is a tail of irrelevant endmembers little correlated with the ground-truth classes.

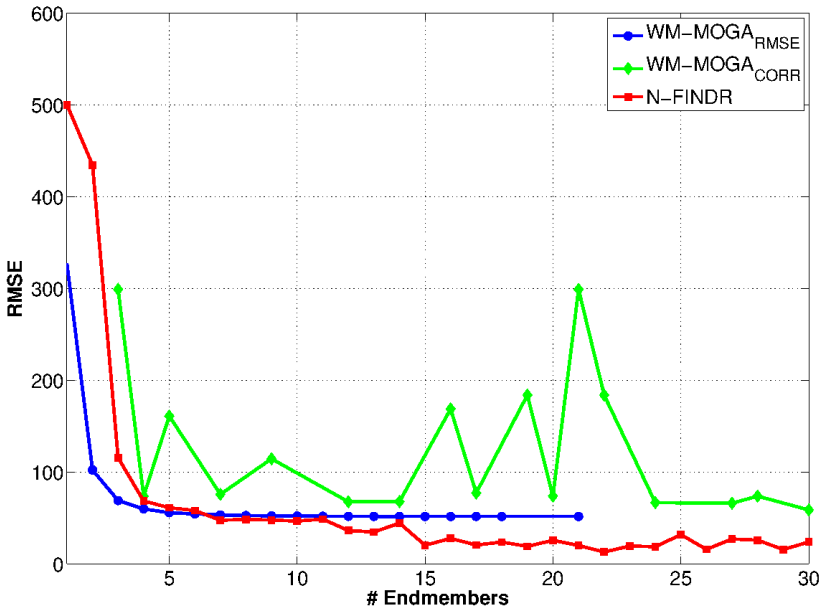


Figure 3.2: Salinas image. Plot of the unmixing residual error versus the number of endmember for the Pareto set of non-dominated solutions found by WM-MOGA and WM-MOGA$_{CORR}$, and the solutions found by the N-FINDR based approach of [38].

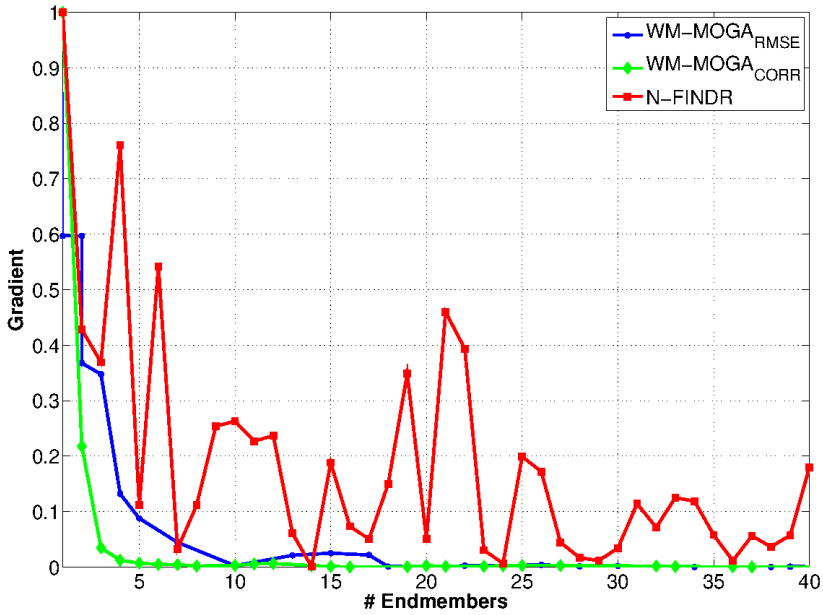


Figure 3.3: Indian Pines image. Relative error $f_{\text{RMSE}}(E^i)/f_{\text{RMSE}}(E^{i-1})$ for the solutions obtained by WM-MOGA and WM-MOGA$_{CORR}$, and N-FINDR in figure 3.1.

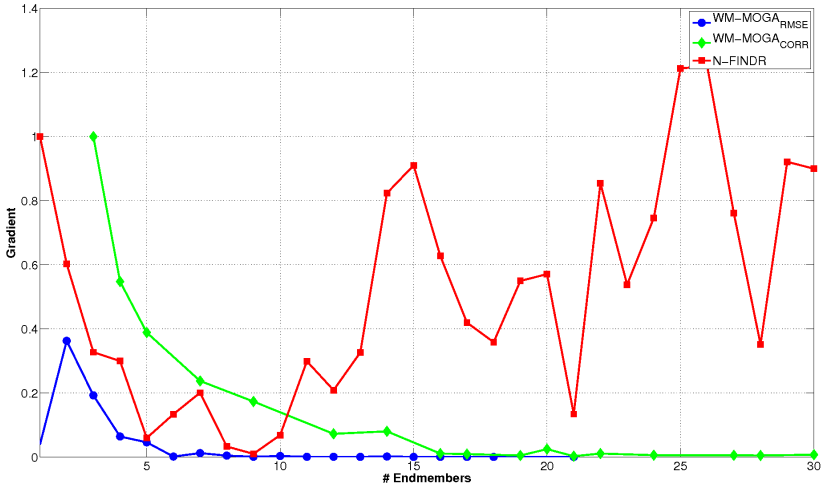
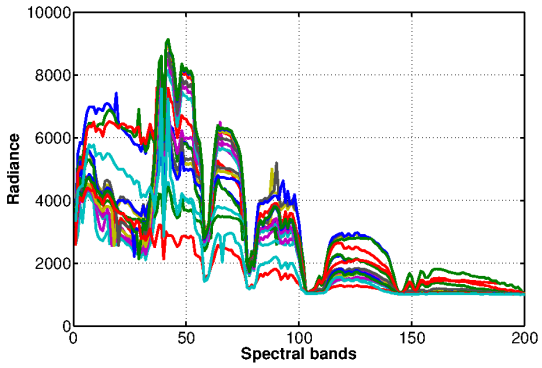
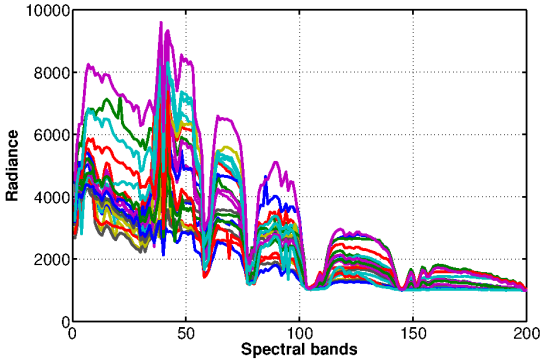


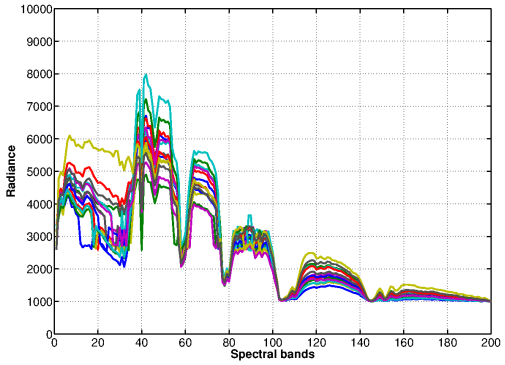
Figure 3.4: Salinas image. Relative error $f_{\text{RMSE}}(E^i)/f_{\text{RMSE}}(E^{i-1})$ for the solutions obtained by WM-MOGA and WM-MOGA$_{CORR}$, and N-FINDR in figure 3.2.



(a)

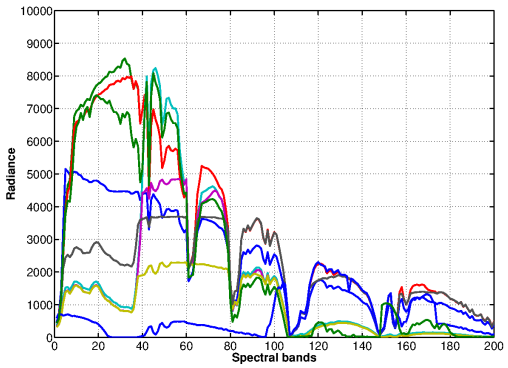


(b)

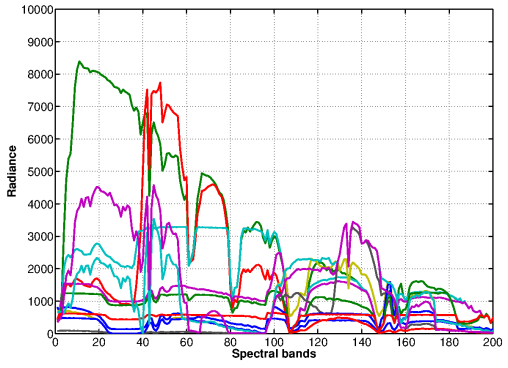


(c)

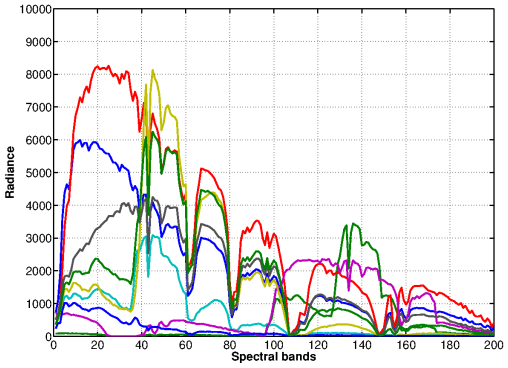
Figure 3.5: Endmembers induced from the Indian Pines scene by the (a) WM-MOGA, (b) WM-MOGA$_{\text{CORR}}$, and (c) N-FINDR approaches following the Occam razor selection strategy. Plot colors are arbitrary. The number of endmembers found by each method is 18, 19, and 14, respectively.



(a)



(b)



(c)

Figure 3.6: Endmembers induced from the Salinas scene by the (a) WM-MOGA, (b) WM-MOGA$_{\mathrm{CORR}}$, and (c) N-FINDR approaches following the Occam razor selection strategy. Plot colors are arbitrary. The number of endmembers found by each method is 9, 12, and 9, respectively.

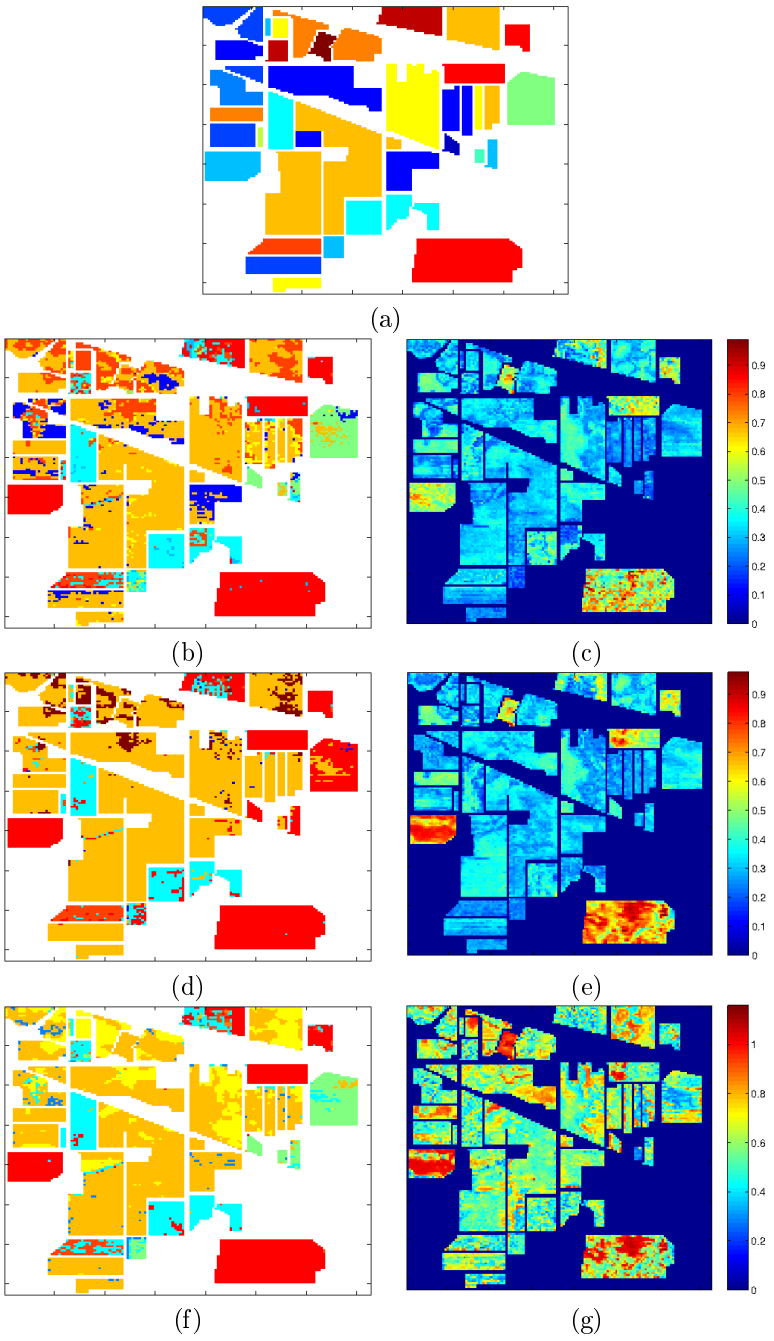


Figure 3.7: (a) Indian Pines ground-truth. Thematic maps obtained by linear combination of the color tags of the ground truth regions with maximal correlation relative to the abundance images (b) WM-MOGA, (d) WM-MOGA$_{CORR}$, (f) N-FINDR. Maximal abundance value per pixel (c) WM-MOGA, (e) WM-MOGA$_{CORR}$, (g) N-FINDR.

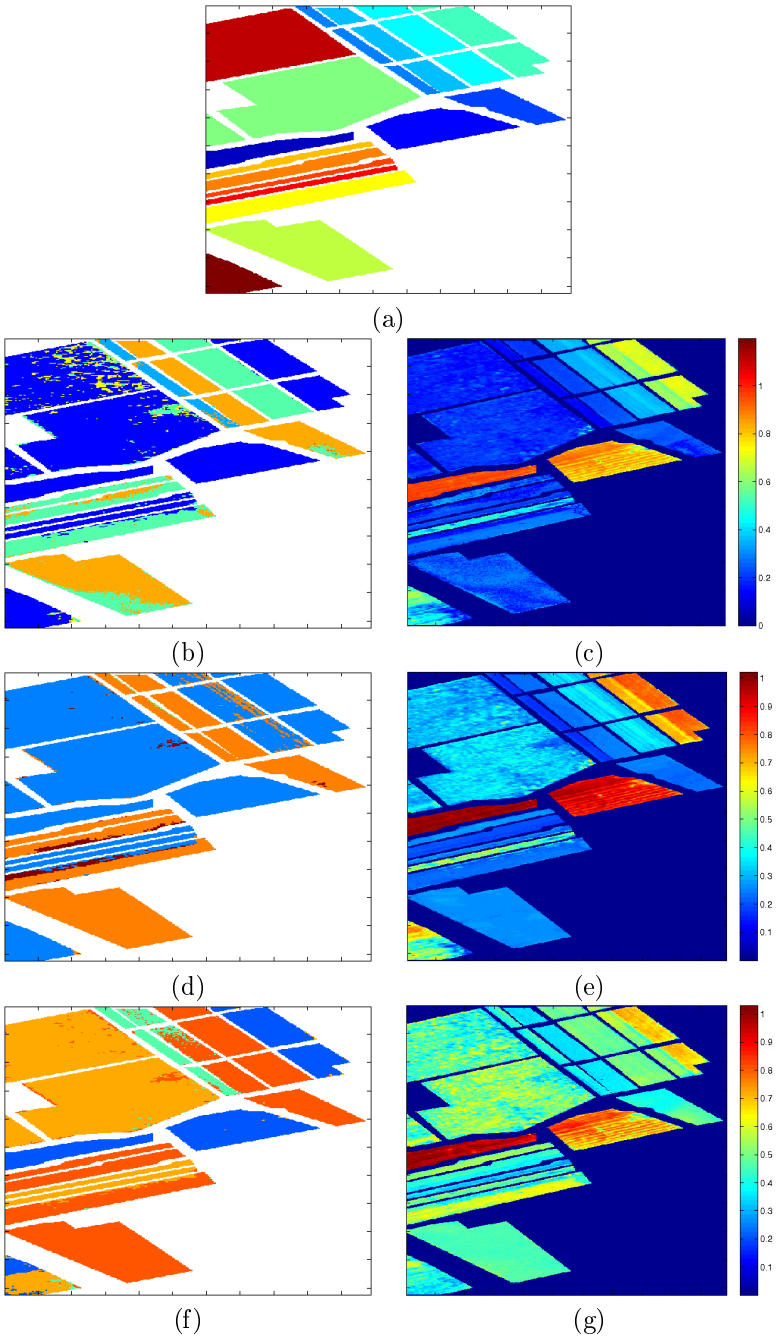
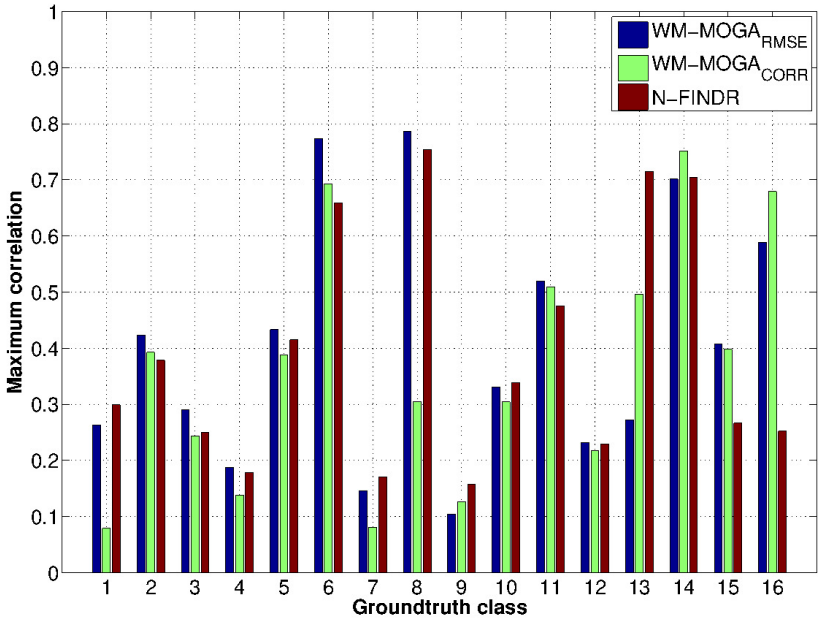
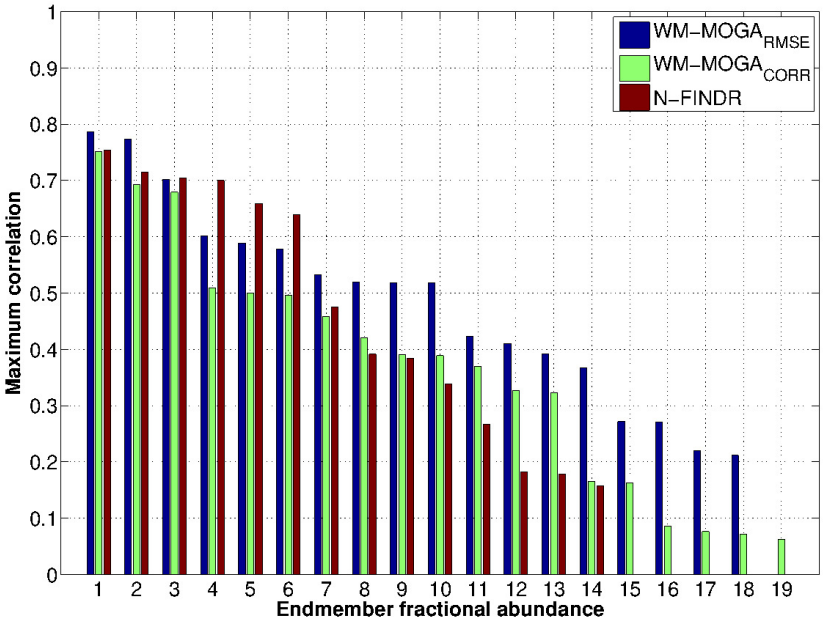


Figure 3.8: (a) Salinas ground-truth. Thematic maps obtained by linear combination of the color tags of the ground truth regions with maximal correlation relative to the abundance images (b) WM-MOGA, (d) WM-MOGA$_{CORR}$, (f) N-FINDR. Maximal abundance value per pixel (c) WM-MOGA, (e) WM-MOGA$_{CORR}$, (g) N-FINDR.

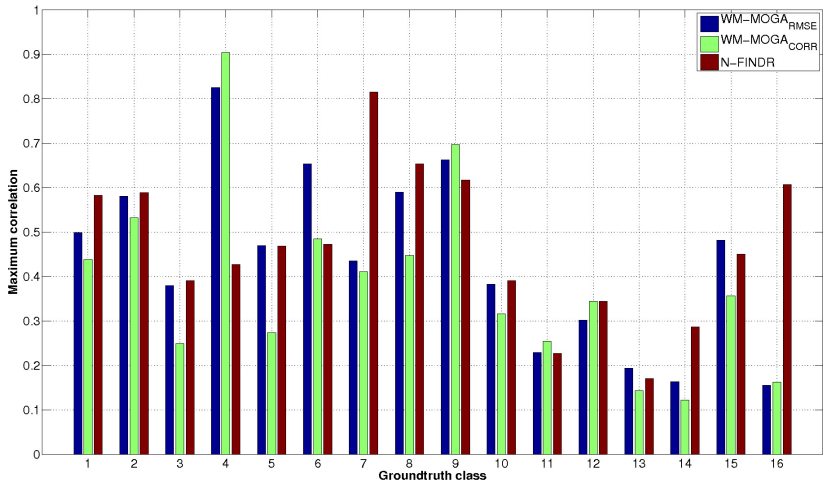


(a)

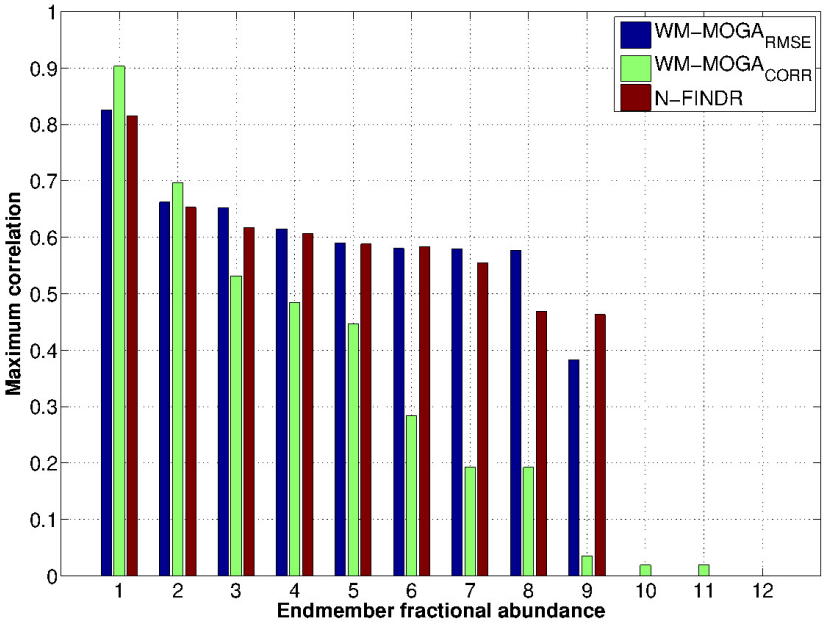


(b)

Figure 3.9: Indian Pines image. Maxima of the correlation coefficients between ground truth classes and induce abundance images. (a) maxima per ground-truth class, (b) maxima per endmember.



(a)



(b)

Figure 3.10: Salinas image. Maxima of the correlation coefficients between ground truth classes and induced abundance images. (a) maxima per ground-truth class, (b) maxima per endmember.

3.7 Conclusions

The WM Algorithm proposed by Ritter at al. [153, 152] is a fast procedure to obtain a set of affine independent vectors which are the vertexes of a convex polytope covering the sample data. Applied to hyperspectral images, WM Algorithm produces a large set of candidate endmembers. This chapter proposes an specific MOGA minimizing the unmixing residual error and the number of endmembers, followed by an Occam razor selection on the Pareto front, to obtain an appropriate set of endmembers tailored to the data. The WM-MOGA compares well to a recent state-of-the-art endmember induction heuristic [38] in terms of the correlation of the induced abundance images with the given ground truth class spatial distribution. Furthermore, we propose an approximation to the MOGA which does not need to compute the linear unmixing based on the individual chromosomes at each generation. This fast process identification of the ground truth classes compares well with the reference heuristic. However, it overestimates the set of endmembers, including some redundant or irrelevant endmembers. Future work may be addressed to improve the fast approach introducing new regularization fitness functions to obtain smaller sets of endmembers of equivalent quality.

Chapter 4

Hyperspectral CBIR systems: general aspects and validation

This chapter introduces the general ideas about content-based image retrieval (CBIR), including its structure and validation issues. Specifically, the validation strategies that will be used in the next chapters are detailed here, including the definition of the validation performance measures and how they are computed depending on the kind of data used, be it synthetic or real data. An special problem is that of the lack of ground truth, which requires special strategies to obtain some system validation results. The approach suggested is to generate equivalent ground truth values from the data through a parallel unsupervised clustering process.

Section 4.1 provides a literature review. Section 4.2 discusses CBIR systems in the remote sensing domain. Section 4.3 comments on the issues of search and the semantic gap. Section 4.4 presents the issues of CBIR system validation, including the performance measures, and the specific use of synthetic and real data. Section 4.5 discusses the validation issue when there is no ground truth available. No conclusion section is given in this chapter.

4.1 Literature review

Traditional Database management systems (DBMS) are text-based information retrieval systems. The use of such systems to search in image databases is known as text-based image retrieval (TBIR). In TBIR systems images and textual metadata are stored together. Metadata describes technical information about the image acquisition, as well as manually annotations describing on the image contents. Google images[1] and Flickr[2] are examples of commercial TBIR applications. Another example is Picture Archiving and Communication Systems (PACS) [90] managing medical image databases. PACS can search for a

[1] `http://www.google.com/imghp`
[2] `http://www.flickr.com/`

patient medical image history and solve queries involving the combination of different information criteria. However, text information does not allow accurate content description. Manual image annotation is expensive and prone to inaccuracies due to discrepancies in human subjective perception. To overcome these disadvantages, content-based image retrieval (CBIR) was introduced in the early 1980s.

CBIR systems [158, 169, 116, 50, 119] retrieve images stored in an image database guided by their intrinsic information contents. For instance, in query-by-example the user defines the database query by giving an image example expecting to retrieve the most similar images found in the database. To this end, image content should be described by feature vectors that could be extracted from the images by means of computer vision and digital image processing techniques. These feature vectors are low-level content descriptors that might be used as image indexes to perform searches in the database. A query corresponds to a point in feature space, consequently image retrieval answering the query consists of finding the images in the database closest in feature space to the query. Therefore, we need to define a similarity/dissimilarity measure in feature space in order to decide which images are closer to the query. Some prototypes were developed along the 1990s such as QBIC [65], Photobook [137], VisualSEEK [170], Virage [80], Netra [117], Viper [173], PicSOM [104], or SIMPLIcity [190].

The research field of CBIR systems has grown exponentially after the year 2000. A search for publications containing the phrase "Image Retrieval" on Google Scholar and the digital libraries of ACM, IEEE and Springer, in the years from 1995 to 2005, results ind "a roughly exponential growth in interest in image retrieval and closely related topics ... [and a] particularly strong growth over the last five years, spanning new techniques, support systems, and application domains." [50]. This effect has reached remote sensing community as well. The interest in fast and intelligent information retrieval systems over collections of remote sensing images is increasing as the volume of available data grows exponentially. A single in-orbit Synthetic Aperture Radar (SAR) sensor collects about 10-100 Gbytes of data per day, giving about 10 Tbytes of imagery data per year. New Earth Observation missions are increasing the number and quality of sensors in Earth orbit allowing to acquire image data in amounts orders of magnitude higher than previously. Current remote sensing retrieval systems offer to their users raw images, thematic maps and ancillary data in response to very precise queries requiring a detailed knowledge of the structure of the stored information. Remote sensing [99, 1] is a crucial source of information to support decision making, so that there is a growing need for more flexible ways to access the information based on the image intrinsic properties, i.e. texture, spectra, shape; and induced semantic contents.

Research on modern CBIR systems has shifted from designing CBIR systems based on low-level feature descriptors of the visual content of images, to CBIR systems capable of dealing with human semantics. In this path, two gaps appears that motivate most of the problems we would encounter, the sensory gap and the semantic gap [169]. The sensory gap is the gap between the object

in the world and the information in a (computational) description derived from a recording of that scene. The semantic gap is the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data has for a user in a given situation. While the former makes recognition from image content challenging due to limitations in recording, the latter brings in the issue of a user's interpretations of pictures and how it is inherently difficult for visual content to capture them [50].

Approaches to CBIR in remote sensing image databases proposed up to now has been focused on panchromatic, SAR or low-dimensional multispectral images. An early approach to the definition of a Remote Sensing CBIR (RS-CBIR) system [85] was formulated over physical models to retrieve images from a multispectral image database. Works in [47, 46, 45, 48] describe the development of an operative RS-CBIR system, the knowledge-driven information mining (KIM) system, to manage and explore large volumes of remote sensing image data. The KIM system evolves from the authors previous works in spatial information retrieval [49, 160] and probabilistic interactive learning [161] to mine SAR and multispectral remote sensing databases. In [45] and [48], authors overview the KIM system from the perspective of human-machine communication (HMC) and human-centered concept (HCC) respectively. Another approach [128], extends PicSOM [104] CBIR system to query remote sensing image databases to detect man-made structures and changes in multispectral image data. Authors in [79] introduce a RS-CBIR system built up from the point of view of Information Theory. They propose an informational similarity measure based on coding length and they use it for the image retrieval of satellite image time series. [159] describes the Satellite Image Matching and Retrieval (SIMR) system, which represents an image by a signature that consists of spectral and spatial attributes at different resolution levels using the quadtree data structure. SIMR employs several distance measures to compute the similarity between two image signatures.

Focusing on the semantic gap in RS-CBIR systems, [56] describes a framework for semantic-enabled knowledge discovery from Earth Observation (EO) data archives modeling the information sources by domain-specific ontologies, which are capable of capturing knowledge structures. Additionally, [174] proposes a semantic-based RS-CBIR system based on ontology and grid technologies. The Global Earth Observation System of Systems (GEOSS) [57] is an end-to-end process that enables the collection and distribution of accurate, reliable EO data, information, products, and services to both suppliers and consumers worldwide. Latent Dirichlet Allocation (LDA) [111] is used to annotate large satellite images using semantic concepts defined by the user, or LDA model to discover semantic rules [21] linking the low-level primitive features with user-defined high-level semantics on different manually created cartographic products.. Other approach[106] uses thematic maps obtained by a machine learning classifier and textural features to search on a multispectral image database. Same authors extend their work and proposes reduce the semantic gap by using a context-sensitive Bayesian network for semantic inference of segmented scenes [110]. The work in [63] addresses relevance feedback in remote sensing

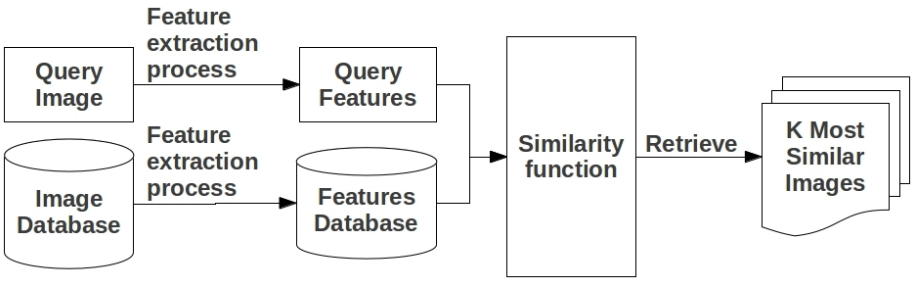


Figure 4.1: Diagram of a common hyperspectral CBIR system.

multispectral CBIR proposing an active learning methodology. Authors give experimental results on high-resolution panchromatic satellite images. Image annotation is also the goal pursued in [19], where a semi-supervised method to auto-annotate and discover unknown semantic classes in multispectral satellite image databases is proposed.

4.2 RS-CBIR systems architecture

Figure 4.1 shows a diagram of a CBIR system which is also appropriate for the Remote Sensing domain (RS-CBIR). There are two main elements in the definition of the RS-CBIR:

(a) the image features and corresponding feature extraction process, and

(b) the similarity measure defined in the feature space guiding the database search.

Identifying a set of salient image features for image indexing and similarity evaluation is a central issue in CBIR systems. Color, shape, texture and spatial relationships among segmented objects are typical features employed for image indexing [116]. Hyperspectral images contain a rich spectral information, therefore it is natural to exploit this information defining spectral image features. The definition of an appropriate dissimilarity function must follow accordingly.

CBIR image domains are classified [169] as narrow and broad. Narrow image domains have limited variability. On the other hand, broad domains tend to have high variability blurring the semantic meanings. RS-CBIR systems often lie in the later category. Furthermore, RS-CBIR systems present added challenges such as data visualization and ground-truth scarcity. Visualization o f hyperspectral image data is not straightforward as they are formed of hundred of contiguous spectral bands. False color images composed by selecting three bands can miss some the spectral relevant information contained in the image. Even expert users have often difficulties to understand remote sensing images by visual inspection, which poses the an additional issue of RS-CBIR systems

validation. Ground-truth data is expensive and prone to errors and thus, its availability is scarce.

In CBIR systems defining the query may be a critical design issue. In RS-CBIR systems the Query By Example(s), where the query is represented by an example image (or a set of images) related to the search target, would be the preferred choice. When the user is only interested in an specific region of the images, he may specify the query by selecting a region of interest (ROI), which is known as the Query By Region(s). However, the *Page Zero Problem* [205] arises when the user lacks an example image to start querying. Conventional Query By Text could be useful in situations where relevant information about images is collected as textual metadata, as for example, geographical coordinates, time of acquisition or sensor type. Text may be used as a complement to the visual query, or may solve the Page Zero Problem.

4.3 Search and the semantic gap

Regarding the kind of search, the following types can be distinguished [169]:

1. Search by category: the user wants to retrieve images belonging the same class.

2. Search by target: the user wants to retrieve images containing an object.

3. Search by association: user has not an specific target in mind and its only aim is to browse through the database.

In [50], authors extend the above search-type-based classification with a user-intent-based classification:

1. Browser: this is a user browsing for pictures with no clear end-goal. A browser's session would consist of a series of unrelated searches. A typical browser would jump across multiple topics during the course of a search session. Her queries would be incoherent and diverse in topic.

2. Surfer: a surfer is a user surfing with moderate clarity of an end-goal. A surfer's actions may be somewhat exploratory in the beginning, with the difference that sub- sequent searches are expected to increase the surfer's clarity of what she wants from the system.

3. Searcher: this is a user who is very clear about what she is searching for in the system. A searcher's session would typically be short, with coherent searches leading to an end-result.

In [58], authors define three levels of queries in CBIR:

- Level 1: retrieval by primitive features such as color, texture, shape or the spatial location of image elements. Typical query is query by example.

- Level 2: retrieval of objects of given type identified by derived features, with some degree of logical inference.
- Level 3: Retrieval by abstract attributes, involving a significant amount of high-level reasoning about the purpose of the objects or scenes depicted.

Levels 2 and 3 together are referred to as semantic image retrieval, and the gap between Levels 1 and 2 as the semantic gap [58], defined in [169] as follows:

Definition 1. The semantic gap is the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation.

To bridge the 'semantic gap' some semantic search strategy may be required. Techniques for semantic gap reduction fall into five categories [116]:

(1) using object ontology to define high-level concepts,

(2) using machine learning tools to associate low-level features with query concepts,

(3) introducing relevance feedback (RF) into retrieval loop for continuous learning of users' intention,

(4) generating semantic templates to support high-level image retrieval, and

(5) making use of both the visual content of images and the textual information.

Relevance feedback (RF) is an on-line processing which tries to learn the user's intentions on the fly [116]. Since every user's need and semantics grasp is different and varies with time, database categorization cannot be fixed. Thus, the total number of classes and the class membership are not available beforehand [204]. Figure 4.2 shows a common diagram flow of a relevance feedback CBIR process. The system returns the set of most similar images and visualizes them so the user can evaluate them as relevant or irrelevant. User's feedback is used then by some machine learning technique to modify the system's similarity measure in order to return a new set of images. This relevance feedback loop iteratively adapts the low-level features-based CBIR system to learn the user's query semantics.

4.4 Hyperspectral CBIR systems validation

There are two key elements of the validation process. First, the performance measures used to compare the diverse instances of the system obtained by different parametrizations. Second, the strategy followed to obtain the reference measures. Namely, how the ground truth references are obtained and used to compute the benchmarking performance measures. For validation experiments using synthetic datasets and real data a similar methodology can be applied. However there is some difference on the kind of hypothesis tested: Synthetic

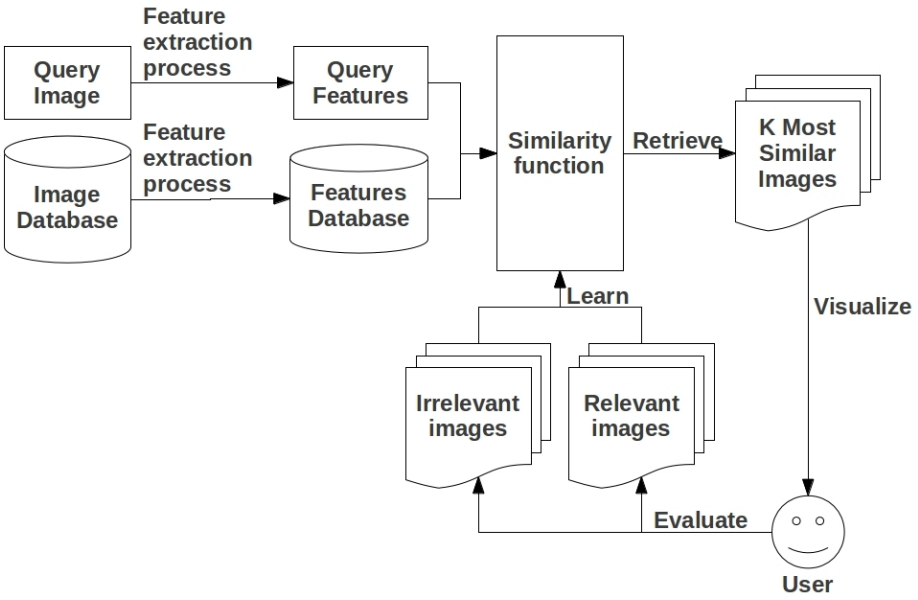


Figure 4.2: Flow diagram of a common relevance feedback CBIR process.

datasets serve to test the system robustness against changes in its internal components as well as against noisy data. Real datasets serve to test the applicability of the CBIR system in a real scenario.

The main handicaps for the evaluation of RS-CBIR systems are the lack of ground truth knowledge (categories) and the users difficulties to evaluate the returned images giving a positive/negative feedback. The former is due to the expensive, tedious and error prone groundtruth gathering process, and it is a well known problem in RS classification [12]. The later is due to difficulties in the generation of adequate visualizations (i.e. hyperspectral images) and the in interpretation by visual inspection. RS-CBIR feedback retrieval requires domain-specific user skills and new interaction methodologies yet to be developed. We sketch a RS-CBIR validation strategy that overcomes some of these problems by giving a quantitative and qualitative measure of RS-CBIR performance using only the RS data inherent structures.

4.4.1 CBIR performance measures

Evaluation metrics from information retrieval field have been adopted to evaluate CBIR systems quality. The two most used evaluation measures are *precision* and *recall* [169, 46]. Precision, p, is the fraction of the returned images that are relevant to the query. Recall, q, is the fraction of returned relevant images respect to the total number of relevant images in the database according to *a priori* knowledge. If we denote T the set of returned images and R the set of

all the images relevant to the query, then

$$p = \frac{|T \cap R|}{|T|} \tag{4.1}$$

$$r = \frac{|T \cap R|}{|R|} \tag{4.2}$$

Precision and recall follow inverse trends when considered as functions of the scope of the query. Precision falls while recall increases as the scope increases. Thus, precision and recall measures are usually given as precision-recall curves for a fixed scope. To evaluate the overall performance of a CBIR system, the Average Precision and Average Recall are calculated over all the query images in the database. For a query of scope k, these are defined as:

$$\bar{p}_k = \frac{1}{N} \sum_{\alpha=1}^{N} p_k(H_\alpha) \tag{4.3}$$

and

$$\bar{r}_k = \frac{1}{N} \sum_{\alpha=1}^{N} r_k(H_\alpha). \tag{4.4}$$

Some generalization plots were proposed by [92]. Generality, defined as $g = \frac{N_\alpha}{N}$, measures the degree of embedding of the relevant class into the database and goes to zero as the size of the database grows. The relative scope is defined as $a = \frac{N_\alpha}{k}$. When $a = 1$ the recall and precision values are equal. The Generality-Precision=Recall plots show the values of the precision at relative scope $a = 1$ for growing g in a logarithmic scale. The aim is to assess the effect of the database growth on the system performance.

The Normalized Rank [129] is a performance measure used to summarize system's performance into an scalar value. The normalized rank for a given image ranking Ω_α, denoted as $\mathrm{Rank}\,(H_\alpha)$, is defined as:

$$\mathrm{Rank}\,(H_\alpha) = \frac{1}{NN_\alpha} \left(\sum_{i=1}^{N_\alpha} \Omega_\alpha^i - \frac{N_\alpha\,(N_\alpha - 1)}{2} \right), \tag{4.5}$$

where N is the number of images in the dataset, N_α is the number of relevant images for the query H_α, and Ω_α^i is the rank at which the i-th image is retrieved. This measure is 0 for perfect performance, and approaches 1 as performance worsens, being 0.5 equivalent to a random retrieval. The average normalized rank, ANR, for the full dataset is given by:

$$\mathrm{ANR} = \frac{1}{N} \sum_{\alpha=1}^{N} \mathrm{Rank}\,(H_\alpha)\,. \tag{4.6}$$

4.4.2 General validation methodology

Algorithm 4.1 specifies the general validation methodology followed in all the experiments that will be reported in the next chapters dealing with RS-CBIR systems. It is composed of three phases. First, we perform off-line the feature extraction of the images in the given dataset. In the second phase, we obtain the CBIR system's performance measures for each image in the dataset. Finally, we calculate an averaged performance of the whole CBIR system by averaging the performance measures obtained for all the images in the dataset.

For each hyperspectral image H_α in the dataset we calculate the dissimilarity measure between H_α and each of the remaining images in the dataset. These dissimilarities are represented as a vector $\mathbf{s}_\alpha = [s_{\alpha 1}, \ldots, s_{\alpha N}]$, where N is the number of images in the dataset and $s_{\alpha,\beta}$ is the dissimilarity between the images H_α and H_β, with $\alpha, \beta = 1, \ldots, N$. The ranking of the dataset relative to the query image $\Omega_\alpha = [\omega_{\alpha,p} \in \{1, \ldots, N\}; p = 1, \ldots, N]$ is defined as the set of image indexes ordered according to increasing values of their corresponding entries in the dissimilarity vector $\mathbf{s}_\alpha$. That is, Ω_α is obtained sorting in increasing order the components of $\mathbf{s}_\alpha$ so that $s_{\alpha,\omega_{\alpha,p}} \leq s_{\alpha,\omega_{\alpha,p+1}}$. Each hyperspectral image H_α is used to query the database $Q_k(H_\alpha)$, where k is the scope of the query taking values in the range $1 \leq k \leq N$. The k most similar (less dissimilar) images H_β in the dataset relative to the image H_α are returned by the system. The set of returned images for query $Q_k(H_\alpha)$ is denoted as $T_k(H_\alpha)$. The set of relevant images $V_k(H_\alpha)$ for the query $Q_k(H_\alpha)$ is obtained from the ground-truth. The way the set of relevant images is defined depends on the strategy used to set the ground-truth. Next, we use the ranking Ω_α and the returned and relevant sets, $T_k(H_\alpha)$ and $V_k(H_\alpha)$, to calculate the precision (equation (4.1)), recall (equation (4.2)) and rank (equation (4.5)) for the query image H_α.

Finally, the performance measures obtained for all the images in the dataset are averaged to calculate the averaged precision (equation (4.3)), averaged recall (equation (4.4)) and averaged normalized rank (equation (4.6)) of the whole CBIR system.

4.4.3 Validation using synthetic datasets

Figure 4.3 shows a diagram of the CBIR validation methodology using synthetic datasets. First, a synthetic dataset must be created by some synthetic image generation process. In order to do that, some groundtruth features should be used and stored in a ground-truth features dataset. By other hand, some feature extraction algorithms would be used to extract the features from each image in the synthetic dataset. These features would be stored in an extracted features dataset as well. Then, we follow the general methodology as explained above where the relevant set of a given query image is defined over the ranking obtained using the groundtruth features.

The validation process compares the ground-truth induced image ordering with the ordering induced by the dissimilarity function over the extracted features. The vector of ground-truth dissimilarities computed using the knowledge

Algorithm 4.1 General validation methodology

1. Extract off-line the features from the dataset images.
2. For each image H_α in the dataset, $\alpha = 1, \ldots, N$:
   (a) Compute the dissimilarities between H_α and all the images in the dataset: $\mathbf{s}_\alpha = [s_{\alpha 1}, \ldots, s_{\alpha N}]$.
   (b) Obtain the rankings from $\mathbf{s}_\alpha$: $\Omega_\alpha = [\omega_{\alpha,p} \in \{1, \ldots, N\}; p = 1, \ldots, N]$ such that $s_{\alpha,\omega_{\alpha,p}} \leq s_{\alpha,\omega_{\alpha,p+1}}$.
   (c) For each scope value k, $k = 1, \ldots, N$:
      i. Return the k most similar images respect to H_α: $T_k(H_\alpha)$.
      ii. Obtain the relevant set for the query $Q_k(H_\alpha)$ in base to the ground-truth: $V_k(H_\alpha)$.
   (d) Obtain the performance measures for the image H_α using Ω_α, $T_k(H_\alpha)$ and $Q_k(H_\alpha)$.
3. Calculate the averaged performance measures for all images.

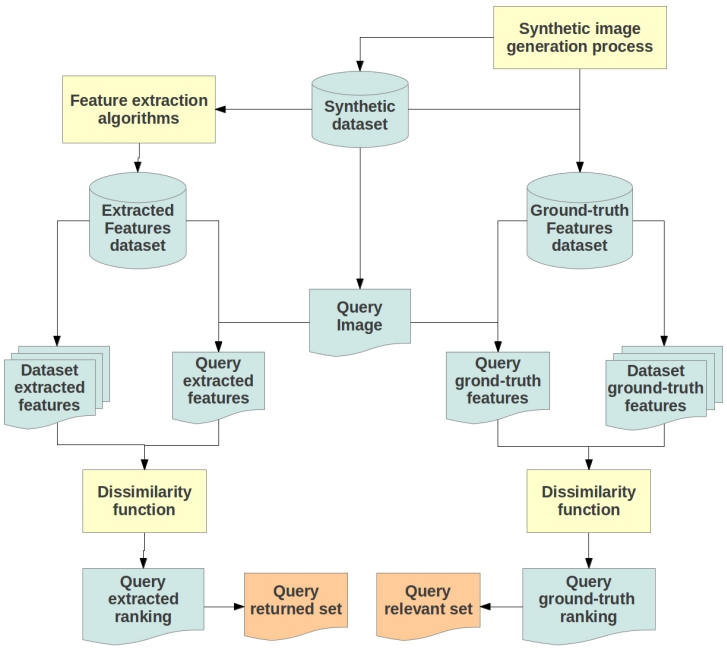


Figure 4.3: CBIR validation using synthetic datasets diagram

of the image ground-truth features is denoted as $\mathbf{s}_\alpha^{\mathrm{GT}}$. The vector of dissimilarities computed using the features extracted by some feature extraction algorithms is denoted $\mathbf{s}_\alpha^{\mathrm{EXT}}$. We distinguish rankings $\Omega_\alpha^{\mathrm{GT}}$ and $\Omega_\alpha^{\mathrm{EXT}}$ corresponding to the dissimilarities computed on the ground-truth and extracted features respectively. The set of returned images $T_k(H_\alpha)$ and the set of relevant images $V_k(H_\alpha)$ for a query $Q_k(H_\alpha)$ are defined as follows:

$$T_k(H_\alpha) = \Omega_{\alpha,k}^{\mathrm{EXT}} = \left[\omega_{\alpha,p}^{\mathrm{EXT}} \text{ s.t. } s_{\alpha,\omega_{\alpha,p}^{\mathrm{EXT}}} \leq s_{\alpha,\omega_{\alpha,k}^{\mathrm{EXT}}}\right], \tag{4.7}$$

$$V_k(H_\alpha) = \Omega_\alpha^{\mathrm{GT}} = \left[\omega_{\alpha,p}^{\mathrm{GT}} \text{ s.t. } s_{\alpha,\omega_{\alpha,p}^{\mathrm{GT}}} \leq t\right], \tag{4.8}$$

where t is a threshold value that can be set either by the k-th ground-truth ranked image, $t = s_{\alpha,\omega_{\alpha,k}^{\mathrm{GT}}}$, or by statistics calculated over the ground-truth dissimilarity vector $\mathbf{s}_\alpha^{\mathrm{GT}}$, for instance $t = \bar{s}_\alpha^{\mathrm{GT}} - 2\sigma_{\mathbf{s}_\alpha^{\mathrm{GT}}}$, such that $\bar{s}_\alpha^{\mathrm{GT}}$ and $\sigma_{\mathbf{s}_\alpha^{\mathrm{GT}}}$ are the mean and standard deviation of $\mathbf{s}_\alpha^{\mathrm{GT}}$ respectively. These definitions allows the cardinality of both returned and relevant sets to be bigger than the scope k.

4.4.4 Validation using real datasets

Figure 4.4 shows a diagram of the CBIR validation methodology using real hyperspectral datasets. First, some feature extraction algorithms would be used to extract the features from each image in the dataset. These features would be stored in an extracted features dataset. By other hand, real images should be assigned to predefined categories by some categorization process. Then, we follow the general methodology as explained above where the relevant set of a given query image is defined over the categories of the query image and the images in the dataset.

The validation process evaluates the ranking induced by the dissimilarity function over the extracted features using the a-priori known image categories. The vector of dissimilarities computed using the features extracted by some feature extraction algorithms is denoted $\mathbf{s}_\alpha^{\mathrm{EXT}}$, and the ranking is denoted as $\Omega_\alpha^{\mathrm{EXT}}$. The set of returned images $T_k(H_\alpha)$ and the set of relevant images $V_k(H_\alpha)$ for a query $Q_k(H_\alpha)$ are defined as follows:

$$T_k(H_\alpha) = \Omega_{\alpha,k}^{EXT} = \left[\omega_{\alpha,p}^{EXT} \text{ s.t. } s_{\alpha,\omega_{\alpha,p}^{IND}} \leq s_{\alpha,\omega_{\alpha,k}^{IND}}\right] \tag{4.9}$$

$$V_k(H_\alpha) = \left[\beta \text{ s.t. } \mathcal{C}(\beta) = \mathcal{C}(\alpha)\right] \tag{4.10}$$

where $\mathcal{C}(\gamma)$ indicates the category to which the patch H_γ belongs. This way, the relevant set for a query image H_α is formed for all patches belonging to its same category $\mathcal{C}(\alpha)$.

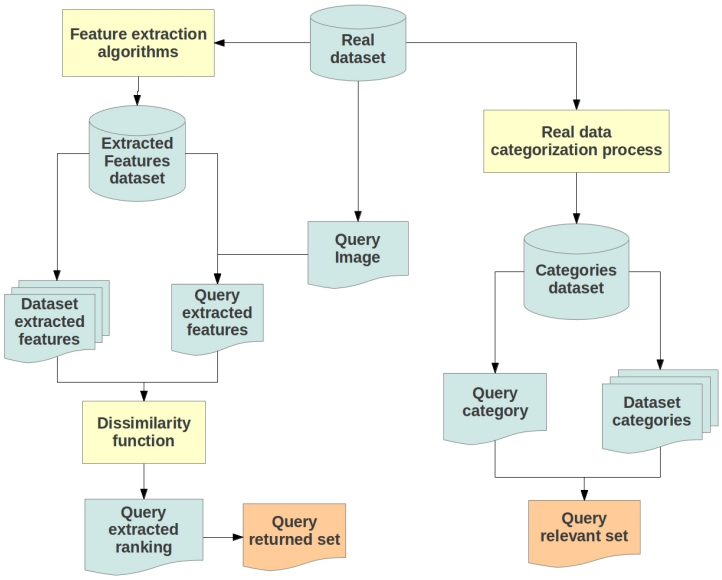


Figure 4.4: CBIR validation using real datasets diagram

4.5 Validation without *a priori* knowledge

CBIR systems are query-dependent, that is, its performance depends not only on the dataset $\mathcal{D}$, but on its ability to suit each user goals, which are specified on the query. Thus, the expected response of an ideal CBIR system to a potential user's query can be seen as a relevant/irrelevant mapping of the dataset $\mathcal{D}$ in an unknown feature space, Ω. We name this ideal mapping as a *potential search map*. The performance of a CBIR system is related to its capacity to elaborate a mapping equivalent to the potential search map for any given query. To do that, the CBIR system should use the user's feedback to iteratively transform the original feature space, Φ, and/or modify the dissimilarity function, so the response of the CBIR system to a query will converge on a map, equivalent to the potential search map corresponding to the query.

The problem of validation without a groundtruth for remote sensing classification assessment has been addressed in [11]. Authors in [11] propose a data-driven thematic map quality assessment strategy, named DAMA, suitable for comparative purposes when competing discrete mapping products are provided with little or no ground truth knowledge. It exploits a large number of implicit reference samples extracted from multiple reference cluster maps generated from unlabeled blocks of the input RS image, that are clustered separately to detect genuine, but small, image details at the cost of little human supervision. Thus, the output consists of unsupervised relative quantitative indexes (unsupervised *map quality* measures, in contrast to traditional supervised *map accuracy* measures) of labeling and segmentation consistency between every competing map

and the set of multiple reference cluster maps.

The goal is to compute labeling and segmentation indexes of the consistency between a map x generated from a digital input image z, and multiple cluster maps generated from z without employing any prior knowledge. The procedure consists of three steps:

1. Locate across raw image z several blocks of unlabeled data, $\{s_{z_i} \subseteq z, i = 1, \ldots, Q\}$, using no prior knowledge and with a minimum of human intervention. These *unlabeled candidate representative raw areas*, s_{z_i}, have to satisfy some heuristic constraints: (a) be sufficiently small so that it is easy to analyze it by clustering algorithms, and (b) contain at least two of the cover types of interest according to photo-interpretation criteria. Each land cover type must appear in one or more blocks, and the set of blocks should be sufficiently large to provide a statistically valid dataset of independent samples and to be representative of all possible variations in each land cover.

2. Each block s_{z_i} is subject to clustering separately, generating Q independent so-called multiple reference cluster maps, $\{x_i^*, i = 1, \ldots, Q\}$.

3. Estimate the labeling (class) and segmentation (spatial) agreement between each reference cluster map x_i^* and the portion of the test map, x_i corresponding to the block.

4. Combine independently the spatial and agreement fidelity results collected by submaps according to empirical (subjective) image quality criteria.

We were inspired by the DAMA strategy to propose a RS-CBIR validation strategy that could deal with the scarcity of labeled data. As DAMA, we propose to perform clustering processes over a dataset $\mathcal{D}$ to discover data inherent structures. However, we propose to use this inherent structures to simulate potential user queries. Each clustering can be used to model potential search maps of a family of queries $\mathcal{Q}$. Then, the simulated potential search maps are used to provide precision and recall measures, that show the RS-CBIR system capacity to solve the family of queries $\mathcal{Q}$ on a dataset $\mathcal{D}$.

A clustering models a family of queries $\mathcal{Q}$ by defining a set of potential search maps. Different clusterings can be obtained to represent different families of queries. It is supposed that potential users queries have some coherence, so the relevant classes are represented by some clustering mapping in an unknown feature space Ω. We do the inverse process, expecting a given clustering of the dataset $\mathcal{D}$ in some feature space, that could be different from the one used on the RS-CBIR system, represents a potential users kind of queries, a query family.

The procedure is as following:

1. Perform a clustering over a dataset $\mathcal{D}$. The clustering process defines a mapping $X^* = \{x_1^*, \ldots, x_n^*\}$, where x_i^* indicates the cluster c_l, $l = 1, \ldots, C$, image i belongs to, where C is the number of clusters found in the clustering process. This clustering represents a family of queries

$\mathcal{Q} = \{Q(x_1), \ldots, Q(x_n)\}$ where each query $Q(x_j)$, $j = 1, \ldots, n$, is a query by example defined over the sample image x_j.

2. Given a RS-CBIR system, calculate its response to each of the queries $Q(x_j) \in \mathcal{Q}$, $Q(x_j) = \{x_{i_1}, \ldots, x_{i_n}\}_{j=1}^{n}$, which can be represented as a matrix $M = \{m_{ij}\}$, $i, j = 1, \ldots, n$, so m_{ij} indicates the i-th most similar image to the query Q_j.

3. Being k the query scope, the set of returned images $T_k(x_j)$ and the set of all the relevant images $R_k(x_j)$ to a query $Q_k(x_j)$ are given by:

$$T_k(x_j) = \left\{\cup_{i=1}^{k} m_{ij}\right\} \tag{4.11}$$

$$R_k(x_j) = \left\{\cup_{x_i^* = l} x_i; x_j \in c_l\right\} \tag{4.12}$$

4. Now the precision and recall measures for the query $Q_k(x_j)$ can be calculated by substituting (4.11),(4.12) in equations (4.1),(4.2). The average of the precision and recall measures estimated by all the queries $Q_k(x_j) \in \mathcal{Q}$ is a quality assessment of the RS-CBIR system response respect to the family of queries $\mathcal{Q}$ on $\mathcal{D}$.

Chapter 5

Hyperspectral CBIR systems: spectral dissimilarity

This chapter proposes a specific Content-Based Image Retrieval (CBIR) system for hyperspectral images exploiting its rich spectral information. The CBIR image features are the endmember signatures obtained from the image data by Endmember Induction Algorithms (EIA). This chapter introduces a dissimilarity measure between hyperspectral images computed over the image induced endmembers, proving that it complies with the axioms of a distance. We provide a comparative discussion of dissimilarity functions, and quantitative evaluation of their relative performances on a large collection of synthetic hyperspectral images, and on a dataset extracted from a real hyperspectral image. Alternative dissimilarity functions considered are the Hausdorff distance and robust variations of it. We assess the CBIR performance sensitivity to changes in the distance between endmembers, the EIA employed, and some other conditions. The proposed hypespectral image distance improves over the alternative dissimilarities in all quantitative performance measures. The visual results of the CBIR on the real image data demonstrate its usefulness for practical applications.

The structure of the Chapter is as follows: Section 5.1provides an introduction to the chapter. In Section 5.2 defines the proposed hyperspectral image dissimilarity, proving that it is a distance. Section 5.3 introduces the dissimilarity functions used for comparison. Section 5.4 gives comparative and sensitivity results computed on the synthetic hyperspectral image databases. Section 5.5 gives results on dataset composed of image blocks extracted from a large real hyperspectral image. Finally, we give some conclusions in section 5.6.

5.1 Introduction

Recalling the ideas in Chapter 4, there are two main elements in the definition of a CBIR system [169]: (a) the image features and the corresponding feature extraction process, and (b) the similarity measure defined on the feature space

guiding the database search. Hyperspectral images contain rich spectral information, therefore it is natural to define spectral image features. As far as we know the only attempts to propose such a spectral image characterization to guide CBIR search in hyperspectral image databases are [141, 186]. Assuming the image formation framework given by the linear mixing model [98], the image pixel spectra are a linear combination of elementary material signatures, called endmembers. Therefore, endmembers are appropriate spectral features providing a global characterization of the image. Endmember Induction Algorithms (EIA) induce the endmember signatures from the image data, providing an autonomous processes for feature extraction which can be applied to each image independently. Thus, we are interested in this chapter in assessing the descriptive power of endmembers for database search, and on the value of Endmember Induction Algorithms (EIA) as feature extraction processes.

The definition of an appropriate dissimilarity function must take into account that the endmembers are global image features, so that metrics or semimetrics defined for pattern matching in images can be less appropriate for this kind of data. Specifically the notion of outlier due to occlusions in shape detection [168], does not apply well to sets of endmembers. We need a dissimilarity measure which decreases when the hyperspectral images have more similar elementary materials, as characterized by their spectral signatures, regardless of their spatial distribution on the image. This chapter introduces a formal distance between hyperspectral images based on spectral information. This hyperspectral image distance is a function of the distances between individual image endmembers. Therefore, the whole approach may be sensitive to the endmember distance and the applied EIA. We compare the proposed distance with the Hausdorff distance and a robust Least Trimmed Squares (LTS) modification of the Hausdorff distance [168].

The issue of validation, or evaluation, of the CBIR approach is not trivial. CBIR systems are desired to handle large amounts of data, however the available labeled information is scarce. One useful approach is to provide an extensive collection of synthetic images whose meaning is well known, i.e. we have perfect knowledge of the ground truth allowing exact computation of the performance measures. Other approach is to perform bootstrapping of image blocks of a large image containing well identified regions that can provide a convenient image block labeling. We have followed both approaches in the experimental validation of the CBIR system constituted by the endmember induction and the proposed distance.

5.2 Proposed spectral dissimilarity

Let it be $E_\alpha = \left\{\mathbf{e}_1^\alpha, \mathbf{e}_2^\alpha, \ldots, \mathbf{e}_{p_\alpha}^\alpha\right\}$ the set of endmembers induced from the hyperspectral image H_α by some endmember induction algorithm (EIA), where p_α is the number of induced endmembers from the α-th image. A dissimilarity function between two hyperspectral images, $s\left(H_\alpha, H_\beta\right)$ is defined on the basis of the distances between their corresponding set of endmembers E_α and E_β. We

compute the Spectral Distance Matrix, $D_{\alpha,\beta}$, whose elements are the distances between each pair of endmembers from each image:

$$D_{\alpha,\beta} = [d_{i,j};\ i = 1,\dots,p_\alpha;\ j = 1,\dots,p_\beta]\,, \tag{5.1}$$

where $d_{i,j}$ is the distance between the endmembers $\mathbf{e}_i^\alpha, \mathbf{e}_j^\beta \in \mathbb{R}^q$.

Here we consider the Euclidean distance, d_E:

$$d_E(\mathbf{e}_1, \mathbf{e}_2) = \|\mathbf{e}_1 - \mathbf{e}_2\|\,, \tag{5.2}$$

and the angular distance, a.k.a. Spectral Angle Mapper (SAM) distance in remote sensing applications [29], d_S:

$$d_S(\mathbf{e}_1, \mathbf{e}_2) = \cos^{-1}\left(\frac{\mathbf{e}_1 \cdot \mathbf{e}_2}{\|\mathbf{e}_1\|\,\|\mathbf{e}_2\|}\right), \tag{5.3}$$

where $\mathbf{e}_1 \cdot \mathbf{e}_2$ denotes the vector inner product.

We compute the vectors of row and column minimal values of $D_{\alpha,\beta}$ respectively denoted as $\mathbf{m}_\alpha^\beta = \left[m_{\alpha,1}^\beta, \dots, m_{\alpha,p_\alpha}^\beta\right]$ and $\mathbf{m}_\beta^\alpha = \left[m_{\beta,1}^\alpha, \dots, m_{\beta,p_\beta}^\alpha\right]$. Their components are computed as

$$m_{\alpha,i}^\beta = \min_{j=1,\dots,p_\beta}\{d_{i,j}\},\ i = 1,\dots,p_\alpha$$

$$m_{\beta,j}^\alpha = \min_{i=1,\dots,p_\alpha}\{d_{i,j}\},\ j = 1,\dots,p_\beta.$$

Definition 2. The spectral dissimilarity between two hyperspectral images, H_α, H_β, is given by the following expression:

$$s(H_\alpha, H_\beta) = \left\|\mathbf{m}_\alpha^\beta\right\| + \left\|\mathbf{m}_\beta^\alpha\right\|. \tag{5.4}$$

The magnitude of the row minimum vector $\left\|\mathbf{m}_\alpha^\beta\right\|$ is a measure of the closeness of the endmembers of H_α to any endmember of H_β. Conversely, The magnitude of the column minimum vector $\left\|\mathbf{m}_\beta^\alpha\right\|$ is a measure of the closeness of the endmembers of H_β to any endmember of H_α. In the following we prove that this dissimilarity is a distance.

Lemma 3. *The dissimilarity function of Definition* (16) *is a semimetric.*

Proof. The dissimilarity function of Definition (16) complies with the axioms required to be a semimetric for any hyperspectral images H_α, H_β:

- Non-negativity: $s(H_\alpha, H_\beta) \geq 0$ because both norms are non-negative $\left\|\mathbf{m}_\alpha^\beta\right\| \geq 0$, $\left\|\mathbf{m}_\beta^\alpha\right\| \geq 0$.
- Identity of indiscernibles: it is restricted to the equivalence of images whose endmembers obtained by the EIA are the same.

  - If $E_\alpha = E_\beta$ then both row and column minimal vectors will be null $\left\|\mathbf{m}_\alpha^\beta\right\| = \left\|\mathbf{m}_\beta^\alpha\right\| = 0$, because the diagonal of $D_{\alpha,\beta}$ will be zero, $d_{i,i} = 0$, then $s(H_\alpha, H_\beta) = 0$. This is true even if the EIA finds a permutation of the endmembers when applied to the same or different images.
  - On the other sense, if $s(H_\alpha, H_\beta) = 0$ then all the components of the row and column minimal vectors must be zero, $m_{\alpha,i} = m_{\beta,j} = 0$ for $i, j = 1, \ldots, p$. Therefore, for each endmember in E_α there is another identical to it in E_β and, conversely, for each endmember in E_β there is another identical to it in E_α. Thus, $E_\alpha = E_\beta$.

- Symmetry: $s(H_\alpha, H_\beta) = s(H_\beta, H_\alpha)$ is immediate if we note that $D_{\alpha,\beta} = (D_{\beta,\alpha})^T$. The row minimal vector of $D_{\alpha,\beta}$ will be identical to the column minimal vector of $(D_{\beta,\alpha})^T$, and the column minimal vector of $D_{\alpha,\beta}$ will be identical to the row minimal vector of $(D_{\beta,\alpha})^T$. Therefore the expressions of $s(H_\alpha, H_\beta)$ and $s(H_\beta, H_\alpha)$ are identical by the commutativity of addition.

□

To prove that the dissimilarity of Definition (16) is a distance we need to prove that it satisfies the triangle inequality. We proceed proving first that the norms of the vectors of minimum values satisfy the triangle inequality. Then, it follows immediately from these results that the dissimilarity satisfies the triangle inequality.

Lemma 4. *Given three hyperspectral images $H_\alpha, H_\beta, H_\gamma$, with corresponding induced sets of endmembers $E_\alpha, E_\beta, E_\gamma$, the following triangle inequality holds*

$$\|\mathbf{m}_\alpha^\gamma\| \leq \left\|\mathbf{m}_\alpha^\beta\right\| + \left\|\mathbf{m}_\beta^\gamma\right\|.$$

Proof. For each $\mathbf{e}_i^\alpha \in E_\alpha$ there is a $\mathbf{e}_{j'}^\beta \in E_\beta$ such that $d\left(\mathbf{e}_i^\alpha, \mathbf{e}_{j'}^\beta\right) = \min_{j=1,\ldots,p_\beta} d\left(\mathbf{e}_i^\alpha, \mathbf{e}_j^\beta\right)$. In the same way, there is a $\mathbf{e}_{k'}^\gamma \in E_\gamma$ such that $d\left(\mathbf{e}_{j'}^\beta, \mathbf{e}_{k'}^\gamma\right) = \min_{k=1,\ldots,p_\gamma} d\left(\mathbf{e}_{j'}^\beta, \mathbf{e}_k^\gamma\right)$. By the triangle inequality of the distance between endmembers we have that:

$$d(\mathbf{e}_i^\alpha, \mathbf{e}_{k'}^\gamma) \leq d\left(\mathbf{e}_i^\alpha, \mathbf{e}_{j'}^\beta\right) + d\left(\mathbf{e}_{j'}^\beta, \mathbf{e}_{k'}^\gamma\right),$$

If $d(\mathbf{e}_i^\alpha, \mathbf{e}_{k'}^\gamma) = \min_{k=1,\ldots,p_\gamma} d(\mathbf{e}_i^\alpha, \mathbf{e}_k^\gamma)$ then the proposition is true. Suppose that there is another $\mathbf{e}_{k^*}^\gamma \in E_\gamma$ such that $d(\mathbf{e}_i^\alpha, \mathbf{e}_{k^*}^\gamma) < d(\mathbf{e}_i^\alpha, \mathbf{e}_{k'}^\gamma)$. This implies that

$$d\left(\mathbf{e}_i^\alpha, \mathbf{e}_{j^*}^\beta\right) + d\left(\mathbf{e}_{j^*}^\beta, \mathbf{e}_{k^*}^\gamma\right) < d\left(\mathbf{e}_i^\alpha, \mathbf{e}_{j'}^\beta\right) + d\left(\mathbf{e}_{j'}^\beta, \mathbf{e}_{k'}^\gamma\right).$$

We have two cases in which this could be true:

- $j^* = j'$ implies that $d\left(\mathbf{e}_{j'}^{\beta}, \mathbf{e}_{k^*}^{\gamma}\right) < d\left(\mathbf{e}_{j'}^{\beta}, \mathbf{e}_{k'}^{\gamma}\right)$ contradicting the fact that we have selected $\mathbf{e}_{k'}^{\gamma}$ as the minimum distance endmember.
- $j^* \neq j'$ implies that $d\left(\mathbf{e}_{i}^{\alpha}, \mathbf{e}_{j^*}^{\beta}\right) < d\left(\mathbf{e}_{i}^{\alpha}, \mathbf{e}_{j'}^{\beta}\right)$ contradicting the fact that we have selected $\mathbf{e}_{j'}^{\beta}$ as the minimum distance endmember.

The triangle inequality $\|\mathbf{m}_{\alpha}^{\gamma}\| \leq \left\|\mathbf{m}_{\alpha}^{\beta}\right\| + \left\|\mathbf{m}_{\beta}^{\gamma}\right\|$ follows directly from the triangle inequalities of all the entries of the distance matrix.

□

Theorem 5. *The dissimilarity function of Definition 16 is a distance.*

Proof. According to Lemma 3 the dissimilarity function of equation (5.4) is a semimetric. We refer to lemma 4 to prove that given three hyperspectral images $H_{\alpha}, H_{\beta}, H_{\gamma}$ the dissimilarity function (5.4) complies with the triangle inequality

$$s\left(H_{\alpha}, H_{\gamma}\right) \leq s\left(H_{\alpha}, H_{\beta}\right) + s\left(H_{\beta}, H_{\gamma}\right).$$

Lemmas 4 shows that triangle inequality holds for the norms $\left\|\mathbf{m}_{\alpha}^{\beta}\right\|$ and $\left\|\mathbf{m}_{\beta}^{\alpha}\right\|$ of the vectors of row and column minimal values, respectively. Furthermore, the elementary triangle inequalities of the endmembers are completely determined and the addition $\left\|\mathbf{m}_{\alpha}^{\beta}\right\| + \left\|\mathbf{m}_{\beta}^{\alpha}\right\| = s\left(H_{\alpha}, H_{\gamma}\right)$ complies with the triangle inequality. Thus, the proposed dissimilarity measure complies with the axioms of a distance. □

Remark 6. We have used in previous works [185, 186] the following definition for the dissimilarity between hyperspectral images represented by their corresponding sets of endmembers:

$$s_w\left(H_{\alpha}, H_{\beta}\right) = \left(\left\|\mathbf{m}_{\alpha}^{\beta}\right\| + \left\|\mathbf{m}_{\beta}^{\alpha}\right\|\right)\left(\left|p_{\alpha} - p_{\beta}\right| + 1\right). \tag{5.5}$$

It is easy to check that this dissimilarity meets the conditions of a semimetric following the reasoning of Lemma 3. However, it does not comply with the triangular inequality as we following prove and therefore, it is not a distance.

Lemma 7. *The triangle inequality does not hold everywhere for dissimilarity function* $s_w\left(H_{\alpha}, H_{\beta}\right)$ *of equation* (5.5).

Proof. The triangle inequality on this dissimilarity

$$s_w\left(H_{\alpha}, H_{\gamma}\right) \leq s_w\left(H_{\alpha}, H_{\beta}\right) + s_w\left(H_{\beta}, H_{\gamma}\right),$$

is rewritten as

$$\delta_{\alpha,\gamma}\left(\left|p_{\alpha} - p_{\gamma}\right| + 1\right) \leq \delta_{\alpha,\beta}\left(\left|p_{\alpha} - p_{\beta}\right| + 1\right) + \delta_{\beta,\gamma}\left(\left|p_{\beta} - p_{\gamma}\right| + 1\right),$$

where $\delta_{\alpha,\beta} = \left\|\mathbf{m}_\alpha^\beta\right\| + \left\|\mathbf{m}_\beta^\alpha\right\|$. Reorganizing we obtain

$$\delta_{\alpha,\gamma} \leq \delta_{\alpha,\beta}\frac{|p_\alpha - p_\beta| + 1}{|p_\alpha - p_\gamma| + 1} + \delta_{\beta,\gamma}\frac{|p_\beta - p_\gamma| + 1}{|p_\alpha - p_\gamma| + 1}, \tag{5.6}$$

because $|p_\alpha - p_\gamma| + 1 \geq 1$. Though the triangle inequality holds for $\delta_{\alpha,\beta}$ as proved in Theorem 5, there may be hyperspectral image triples where

$$\frac{|p_\alpha - p_\beta| + 1}{|p_\alpha - p_\gamma| + 1} \ll 1 \text{ or } \frac{|p_\beta - p_\gamma| + 1}{|p_\alpha - p_\gamma| + 1} \ll 1,$$

breaking the inequality of equation (5.6). Therefore, the triangle inequality does not hold everywhere for this dissimilarity. □

5.3 Other dissimilarity functions

Here we introduce the Hausdorff distance and variations of it we have found on the literature [168]. We reformulate them in terms of sets of endmembers and discuss on their usefulness for the problem of comparing hyperspectral images by their induced spectral information.

Definition 8. The Hausdorff distance between a pair of hyperspectral images is defined as follows:

$$s_H\left(H_\alpha, H_\beta\right) = \max\left(\max\left(\mathbf{m}_\alpha^\beta\right), \max\left(\mathbf{m}_\beta^\alpha\right)\right), \tag{5.7}$$

where H_α,H_β are hyperspectral images; and, $\mathbf{m}_\alpha^\beta$, $\mathbf{m}_\beta^\alpha$ are the vectors of row and column minimal values of $D_{\alpha,\beta}$ as it has been defined in equation (5.1).

Hausdorff distance is used for the matching of point clouds in arbitrary spaces, with applications in pattern recognition. For CBIR it is applied on the image features, allowing for comparison among images with different feature set sizes. Compared with the distance of Definition 16, the Hausdorff distance disregards information that can be relevant in our CBIR application in a way that we can not formalize easily, but that can be illustrated with an example.

Example 9. Assume that the endmember distances are normalized in $[0, 1]$ so that 1 is extreme dissimilarity. Let us have hyperspectral images H_α, H_β, and H_γ, such that $\mathbf{m}_\alpha^\beta = [0.9, 0.9, 0.9]$, $\mathbf{m}_\beta^\alpha = [0.9, 0.9]$, $\mathbf{m}_\alpha^\gamma = [0.9, 0.1, 0.1]$, $\mathbf{m}_\gamma^\alpha = [0.1, 0.1]$. Then, the similarity between images H_β, and H_γ relative to H_α will be the same according to the Hausdorff distance: $s_H\left(H_\alpha, H_\beta\right) = s_H\left(H_\alpha, H_\gamma\right) = 0.9$. However, the spectral content of H_α is more similar to H_γ than to H_β since they share two endmembers, while H_α and H_β do not share anyone. The proposed distance of Definition 16 gives a result according to our intuition $s\left(H_\alpha, H_\beta\right) = 2.8316 \gg s\left(H_\alpha, H_\gamma\right) = 1.0525$.

Definition 10. The robust M-estimation Hausdorff dissimilarity between a pair of hyperspectral images is defined as follows:

$$s_\tau\left(H_\alpha, H_\beta\right) = \max\left(h_\tau\left(\mathbf{m}_\alpha^\beta\right), h_\tau\left(\mathbf{m}_\beta^\alpha\right)\right), \tag{5.8}$$

where the robust M-estimation of the distance between endmembers is given by:

$$h_\tau\left(\mathbf{m}_\alpha^\beta\right) = \frac{1}{p_\alpha}\sum_{i=1}^{p_\alpha}\rho_\tau\left(m_{\alpha,i}^\beta\right) \text{ and } h_\tau\left(\mathbf{m}_\beta^\alpha\right) = \frac{1}{p_\beta}\sum_{j=1}^{p_\beta}\rho_\tau\left(m_{\beta,j}^\alpha\right).$$

The function $\rho_\tau(\cdot)$ is a thresholding function specified as follows:

$$\rho_\tau(x) = \begin{cases} x, & x \leq \tau \\ \tau, & x > \tau \end{cases}.$$

Notice that $m_{\alpha,i}^\beta, m_{\beta,j}^\alpha \geq 0$ and therefore, we do not need to compute $|x|$ as in [168]. The parameter τ specifies the minimum distance for an endmember to be considered as an outlier. We following prove that this dissimilarity is a semimetric but not a distance since it does not comply with triangle inequality.

Lemma 11. *The dissimilarity function $s_\tau\left(H_\alpha, H_\beta\right)$ of equation (5.8) is a semimetric.*

Proof. The dissimilarity function $s_\tau\left(H_\alpha, H_\beta\right)$ of equation (5.8) complies with the axioms required for a semimetric for any hyperspectral images H_α, H_β:

- Non-negativity: $s_\tau\left(H_\alpha, H_\beta\right) \geq 0$ since $m_{\alpha,i}^\beta, m_{\beta,j}^\alpha \geq 0$ for all $i = 1, \ldots, p_\alpha$ and $j = 1, \ldots, p_\beta$ and, there is not any operation in the computation of $s_\tau\left(H_\alpha, H_\beta\right)$ that may produce a negative outcome.
- Identity of indiscernibles: it is restricted to the equivalence of images whose endmembers obtained by the EIA are the same.
  - If $E_\alpha = E_\beta$ then $m_{\alpha,i}^\beta = m_{\beta,j}^\alpha = 0$, and consequently $s_\tau\left(H_\alpha, H_\beta\right) = 0$. This is true even if the EIA finds a permutation of the endmembers when applied to the same or different images.
  - On the other sense, if $s_\tau\left(H_\alpha, H_\beta\right) = 0$ then all $m_{\alpha,i}^\beta = m_{\beta,j}^\alpha = 0$. Therefore, for each endmember in E_α there is another identical to it in E_β and, conversely, for each endmember in E_β there is another identical to it in E_α. Thus, $E_\alpha = E_\beta$.
- Symmetry: is immediate if we note that $D_{\alpha,\beta} = \left(D_{\beta,\alpha}\right)^T$: $s_\tau\left(H_\alpha, H_\beta\right) = \max\left(h_\tau\left(\mathbf{m}_\alpha^\beta\right), h_\tau\left(\mathbf{m}_\beta^\alpha\right)\right) = \max\left(h_\tau\left(\mathbf{m}_\beta^\alpha\right), h_\tau\left(\mathbf{m}_\alpha^\beta\right)\right) = s_\tau\left(H_\beta, H_\alpha\right)$

□

Lemma 12. *The triangle inequality does not hold everywhere for dissimilarity function $s_\tau\left(H_\alpha, H_\beta\right)$ of equation (5.8).*

Proof. The triangle inequality on this dissimilarity

$$s_\tau\left(H_\alpha, H_\gamma\right) \leq s_\tau\left(H_\alpha, H_\beta\right) + s_\tau\left(H_\beta, H_\gamma\right),$$

is rewritten as

$$\max\left(h_\tau\left(\mathbf{m}_\alpha^\gamma\right), h_\tau\left(\mathbf{m}_\gamma^\alpha\right)\right) \leq \max\left(h_\tau\left(\mathbf{m}_\alpha^\beta\right), h_\tau\left(\mathbf{m}_\beta^\alpha\right)\right) + \max\left(h_\tau\left(\mathbf{m}_\beta^\gamma\right), h_\tau\left(\mathbf{m}_\gamma^\beta\right)\right).$$

From Lemma 4, for each $\mathbf{e}_i^\alpha \in E_\alpha$ there is a $\mathbf{e}_{j'}^\beta \in E_\beta$ such that $d\left(\mathbf{e}_i^\alpha, \mathbf{e}_{j'}^\beta\right) = \min_{j=1,\ldots,p_\beta} d\left(\mathbf{e}_i^\alpha, \mathbf{e}_j^\beta\right) = m_{\alpha,i}^\beta$. In the same way, there is a $\mathbf{e}_{k'}^\gamma \in E_\gamma$ such that $d\left(\mathbf{e}_{j'}^\beta, \mathbf{e}_{k'}^\gamma\right) = \min_{k=1,\ldots,p_\gamma} d\left(\mathbf{e}_{j'}^\beta, \mathbf{e}_k^\gamma\right) = m_{\beta,j'}^\gamma$. By the triangle inequality of the distance between endmembers we have that $d\left(\mathbf{e}_i^\alpha, \mathbf{e}_{k'}^\gamma\right) \leq d\left(\mathbf{e}_i^\alpha, \mathbf{e}_{j'}^\beta\right) + d\left(\mathbf{e}_{j'}^\beta, \mathbf{e}_{k'}^\gamma\right)$. We proved in Lemma 4 that $d\left(\mathbf{e}_i^\alpha, \mathbf{e}_{k'}^\gamma\right) = \min_{k=1,\ldots,p_\gamma} d\left(\mathbf{e}_i^\alpha, \mathbf{e}_k^\gamma\right) = m_{\alpha,i}^\gamma$. Therefore, the triangle inequality $m_{\alpha,i}^\gamma \leq m_{\alpha,i}^\beta + m_{\beta,j'}^\gamma$ holds. Taking into account that ρ_τ is monotonically increasing, we have $\rho_\tau\left(m_{\alpha,i}^\gamma\right) \leq \rho_\tau\left(m_{\alpha,i}^\beta\right) + \rho_\tau\left(m_{\beta,j'}^\gamma\right)$. Adding all the inequalities for the endmembers in E_α we obtain:

$$\sum_{i=1}^{p_\alpha} \rho_\tau\left(m_{\alpha,i}^\gamma\right) \leq \sum_{i=1}^{p_\alpha} \rho_\tau\left(m_{\alpha,i}^\beta\right) + \sum_{j' \in J} \rho_\tau\left(m_{\beta,j'}^\gamma\right),$$

where $J = \left\{ j' \,\middle|\, d\left(\mathbf{e}_i^\alpha, \mathbf{e}_{j'}^\beta\right) = m_{\alpha,i}^\beta; i = 1, \ldots, p_\alpha \right\}$. We allow J to be a superset, i.e. allowing repeating elements corresponding to the situation of several endmembers in E_α mapped to the same endmember in E_β. Hence, we can not prove

$$h_\tau\left(\mathbf{m}_\alpha^\gamma\right) \leq h_\tau\left(\mathbf{m}_\alpha^\beta\right) + h_\tau\left(\mathbf{m}_\beta^\gamma\right),$$

unless $h_\tau\left(\mathbf{m}_\beta^\gamma\right) \geq \sum_{j' \in J} \rho_\tau\left(m_{\beta,j'}^\gamma\right)$. We have three cases:

1. $\sum_{j' \in J} \rho_\tau\left(m_{\beta,j'}^\gamma\right) = h_\tau\left(\mathbf{m}_\beta^\gamma\right)$. This condition is true only if $|J| = p_\beta$ and there are no repeated elements in J.
2. $\sum_{j' \in J} \rho_\tau\left(m_{\beta,j'}^\gamma\right) < h_\tau\left(\mathbf{m}_\beta^\gamma\right)$. This condition is true in the following situations:
   (a) $|J| < p_\beta$ and there are no repeated elements in J, or the repeated elements are small enough and do not add up above $h_\tau\left(\mathbf{m}_\beta^\gamma\right)$.
   (b) $|J| \geq p_\beta$ and the repeated elements in J are small enough and do not add up above $h_\tau\left(\mathbf{m}_\beta^\gamma\right)$.

3. $\sum_{j' \in J} \rho_\tau \left(m^\gamma_{\beta,j'} \right) > h_\tau \left(\mathbf{m}^\gamma_\beta \right)$. This condition is true when the repeated elements in J are large enough to add up above $h_\tau \left(\mathbf{m}^\gamma_\beta \right)$, i.e. repetitions of the maximum value, whatever the relation between J and p_β.

In cases 1 and 2, the triangle inequality holds. However, in case 3 it does not hold. There is extra condition on the data that prevents case 3 from happening. Therefore, we conclude that the triangle inequality does not hold everywhere in data domain.

Following a similar reasoning we find that the triangle inequality

$$h_\tau \left(\mathbf{m}^\alpha_\gamma \right) \leq h_\tau \left(\mathbf{m}^\alpha_\beta \right) + h_\tau \left(\mathbf{m}^\beta_\gamma \right)$$

does not hold everywhere and consequently,

$$s_\tau \left(H_\alpha, H_\gamma \right) \leq s_\tau \left(H_\alpha, H_\beta \right) + s_\tau \left(H_\beta, H_\gamma \right),$$

does not neither. Thus, the dissimilarity $s_\tau \left(H_\alpha, H_\gamma \right)$ is not a metric. □

Definition 13. The robust least trimmed square (LTS) Hausdorff dissimilarity between a pair of hyperspectral images is defined as follows:

$$s_L \left(H_\alpha, H_\beta \right) = \max \left(h_L \left(\mathbf{m}^\beta_\alpha \right), h_L \left(\mathbf{m}^\alpha_\beta \right) \right), \tag{5.9}$$

where the robust LTS estimation of the distance between endmembers is given by:

$$h_L \left(\mathbf{m}^\beta_\alpha \right) = \frac{1}{Lp_\alpha} \sum_{i=1}^{Lp_\alpha} o \left(\mathbf{m}^\beta_\alpha \right)_i \text{ and } h_L \left(\mathbf{m}^\alpha_\beta \right) = \frac{1}{Lp_\beta} \sum_{j=1}^{Lp_\beta} o \left(\mathbf{m}^\alpha_\beta \right)_j .$$

The function $o \left(\mathbf{m}^\beta_\alpha \right)_i$ denotes the i-th element of vector $\mathbf{m}^\beta_\alpha$ after ordering its components in ascending order, i.e. $o \left(\mathbf{m}^\beta_\alpha \right)_1 \leq o \left(\mathbf{m}^\beta_\alpha \right)_2 \leq \ldots \leq o \left(\mathbf{m}^\beta_\alpha \right)_{p_\alpha}$. The function $o \left(\mathbf{m}^\alpha_\beta \right)_j$ is analogous for vector $\mathbf{m}^\alpha_\beta$. The parameter $L \in [0, 1]$ specifies the fraction of components of the minimum vectors $\mathbf{m}^\beta_\alpha$ and $\mathbf{m}^\alpha_\beta$ to be taken into account. We following prove that this dissimilarity is a semimetric but not a distance since it does not comply with the triangle inequality.

Lemma 14. *The dissimilarity function $s_L \left(H_\alpha, H_\beta \right)$ of equation (5.8) is a semimetric.*

Proof. The proof is identical to that of Lemma (11). □

Lemma 15. *The triangle inequality does not hold everywhere for dissimilarity function $s_L \left(H_\alpha, H_\beta \right)$ of equation (5.9).*

Proof. The triangle inequality on this dissimilarity

$$s_L \left(H_\alpha, H_\gamma \right) \leq s_L \left(H_\alpha, H_\beta \right) + s_L \left(H_\beta, H_\gamma \right),$$

is rewritten as

$$\max\left(h_L\left(\mathbf{m}_\alpha^\gamma\right), h_L\left(\mathbf{m}_\gamma^\alpha\right)\right) \leq \max\left(h_L\left(\mathbf{m}_\alpha^\beta\right), h_L\left(\mathbf{m}_\beta^\alpha\right)\right) + \max\left(h_L\left(\mathbf{m}_\beta^\gamma\right), h_L\left(\mathbf{m}_\gamma^\beta\right)\right).$$

Following the reasoning in Lemma 12 we obtain the following inequality:

$$\sum_{i=1}^{p_\alpha} m_{\alpha,i}^\gamma \leq \sum_{i=1}^{p_\alpha} m_{\alpha,i}^\beta + \sum_{j'\in J} m_{\beta,j'}^\gamma,$$

where $J = \left\{j' \middle| d\left(\mathbf{e}_i^\alpha, \mathbf{e}_{j'}^\beta\right) = m_{\alpha,i}^\beta; i = 1,\ldots,p_\alpha\right\}$. We allow J to be a superset, i.e. allowing repeating elements corresponding to the situation of several endmembers in E_α mapped to the same endmember in E_β. Without loss of generality, we consider $L = 1$, then we have

$$h_L\left(\mathbf{m}_\alpha^\gamma\right) \leq h_L\left(\mathbf{m}_\alpha^\beta\right) + \sum_{j'\in J} m_{\beta,j'}^\gamma.$$

Hence, we can not prove

$$h_L\left(\mathbf{m}_\alpha^\gamma\right) \leq h_L\left(\mathbf{m}_\alpha^\beta\right) + h_L\left(\mathbf{m}_\beta^\gamma\right),$$

unless $h_\tau\left(\mathbf{m}_\beta^\gamma\right) \geq \sum_{j'\in J} m_{\beta,j'}^\gamma$. It is easy to find a configuration of the set of endmembers where this condition is violated. For instance, when $|J| = p_\beta$ and there are several repetitions of the greatest value of $\mathbf{m}_\beta^\gamma$ in J. Therefore, the triangle inequality does not hold everywhere even when $L = 1$. Following a similar reasoning we find that the triangle inequality

$$h_L\left(\mathbf{m}_\gamma^\alpha\right) \leq h_L\left(\mathbf{m}_\beta^\alpha\right) + h_L\left(\mathbf{m}_\gamma^\beta\right)$$

does not hold everywhere and consequently,

$$s_L\left(H_\alpha, H_\gamma\right) \leq s_L\left(H_\alpha, H_\beta\right) + s_L\left(H_\beta, H_\gamma\right),$$

does not hold neither. Thus, the dissimilarity $s_\tau\left(H_\alpha, H_\gamma\right)$ is not a metric. □

Both robust M-estimation and LTS Hausdorff semimetrics are designed to cope with outlier points in image pattern matching applications looking for point correspondences [168], where the outliers come from the occurrence of occlusions inducing confusion between different shapes or background noise. Therefore, these dissimilarities may be expected to be biased to find partial matchings of the data points. From the point of view of the problem at hands these dissimilarities have two inconveniences. First, endmembers are global descriptors of the image content, not strictly associated with an image coordinate or point. Therefore, the concept of outlier as such does not apply well in our CBIR system: outlier endmembers represent a wide variation of a holistic image

feature, i.e. the appearance of a different material, regardless of image region distributions. Diminishing or bounding the effect of widely different endmembers forces the similarity between images containing very disparate materials. Second, the estimation and validation of the appropriate parameter settings, either the threshold τ or the order statistic L, implies a delicate and complex computational process, which does not add general value to the approach because any setting will often be specific of the particular data domain.

5.4 Experiments on synthetic data

Following the procedures detailed in Appendix C we generated three sets of candidate ground-truth endmembers, with 5, 10 and 20 endmembers each, representing an increasing diversity of the materials that can be found in the scene. We denote the datasets generated from each set of candidate ground-truth endmembers pool as 5-E, 10-E and 20-E datasets respectively. We also defined three categories of image spatial size, with images having 64×64, 128×128 and 256×256 pixels. Increasing image size, we increase the available information and the computational cost. We have synthesized a total of 18000 hyperspectral images divided in nine datasets of 2000 images each.

5.4.1 Experimental methodology

We have performed independent query computational experiments over each of the nine synthetic hyperspectral image datasets using the distance of Definition 16, hereafter denoted as "Grana" in the plots and tables. The experiments are of two kinds. First we test the sensitivity of the approach to the distance between individual endmembers, and the sensitivity to the EIA used to induce the endmembers. We test the Euclidean distance of equation (6.3), and the SAM distance of equation (6.4). The EIA tested are the N-FINDER and the EIHA algorithms. Second, we compare the results of the proposed "Grana" distance with those obtained computing the Hausdorff distance [50, 100] on the same image features (the endmembers).

The validation process compares the ground-truth induced image ordering with the ordering induced by the dissimilarity function. The vector of ground-truth dissimilarities computed using the knowledge of the image ground truth endmembers is denoted $\mathbf{s}_\alpha^{GT}$. The vector of dissimilarities computed using the endmembers induced by one of the EIAs (either N-FINDER or EIHA) is denoted $\mathbf{s}_\alpha^{IND}$. The ground-truth dissimilarity $s_{\alpha,\beta}^{IND}$ between the images H_α and H_β is computed using either the spectral dissimilarity function of Definition 16 or the Hausdorff distance of Definition 8. Computation of performance measures recall and precision are computed as described in Chapter 4 using the query scope to determine the size of the set of relevant images.

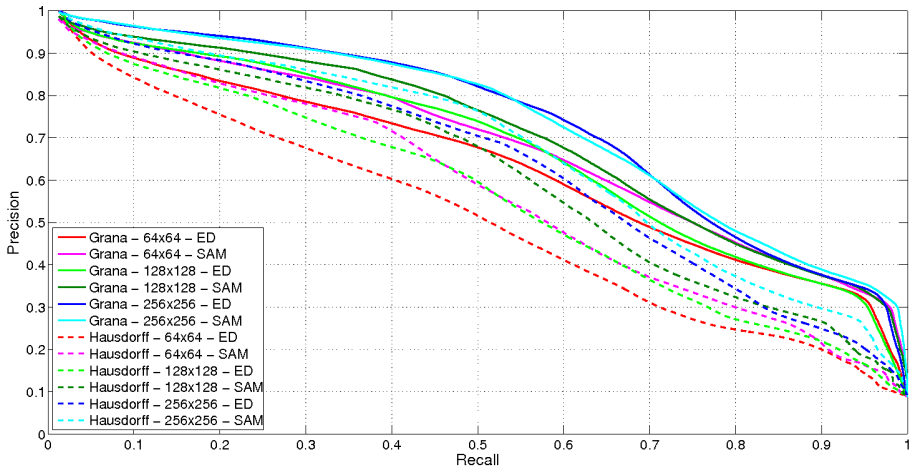


Figure 5.1: Precision and Recall results for the 5-E dataset using the EIHA endmember induction algorithm for the Grana distance and the Hausdorff distance.

5.4.2 Performance results

In the figures and tables below we call "Grana" the distance of Definition 16. We performed the following computational experiments: for each dataset of 2000 synthetic images we computed independently the Precision and Recall obtained using the endmembers induced by the alternative EIA (N-FINDER and EIHA) and alternative distances (Euclidean and SAM), for the Grana and Hausdorff distances. We present these results in figures 5.1-5.6. Each figure corresponds to a dataset and EIA. Each figure shows twelve precision-recall curves corresponding to each combination of the dissimilarity function (Grana or Hausdorff), spatial image size (64×64, 128×128 or 256×256) and endmember distance (Euclidean or SAM). Tables 5.1, 5.2 and 5.3 show the area under the precision-recall curves and the average normalized rank for each of the 5-D, 10-D and 20-D datasets.

5.4.3 Discussion of results

The discussion of the results is structured into a series of relevant questions.

Classification results Can the proposed CBIR approach discover the underlying classification of the images induced by the ground truth endmembers?. Artificial classes could be defined by images having a zero distance on the ground truth endmembers. These classes will be small sized, therefore to determine if the CBIR really uncovers the ground-truth classification we need to look at the values of Precision for low Recall values, which are pretty high in all the plots. Increasing the query scope many irrelevant images are forced into the answer of the query, reducing Precision and increasing Recall. For all combinations of the

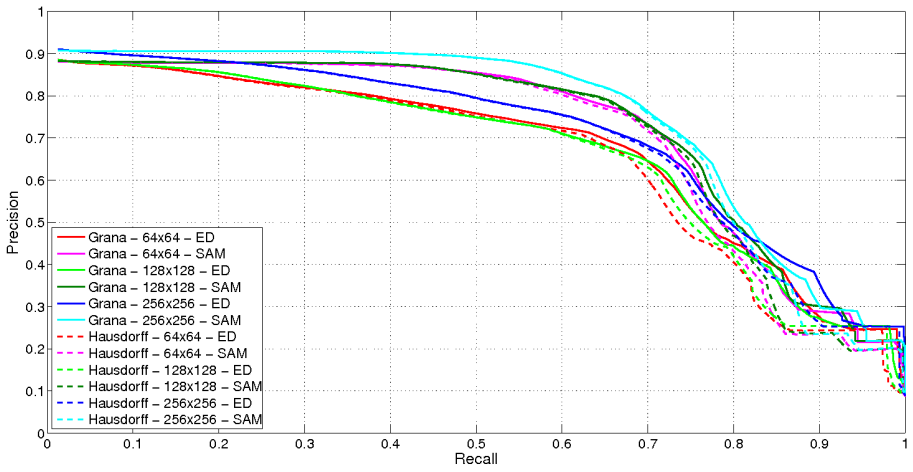


Figure 5.2: Precision and Recall results for the 5-E dataset using the N-FINDER endmember induction algorithm for the Grana distance and the Hausdorff distance.

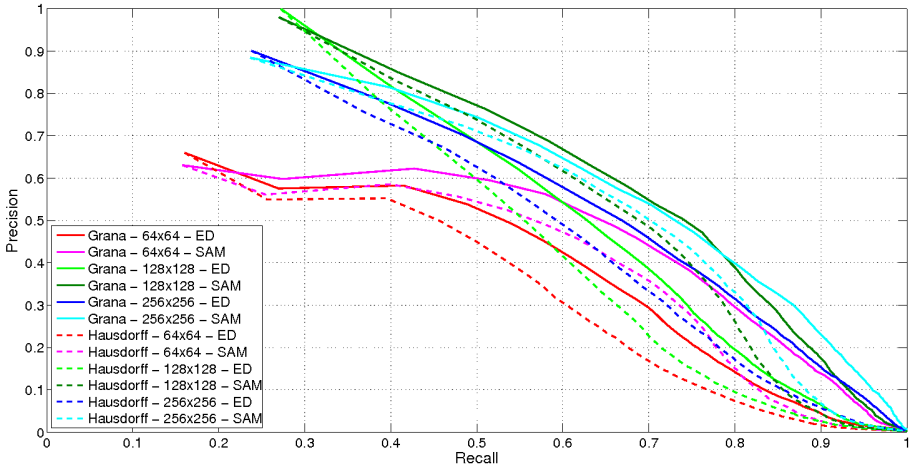


Figure 5.3: Precision and Recall results for the 10-E dataset using the EIHA endmember induction algorithm for the Grana distance and the Hausdorff distance.

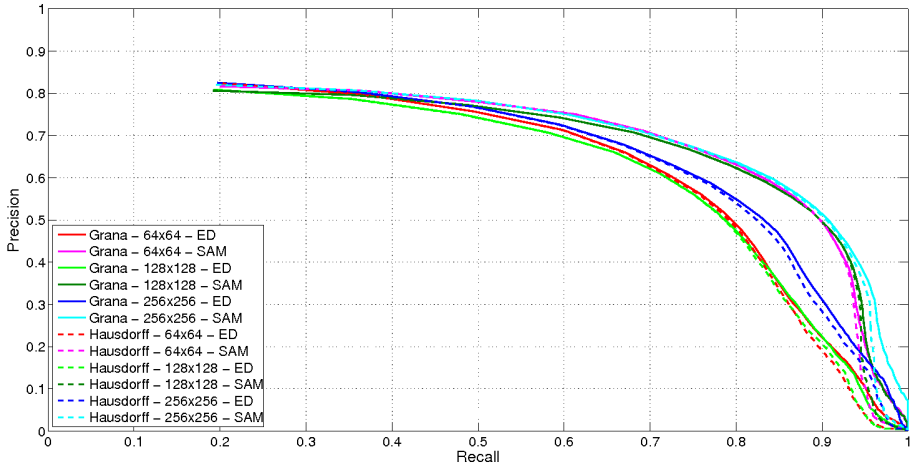


Figure 5.4: Precision and Recall results for the 10-E dataset using the N-FINDER endmember induction algorithm for the Grana distance and the Hausdorff distance.

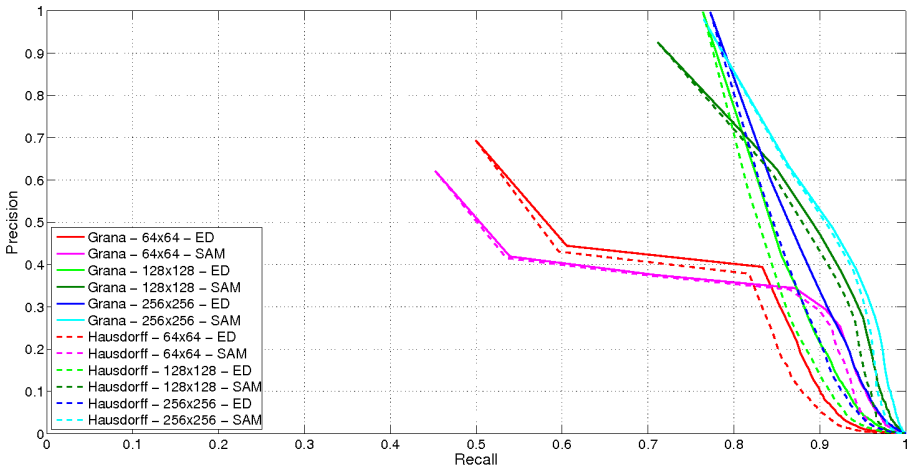


Figure 5.5: Precision and Recall results for the 20-E dataset using the EIHA endmember induction algorithm for the Grana distance and the Hausdorff distance.

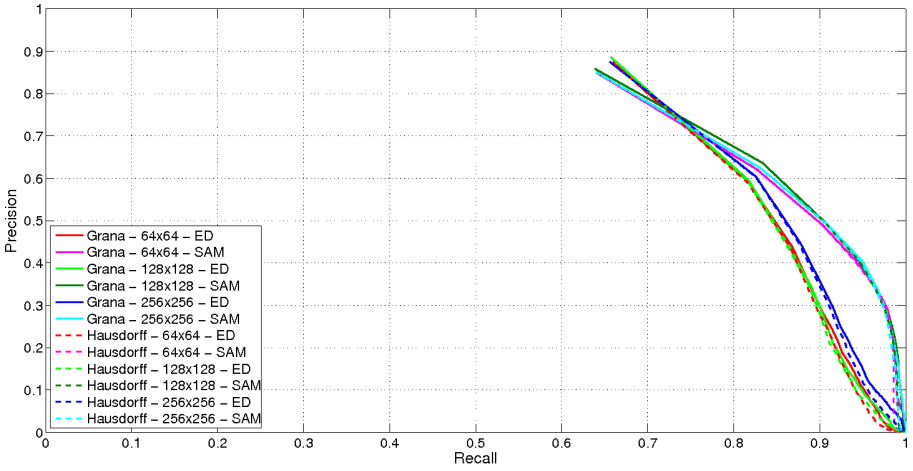


Figure 5.6: Precision and Recall results for the 20-E dataset using the N-FINDER endmember induction algorithm for the Grana distance and the Hausdorff distance.

Spatial size	Distance	EIA	AG	AH	RG	RH
64×64	Euclidean	EIHA	0,6365	0,5097	0,0376	0,1014
64×64	Euclidean	N-FINDER	0,6747	0,6538	0,0250	0,0433
64×64	SAM	EIHA	0,6805	0,5713	0,0228	0,0741
64×64	SAM	N-FINDER	0,7202	0,705	0,0220	0,0329
128×128	Euclidean	EIHA	0,6722	0,5587	0,0355	0,0808
128×128	Euclidean	N-FINDER	0,6701	0,6592	0,0293	0,0384
128×128	SAM	EIHA	0,7018	0,6115	0,0219	0,0582
128×128	SAM	N-FINDER	0,727	0,7134	0,0190	0,0299
256×256	Euclidean	EIHA	0,734	0,6344	0,0215	0,0528
256×256	Euclidean	N-FINDER	0,7117	0,701	0,0161	0,0249
256×256	SAM	EIHA	0,7362	0,6661	0,0153	0,0379
256×256	SAM	N-FINDER	0,7561	0,7438	0,0154	0,0261

Table 5.1: Area under the precision-recall curve and average normalized rank for the 5-E dataset. AG: Area under PR curve for Grana distance, AH: Area under PR curve for Hausdorff distance, RG: average normalized Rank for Grana distance, RH: average normalized Rank for Hausdorff distance.

Spatial size	Distance	EIA	AG	AH	RG	RH
64×64	Euclidean	EIHA	0,4205	0,3716	0,0256	0,0567
64×64	Euclidean	N-FINDER	0,6467	0,6404	0,0080	0,0246
64×64	SAM	EIHA	0,4809	0,4252	0,0125	0,0449
64×64	SAM	N-FINDER	0,7023	0,6981	0,0036	0,0146
128×128	Euclidean	EIHA	0,6152	0,5578	0,0230	0,0454
128×128	Euclidean	N-FINDER	0,6346	0,6301	0,0118	0,0260
128×128	SAM	EIHA	0,6818	0,6372	0,0102	0,0383
128×128	SAM	N-FINDER	0,6948	0,6909	0,0033	0,0140
256×256	Euclidean	EIHA	0,6086	0,5496	0,0101	0,0271
256×256	Euclidean	N-FINDER	0,6706	0,6644	0,0054	0,0149
256×256	SAM	EIHA	0,6464	0,6065	0,0062	0,0286
256×256	SAM	N-FINDER	0,7127	0,7062	0,0017	0,0078

Table 5.2: Area under the precision-recall curve and average normalized rank for the 10-E dataset. AG: Area under PR curve for Grana distance, AH: Area under PR curve for Hausdorff distance, RG: average normalized Rank for Grana distance, RH: average normalized Rank for Hausdorff distance.

Spatial size	Distance	EIA	AG	AH	RG	RH
64×64	Euclidean	EIHA	0,521	0,507	0,0140	0,0249
64×64	Euclidean	N-FINDER	0,7419	0,7386	0,0069	0,0111
64×64	SAM	EIHA	0,4747	0,4686	0,0045	0,0151
64×64	SAM	N-FINDER	0,7562	0,7549	0,0013	0,0058
128×128	Euclidean	EIHA	0,8485	0,8349	0,0079	0,0135
128×128	Euclidean	N-FINDER	0,7496	0,7473	0,0060	0,0067
128×128	SAM	EIHA	0,8173	0,8086	0,0026	0,0105
128×128	SAM	N-FINDER	0,7679	0,7666	<0.0001	0,0037
256×256	Euclidean	EIHA	0,8667	0,8498	0,0039	0,0097
256×256	Euclidean	N-FINDER	0,7521	0,7493	0,0020	0,0037
256×256	SAM	EIHA	0,8815	0,8763	0,0019	0,0088
256×256	SAM	N-FINDER	0,7607	0,759	<0.0001	0,0036

Table 5.3: Area under the precision-recall curve and average normalized rank for the 20-E dataset. AG: Area under PR curve for Grana distance, AH: Area under PR curve for Hausdorff distance, RG: average normalized Rank for Grana distance, RH: average normalized Rank for Hausdorff distance.

CBIR system parameter values corresponding to plots in the figures 5.1 to 5.6 the system is able to discover quite efficiently the underlying classification.

Endmember distance robustness Is the CBIR system robust to the choice of the distance computed between individual endmembers?. From the examination of tables 5.1, 5.2 and 5.3 the SAM distance consistently gives better results than the Euclidean distance: (a) the area under the precision-recall curve is greater for SAM than for the Euclidean distance, all other parameters equal, (b) the ANR is consistently lower for SAM than for Euclidean distance. There is, thus, some sensitivity of the CBIR system to the endmember distance, but with less impact than other experimental factors. This effect can be due to the scale robustness of SAM which allow to cope better with the imprecise recovery of the endmembers by the EIA.

EIA sensitivity Is the CBIR system sensitive to the EIA used for endmember induction?. Figures 5.1, 5.3 and 5.5 show the Precision-Recall curves obtained on the endmembers induced by the EIHA, while figures 5.2, 5.4 and 5.6 show the Precision-Recall curves obtained on the endmembers induced by N-FINDER. Pairing corresponding plots with the same set of candidate endmembers, image size, endmember distance and image distance gives a qualitative impression of the CBIR sensitivity to EIA. We find them very similar regardless of the EIA used, therefore a qualitative insensitivity of the CBIR approach on the EIA can be declared. For a more quantitative assessment, examining tables 5.1, 5.2 and 5.3 we find that the N-FINDER performs better for smaller images and smaller sets of candidate endmembers. The EIHA improves N-FINDER for the largest spatial size images and the biggest set of candidate endmembers. However, we do not find these effects strong enough to claim that the CBIR must employ a specific EIA.

Effect of image size Is there any effect of the image size on the CBIR system?. The examination of tables 5.1, 5.2 and 5.3 show that results improve with the size of the image for both scalar performance measures. Change in image size induces changing the relative results of the EIA considered. The effect of image size is very clearly appreciated in the Precision-Recall curves in figures 5.1 to 5.6 gathering plots corresponding to the same image size regardless of all remaining experimental parameters. Smaller images give lower Precision-Recall curves. Increasing the image size the curves grow, but the effect does saturate. This effect decreases as the size of the endmember pool grows.

Ground truth diversity Is there any effect in the CBIR system performance due to the ground-truth diversity measured by the size of the underlying set of candidate endmembers?. We expect some effect because if the diversity of candidate endmembers is higher, the image semantic domain may be considered as broader and the expected size of the query relevant set is smaller. We find a strong qualitative effect in the figures 5.1 to 5.6, regardless of other parameters.

Minimum Recall values found in the experiments grow with the size of the set of candidate endmembers. In tables 5.1, 5.2 and 5.3 the area under the Precision-Recall curves grows accordingly. The effect is less strong for the ANR than for the area under the Precision-Recall curve.

Comparing distances Does the proposed distance of Definition (16) improve over the conventional Hausdorff distance?. If we examine the figures 5.1-5.6 we can appreciate in all of them that the Precision-Recall curve corresponding to the Hausdorff distance falls below the one corresponding to the Grana distance. If we consider the scalar performance results given in tables 5.1, 5.2 and 5.3 we find a clearer confirmation of this effect. The area under the Precision-Recall curve is consistently greater for the Grana distance than for the Hausdorff distance, under all combinations of experimental factors. At the same time, the ANR is consistently smaller for the Grana distance than for the Hausdorff distance, with no sensitivity to the size of the set of candidate endmembers, while the area under the Precision-Recall curve shows less improvement for the large set of candidate endmembers.

5.5 Experiments on real data

Here we test the proposed Spectral CBIR system over the HyMAP dataset of real hyperspectral images described in Appendix D to validate the system usage in a real scenario. We report comparative results of spectral CBIR systems using the proposed spectral dissimilarity function of Definition 16, hereafter named as "Grana" dissimilarity, the Hausdorff and the robust LTS Hausdorff dissimilarity functions. The groundtruth is given by the a-priori categorization made by visual inspection, and the set of relevant images is composed in a different way to previous experiments using synthetic data. The image features are composed of the endmembers induced by either the EIHA or the N-FINDR algorithms described in Appendix B. The computation of the CBIR performance measures is done according to the description in Chapter 4.

5.5.1 Performance results

Table 5.4 contains some statistics of the dataset: the number of image blocks *per* class, and the mean and standard deviation of the number of endmembers induced from the data. Note the high standard deviations of the numbers of endmembers induced within the classes.

Precision recall Figures 5.7 and 5.8 show the average Precision-Recall curves for the EIHA and N-FINDR endmember induction algorithms, respectively, using the complete database of image blocks extracted from the DLR HyMAP image. The plots are the average of the Precision and Recall for the three main classes: *Forest*, *Fields*, *Urban*. The *Others* and *Mixed* classes are nuisance

	EIHA		NFINDR		#blocks
class	mean	std.dev.	mean	std.dev.	
Forests	6.84	2.10	7.30	1.25	39
Fields	7.75	2.67	6.16	1.56	160
Urban	7.58	2.14	4.33	1.09	24
Mixed	6.92	2.36	5.83	1.71	102
Others	5.82	1.74	5.31	1.65	35

Table 5.4: Statistics of the dataset constituted by the image blocks extracted from the real hyperspectral image.

classes. The plots correspond to the different dissimilarity measures considered: The Grana distance, the Hausdorff distance and the robust LTS Hausdorff dissimilarity. The Grana distance performs better than the remaining dissimilarities, regardless of the endmember distance (Euclidean or SAM) and endmember induction algorithm (EIHA or NFINDR). The robust LTS Hausdorff dissimilarity performs better with the NFINDR endmembers than with the EIHA, contrary to the Hausdorff distance.

ANR We plot in Figure 5.9 the average ANR over the cover classes weighted by the number of samples of each class obtained using the robust LTS Hausdorff distance for varying values of its L parameter. The value of L has strong influence on the ANR with some trend of improvement towards the higher range of values in some instances of the plot. However, optimal values appear almost anywhere. There is a definite strong effect of the endmember induction algorithm and the endmember distance, the ANR improves with NFINDR and SAM.

Tables 5.6 and (5.5) show the Average Normalized Rank (ANR) values for the Grana and the Hausdorff distances, and the robust LTS Hausdorff dissimilarity, for the EIHA and NFINDR endmember induction algorithms respectively. Overall, the Grana distance obtains the best average ANR. We specify the ANR for each image block class and combination of distances. There is no definite effect of the endmember distance, though there are some instances where the Euclidean distance between endmembers show a dramatic increase in value (worsening). The effect of the feature distance on each class' ANR has some big variations. For the *Forest* and *Mixed* classes, the robust LTS Hausdorff semimetric improves the Hausdorff distance, while the contrary happens in the *Fields*, *Urban* and *Others* classes. The worst results correspond to the *Mixed* and *Others* classes, which can be attributed to their heterogeneity and high confusion with the other classes in terms of their endmembers, because these classes contain heterogeneous regions. We do not appreciate a clear bias of the ANR measure against small sized classes as claimed in [120]. The smaller classes, *Forest*, *Urban* and *Others*, have different average ANR values. The worst is the *Others* class, but the others have values comparable to the most abundant classes. In fact, the *Mixed* class is very large and has worse results

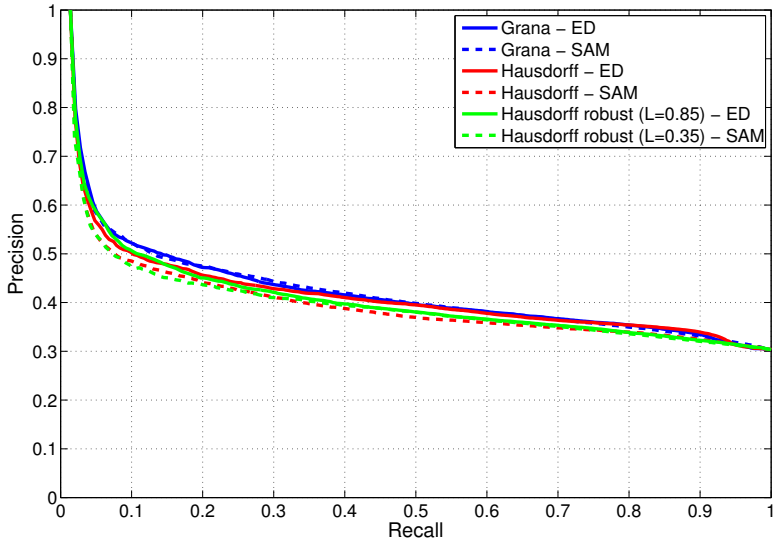


Figure 5.7: Precision-Recall on the real image database for the dissimilarities considered. Endmembers induced by EIHA

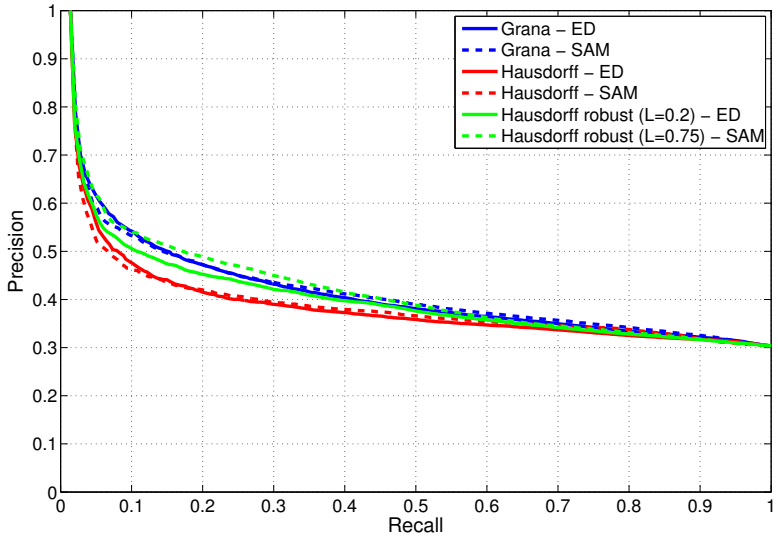


Figure 5.8: Precision-Recall on the real image database for the dissimilarities considered. Endmembers induced by N-FINDR

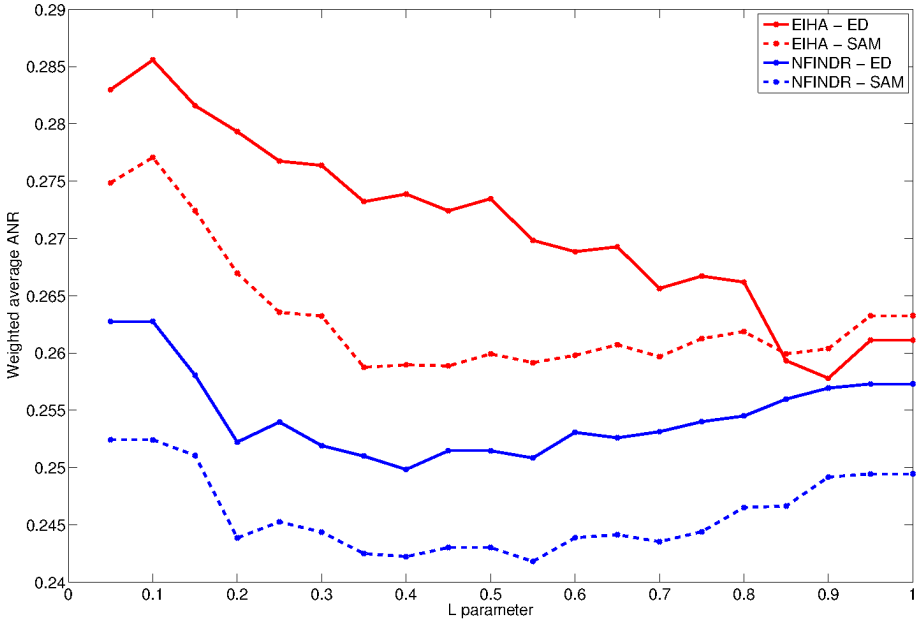


Figure 5.9: Weighted average ANR as function of the robust Hausdorff LTS dissimilarity parameter L.

than *Forest* and *Urban*.

Generality Figures 5.10 and 5.11 show the average Precision=Recall-Generality curves for the EIHA and N-FINDR endmember induction methods respectively. The abscissa scale is logarithmic following the standard suggested in [92] though the database is not increased doubling its size, but adding the *Mixed* and *Others* classes to the *Fields*, *Forests*, and *Urban* classes, providing three points of the generality plot *per* tested distance. The results show that the Grana distance is the most robust dissimilarity function. A surprising result is the dramatic decrease of the robust LTS Hausdorff dissimilarity on the endmembers induced by NFINDR. In general, the performance decrease with the size of the database is less for NFINDR than for EIHA induced endmembers.

Figures 5.12 show the response to a *Fields* image block query on the basis of the EIHA induced endmembers, given by the first ten ranked image blocks. The query image contains a small road crossing the fields with diverse degrees of growth covering. The best collection of responses is figure 5.12(d) provided by the SAM endmember distance and the Grana distance. Overall, responses to this query are rather natural, even images containing some building or big road features, have big regions corresponding to fields.

Figure 5.13 show the response to a *Urban* image block query on the basis of the NFINDR induced endmembers, given by the first ten ranked image blocks.

		Forest	Fields	Urban	Mixed	Others	Average
G	ED	0.16	0.16	0.40	0.37	0.37	**0.25**
	SAM	0.20	0.17	**0.22**	**0.36**	0.36	**0.25**
HD	ED	0.22	**0.15**	0.29	0.39	0.38	0.26
	SAM	0.25	0.19	0.25	0.37	0.40	0.27
HD-LTS	ED (L=0.9)	**0.15**	0.19	0.29	**0.36**	0.39	0.26
	SAM (L=0.35)	0.20	0.20	**0.22**	**0.36**	**0.32**	0.26

Table 5.5: ANR values for real image database (optimal column values in bold). Endmembers computed with EIHA. Dissimilarities: Grana (G), Hausdorff distance (HD), robust LTS Hausdorff dissimilarity (HD-LTS).

		Forest	Fields	Urban	Mixed	Others	Average
G	ED	0.11	**0.20**	0.14	0.36	0.43	0.25
	SAM	0.10	**0.20**	**0.12**	0.35	0.39	**0.23**
HD	ED	0.21	0.21	0.22	0.38	0.46	0.29
	SAM	0.19	**0.20**	0.20	0.38	0.43	0.28
HD-LTS	ED (L=0.4)	**0.09**	0.22	0.16	**0.34**	0.37	0.25
	SAM (L=0.55)	**0.09**	0.21	0.14	**0.34**	**0.34**	0.24

Table 5.6: ANR values for real image database (optimal column values in bold) endmembers computed with N-FINDR. Dissimilarities: Grana (G), Hausdorff distance (HD), robust LTS Hausdorff dissimilarity (HD-LTS).

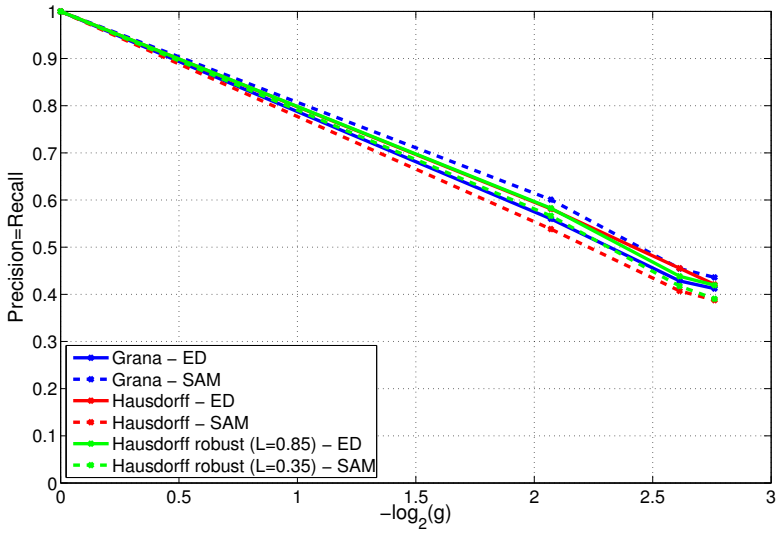


Figure 5.10: Generality-Precision=Recall plot for the real image database. Endmembers induced by EIHA.

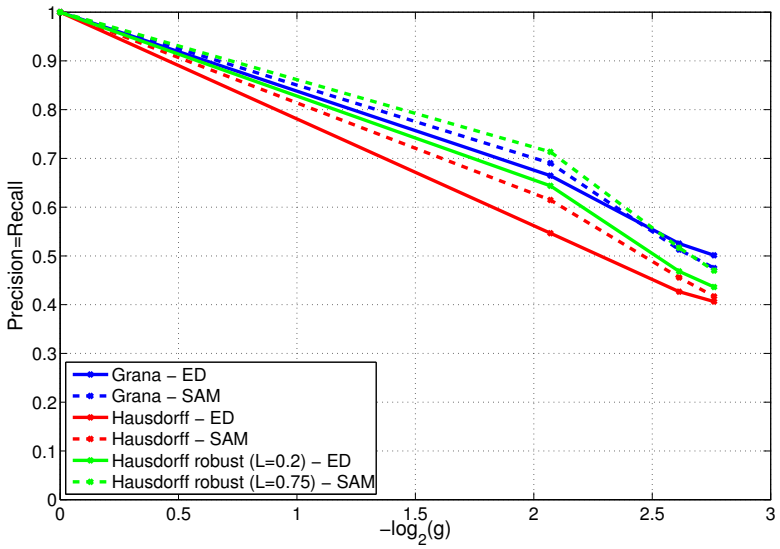


Figure 5.11: Generality-Precision=Recall plot for the real image database. Endmembers induced by N-FINDR

The urban areas contain some trees and green regions, therefore there is some confusion between this class and the vegetation classes. All responses contain several images corresponding to *Fields* class. The Hausdorff dissimilarities contain more such confusions than the Grana distance. In fact, the best response is figure 5.13(d) corresponding to the Grana distance.

5.6 Conclusions

In this chapter we introduce a feature extraction process and a distance between hyperspectral images which can be used for CBIR system on databases of hyperspectral images. Image features are the endmembers induced from the image data by some Endmember Induction Algorithm (EIA). The distance between images is computed over these induced endmembers. We have tested the sensitivity of the CBIR system defined on this distance to the EIA used, the individual endmember distance, the image size, and the diversity of the underlying ground-truth on a large database of synthetic hyperspectral images. We find the CBIR system to be sensitive to the ground-truth diversity and the image size. We find that the proposed distance improves the Hausdorff distance for the same CBIR task. Besides, we have tested the CBIR system on a dataset built from a large real hyperspectral image, specifically we find that the proposed distance improves over the Hausdorff distance, and a robust LTS Hausdorff semimetric. Overall, the results in this chapter, specially those obtained on the real image

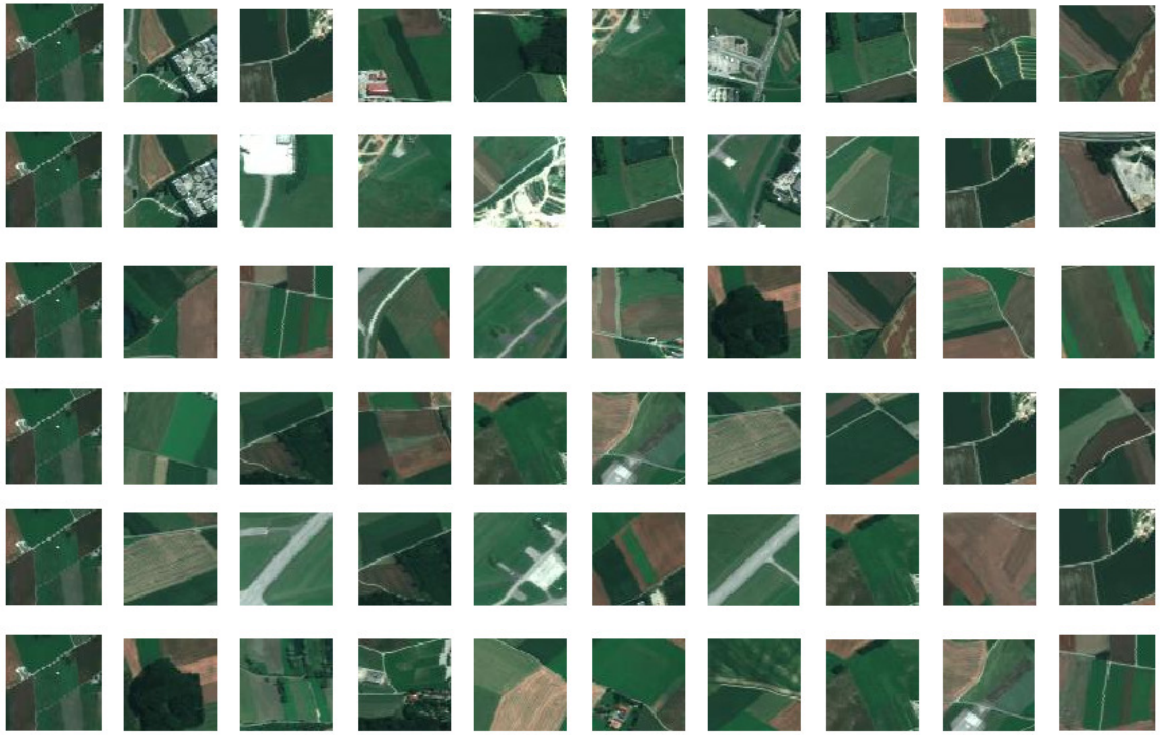


Figure 5.12: Response to a query '*Fields*' image block, scope $k = 10$, using EIHA method for endmember extraction. (a) to (c) using the Euclidean distance, (d) to (f) using the SAM distance. Dissimilarities: (a), (d) Grana, (b), (e) Hausdorff, (c), (f) robust LTS Hausdorff.

Figure 5.13: Response to query '*Urban*' image block, scope $k = 10$, using NFINDR method for endmember extraction. (a) to (c) using the Euclidean distance, (d) to (f) using the SAM distance. Dissimilarities: (a), (d) Grana, (b), (e) Hausdorff, (c), (f) robust LTS Hausdorff.

dataset, confirm that the choice of spectral features, specifically the image endmembers, provide a meaningful CBIR system for hyperspectral images. The query answers are easy to interpret as relevant by the human user, even for the untrained eye.

Chapter 6

Hyperspectral CBIR systems: spectral-spatial dissimilarity

This Chapter introduces a novel Content-Based Image Retrieval (CBIR) system for hyperspectral image databases using both spectral and spatial features computed following an unsupervised unmixing process. The system allows the user to retrieve hyperspectral images containing materials similar to the query image, and in a similar proportion. We provide validation results using both synthetic hyperspectral datasets and real hyperspectral data.

The contents of the chapter are the following. Section 6.1 gives a brief introduction. Section 6.2 provides a description of the Spectral-Spatial CBIR system. Section 6.3 gives the validation results using synthetic hyperspectral images. Section 6.4 gives validation results using a real hyperspectral dataset. Section 6.5 gives our conclusions and directions for further work.

6.1 Introduction

The approach introduced in Chapter 5 has one shortcoming: can not discriminate among images with the same induced endmembers but very different spatial distributions. The use of spatial information aims to overcome this problem. As in Chapter 5, the set of endmembers obtained from the image by an Endmember Induction Algorithm provides the image spectral features. Spatial features are computed as abundance image statistics. Both kinds of information are functionally combined into a dissimilarity measure between two hyperspectral images guiding the search for answers to database queries.

The proposed Spectral-Spatial CBIR system for hyperspectral imagery relies on the linear mixing formulation [98] to extract the Spectral-Spatial features that drive the retrieval process. We conduct experiments using two well known EIAs, the N-FINDER [193] and the Fast Iterative Pixel Purity Index (FIPPI) [31], and a fast and light algorithm based on lattice computing, the Incremental Lattice Strong Independence Algorithm (ILSIA) [68]. All these EIA

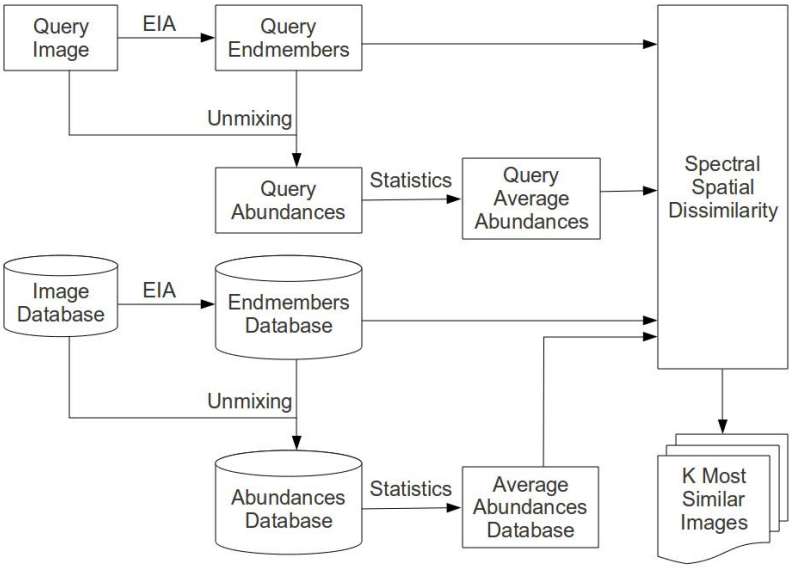


Figure 6.1: Spectral-Spatial CBIR system's schema

are described in Appendix B. We also used the Harsanyi-Farrand-Chang (HFC) virtual dimensionality method [30] to estimate the number of endmembers when required, and the Fully Constrained Least Squares Unmixing (FCLSU) method [105] to estimate the fractional abundances.

6.2 Spectral-Spatial CBIR system

Figure 6.1 shows the Spectral-Spatial CBIR system scheme. The core of the CBIR system is the Spectral-Spatial dissimilarity function between two hyperspectral images by means of their spectral and spatial features. The system interacts with a feature database where the Spectral-Spatial features of the images are stored. These features have been previously extracted by offline application of an Endmember Induction Algorithm (EIA) and a spectral unmixing method using the endmembers extracted by the EIA from the image. System interrogation is done using a query example approach. First, the query example is processed to extract its Spectral-Spatial features and second, it is compared to the images in the database using the Spectral-Spatial dissimilarity measure. A ranking of the images in the database is elaborated by ascending order of dissimilarity to the query. Finally, the system returns the k images in the database corresponding to the first k ranking positions, where k is known as the query's *scope*.

Let $E_\alpha = \left\{\mathbf{e}_1^\alpha, \mathbf{e}_2^\alpha, \ldots, \mathbf{e}_{p_\alpha}^\alpha\right\}$ be the set of endmembers induced from the

hyperspectral image H_α by some endmember induction algorithm (EIA), where p_α is the number of induced endmembers from the α-th image; and $\bar{\Phi}_\alpha = \left\{\bar{\phi}_1^\alpha, \bar{\phi}_2^\alpha, \ldots, \bar{\phi}_{p_\alpha}^\alpha\right\}$ the normalized average abundances of H_α, where $\bar{\phi}_i^\alpha$ is defined as:

$$\bar{\phi}_i^\alpha = \frac{1}{M}\sum_{m=1}^{M}\frac{\phi_i^\alpha(m)}{\sum_{l=1}^{p_\alpha}\phi_l^\alpha(m)},\ i=1,\ldots,p_\alpha, \tag{6.1}$$

being $\phi_i^\alpha(m)$ the i-th endmember fractional abundance for a pixel m, and M the total number of pixels in the image. Given two images, H_α and H_β, we compute the Spectral Distance Matrix, $D_{\alpha,\beta}$, as

$$D_{\alpha,\beta} = [d_{i,j};\ i=1,\ldots,p_\alpha;\ j=1,\ldots,p_\beta], \tag{6.2}$$

whose elements d_{ij} are the distances between the endmembers $\mathbf{e}_i^\alpha, \mathbf{e}_j^\beta \in \mathbb{R}^q$ of each image, for instance the Euclidean distance, d_E:

$$d_E(\mathbf{e}_1, \mathbf{e}_2) = \|\mathbf{e}_1 - \mathbf{e}_2\|, \tag{6.3}$$

or the angular distance, aka Spectral Angle Mapper (SAM) distance in remote sensing applications [29], d_S:

$$d_S(\mathbf{e}_1, \mathbf{e}_2) = \cos^{-1}\left(\frac{\mathbf{e}_1 \cdot \mathbf{e}_2}{\|\mathbf{e}_1\|\,\|\mathbf{e}_2\|}\right), \tag{6.4}$$

where $\mathbf{e}_1 \cdot \mathbf{e}_2$ denotes the vector inner product.

Definition 16. The spectral-spatial dissimilarity between two hyperspectral images, H_α, H_β, is given by the following expression:

$$s(H_\alpha, H_\beta) = \sum_{i,j} r_{i,j} d_{i,j}, \tag{6.5}$$

where $d_{i,j}$ is the spectral distance between endmembers $\mathbf{e}_i^\alpha$ and $\mathbf{e}_j^\beta$, and r_{ij} is the significance associated to $d_{i,j}$, $i=1,\ldots,p_\alpha$; $j=1,\ldots,p_\beta$.

The formulation of the measure needs a definition for the significance matrix $R_{\alpha,\beta} = [r_{i,j};\ i=1,\ldots,p_\alpha;\ j=1,\ldots,p_\beta]$. We follow the *most similar highest priority* (MSHP) principle [107], making use of the normalized average abundances $\bar{\Phi}_\alpha$ and $\bar{\Phi}_\beta$. The average abundances represent "significance credits" assigned to the endmember spectral distances by Algorithm 6.1.

The algorithm for the assignment of significance credits starts by initializing the set of all possible endmember pairs, $\mathcal{M} = \{(i,j) :\ i=1,\ldots,p_\alpha;\ j=1,\ldots,p_\beta\}$, and the set of previously selected endmember pairs, $\mathcal{L} = \{\}$ (Steps 1, 2). In each subsequent iteration, the algorithm first selects the pair of endmembers $(i,j) :\ i=1,\ldots,p_\alpha;\ j=1,\ldots,p_\beta$, with minimum spectral distance from the set of available pairs, $(i',j') = \arg\min_{i,j} d_{i,j},\ (i,j) \in \mathcal{M} - \mathcal{L}$, (Step 3). Let (i',j') denote the selected pair. Second, the value of the minimum of the average fractional abundances is assigned as the pair's corresponding significance,

Algorithm 6.1 Significance credit assignment algorithm.

1. Set $\mathcal{L} = \{\}$.
2. Denote $\mathcal{M} = \{(i,j) : \ i = 1, \ldots, m_\alpha; \ j = 1, \ldots, m_\beta\}$.
3. Choose the pair $(i.j)$ with minimum $d_{i,j}$ for all $(i,j) \in \mathcal{M} - \mathcal{L}$. Label the corresponding (i,j) as (i',j').
4. $r_{i',j'} = \min\left\{\bar{\phi}_{i'}^{\alpha}, \bar{\phi}_{j'}^{\beta}\right\}$.
5. If $\bar{\phi}_{i'}^{\alpha} < \bar{\phi}_{j'}^{\beta}$, set $r_{i',j} = 0$, for all $j \neq j'$; otherwise, set $r_{i,j'} = 0$, for all $i \neq i'$.
6. If $\bar{\phi}_{i'}^{\alpha} < \bar{\phi}_{j'}^{\beta}$, set $\bar{\phi}_{i'}^{\alpha} = 0$ and $\bar{\phi}_{j'}^{\beta} = \bar{\phi}_{j'}^{\beta} - \bar{\phi}_{i'}^{\alpha}$; otherwise, set $\bar{\phi}_{j'}^{\beta} = 0$ and $\bar{\phi}_{i'}^{\alpha} = \bar{\phi}_{i'}^{\alpha} - \bar{\phi}_{j'}^{\beta}$.
7. $\mathcal{L} = \mathcal{L} + \{(i',j')\}$.
8. If $\sum_{i=1}^{p_\alpha} \bar{\phi}_i^{\alpha} > 0$ and $\sum_{j=1}^{p_\beta} \bar{\phi}_j^{\beta} > 0$, go to step 3; otherwise, stop.

$r_{i'j'} = \min\left\{\bar{\phi}_{i'}^{\alpha}, \bar{\phi}_{j'}^{\beta}\right\}$ (Step 4). Notice that the average abundances are always equal to or greater than zero. If $\bar{\phi}_{i'}^{\alpha} < \bar{\phi}_{j'}^{\beta}$, then the elements of significance matrix row i' are set to zero; otherwise, the elements of significance matrix column j' are set to zero (Step 5). Then, the pool of significance credits is reduced. If $\bar{\phi}_{i'}^{\alpha} < \bar{\phi}_{j'}^{\beta}$, then set $\bar{\phi}_{i'}^{\alpha} = 0$ and $\bar{\phi}_{j'}^{\beta} = \bar{\phi}_{j'}^{\beta} - \bar{\phi}_{i'}^{\alpha}$; otherwise set $\bar{\phi}_{j'}^{\beta} = 0$ and $\bar{\phi}_{i'}^{\alpha} = \bar{\phi}_{i'}^{\alpha} - \bar{\phi}_{j'}^{\beta}$ (Step 6). Finally, (i',j') is added to the set of previously selected pairs, $\mathcal{L}$ (Step 7). When the stopping condition, $\sum_{i=1}^{p_\alpha} \bar{\phi}_i^{\alpha} = 0$ or $\sum_{j=1}^{p_\beta} \bar{\phi}_j^{\beta} = 0$, is met the algorithm ends; otherwise, a new iteration starts (Step 8).

6.3 Validation using synthetic datasets

Following the procedures detailed in Appendix C we generated two sets of 1000 hyperspectral images using the simulated using Legendre polynomials approach for the abundance images, each set having a specific domain spatial size, $D^{(64)}$ denotes 64×64 pixel images, $D^{(128)}$, denotes images of size 128×128 pixels. In addition to $D^{(64)}$ and $D^{(128)}$ datasets, hereafter named as the clean datasets, six more datasets have been built by adding random Gaussian noise to each of the clean dataset images, resulting in noise images with signal to noise ratios (SNR) 30dB, 40dB and 50dB. We denote clean dataset as D_o, and the noisy datasets as D_{30dB}, D_{40dB} and D_{50dB}. Thereby, we have in total eight synthetic hyperspectral datasets: $D_o^{(64)}$, $D_{30dB}^{(64)}$, $D_{40dB}^{(64)}$, $D_{50dB}^{(64)}$, $D_o^{(128)}$, $D_{30dB}^{(128)}$, $D_{40dB}^{(128)}$ and $D_{50dB}^{(128)}$.

6.3.1 Experimental methodology

We have performed independent experiments over each of the eight hyperspectral synthetic datasets, following the methodology explained in Chapter 4. Let us distinguish between $\mathbf{s}_\alpha^{\mathrm{GT}}$, the vector of dissimilarities computed using the known ground truth endmembers and fractional abundances used to synthesize the images, and $\mathbf{s}_\alpha^{\mathrm{IND}}$, the vector of dissimilarities computed using the endmembers induced by one of the EIAs (either ILSIA, N-FINDER or FIPPI) and their corresponding abundances estimated by the FCLSU algorithm. Computation of performance measures recall and precision are computed as described in Chapter 4 using the mean and standard deviation of the ground truth to determine the size of the set of relevant images.

6.3.2 Performance results

Table 6.1 shows the average and standard deviation of the number of endmembers induced by the different EIAs from the $D^{(64)}$ and $D^{(128)}$ datasets. We have estimated the induced endmember values for the clean and noisy datasets in four separated groups, corresponding to the number m of ground truth endmembers used for the image synthesis. We can see the number of endmembers induced by N-FINDR algorithm, corresponding to the number of endmembers estimated by HFC method, is close to the actual number of endmembers in the image even in noisy datasets. FIPPI algorithm, which uses HFC estimation as an initial value, finally induces a lower number of endmembers and it is more sensitive to noise. ILSIA algorithm, which does not require a-priori knowledge of the number of endmembers in the image, underestimates the number of actual endmembers and it is also sensible to noise conditions.

Precision recall Figures 6.2, 6.3, 6.4 and 6.5 show the plots of the precision-recall curves for the noise free dataset and the 30dB, 40dB and 50dB datasets respectively. It can be appreciated that for low levels of noise the performance is similar to the noise free case. In general, the plots show a rather high insensitivity to the choice of EIA and individual endmember spectral distance, because corresponding curves are not very different, except in the limit of low recall values where a clear sensitivity to the individual endmember distance is made apparent. The slightly better performance of N-FINDR can be explained by the greater number of endmembers induced by this algorithm compared to ILSIA and FIPPI algorithms. However, the little differences between algorithms performance suggests that the fewer endmembers induced by ILSIA and FIPPI contain the most relevant information for image comparison. The Euclidean distance systematically gets better results than the SAM distance, giving higher precision at the same recall value. Increasing recall value range reverses the picture, so that the SAM distance improves systematically over the Euclidean distance in the noise free case. One effect of the noise is the cancellation of this effect. For the highest noise (30dB) the SAM never improves the Euclidean distance. Other effect of the noise is the attenuation of the effect of the EIA

	$D_o^{(64)}$	$D_{30dB}^{(64)}$	$D_{40dB}^{(64)}$	$D_{50dB}^{(64)}$
ILSIA ($m=2$)	2.03 ± 0.17	2.98 ± 0.61	3.54 ± 1.01	4.03 ± 1.27
ILSIA ($m=3$)	2.02 ± 0.13	3.13 ± 0.65	3.57 ± 1.04	3.88 ± 1.31
ILSIA ($m=4$)	2.06 ± 0.23	3.13 ± 0.77	3.55 ± 1.01	3.66 ± 1.24
ILSIA ($m=5$)	2.08 ± 0.28	3.17 ± 0.75	3.66 ± 1.23	3.52 ± 1.29
N-FINDR ($m=2$)	2.00 ± 0.06	1.98 ± 0.14	2.00 ± 0.06	2.00 ± 0.06
N-FINDR ($m=3$)	2.88 ± 0.33	2.74 ± 0.48	2.88 ± 0.33	2.88 ± 0.33
N-FINDR ($m=4$)	3.69 ± 0.52	3.10 ± 0.78	3.61 ± 0.60	3.68 ± 0.52
N-FINDR ($m=5$)	4.44 ± 0.70	3.42 ± 0.91	4.13 ± 0.83	4.40 ± 0.73
FIPPI($m=2$)	2.18 ± 0.62	3.14 ± 0.68	3.23 ± 0.63	3.15 ± 0.65
FIPPI ($m=3$)	2.32 ± 0.73	3.85 ± 1.00	3.92 ± 0.93	3.75 ± 1.03
FIPPI($m=4$)	2.26 ± 0.72	4.13 ± 1.23	4.25 ± 1.17	4.06 ± 1.25
FIPPI($m=5$)	2.38 ± 0.78	4.32 ± 1.30	4.42 ± 1.39	4.20 ± 1.36

(a)

	$D_o^{(128)}$	$D_{30dB}^{(128)}$	$D_{40dB}^{(128)}$	$D_{50dB}^{(128)}$
ILSIA ($m=2$)	2.06 ± 0.23	3.36 ± 0.73	4.02 ± 1.21	4.83 ± 1.70
ILSIA ($m=3$)	2.02 ± 0.15	3.54 ± 0.91	4.11 ± 1.35	4.51 ± 1.49
ILSIA ($m=4$)	2.09 ± 0.29	3.52 ± 0.80	4.32 ± 1.35	4.66 ± 1.76
ILSIA ($m=5$)	2.10 ± 0.30	3.72 ± 0.95	4.20 ± 1.41	4.23 ± 1.69
N-FINDR ($m=2$)	2.00 ± 0.00	2.00 ± 0.00	2.00 ± 0.00	2.00 ± 0.00
N-FINDR ($m=3$)	2.94 ± 0.24	2.84 ± 0.39	2.93 ± 0.26	2.94 ± 0.24
N-FINDR ($m=4$)	3.84 ± 0.37	3.49 ± 0.63	3.80 ± 0.41	3.84 ± 0.37
N-FINDR ($m=5$)	4.63 ± 0.58	3.69 ± 0.82	4.36 ± 0.77	4.60 ± 0.59
FIPPI($m=2$)	2.17 ± 0.59	3.20 ± 0.57	3.20 ± 0.55	3.17 ± 0.57
FIPPI ($m=3$)	2.21 ± 0.58	4.00 ± 0.95	3.95 ± 0.94	3.83 ± 0.96
FIPPI($m=4$)	2.27 ± 0.68	4.58 ± 1.26	4.45 ± 1.31	4.21 ± 1.26
FIPPI($m=5$)	2.43 ± 0.81	4.64 ± 1.31	4.83 ± 1.42	4.23 ± 1.33

(b)

Table 6.1: Sample mean and sample standard deviation of the number of endmembers induced from the synthetic datasets relative to the EIA and the number of endmembers, m, used for the image synthesis: (a) $D^{(64)}$, (b) $D^{(128)}$.

	Averaged Normalized Rank (ANR)					
	ILSIA		N-FINDR		FIPPI	
Dataset	ED	SAM	ED	SAM	ED	SAM
$D_o^{(64)}$	**0.043**	0.053	0.050	0.058	0.064	0.074
$D_{30dB}^{(64)}$	0.045	0.101	**0.043**	0.097	**0.043**	0.099
$D_{40dB}^{(64)}$	0.042	0.057	**0.036**	0.048	0.037	0.052
$D_{50dB}^{(64)}$	0.042	0.053	**0.038**	0.045	0.041	0.051

(a)

	Averaged Normalized Rank (ANR)					
	ILSIA		N-FINDR		FIPPI	
Dataset	ED	SAM	ED	SAM	ED	SAM
$D_o^{(128)}$	0.047	0.057	**0.044**	0.050	0.068	0.077
$D_{30dB}^{(128)}$	0.044	0.086	0.044	0.088	**0.038**	0.082
$D_{40dB}^{(128)}$	0.044	0.056	**0.035**	0.043	0.039	0.049
$D_{50dB}^{(128)}$	0.043	0.052	**0.033**	0.038	0.040	0.047

(b)

Table 6.2: ANR results for synthetic datasets: (a) $D^{(64)}$ (b) $D^{(128)}$.

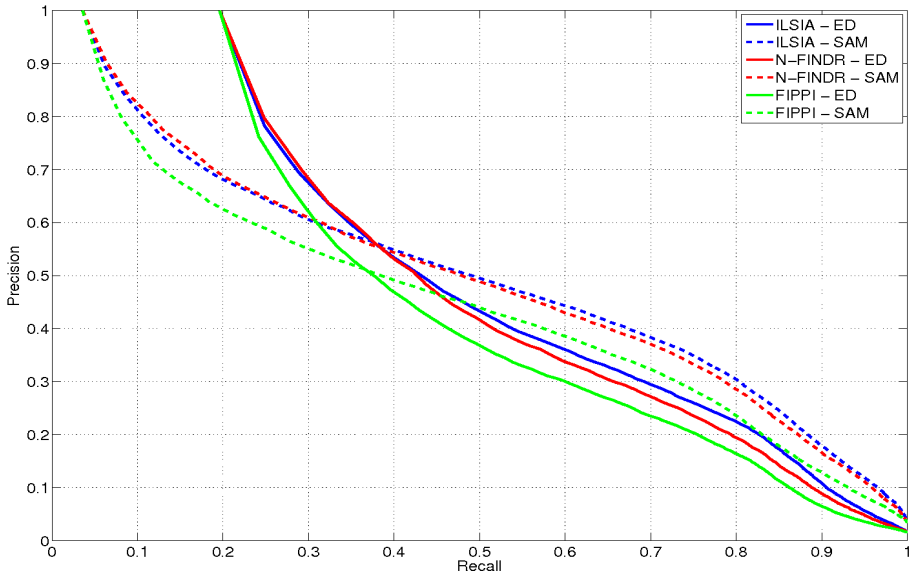
chosen. In the noise free data set, ILSIA and N-FINDR show a small improvement on the FIPPI algorithm, however, these relations change according to the noise level, being attenuated at the highest noise level.

The differences between the performance results on datasets $D^{(64)}$ and $D^{(128)}$ are almost negligible showing that the spatial size by itself, without an increasing in the number of distinct materials, does not modify the proposed Spectral/Spatial CBIR system performance.

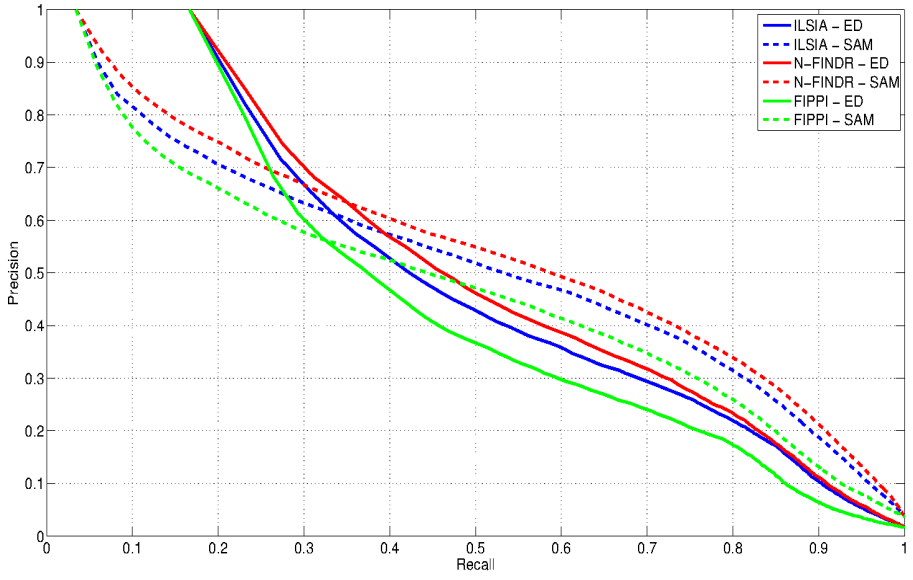
ANR Table 6.2 gives the Averaged Normalized Rank values obtained. The results confirm the conclusions from the figures. The SAM distance is most affected by noise, showing the worse results for the noisiest data. The Euclidean distance is much more robust relative to noise. Comparing the effect of the EIA chosen, the ILSIA gives the best result in the noise free data, and the differences between algorithms are attenuated for the noisiest data.

6.4 Validation on real data

Here we test the proposed Spectral-Spatial CBIR system over the HyMAP dataset of real hyperspectral images described in Appendix D to validate the system usage in a real scenario. We also compare it to the system proposed in [141], hereafter named as Plaza's CBIR system. We perform three experiments to validate the use of the proposed Spectral-Spatial CBIR system. In first experiment we tested the system using the patches belonging to the three main categories: Forests, Fields and Urban Areas. In second experiment we added

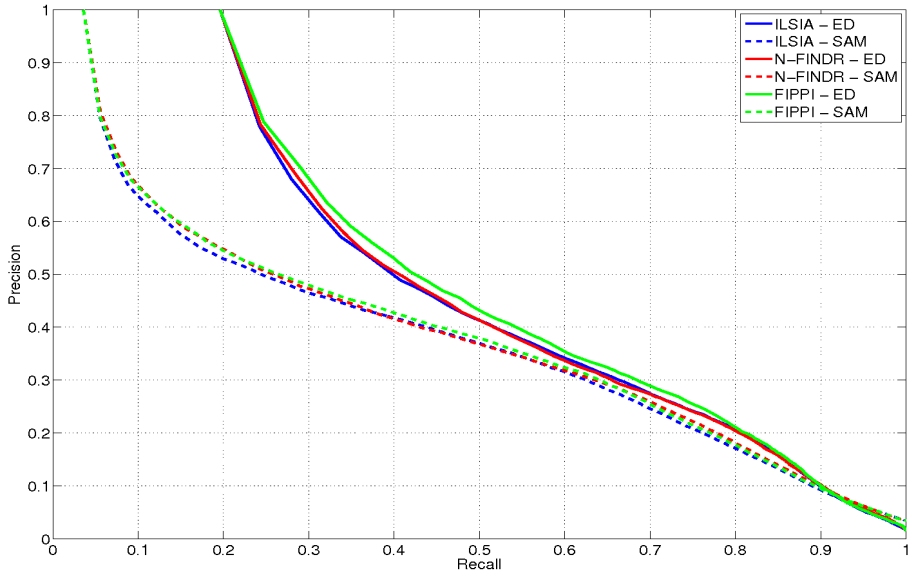


(a)

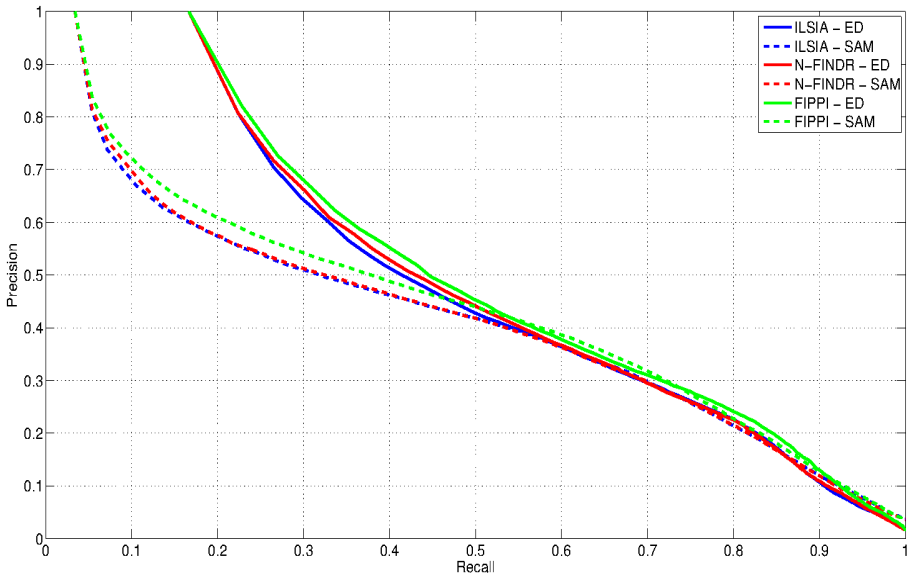


(b)

Figure 6.2: Precision-recall curves for D_o synthetic datasets: (a) $D_o^{(64)}$ (b) $D_o^{(128)}$.

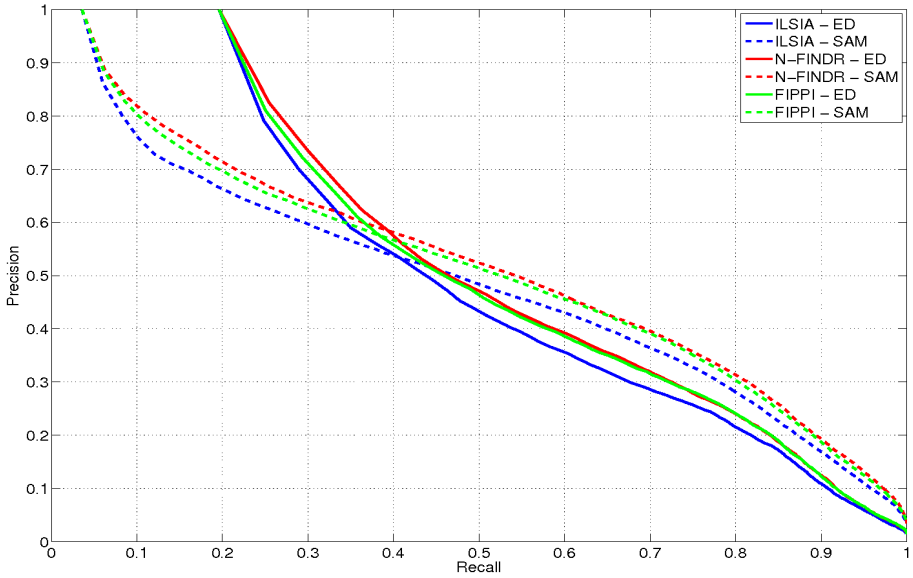


(a)

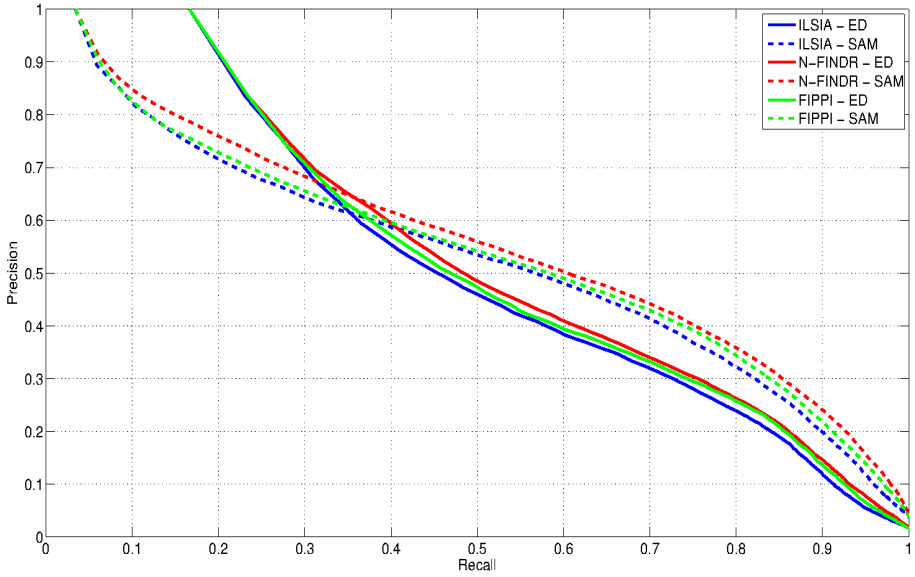


(b)

Figure 6.3: Precision-recall curves for D_{30dB} synthetic datasets: (a) $D_{30dB}^{(64)}$ (b) $D_{30dB}^{(128)}$.

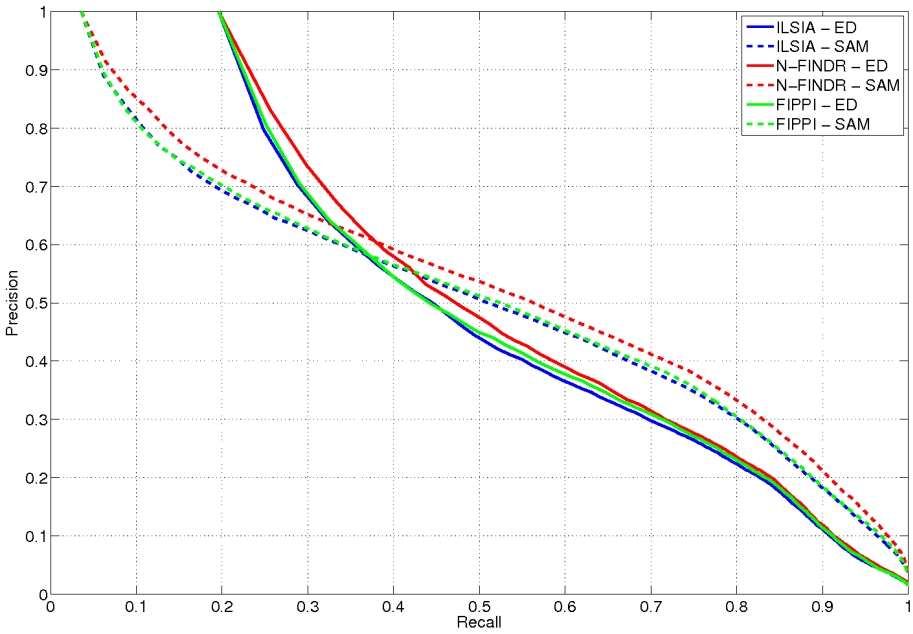


(a)

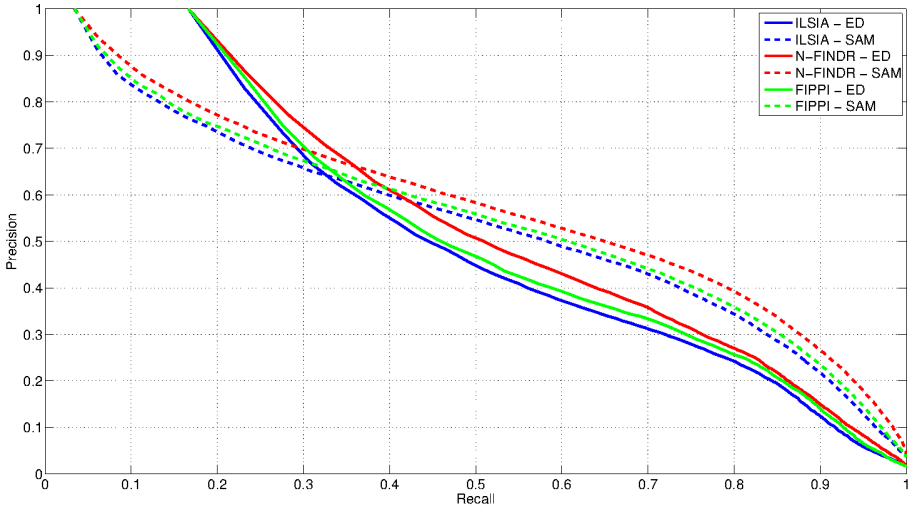


(b)

Figure 6.4: Precision-recall curves for D_{40dB} synthetic datasets: (a) $D_{40dB}^{(64)}$ (b) $D_{40dB}^{(128)}$.



(a)



(b)

Figure 6.5: Precision-recall curves for D_{50dB} synthetic datasets: (a) $D_{50dB}^{(64)}$ (b) $D_{50dB}^{(128)}$.

	ILSIA	N-FINDR	FIPPI
Forests	3.85 ± 1.16	7.31 ± 1.26	5.64 ± 1.31
Fields	2.99 ± 0.80	6.16 ± 1.57	4.42 ± 1.28
Urban Areas	2.75 ± 0.61	4.33 ± 1.09	3.50 ± 1.06
Mixed	2.98 ± 0.80	5.83 ± 1.71	4.26 ± 1.50
Others	2.77 ± 0.60	5.31 ± 1.66	3.94 ± 1.35

Table 6.3: Sample mean and sample standard deviation of the number of endmembers induced from the real HyMap categories relative to the EIA.

patches from the fourth category: Mixed. Finally, in third experiment we used the full patches database. The image features are composed of the endmembers induced by either the EIHA or the N-FINDR algorithms described in Appendix B. The computation of the CBIR performance measures is done according to the description in Chapter 4.

6.4.1 Performance results

Table 6.3 shows the average and standard deviation of the number of endmembers induced by the three EIAs for each of the five categories. The results coincide with the values obtained over the synthetic data in the sense that N-FINDR algorithm induces the most number of endmembers and ILSIA algorithm the least. We can also see that the average number of endmembers induced for Forests and Fields patches is slightly greater than for the other categories.

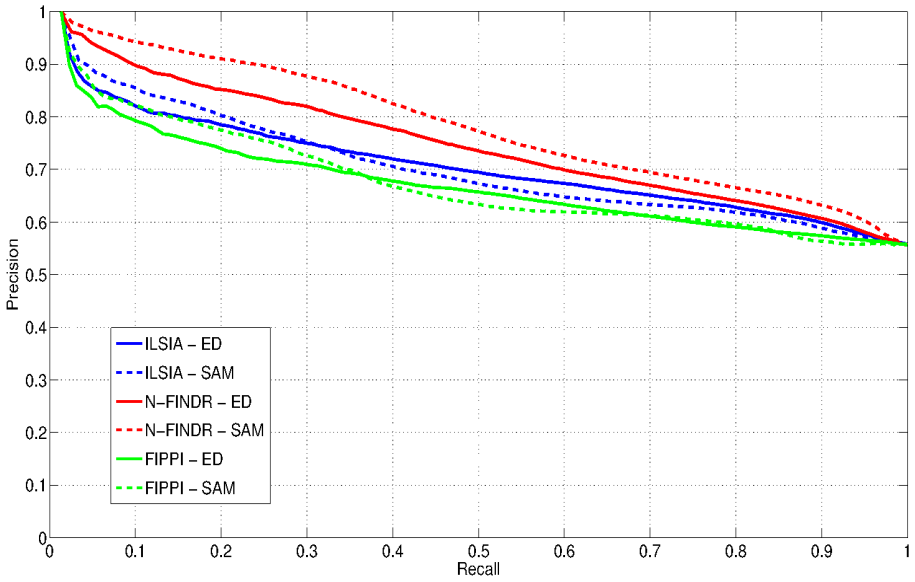
Precision-recall Figure 6.6 shows the precision-recall curve for the first experiment, with varying EIA and endmember distance. Surprisingly, for the proposed Spectral-Spatial CBIR system the best result correspond to the SAM distance, when using the N-FINDR endmember induction. For low recall values, the differences between algorithms and distances are negligible. In all combinations, the precision-recall curves are very high, showing that the approach is feasible for real-life applications. In table 7.3 we have the ANR results for the first experiment. Notice that most values when computing the SAM among endmembers induced by N-FINDR are below 0.1. These results strengthens the evidence from synthetic datasets experiments that N-FINDR performs slightly better because it induces more endmembers than the competing EIAs, but that the little differences in performance suggest that those additional endmembers contains redundant information. However, that is not the case of Urban Areas category where N-FINDR performs much better combined with the SAM distance. These could be because urban areas spectral signatures vary significantly more in amplitude than forests and fields due to illumination conditions. Thus, N-FINDR successes in finding endmembers containing this spectral information and SAM distance performs better than Euclidean distance to assess this amplitude difference.

Increasing categories Adding new categories, the performance of the system degrades gracefully, as shown in figures 6.7 and 6.8, maintaining high precision for low recall values. The ANR results in tables 7.4 and 7.5 still provide the best results for SAM distance. This must be due to the magnitude normalization performed by the SAM that removes some illumination effects that are stronger for the Euclidean distance. In all cases, the non-homogeneous categories, such as Urban Area, Mixed and Other, are the most difficult to retrieve, and the inclusion of new categories does not help to improve retrieval figures. In figures 6.9 and 6.10 we show examples of a system's response for a query defined over a Forest category patch for both, the proposed Spectral-Spatial CBIR system and the Plaza's CBIR system respectively.

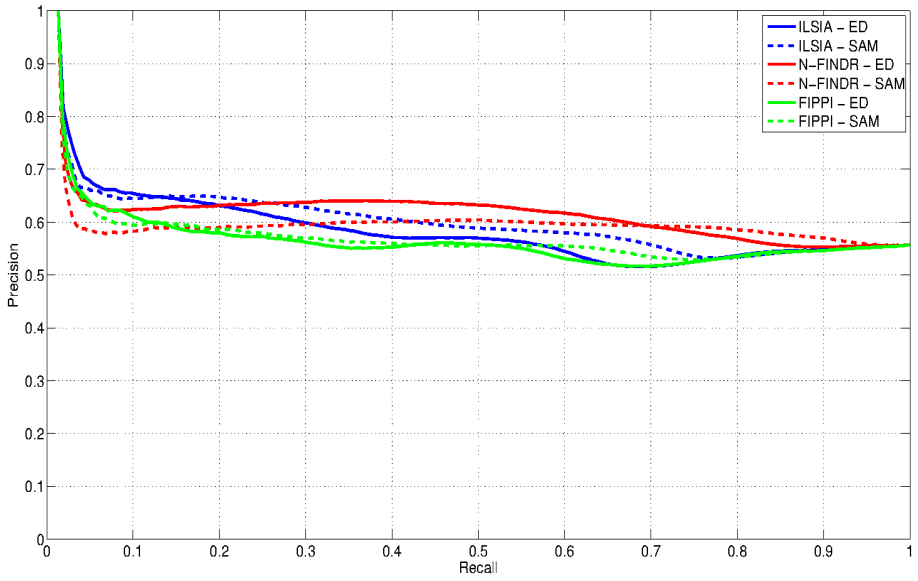
Comparing to the competing Plaza's CBIR system, the proposed Spectral-Spatial CBIR system improves it, achieving the former a better performance only for the Mixed category in experiment 3.

6.5 Conclusions

This chapter introduces a spatial-spectral CBIR system, providing validation results on synthetic and real hyperspectral data. To validate our approach we have followed a rigorous methodological framework using both synthetic and real datasets. We have also performed comparisons with an early CBIR system with results improved by our system. The results on the synthetic datasets demonstrate the system robustness against noise and changes in the choice of endmember induction algorithm and endmember distances. The results on real data confirm the usefulness of the proposed system. The system performance does not degrade when the number of categories in the real data increase. From a qualitative point of view, it can be appreciated on the real data that the spectral-spatial system provides more coherent spatial answers that the approach proposed in Chapter 5. That is, images with similar spatial distributions are clearly preferred as intended.



(a)



(b)

Figure 6.6: Precision-recall curves for HyMAP experiment 1: (a) Spectral-Spatial CBIR (b) Plaza's CBIR.

	Averaged Normalized Rank (ANR)					
	ILSIA		N-FINDR		FIPPI	
Category	ED	SAM	ED	SAM	ED	SAM
Forests	0.115	0.069	0.083	**0.023**	0.143	0.082
Fields	0.093	0.109	0.090	**0.079**	0.119	0.128
Urban Areas	0.334	0.250	0.152	**0.010**	0.255	0.220
Average	0.181	0.143	0.108	**0.068**	0.172	0.143

(a)

	Averaged Normalized Rank (ANR)					
	ILSIA		N-FINDR		FIPPI	
Category	ED	SAM	ED	SAM	ED	SAM
Forests	0.140	0.192	0.197	0.301	**0.136**	0.191
Fields	0.170	0.154	0.136	**0.133**	0.178	0.171
Urban Areas	0.374	0.318	0.316	0.298	0.322	**0.268**
Average	0.228	0.221	0.216	0.244	0.212	**0.210**

(b)

Table 6.4: ANR results for HyMAP experiment 1: (a) Spectral-Spatial CBIR (b) Plaza's CBIR.

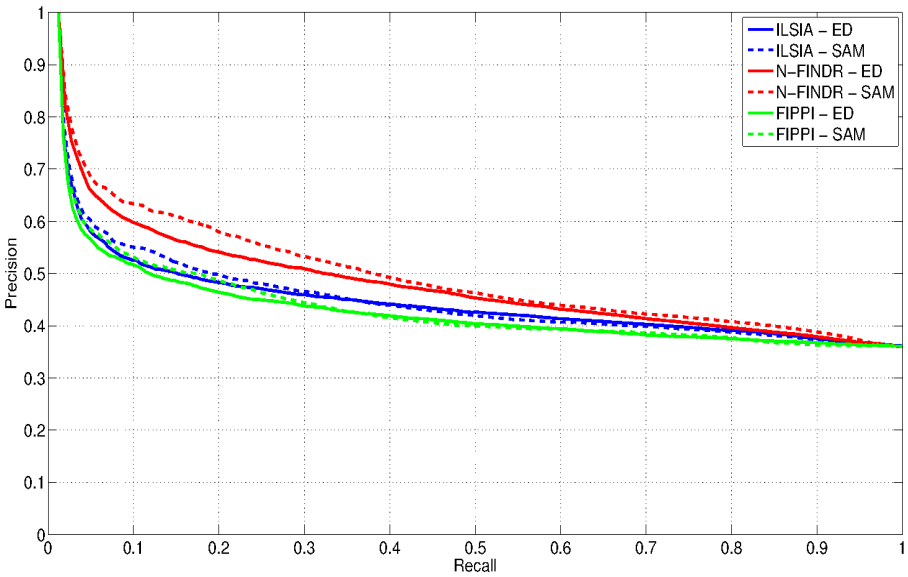
	Averaged Normalized Rank (ANR)					
	ILSIA		N-FINDR		FIPPI	
Category	ED	SAM	ED	SAM	ED	SAM
Forests	0.146	0.097	0.115	**0.054**	0.169	0.113
Fields	0.194	0.213	0.187	**0.180**	0.219	0.227
Urban Areas	0.338	0.255	0.156	**0.108**	0.253	0.220
Mixed	0.356	**0.340**	0.352	0.348	0.364	0.359
Average	0.259	0.226	0.203	**0.173**	0.251	0.230

(a)

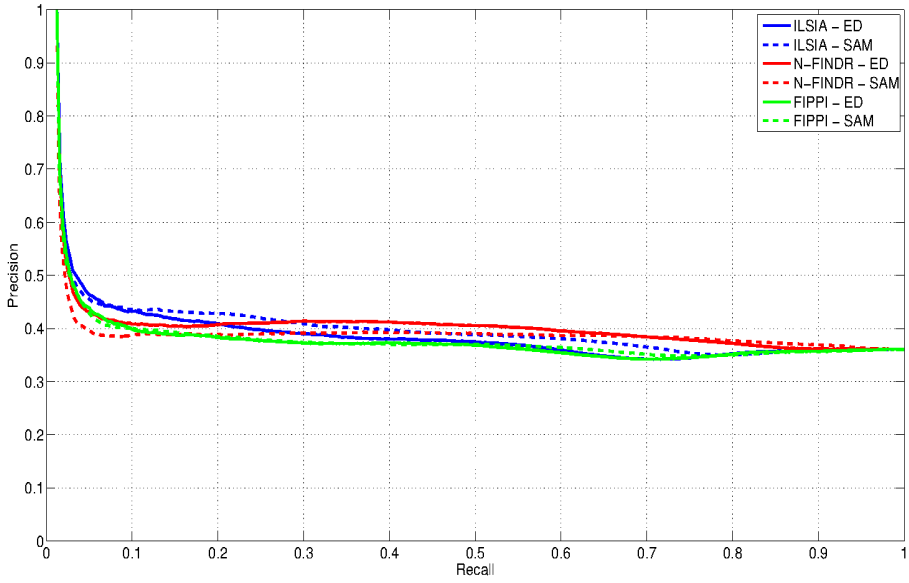
	Averaged Normalized Rank (ANR)					
	ILSIA		N-FINDR		FIPPI	
Category	ED	SAM	ED	SAM	ED	SAM
Forests	0.176	0.230	0.242	0.329	**0.160**	0.214
Fields	0.275	0.261	0.239	**0.236**	0.283	0.278
Urban Areas	0.405	0.350	0.335	0.324	0.345	**0.287**
Mixed	0.337	**0.320**	0.333	0.341	0.341	0.338
Average	0.298	0.290	0.287	0.307	0.282	**0.279**

(b)

Table 6.5: ANR results for HyMAP experiment 2: (a) Spectral-Spatial CBIR (b) Plaza's CBIR.

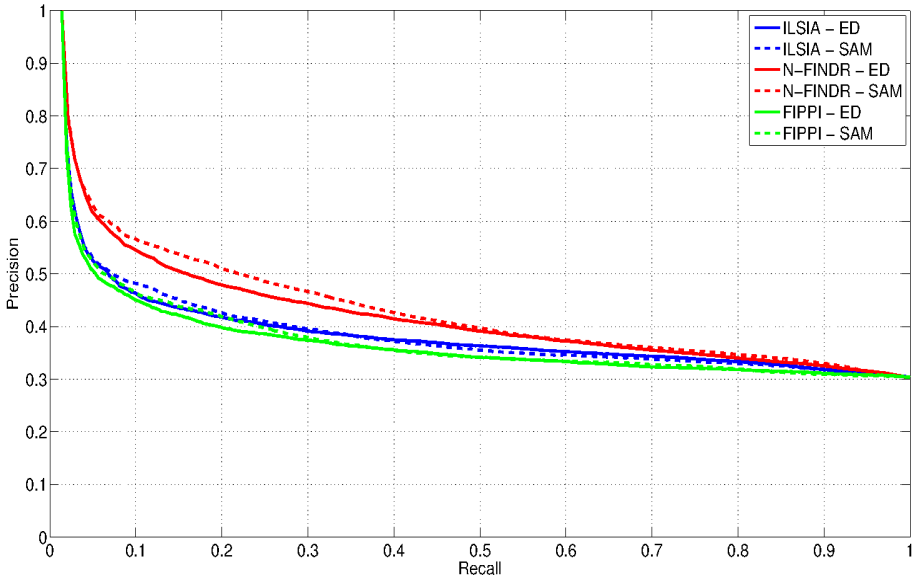


(a)

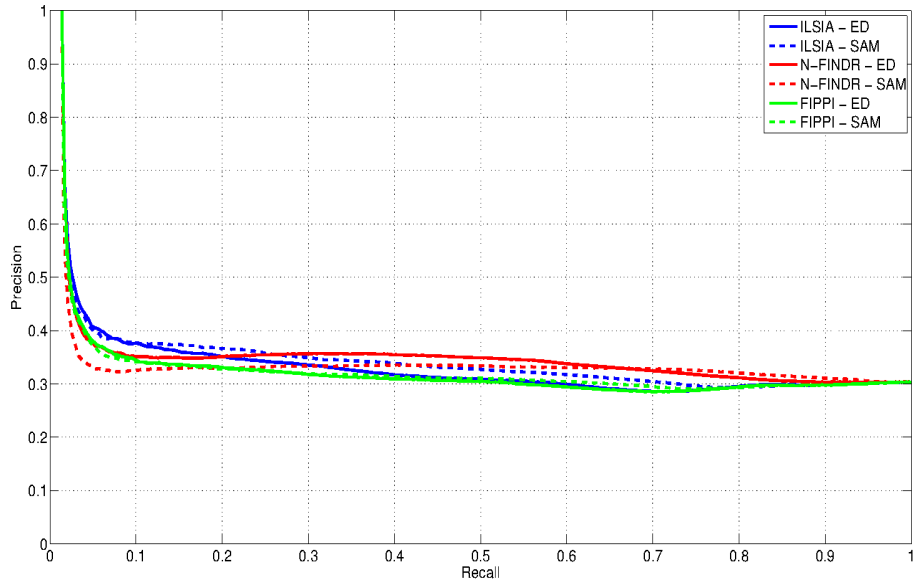


(b)

Figure 6.7: Precision-recall curves for HyMAP experiment 2: (a) Spectral-Spatial CBIR (b) Plaza's CBIR.



(a)



(b)

Figure 6.8: Precision-recall curves for HyMAP experiment 3: (a) Spectral-Spatial CBIR (b) Plaza's CBIR.

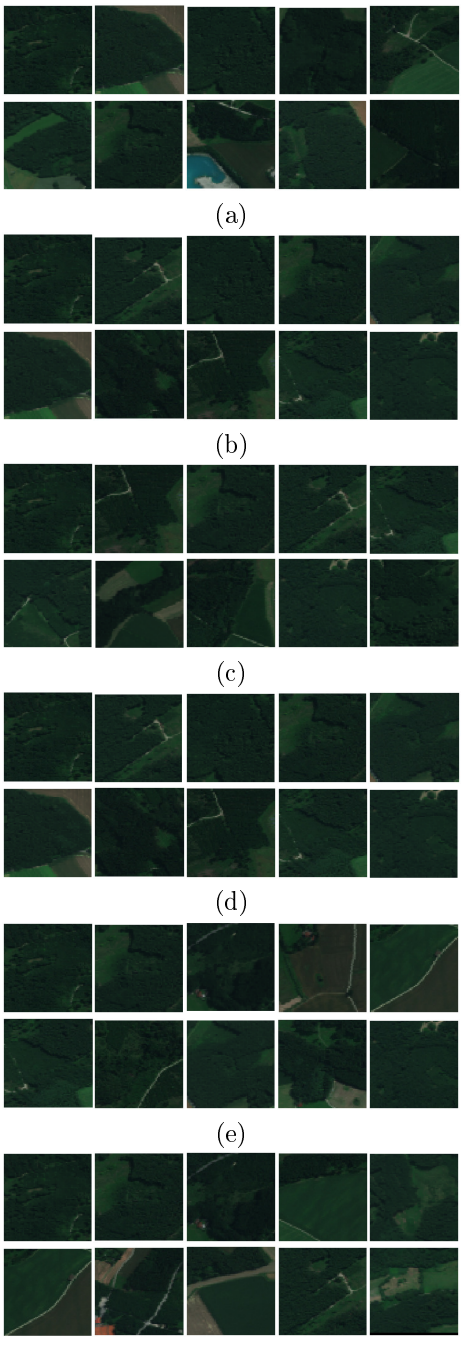
	Averaged Normalized Rank (ANR)					
	ILSIA		N-FINDR		FIPPI	
Category	ED	SAM	ED	SAM	ED	SAM
Forests	0.139	0.095	0.109	**0.051**	0.166	0.113
Fields	0.211	0.228	0.197	**0.196**	0.236	0.242
Urban Areas	0.336	0.261	0.155	**0.119**	0.250	0.224
Mixed	0.360	**0.347**	0.352	0.356	0.372	0.368
Other	0.460	0.466	0.455	**0.404**	0.482	0.463
Average	0.301	0.280	0.254	**0.225**	0.301	0.282

(a)

	Averaged Normalized Rank (ANR)					
	ILSIA		N-FINDR		FIPPI	
Category	ED	SAM	ED	SAM	ED	SAM
Forests	0.179	0.233	0.240	0.326	**0.162**	0.217
Fields	0.304	0.287	0.260	**0.256**	0.313	0.304
Urban Areas	0.419	0.361	0.341	0.334	0.357	**0.295**
Mixed	0.360	**0.341**	0.348	0.354	0.366	0.359
Other	**0.352**	0.378	0.403	0.412	0.357	0.387
Average	0.323	0.320	0.319	0.336	**0.311**	0.313

(b)

Table 6.6: ANR results for HyMAP experiment 3: (a) Spectral-Spatial CBIR (b) Plaza's CBIR.



(a) (b) (c) (d) (e) (f)

Figure 6.9: Images retrieved by the proposed Spectral/Spatial CBIR system for a Forest query example from experiment 3: (a) ILSIA + ED (b) ILSIA + SAM (c) N-FINDR + ED (d) N-FINDR + SAM (e) FIPPI + ED (f) FIPPI + SAM.

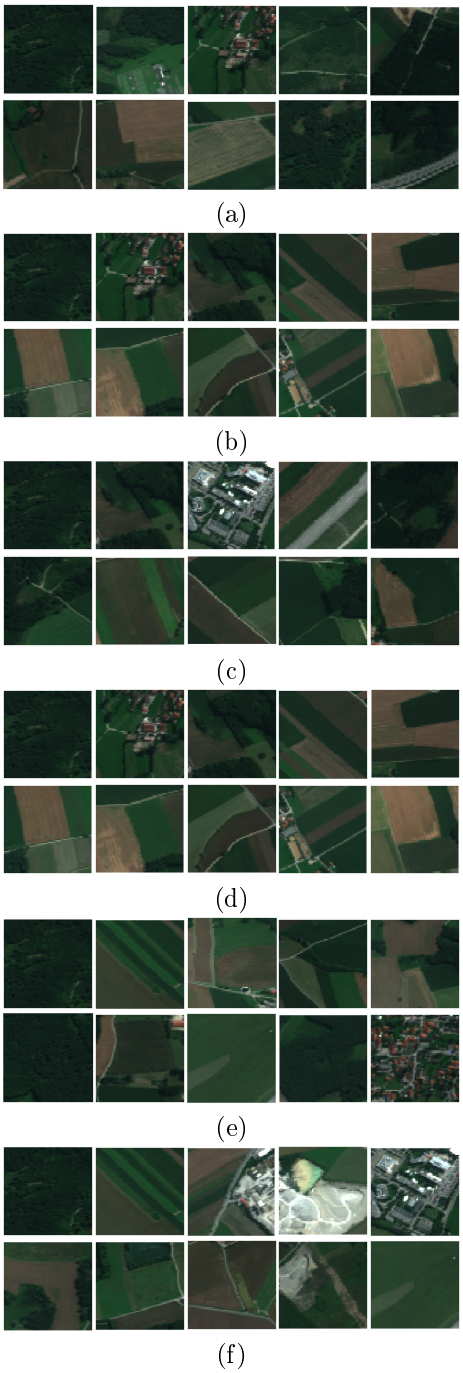


Figure 6.10: Images retrieved by Plaza's CBIR system for a Forest query example from experiment 3: (a) ILSIA + ED (b) ILSIA + SAM (c) N-FINDR + ED (d) N-FINDR + SAM (e) FIPPI + ED (f) FIPPI + SAM.

Chapter 7

Hyperspectral CBIR system: dictionary dissimilarity

The aim of this chapter is to introduce CBIR systems based on parameter-less and abstract distances, which can be computed on hyperspectral images or any other kind of signals. Such kind of distances open the way for the construction of multi-modal CBIR systems, which can include diverse remote sensing informations. The starting point is the normalized compression distance (NCD) which is a computable approximation to the normalized information distance (NID). However, the computational cost of NCD for CBIR is prohibitive. Distances based on dictionaries induced by dictionary-based compressors, may overcome these problem. In this chapter, these distances are proposed and tested for hyperspectral image CBIR, comparing the Normalized Dictionary Distance (NDD) and the Fast Dictionary Distance (FDD) against the NCD over different datasets of hyperspectral images.

The chapter structure is as follows: Section 7.1 gives a short introduction. Sections 7.2 reviews the normalized compression distance. Section 7.3 reports a classification experiment using NCD. Section 7.4 reviews the dictionary distances FDD and NDD. Section 7.5 introduces the hyperspectral CBIR system based on dictionary distances, and some experimental results on real data. Finally, we present some conclusions in Section 7.6.

7.1 Introduction

Kolmogorov complexity lies in the core of *algorithmic information theory* [27, 172] that focuses on the information of individual signals, an approach completely different to classical Shannon's probabilistic approach to information theory [166]. The normalized information distance (NID) [15] is an universal metric distance based on Kolmogorov complexity [109]. However, Kolmogorov complexity is not computable in the Turing sense. The normalized compression distance (NCD) [108] is a computable distance that approximates NID by using

normal compressors. There has been an increasing interest in using NCD for pattern recognition [192] and in the last years NCD has been successfully applied to different pattern recognition problems including remote sensing [26, 25].

The NCD approach to pattern recognition is parameter-free (except for the compressor's internal parameters configuration) avoiding to tune up parameters to realize operative implementations of CBIR systems. Moreover, it does not require any feature extraction process. However, the use of NCD in a CBIR system demands a high computational cost due to the need of performing the compression of the concatenations of the query image to each of the images in the database. The use of dictionary distances [118, 25] has been proposed to provide an approximation to NCD when computational cost is an issue. Thus, we propose a CBIR system based on dictionary distances for the mining of remote sensing large collections of hyperspectral images. We compare the use of dictionaries to the use of NCD in three datasets of real hyperspectral images. Results validate the proposed dictionary-based hyperspectral CBIR system.

7.2 Normalized Compression Distance

The *conditional Kolmogorov complexity* of a signal x given a signal y, $K(x|y)$, is the length of the shortest program running in an universal Turing machine, that outputs x when fed with input y. The *Kolmogorov complexity* of x, $K(x)$, is the length of the shortest program that outputs x when fed with the empty signal λ, that is, $K(x) = K(x|\lambda)$. The *information distance*, $E(x,y)$, is an universal metric distance defined as the length of the shortest binary program in a Turing sense that, from input x outputs y, and from input y outputs x. It is formulated as:

$$E(x,y) = \max\{K(x|y), K(y|x)\}. \tag{7.1}$$

The *normalized information distance*, $NID(x,y)$, is defined as:

$$NID(x,y) = \frac{E(x,y)}{\max\{K(x), K(y)\}}. \tag{7.2}$$

The NID is sometimes known as the *similarity metric* due to its universality property. Here, universality means that the NID is a lower bound for any distance $D(x,y)$ defined on the object's space, i.e. $E(x,y) \leq D(x,y)$, up to an additive constant depending on D but not on x and y. However, $NID(x,y)$ is a function of the Kolmogorov complexity which is non-computable in the Turing sense.

The *normalized compression distance*, $NCD(x,y)$, is a computable version of equation (7.2) based on a given compressor C. It is defined as:

$$NCD(x,y) = \frac{C(xy) - \min\{C(x), C(y)\}}{\max\{C(x), C(y)\}}, \tag{7.3}$$

where $C(\cdot)$ is the length of the signal compressed by using compressor C, and xy is the signal resulting of the concatenation of signals x and y. If the compressor C

is normal, then the NCD is a quasi-universal similarity metric. In the limit when $C(\cdot) = K(\cdot)$, the $NCD(x,y)$ becomes "universal". The distance $NCD(x,y)$ differs from distance $NID(x,y)$ in three aspects [42]:

(a) The universality of $NID(x,y)$ holds only for indefinitely long sequences x,y. When dealing with sequences of finite length n, universality holds only for normalized admissible distances computable by programs whose length is logarithmic in n.

(b) The Kolmogorov complexity is not computable, and it is impossible to know the degree of approximation of $NCD(x,y)$ with respect to $NID(x,y)$.

(c) To calculate the $NCD(x,y)$ only an standard lossless compressor C is needed.

Although better compression implies a better approximation to Kolmogorov complexity, this may not be true for $NCD(x,y)$. A better compressor may not improve compression for all items in the same proportion. Experiments show that differences are not significant if the inner requirements of the underlying compressor C are not violated.

7.3 A classification experiment based on NCD

Figure 7.1 shows a block diagram of the proposed experiment to show the power of NCD for classification problems. The goal is to calculate a distance between two hyperspectral images without the need of selecting/extracting features or parameter tuning. First, hyperspectral images are converted to strings by concatenating their pixels' spectra. To compare two hyperspectral images H_α and H_β, their string representations x_α and x_β, respectively, are compressed by a loss-less compressor C, and their individual compression factors $C(x_\alpha)$ and $C(x_\beta)$ are calculated. Then, x_α and x_β are concatenated and compressed to calculate $C(x_\alpha, x_\beta)$. Now, we can measure the distance between the two images, H_α and H_β, using the Normalized Compression Distance, $NCD(x_\alpha, x_\beta)$ as it was defined in equation (7.3). The matrix of NCD distances between pairs of hyperspectral images from a hyperspectral dataset is:

$$D = \{d_{i,j}\}\,;\; i,j = 1,..,N, \tag{7.4}$$

where N is the number of images in the dataset, and $d_{i,j}$ is the NCD distance between images H_i and H_j. Then, the matrix D of NCD distances can be used as the input to a classifier.

7.3.1 Performance results

The experiments are performed on the HyMAP data described in Appendix D to show the value of NCD for hyperspectral clustering/classification, and we compared the results to other methodologies found on the literature.

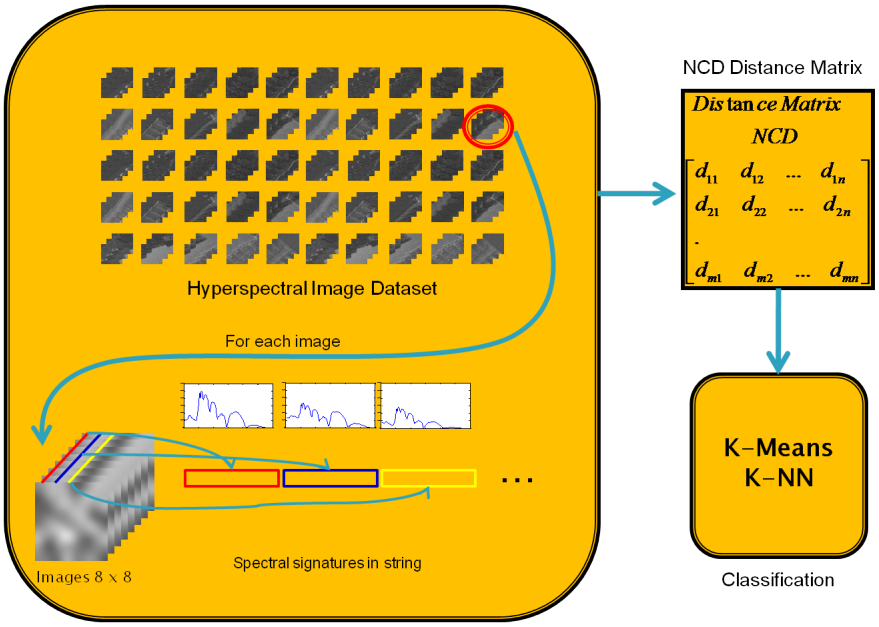


Figure 7.1: Block diagram of the proposed methodology

We took the original image and divided it into several patches of 8×8 pixels, and we selected by visual inspection 130 among them corresponding to three different classes: buildings (30 patches), fields (50 patches) and forests (50 patches). We transformed these patches into strings as it has been described above, and we calculated the matrix of distances D (7.4) between all image pairs using the NCD function (7.3). Then, we used the distance matrix D as the input to well known algorithms: K-Means (unsupervised clustering) and K-NN (supervised classification).

We compared the results obtained with the proposed NCD-based methodology to results obtained using the average patch radiance and the induced endmembers characterization:

- The average patch radiance $\bar{r}_i$ for a patch image H_i is given by $\bar{r}_i = \frac{1}{N}\sum_{j=1}^{N} H_i^{(j)}$, where $H_i^{(j)}$ is the j-th pixel belonging to patch i, and N is the total number of pixels in the patch ($N = 64$ in our case).

- In the endmembers approach, we used EIHA algorithm to induce a set of endmembers $E_i = (\mathbf{e}_{i1}, \ldots, \mathbf{e}_{ip_i})$ for a patch H_i where p_i is the number of endmembers obtained from the patch H_i. Then, we calculated an endmember distance matrix analogous to the NCD distance matrix used as input to the classifiers by the spectral dissimilarity function.

All the experiments have been run using a K-Fold resampling with 10 folds. Tables 7.1 and 7.2 show the confusion matrix, the overall accuracy and the

NCD distances	Buildings	Fields	Forests	Total
Buildings	30	0	0	30
Fields	5	32	13	50
Forests	0	13	37	50

Overall accuracy: 76.15%. *KHAT*: 74.81%.

Average radiance	Buildings	Fields	Forests	Total
Buildings	26	4	0	30
Fields	12	38	0	50
Forests	0	0	50	50

Overall accuracy: 87.69%. *KHAT*: 87.14%.

Endmembers	Buildings	Fields	Forests	Total
Buildings	28	2	0	30
Fields	12	38	0	50
Forests	0	0	50	50

Overall accuracy: 89.23%. *KHAT*: 88.69%.

Table 7.1: Classification Results using the unsupervised K-Means algorithm

KHAT index for the K-Means and K-NN algorithms respectively.

7.4 Dictionary distances

The use of NCD (7.3) for CBIR entails an unfordable cost due to the requirement of compressing the concatenated signals, $C(xy)$. This Thesis proposes distances based on the codewords of the dictionaries extracted by means of dictionary-based compressors, such as the LZW for text strings. This dictionary approach only requires set theory operators to calculate the distance between two signals based on the dictionaries extracted from the signals by the compression algorithm. Thus, dictionary distances are suitable for mining large image databases where the dictionaries of the images in the database can be extracted off-line.

Given a signal x, a dictionary-based compression algorithm looks for patterns in the input sequence from signal x. These patterns, called *words*, are subsequences of the incoming sequence. The compression algorithm produces a set of unique words called *dictionary*. The dictionary extracted from a signal x is hereafter denoted as $D(x)$, with $D(\lambda) = \emptyset$ only if λ is the empty signal. The union and intersection of the dictionaries extracted from signals x and y are denoted as $D(x \cup y)$ and $D(x \cap y)$ respectively. The dictionaries satisfy the following properties [118]:

1. Idempotency: $D(x \cup x) = D(x)$.
2. Monotonicity: $D(x \cup y) \geq D(x)$.
3. Symmetry: $D(x \cup y) = D(y \cup x)$.
4. Distributivity: $D(x \cup y) + D(z) \leq D(x \cup z) + D(y \cup z)$.

NCD distances	Buildings	Fields	Forests	Total
Buildings	30	0	0	30
Fields	0	47	3	50
Forests	0	8	42	50

Overall accuracy: 91.54%. *KHAT*: 91.06%.

Average radiance	Buildings	Fields	Forests	Total
Buildings	26	4	0	30
Fields	2	48	0	50
Forests	0	0	50	50

Overall accuracy: 95.38%. *KHAT*: 95.18%.

Endmembers	Buildings	Fields	Forests	Total
Buildings	22	7	1	30
Fields	2	48	0	50
Forests	1	0	49	50

Overall accuracy: 91.53%. *KHAT*: 91.28%.

Table 7.2: Classification results using the supervised K-NN algorithm

We have found two dictionary distance functions in the literature, the Normalized Dictionary Distance (NDD) [118] and the Fast Dictionary Distance (FDD) [25]:

$$NDD\left(x,y\right)=\frac{D\left(x\cup y\right)-\min\left\{D\left(x\right),D\left(y\right)\right\}}{\max\left\{D\left(x\right),D\left(y\right)\right\}}, \tag{7.5}$$

$$FDD\left(x,y\right)=\frac{D\left(x\right)-D\left(x\cap y\right)}{D\left(x\right)}. \tag{7.6}$$

NDD and FDD are both normalized admissible distances satisfying the metric inequalities. Thus, they result in a non-negative number in the interval $[0,1]$, being zero when the compared files are equal and increasing up to one as the files are more dissimilar.

7.5 Dictionary CBIR system

Figure 7.2 shows the Hyperspectral CBIR system scheme based on dictionaries. The core of the CBIR system is the dictionary distance between two hyperspectral images by means of their previously extracted dictionaries. The system interacts with a dictionary database where the images dictionaries are stored. These dictionaries have been previously extracted by off-line application of a dictionary-based compression algorithm. System interrogation is done using a query example approach. Firstly, the query example is processed to extract its dictionary and secondly, it is compared to the images in the database using the dictionary distance. A ranking of the images in the database is elaborated by ascending order of dissimilarity (ascending distance) to the query. Finally, the

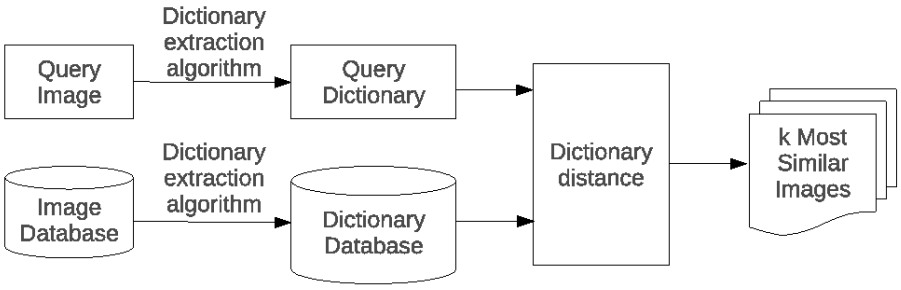


Figure 7.2: Hyperspectral CBIR based on dictionary schema

system returns the k images in the database corresponding to the first k ranking positions, where k is known as the query's *scope*.

7.5.1 Experimental methodology

We independently test the NCD (7.3), the NDD (7.5) and the FDD (7.6) in three experiments corresponding to different HyMAP datasets. First dataset includes the Forests, Fields and Urban Areas categories. Second dataset adds the Mixed category. In the third dataset all images are considered and so, we included the Others category. Each hyperspectral image is first converted to a text file in two ways: pixel-wise and band-wise. Given that a image in a dataset is 64×64 pixels size and has 125 bands, in the pixel-wise ordering the text file is built concatenating the pixels of the images in a zig-zag way, where a pixel is a 125-components vector. In the band-wise ordering the text file is built concatenating the bands of the image, where a band is reordered in zig-zag to form a 64^2-components vector. The NDD and FDD are calculated using the dictionaries extracted by the LZW compression algorithm. The NCD is calculated by CompLearn[1] software using default options, that is BZLIB compressor. The performance measures are computed as described in Chapter 4 for the evaluation of CBIR systems on real data, as have already been used in Chapters 5 and 6.

7.5.2 Performance results

Precision-recall Figures 7.3-7.5 show the precision-recall curves for experiments 1, 2 and 3 respectively. In each figure six precision-recall curves are drawn, corresponding to the three compared distances, NDD, FDD and NCD, applied to the datasets converted into text strings using pixel-wise and band-wise orderings. In all the experiments NDD outperforms the other distances independently of the image to text string conversion ordering used. NCD outperforms FDD showing that the lack of a normalization factor in the FDD is an important issue,

[1] http://www.complearn.org

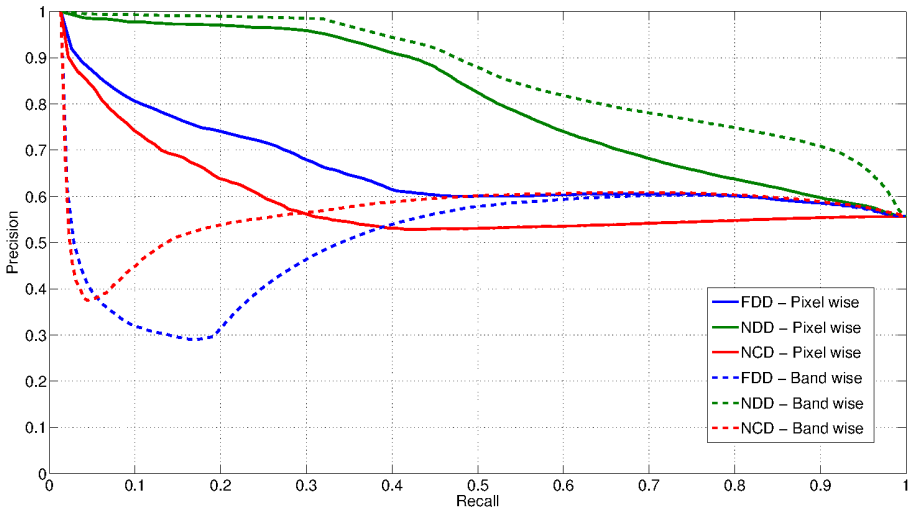


Figure 7.3: Precision-recall curves for HyMAP experiment 1.

affecting the performance of the retrieval system. Furthermore, we expected the band-wise ordering to perform better than the pixel-wise ordering due to the high correlation on consecutive bands. Accordingly, the band-wise NDD gives the best performance in all the experiments. However, surprisingly, the band-wise ordering shows a bad performance for low recall values using FDD and NCD, improving as the recall values increase up to performances similar to the pixel-wise ordering. In general, the performance decreases smoothly as we include hardest categories, 'Mixed' category in experiment 2 and 'Others' category in experiment 3, yielding still good precision-recall values for the NDD function. Also, NCD presents a general lower precision compare to dictionary-based distances, although its performance decreases more slowly than the performances of NDD and FDD as we add more difficult categories.

ANR Tables 7.3-7.5 show the Average Normalized Rank (ANR) for the experiments 1, 2 and 3 respectively. ANR results confirms the average outperform of NDD over FDD and NCD, although FDD slightly outperforms NDD in some cases. Interestingly, ANR can partially explain the effect in the FDD and NCD precision-recall curves using band-wise ordering for low recall values, as it shows FDD is having problems retrieving the 'Fields' category and NCD is having problems retrieving the 'Forests' and 'Urban Areas' categories. Further experiments must be conduced to give a better explanation to why band-wise ordering affects so much FDD and NCD performance.

	ANR		
Category	**FDD**	**NDD**	**NCD**
Forests	0.015	**0.010**	0.129
Fields	0.143	**0.090**	0.180
Urban Areas	**0.005**	**0.005**	0.086
Average	0.055	**0.035**	0.132

(a) Pixel-wise ordering

	ANR		
Category	**FDD**	**NDD**	**NCD**
Forests	0.073	**0.014**	0.292
Fields	0.159	**0.038**	0.118
Urban Areas	**0.004**	**0.004**	0.668
Average	0.079	**0.019**	0.359

(b) Band-wise ordering

Table 7.3: ANR results for HyMAP experiment 1.

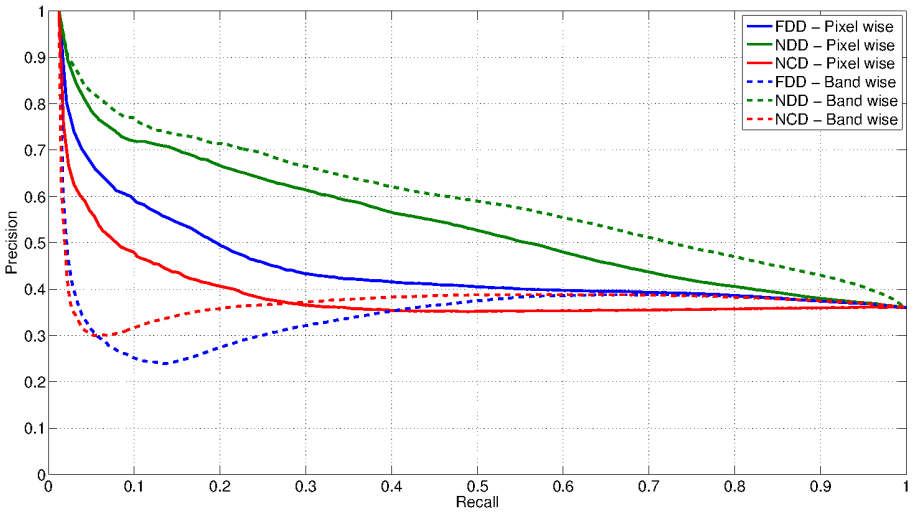


Figure 7.4: Precision-recall curves for HyMAP experiment 2.

	ANR		
Category	**FDD**	**NDD**	**NCD**
Forests	0.069	**0.053**	0.168
Fields	0.283	**0.210**	0.299
Urban Areas	**0.011**	0.012	0.108
Mixed	**0.223**	0.236	0.311
Average	0.146	**0.128**	0.222

(a) Pixel-wise ordering

	ANR		
Category	**FDD**	**NDD**	**NCD**
Forests	0.130	**0.064**	0.310
Fields	0.316	**0.142**	0.219
Urban Areas	**0.005**	0.006	0.681
Mixed	**0.219**	0.226	0.359
Average	0.167	**0.109**	0.392

(b) Band-wise ordering

Table 7.4: ANR results for HyMAP experiment 2.

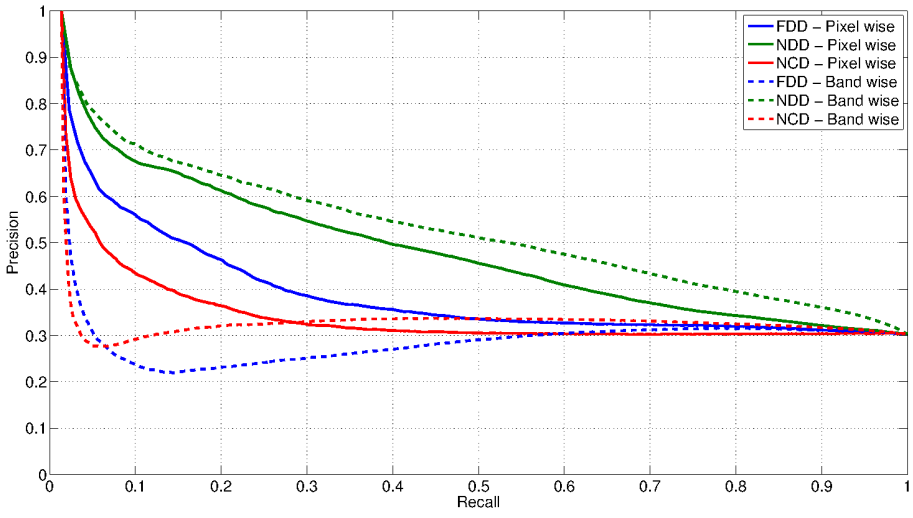


Figure 7.5: Precision-recall curves for HyMAP experiment 3.

	ANR		
Category	**FDD**	**NDD**	**NCD**
Forests	0.065	**0.049**	0.162
Fields	0.323	**0.235**	0.315
Urban Areas	**0.011**	0.013	0.107
Mixed	**0.246**	0.254	0.318
Others	**0.197**	0.232	0.425
Average	0.169	**0.156**	0.266

(a) Pixel-wise ordering

	ANR		
Category	**FDD**	**NDD**	**NCD**
Forests	0.130	**0.061**	0.304
Fields	0.369	**0.164**	0.226
Urban Areas	0.006	**0.008**	0.674
Mixed	0.254	0.250	0.360
Others	**0.177**	0.210	0.570
Average	0.187	**0.139**	0.427

(b) Band-wise ordering

Table 7.5: ANR results for HyMAP experiment 3.

7.6 Conclusions

This chapter introduces a CBIR System for hyperspectral databases based on dictionary distances. The use of a parameter-free approach based on the Normalized Compression Distance (NCD) is not feasible in general due to the computational cost of performing the compression of the signals resulting from the concatenations of the query image and each image in the database. The dictionary distances approach solves the computational cost problem by approximating NCD using dictionaries extracted offline from each of the database images. Results using real hyperspectral datasets show that the Normalized Dictionary Distance (NDD) outperforms the Fast Dictionary Distance (FDD) and the NCD. We also show that in order to extract the dictionaries (or compress the signals for the NCD) the arrangement of the image data in the conversion of the image to a text file affects severely the performance of the FDD and NCD similarity functions. Generally, we can conclude that the presented results validate the use of dictionaries for hyperspectral image retrieval.

Chapter 8

Hyperspectral CBIR system: relevance feedback by dissimilarity spaces

This chapter proposes a novel relevance feedback (RF) methodology specifically suited for hyperspectral Content-Based Image Retrieval (CBIR) systems using dissimilarity spaces instead of feature spaces to solve this problem. We validated the proposed RF for hyperspectral CBIR systems over a real dataset with very promising results.

The chapter structure is as follows. Section 8.1 gives a short introduction. Section 8.2 introduces the proposed RF based on dissimilarity spaces. Section 8.3 define the experimental methodology. Section 8.4 reports the results of an experiment on real hyperspectral image data. Finally, we contribute with some conclusions in section 8.5.

8.1 Introduction

In previous chapters we have dealt with various feature extraction and subsequent distances defined between hyperspectral images in feature spaces expanding previous research in hyperspectral CBIR systems defined over dissimilarity functions. These dissimilarity functions were based either on spectral and spatial features extracted by spectral unmixing techniques or in dictionaries extracted by dictionary-based compressors. The spectral-spatial and the dictionary feature space as well as the dissimilarity functions built upon them were not suitable for direct application in common machine learning techniques.

It is possible to treat dissimilarity functions as kernels in order to use them in kernel-based algorithms such as Support Vector Machines (SVM). However, these dissimilarity functions do not comply often with valid kernel conditions [135]. Authors in [134] propose the definition of dissimilarity spaces as an alter-

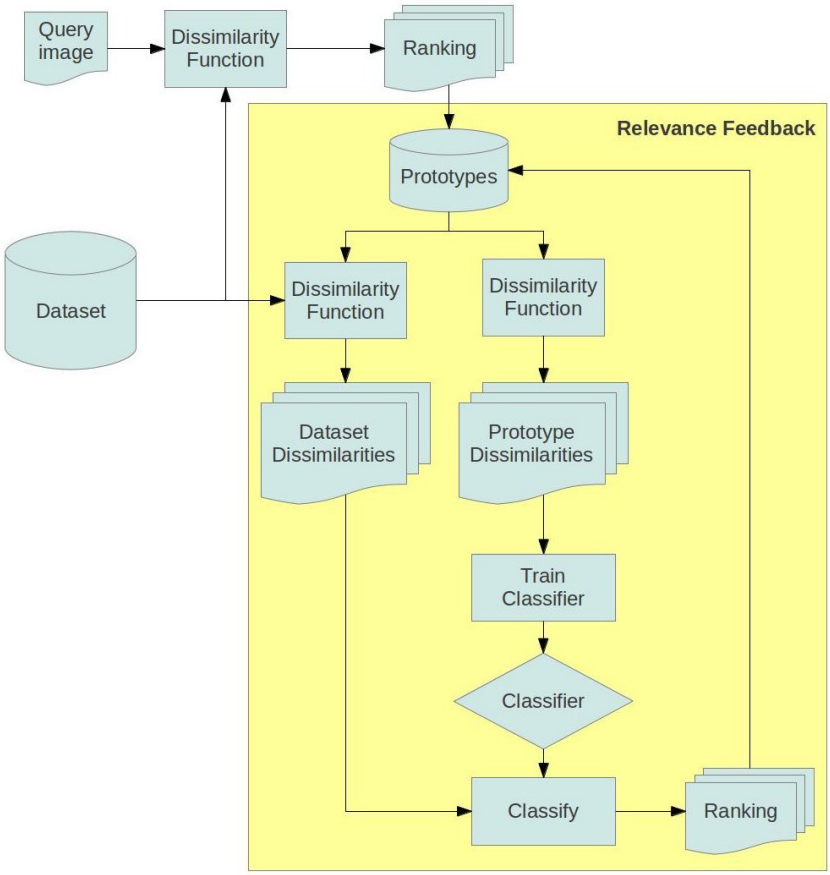


Figure 8.1: CBIR system diagram with the relevance feedback based on dissimilarity space.

native to feature spaces. In dissimilarity spaces some data instances are used as reference points, aka prototypes. These prototypes define a dissimilarity space wherein each dimension coordinates represent the dissimilarity to a prototype. This way a data point is represented in the dissimilarity space by their dissimilarity to the prototypes. The dissimilarity space can be used as a feature space so all the available potential of machine learning techniques can be used. In this paper we propose the use of dissimilarity spaces to define a relevance feedback methodology for hyperspectral CBIR making use of the already available dissimilarity functions.

8.2 Relevance feedback on dissimilarity space

Figure 8.1 shows a diagram of the proposed approach.

- First, a query, $Q_l(H_\alpha)$, is defined following the query-by-image approach. H_α denotes the query image, and $l \in \mathbb{Z}^+$ is the scope of the query, i.e. the number of images that should be retrieved to answer the query.

- Every image H_β in the dataset is compared to the query image by computing some dissimilarity function $s(H_\alpha, H_\beta) = s_{\alpha,\beta}$. These query dissimilarities are aggregated as a vector $\mathbf{s}_\alpha = [s_{\alpha,1}, \ldots, s_{\alpha,N}]$, where N is the number of images in the dataset.

- We sort in increasing order the components of $\mathbf{s}_\alpha$, and the resulting shuffled image indexes constitute the ranking $\Omega_\alpha = [\omega_q \in \{1, \ldots, N\}]$, $q = 1, \ldots, N$, so that $s_{\alpha,\omega_q} \leq s_{\alpha,\omega_{q+1}}$. The ranking Ω_α denotes the zero ranking of the dataset images respect to the query image H_α, that is, an initial ranking obtained directly from the dissimilarity function before entering the relevance feedback process, usually known as the *zero query*. We'll denote this zero ranking as Ω_α^0 to reinforce this notion.

- The first l images of the zero ranking are retrieved for the user's evaluation. These images form the set of prototypes,

$$P = \{H_i\}_{i=1}^{p}, i \in \{1, \ldots, N\},$$

  which with the relevance feedback process starts. Each prototype is used as a reference point the rest of the prototypes and the dataset images would be compared to. Thus, the dissimilarity to a prototype image represents a dimension in the dissimilarity space of the prototypes, denoted as $\mathcal{P}$.

- This dissimilarity space of the prototypes $\mathcal{P}$ is used to perform relevance feedback, calculating the dissimilarities among prototype images, and using them as feature vectors to train a two-class classifier, where prototypes are labeled as belonging to the query class, $\mathcal{C}^+$, the *positive class*, or to its complement, the *negative class*, $\mathcal{C}^-$.

- For each image H_β in the dataset we calculate the dissimilarity vector

$$\mathbf{s}_\beta = [s_{\beta,i}]_{i=1}^{p},$$

  where $s_{\beta,i}$ denotes the dissimilarity between dataset image H_β and prototype image H_i. The dissimilarity vector, $\mathbf{s}_\beta$, represents geometrically a point in P and would be used as the input to the classifier once trained.

- The classifier will return a probability of the image H_β belonging to the query class,

$$p_\beta(\mathcal{C}^+) = 1 - p_\beta(\mathcal{C}^-).$$

  The probabilities given by the classifier are represented as a vector $\mathbf{p}_\alpha = [p_1, \ldots, p_N]$, where N is the number of images in the dataset and p_β is a simplified notation for $p_\beta(\mathcal{C}^+)$, with $\beta = 1, \ldots, N$.

- We sort in increasing order the components of $\mathbf{p}_\alpha$, and the resulting shuffled image indexes constitute the ranking $\Omega_\alpha^t = [\omega_q \in \{1, \ldots, N\}]$, $q = 1, \ldots, N$, so that $p_{\omega_q} \leq p_{\omega_{q+1}}$. The superscript t in Ω_α^t denotes the iteration number of the relevance feedback process, being t a positive integer, $t > 0$. The first l images in Ω_α^t are incorporated to the prototypes set if they were not already present, and a new iteration begins.

- The relevance feedback process ends when the user is satisfied, a maximum number of iterations, $t_{\max}$, is achieved, or no new images are being incorporated to the prototypes set.

8.3 Experimental methodology

We run five retrieval experiments over the HyMAP dataset using a k-NN classifier with $k = \{1, 3, 5, 10, 20\}$. The Normalized Dictionary Distance (NDD) was used to compare the hyperspectral images and a scope, $l = 20$, was selected. Each of the 360 patches was a-priori labeled as belonging to one of the five categories defined above, so user's evaluation was not required and the experiment was fully automatized. The maximum number of retrieval feedback iterations was set to $t_{\max} = 10$.

We calculated the Rank (H_α) for each of the images in the dataset and then we calculated the overall normalized rank (ONR) and the average normalized rank (ANR) defined as follows:

$$ONR = \frac{1}{N} \sum_{\alpha=1}^{N} Rank\,(H_\alpha)\,, \tag{8.1}$$

$$ANR = \sum_{c=1}^{C} \sum_{\alpha=1}^{N} I(H_\alpha, c) Rank\,(H_\alpha)\,. \tag{8.2}$$

where C is the number of categories, $C = 5$ in our experiments, and $I\,(H_\alpha, c)$ is an indicator function that returns 1 if the image H_α belongs to the class c and 0 otherwise. The ONR is the averaged normalized rank over all the dataset, and the ANR is the class-averaged normalized rank.

8.4 Performance results

Figure 8.2 and 8.3 show the overall normalized (8.1) and the average normalized (8.2) results, respectively, for the proposed relevance feedback by dissimilarity spaces using different k-NN classifiers. As it can be seen there is a continuous overall and average improvement along the iterations. Black flat line in both figures denote the ONR and ANR values calculated for the response given by the system to the Zero-query using NDD dissimilarity only. For all cases the relevance feedback response improves the Zero-query response after iterations 2-3.

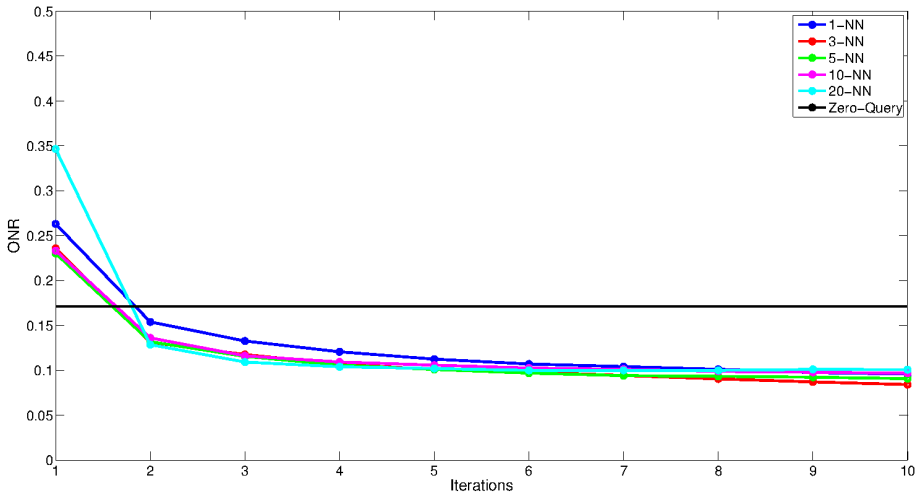


Figure 8.2: Overage normalized rank results.

	Class-specific ANR				
	Zero-Query	RF ($t = 1$)	RF ($t = 3$)	RF ($t = 5$)	RF ($t = 10$)
Forests	0.061	0.212	0.094	0.073	0.050
Fields	0.164	0.174	0.066	0.055	0.040
Urban Areas	0.008	0.108	0.004	0.001	0.002
Mixed	0.250	0.313	0.201	0.184	0.164
Others	0.210	0.400	0.209	0.167	0.140

Table 8.1: Comparison of class-specific ANR values between the Zero-query NDD response and relevance feedback (RF) response at different iterations (t).

The best result is achieved by the 3-NN classifier, whose class-specific results are given in figure 8.4. All the classes improve their performance by the proposed relevance feedback method. The Forest and Fields categories reach ANR values close to 0.05 whilst Urban Areas category approximates a perfect retrieval. The more complicated 'Mixed' and 'Others' categories yields ANR values above 0.15 still improving around 0.2 points in the relevance feedback process. Table 8.1 compares class-specific ANR values for the Zero-query respect to the relevance feedback at different iteration steps using the 3-NN classifier. We can see that relevance feedback improves class-specific ANR values for all classes, most of them at early iterations.

8.5 Conclusions

The chapter presented a relevance feedback method for hyperspectral CBIR systems using dissimilarity spaces. Defining a relevance feedback process for

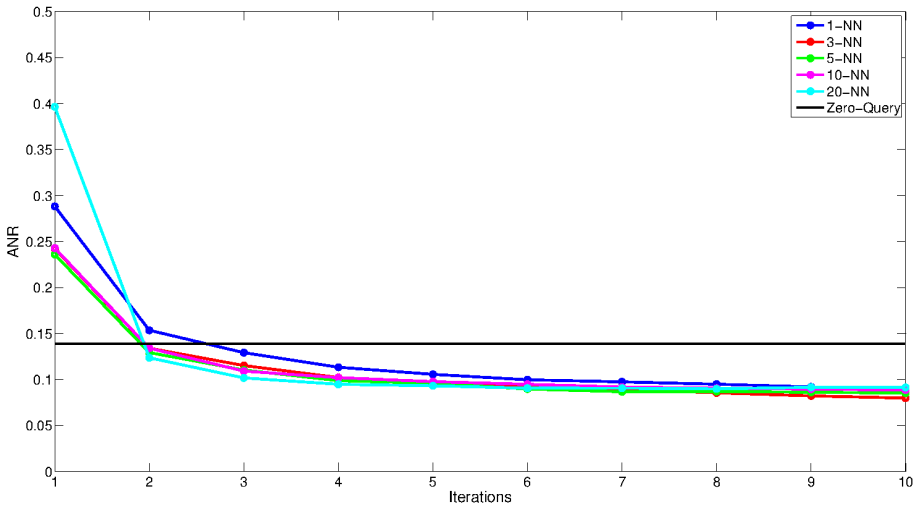


Figure 8.3: Average normalized rank results.

hyperspectral CBIR systems is not easy as most of the available hyperspectral CBIR systems rely on dissimilarity functions that do not fulfill conditions to be used in common machine learning techniques. The proposed approach expands the available dissimilarity-based hyperspectral CBIR systems on the literature in a simple way by defining a new space, the dissimilarity space, as opposed to the usual feature space. The proposed approach achieved very good results in the experiments.

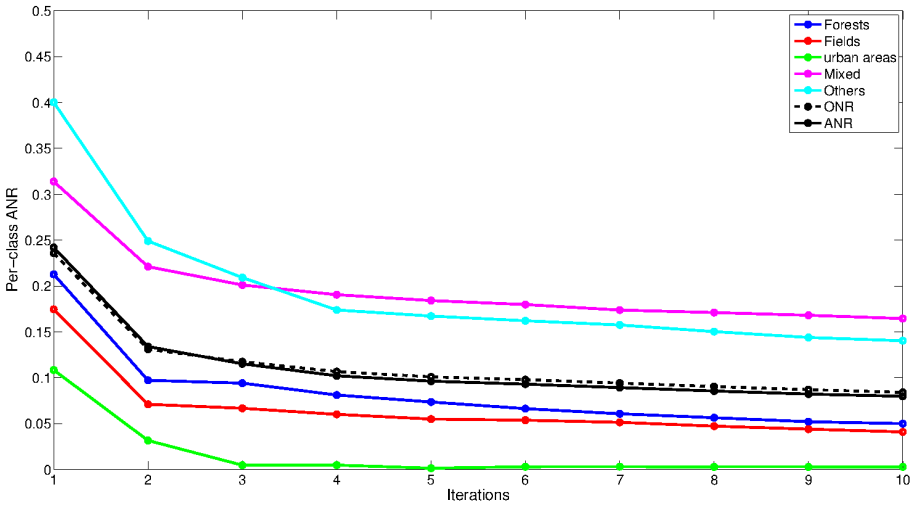


Figure 8.4: Per-class average normalized rank results.

Chapter 9

Multivariant morphology by LAAMs and hyperspectral segmentation

Mathematical Morphology has been a successful image processing technology for binary valued and grayscale images, however it has some open research tracks dealing with multivariate images, specifically with hyperspectral images. The difficulty to define an appropriate order on the multivariate data sets, i.e. the set of pixel spectra, hinders the definition of appropriate morphological operators and filters. One approach that shows promise is the definition of reduced orderings based on classification outputs. This chapter follows this approach to define reduced supervised orderings based on the recall distance of Lattice Auto-Associative Memories (LAAM). A one-side supervised ordering uses only one set of training data set associated to a foreground class, a background/foreground supervised ordering uses two training data sets one for each relevant class. The approach is first tested on RGB images, which are easy to interpret and, thus, allow for a qualitative evaluation. Results on hyperspectral images consists in the spatial-spectral classification using a watershed region segmentation based on the supervised orderings. Pixel spectra are individually classified by conventional Support Vector Machines (SVM). The spatial process is based on the assumption that watershed regions are composed of homogeneous pixels that must belong to the same class, therefore disagreements are solved by majority voting inside the watershed region.

The contents of the chapter are as follows: Section 9.1 introduces relevant concepts of Mathematical Morphology. Section 9.2 introduces the specific h-supervised orderings based on LAMM recall distance. Section 9.3 discusses the unsupervised selection of training samples. Section 9.4 reports experimental results on hyperspectral images. Finally, section 9.5 gives some conclusions on the work of this chapter.

9.1 Mathematical Morphology and Lattice Theory

Mathematical Morphology [163, 82, 164, 66, 171] has been very successful defining image operators and filters for gray-scale and binary images. Lattice Theory [18, 76, 77] gives the most general theoretical background for Mathematical Morphology [157, 84].

A binary relation ϱ satisfying reflexivity, antisymmetry and transitivity properties is called an *order,* denoted by $\leq$. A non-empty set P endowed with an order relation, $\langle P; \leq \rangle$, is a *partially-ordered set* or *poset*, denoted by $\mathcal{P}$. A poset $\mathcal{L} = \langle L; \leq \rangle$ is a *lattice* if an infimum and a supremum exist for any pair of elements of L. Formally:

Definition 17. A poset $\mathcal{L} = \langle L; \leq \rangle$ is a lattice if $\inf H$ and $\sup H$ exist for any finite non-empty subset $H \subseteq L$.

The lattice Duality Principle states that if $\mathcal{L} = \langle L; \leq \rangle$ is a lattice the dual $\mathcal{L}^{\delta} = \langle L; \geq \rangle$ is also a lattice. A lattice can also be described in an algebraic form by setting $a \wedge b = \inf \{a, b\}$ and $a \vee b = \sup \{a, b\}$. Thus, the algebra $\mathcal{L} = \langle L, \wedge, \vee \rangle$ is a lattice. An important type of lattices is that of *complete lattices.*

Definition 18. A lattice $\mathcal{L} = \langle L; \leq \rangle$ is complete iff $\inf H$ and $\sup H$ exist for any subset $H \subseteq L$.

A complete lattice has both a smallest element called *bottom*, denoted as $\bot$, and a greatest element called top, denoted as $\top$. All finite lattices are complete lattices. Morphological operations can be described as mappings between complete lattices. From now on, we denote complete lattices by the symbols $\mathbb{L}$ and $\mathbb{M}$.

Definition 19. For every subset $Y \subseteq \mathbb{L}$ an *erosion* is a mapping $\varepsilon : \mathbb{L} \to \mathbb{M}$ that commutes with the infimum operation, $\varepsilon \left(\bigwedge Y \right) = \bigwedge_{y \in Y} \varepsilon \left(y \right)$. Similarly, a *dilation* is a mapping $\delta : \mathbb{L} \to \mathbb{M}$ that commutes with the supremum operation, $\delta \left(\bigvee Y \right) = \bigvee_{y \in Y} \delta \left(y \right)$.

Definition 20. *Anti-erosion* operator, $\bar{\varepsilon} : \mathbb{L} \to \mathbb{M}$, and *anti-dilation* operator, $\bar{\delta} : \mathbb{L} \to \mathbb{M}$, are defined as mappings holding $\bar{\varepsilon} \left(\bigwedge Y \right) = \bigvee_{y \in Y} \bar{\varepsilon} \left(y \right)$ and $\bar{\delta} \left(\bigvee Y \right) = \bigwedge_{y \in Y} \bar{\delta} \left(y \right)$ respectively.

Any mapping Ψ between complete lattices $\mathbb{L}$ and $\mathbb{M}$ can be expressed in terms of supremums and infimums of these four morphological operators [10].

9.1.1 Multivariant Mathematical Morphology

The extension of Mathematical Morphology to multivariate images is not straightforward since pixels are (high dimensional) vectors without a defined natural

total order. One way to define a morphology on such kind of sets is by constructing a surjective mapping into a lattice $h : \mathbb{R}^p \to \mathcal{L}$ so that we can assume an ordering induced by this mapping. The h-ordering $\leq_h$ is defined as:

$$r \leq_h r' \Leftrightarrow h(r) \leq h(r'); \forall r, r' \in \mathbb{R}^p$$

Some authors [5, 187] define h-ordering on the basis of a supervised classifier trained with some pixel values. Discriminant function values and class a posteriori probabilities provide the surjective mapping h.

Definition 21. A h-supervised ordering for a non-empty set X [5, 187] is a h-ordering that satisfies the conditions $h(b) = \bot$, $\forall b \in B$, and $h(f) = \top$, $\forall f \in F$, where $B, F \subset X$ are subsets of X such that $B \cap F = \emptyset$.

In a recent work [183] we proposed a supervised ordering based on Lattice Auto-Associative Memories (LAAMs) [151, 149], hereafter named as *LAAM-supervised ordering* keeping multivariate morphology under the general setting of Lattice algebra. All the required calculations are defined in base to Lattice algebra operators ($\vee$, $\wedge$ and $+$) and therefore, LAAM-supervised ordering is faster and with less computational burden.

However, the LAAM-supervised ordering do not ensure $h(b) = \bot$, $\forall b \in B$, and $h(f) = \top$, $\forall f \in F$. Here we introduce an *absolute* LAAM-supervised ordering which does comply with the supervised ordering definition. Furthermore, foreground and background training sets need to be defined to run a supervised ordering method. In [5] and [183] authors set B and F manually by selecting data points from the ground-truth. However, a-priori information of remote sensing data is scarce, expensive and unreliable [11]. Furthermore, when large images are processed automatic methodologies are required. In this chapter, we propose an automatic methodology for the definition of background and foreground sets based on the use of endmember induction algorithms (EIAs). The proposed methodology does not require any a-priori information over the data, such as a ground-truth labeling.

9.2 LAAM-Supervised Ordering

9.2.1 LAAM h-mapping

This section introduces a definition of a mapping h based on LAAMs. Following [177] we propose the Chebyshev distance to the LAAM recall as the mapping h inducing a supervised ordering in multivariate data sets. Given a multivariate data vector $\mathbf{c} \in \mathbb{R}^n$ and a non-empty training set $X = \{\mathbf{x}_i\}_{i=1}^K$, $\mathbf{x}_i \in \mathbb{R}^n$ for all $i = 1, \ldots, K$, the LAAM h-mapping is:

$$h_X(\mathbf{c}) = \zeta\left(\mathbf{x}^{\#}, \mathbf{c}\right), \tag{9.1}$$

where $\mathbf{x}^{\#} \in \mathbb{R}^n$ is the recall of vector $\mathbf{c}$ from either the erosive memory M_{XX} or the dilative memory W_{XX},

$$\mathbf{x}^{\#} = M_{xx} \boxslash \mathbf{c} = W_{xx} \boxbslash \mathbf{c}.$$

Function $\zeta(\mathbf{a}, \mathbf{b})$ denotes the Chebyshev distance between two vectors, given by the greatest absolute difference between the vectors' components:

$$\zeta(\mathbf{a}, \mathbf{b}) = \bigvee_{i=1}^{n} |a_i - b_i| .$$

This distance can be computed in the lattice algebra framework as follows:

$$\zeta(\mathbf{a}, \mathbf{b}) = (\mathbf{a}^* \boxtimes \mathbf{b}) \vee (\mathbf{b}^* \boxtimes \mathbf{a}) , \tag{9.2}$$

where $\mathbf{a}^*$ and $\mathbf{b}^*$ denotes the conjugate of $\mathbf{a}$ and $\mathbf{b}$ respectively.

9.2.2 One-side LAAM-supervised ordering

Using the LAAM-based h-mapping of equation (9.1), the one-side LAAM-based supervised h-ordering, denoted $\leq_X$ is defined as follows:

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \ \mathbf{x} \leq_X \mathbf{y} \Longleftrightarrow h_X(\mathbf{x}) \leq h_X(\mathbf{y}) . \tag{9.3}$$

The one-side LAAM-supervised ordering generates a complete lattice, denoted $\mathbb{L}_X$, whose bottom element is the set of fixed points of M_{XX} and W_{XX}, $\perp_X = \mathcal{F}(X)$, and the the top element is $\top_X = +\infty$, allowing to define erosion operators that enlarge the points close to the set of fixed points of X in a Chebyshev sense, and dilation operators that shrink those points in favor of points that lie far from $\mathcal{F}(X)$.

9.2.3 Background/Foreground LAAM-supervised orderings

It is possible to define a supervised ordering using two sets of data points, B and F, such that $B \cap F = \emptyset$, [187] defining 'background' and 'foreground' training sets, respectively. The intuition is that, erosion operators will favor points close to the background, and dilation operators will favor points close to the foreground. In order to do that we independently calculate the LAAM-based h-mapping of equation (9.1) using B and F, respectively denoted as h_B and h_F. Given a vector $\mathbf{c} \in \mathbb{R}^n$, and the training sets B and F, we define *relative* and *absolute LAAM-supervised orderings* using $h_B(\mathbf{c})$ and $h_F(\mathbf{c})$:

- In the relative LAAM-supervised ordering, denoted as $\leq_r$, we combine both h_B and h_F in a single h-function, denoted h_r, by $h_r = h_F - h_B$. The ordering is then given by:

  $$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \ \mathbf{x} \leq_r \mathbf{y} \Longleftrightarrow h_r(\mathbf{x}) \leq h_r(\mathbf{y}) \tag{9.4}$$

  The relative LAAM-supervised ordering generates a complete lattice, $\mathbb{L}_r$, with bottom and top elements defined as $\perp = -\infty$ and $\top = +\infty$ respectively. Those vectors that lie equidistant in a Chebyshev sense from both the foreground and the background fixed points conforms an ordering boundary denoted C_r. Formally, C_r denotes a set of points such that for all $\mathbf{c} \in C_r$ holds that $h_r(\mathbf{c}) = 0$.

- In the absolute LAAM-supervised ordering, denoted as $\leq_a$, we treat separately the data vectors belonging to the background from those belonging to the foreground. We say that the data vector $\mathbf{c}$ belongs to the background if it is closer in a Chebyshev sense to the set of fixed points of B than to the set of fixed points of F, that is, $\mathbf{c} \in B$, if $h_B(\mathbf{c}) \leq h_F(\mathbf{c})$. Otherwise, we say that $\mathbf{c}$ belongs to the foreground, $\mathbf{c} \in F$. Now we can define the absolute ordering based on h_B and h_F as:

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n,\ \mathbf{x} \leq_a \mathbf{y} \Longleftrightarrow \begin{cases} h_B(\mathbf{x}) \leq h_B(\mathbf{y}) & \text{if } x, y \in B \\ x \in B \text{ and } y \in F & \\ h_F(\mathbf{y}) \leq h_F(\mathbf{x}) & \text{if } x, y \in F \end{cases} \tag{9.5}$$

  The absolute LAAM-supervised ordering generates a complete lattice, $\mathbb{L}_a$, where the bottom element is the set of fixed points of the background set, $\perp_a = \mathcal{F}(B)$, and the top element is the set of fixed points of the foreground, $\top_a = \mathcal{F}(F)$.

9.2.4 Disambiguation

Supervised orderings can produce ambiguities during morphological operations since h-functions are not necessarily injective and thus the induced h-ordering might be not a total order. Thus, when we need to differentiate among the members of the equivalence classes $\mathcal{L}[z] = \{\mathbf{x} \in X | h(\mathbf{x}) = z\}$, some additional criterion have to be considered to impose an ordering on them. In the literature, this disambiguation criterion is usually a lexicographical order.

9.2.5 Beucher morphological gradient

The morphological gradient was defined by Serge Beucher in his Thesis [17] as the arithmetic difference between the dilated and eroded images with the elementary structuring element S:

$$g(I) = \delta(I) - \varepsilon(I). \tag{9.6}$$

The Beucher gradient, $g(I)$, represents the maximum variation of the image f intensities within the neighborhood defined by the structuring element S. Beucher gradient is well defined in binary or gray-level images. For multivariate data it may be defined component-wise as:

$$g(I) = \sum_{i=1}^{n} (\delta(I_i) - \varepsilon(I_i)) \tag{9.7}$$

where I_i denotes the i-th band of the image and n is the total number of bands of I. It can also be defined on the h-supervised ordering induced morphological operators:

$$g(I) = h\left(\delta_{\{B,F\}}(I)\right) - h\left(\varepsilon_{\{B,F\}}(I)\right), \tag{9.8}$$

where $\delta_{\{B,F\}}(I)$ and $\varepsilon_{\{B,F\}}(I)$ are the dilation and erosion operator induced by the h-supervised ordering. The supervised Beucher gradient resulting of the use of the proposed LAAM_X and LAAM_r supervised orderings can be defined directly using the respective h_X and h_r functions:

$$g_X(I) = h_X(\delta(I)) - h_X(\varepsilon(I)), \tag{9.9}$$

$$g_r(I) = h_r(\delta(I)) - h_r(\varepsilon(I)). \tag{9.10}$$

The absolute LAAM-supervised ordering, $g_a(I)$, is more complicated as it is defined depending on that the pixels belong to the 'background' or the 'foreground' classes:

$$\forall \mathbf{x} \in I, g_a(\mathbf{x}) = \begin{cases} h_B(\delta(\mathbf{x})) - h_B(\varepsilon(\mathbf{x})) & \text{if } \delta(\mathbf{x}), \varepsilon(\mathbf{x}) \in B \\ h_F(\varepsilon(\mathbf{x})) - h_F(\delta(\mathbf{x})) & \text{if } \delta(\mathbf{x}), \varepsilon(\mathbf{x}) \in F \\ h_F(\delta(\mathbf{x})) - h_B(\varepsilon(\mathbf{x})) + h_B(\delta(\mathbf{x})) - h_F(\varepsilon(\mathbf{x})) & \text{otherwise} \end{cases} \tag{9.11}$$

9.2.6 RGB example

We exemplify on a synthetic RGB image the use of the proposed LAAM-supervised orderings. Figure 9.1 shows the original synthetic RGB image with a spatial dimensionality of 375×365 pixels. Three points, $\mathbf{x}_r$, $\mathbf{x}_g$ and $\mathbf{x}_b$, were selected corresponding to red, green and blue pixels respectively. We compare the use of the proposed LAAM-supervised orderings to the component-wise ordering and a lexicographic ordering. For the one-side LAAM-supervised ordering we set the training set to the blue pixel, $X = \{\mathbf{x}_b\}$. For the absolute and relative background/foreground LAAM-supervised orderings we set the background and foreground training sets to the green and blue pixels respectively, $B = \{\mathbf{x}_g\}$ and $F = \{\mathbf{x}_b\}$. For the lexicographic ordering we prioritized the red band over the green band, and the green band over the blue band, that is $\text{R} \vdash \text{G} \vdash \text{B}$.

We performed an erosion ($\varepsilon(G)$) and a dilation ($\delta(G)$) morphological operations of the original image using the different orderings and then, we calculated gradients obtained by the absolute difference between the dilated and eroded images, $\Delta(G) = |\delta(G) - \varepsilon(G)|$. Results are shown in figure 9.2. The apparition of new colors in the eroded and dilated images, not existing in the original image, can be appreciated for the component-wise ordering. This undesirable effect does not appear in the proposed LAAM-supervised orderings. Also, the proposed LAAM-supervised ordering obtains a better defined gradient than the lexicographical ordering due to the use of all three spectral bands. Summarizing, the proposed LAAM-supervised orderings perform similarly to the component-wise ordering and better than the lexicographic ordering, since they use all the spectral bands; and, in addition, solve the false color problem typical of the component-wise approach.

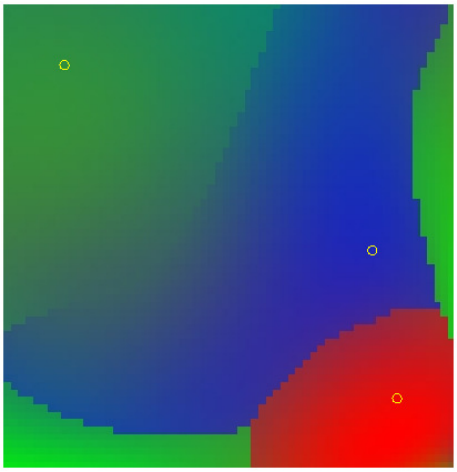


Figure 9.1: Synthetic RGB image. The yellow circles mark the red, green and blue sample points.

9.3 Unsupervised selection of training sets

Any supervised ordering formulation, such as the LAAM-supervised orderings presented above, requires the user to provide training sets. Sometimes this is not possible, maybe due to a lack of ground-truth or because huge amounts of data are going to be processed and an automatic and unsupervised process is required. Here, we propose an automatic and unsupervised methodology to select the training sets, either X for the basic LAAM-supervised ordering or F and B for the absolute and relative LAAM-supervised orderings. In both cases we make use of an EIA to induce a set of endmembers $E = \{\mathbf{e}_i\}_{i=1}^{p}$ from H. Then, we calculate the distances, d_{ij}, for any pair $\mathbf{e}_i, \mathbf{e}_j \in E$, $i, j = 1, \ldots, p$, $i \neq j$. Any distance function can be used as the Chebyshev distance, the spectral angular map (SAM) distance or the Euclidean distance. We give two criteria in base to the pair distances d_{ij} to select the spectra forming the training sets X and F, B respectively:

- To set the training set X we select the endmember $\mathbf{e}_k^o \in E$ that minimizes the average distance to the rest of endmembers:

$$\mathbf{e}_k^o = \arg\min_k \left(\frac{1}{p-1} \sum_{i \neq k} d_{ik} \right), \ i = 1, \ldots, p \tag{9.12}$$

- To set the training sets F and B we select the two endmembers $\mathbf{e}_i^o, \mathbf{e}_j^o \in E$ with maximum SAM distance, that is:

$$\mathbf{e}_i^o, \mathbf{e}_j^o = \arg\max_{i,j} \left(d_{ij} \right) \tag{9.13}$$

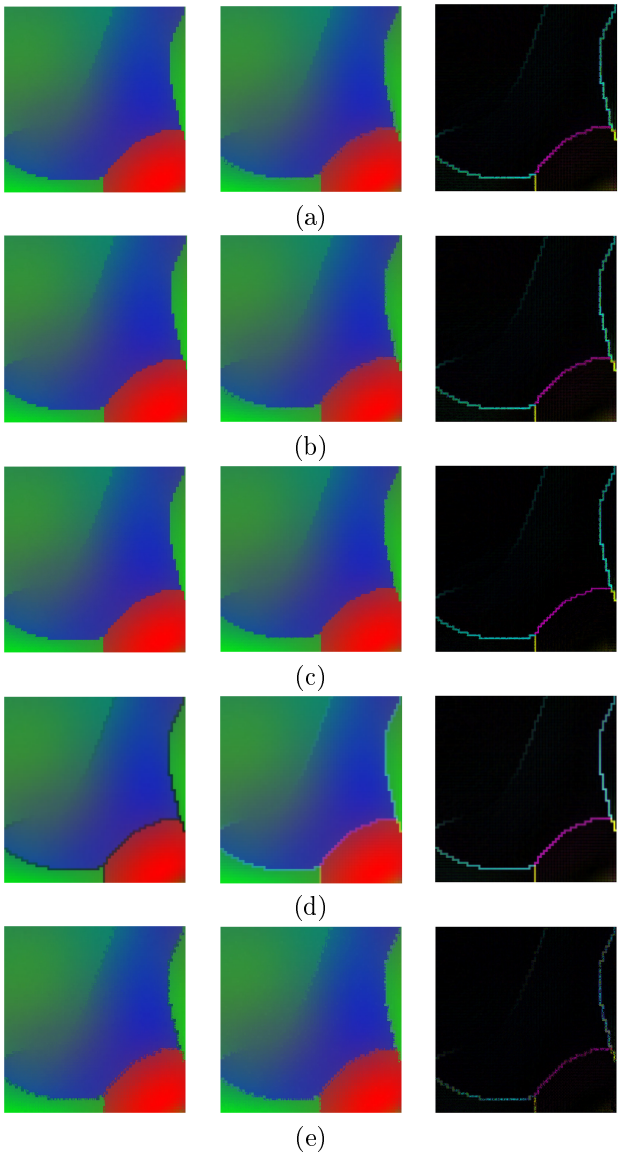


Figure 9.2: Results of applying morphological operations over the synthetic RGB image, G. Columns show: (left) erosion, $\varepsilon(G)$, (middle) dilation, $\delta(G)$, (right) gradients. Rows indicate the ordering used: (a) basic LAAM-supervised ordering setting with $X = \{\mathbf{x}_b\}$, (b) absolute LAAM-supervised ordering with $B = \{\mathbf{x}_g\}$ and $F = \{\mathbf{x}_b\}$, (c) relative LAAM-supervised ordering with $B = \{\mathbf{x}_g\}$ and $F = \{\mathbf{x}_b\}$ (d) component-wise ordering, (e) lexicographical ordering (R $\vdash$ G $\vdash$ B). In all cases a flat 5×5 square structural element was used.

We can set each $\mathbf{e}_i^o, \mathbf{e}_j^o$ randomly to F and B given that LAAM-supervised orderings are dual.

9.3.1 RGB example

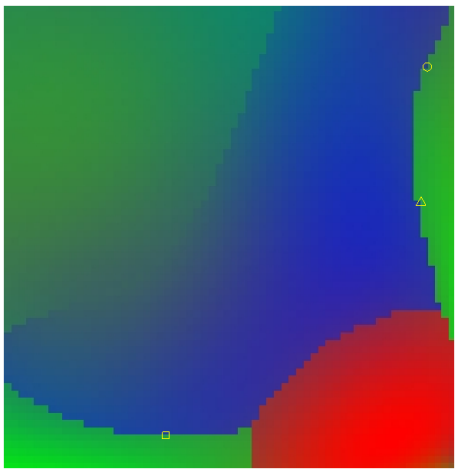
As before we made use of the synthetic RGB image 9.1 to exemplify the use of the proposed automatic and unsupervised methodology to select the background and foreground training sets to perform LAAM-supervised orderings. We induced the endmembers from the original image by applying the Incremental Lattice Strong Independence Algorithm (ILSIA) described in Appendix B, and we calculate the between endmembers distances using the SAM distance. Figure 9.3 shows the endmember induced by the ILSIA algorithm. Endmember 3 (pixel marked as a yellow square) was the one minimizing the average SAM distance to the rest of endmembers (9.12) and so was assigned to the basic training set. Endmembers 1 and 2 (yellow circle and triangle respectively) were the ones with maximum SAM distance and were assigned to foreground and background sets. Figure 9.4 show the results of applying morphological erosion and dilation operators, and the resulting Beucher gradient, using the proposed LAAM-supervised orderings by the automatically selected training sets. We observe visually that the results are similar to the ones obtained by manually selecting the training sets.

9.4 Experiments with hyperspectral images

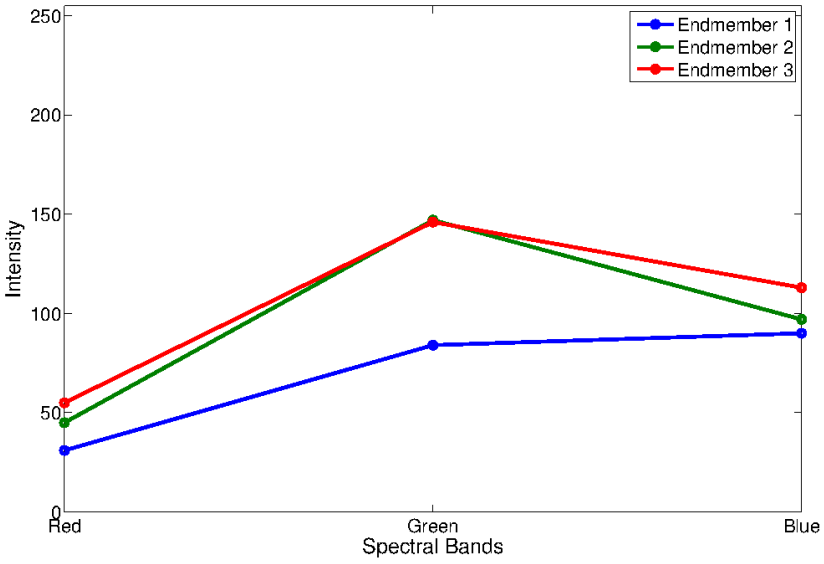
9.4.1 Methodology

We have shown with a synthetic RGB image that the proposed LAAM-supervised orderings are capable of using all the spectral information without incurring in the false color problem. Here we give experimental evidence that the proposed automatic and unsupervised methodology to build up the training sets together to the proposed LAAM-supervised ordering algorithms are suitable for their use in morphological filters oriented to machine learning techniques over multivariate data such as the classification of hyperspectral images. We applied the contextual classification approach for hyperspectral images presented in [179] over two well-known hyperspectral scenes, comparing the performing of a pixel-wise SVM to the result of the contextual SVM using watershed segmentation.

Figure 9.5 shows a diagram of the experimental methodology. First, a pixel-wise SVM is performed over the hyperspectral training set obtaining a SVM classification map. Then, the SVM classification map is combined with a watershed segmentation to improve the pixel-wise results. The three proposed LAAM-supervised orderings and a component-wise ordering are independently applied to get a Beucher gradient that will be used to obtain the watershed segmentations. Each watershed segmentation is combined with the pixel-wise SVM classification map using both the proposed methodologies by authors in [179], named WHEDS and NWHEDS. Both methodologies assign to each region in



(a)



(b)

Figure 9.3: Endmembers induced by the ILSIA algorithm from the synthetic RGB image: (a) coordinates (in yellow) of the endmembers in the RGB image, (b) spectral signatures of the endmembers. The triangle point corresponds to the endmember 1, the circle to the endmember 2, and the square to endmember 3.

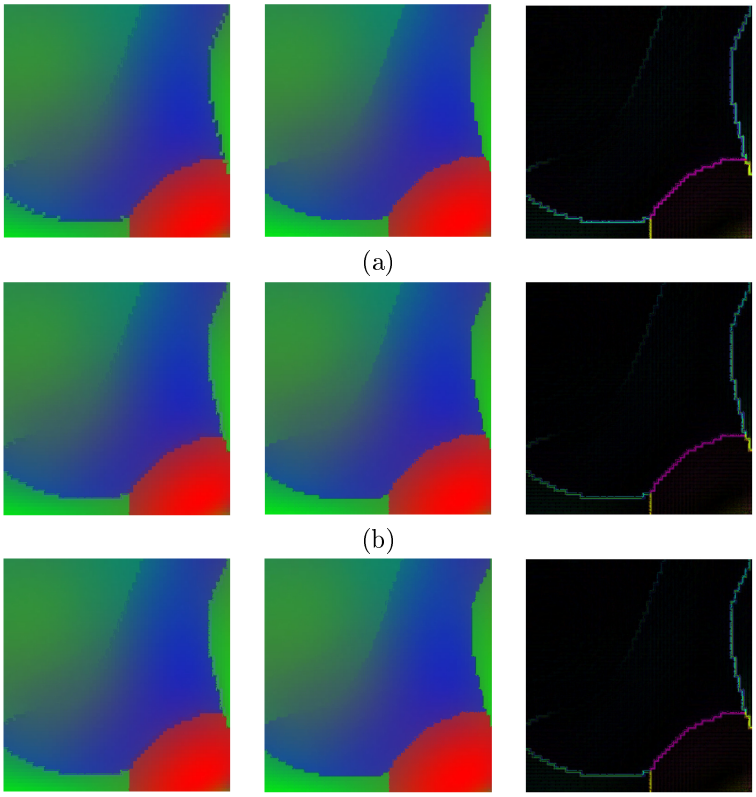


Figure 9.4: Results of applying morphological operations over the synthetic RGB image, G, using the proposed unsupervised methodology to automatically set the training sets, X and F, B. Columns show: (left) erosion, $\varepsilon(G)$, (middle) dilation, $\delta(G)$, (right) gradients, $\Delta(G) = |\delta(G) - \varepsilon(G)|$. Rows indicate the ordering used: (a) basic LAAM-supervised ordering, (b) absolute LAAM-supervised ordering, (c) relative LAAM-supervised ordering. In all cases a flat 5×5 square structural element was used.

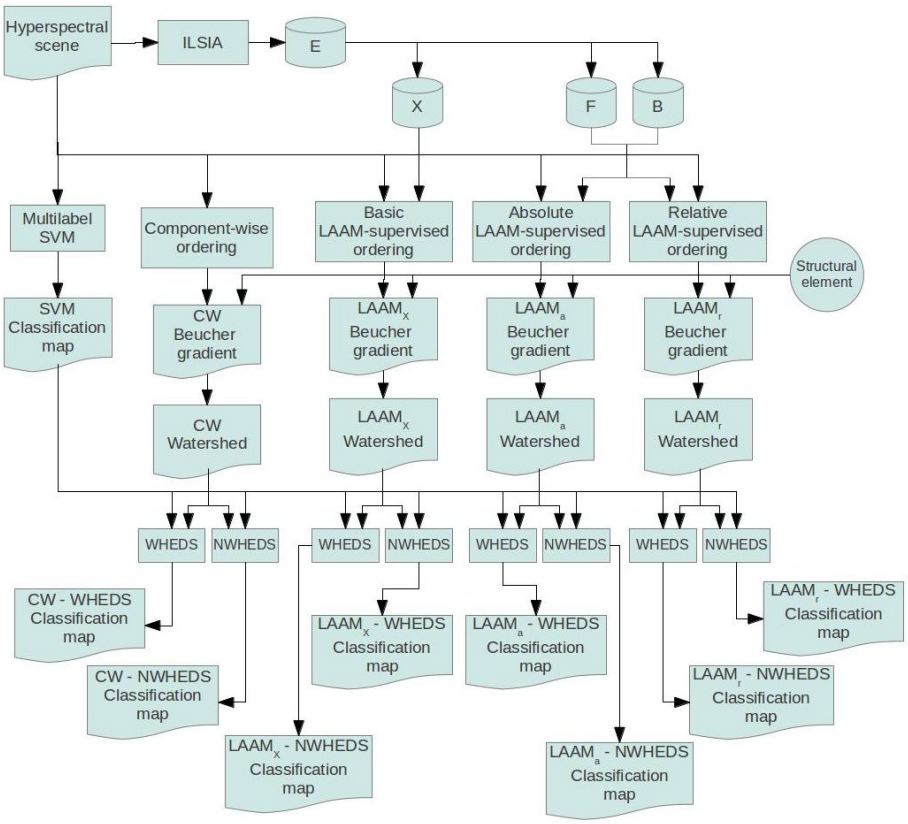


Figure 9.5: Experimental methodology diagram.

the watershed segmentation the label with majority presence in the corresponding pixel-wise SVM classification map region. WHEDS and NWHEDS differ in the use of the watersheds pixels (the pixels dividing the watershed regions). In WHEDS, each watershed pixel is assigned to the watershed region whose median vector is closer to the watershed pixel vector before calculating the majority label on the region. In NWHEDS however, the watershed pixels keep the class label assigned by the pixel-wise classification.

For all the following experiments the LIBSVM library[1] has been used to run the multi-class (one versus all) pixel-wise SVM with a Radial Basis Function (RBF) kernel. A 5-fold cross validation has been used to select the model parameters C and γ. The ILSIA algorithm has been used to induce the endmembers for the selection of the training sets used in the LAAM-supervised orderings. The MATLAB standard implementation of the watershed algorithm [125] was used in all the experiments.

[1] http://www.csie.ntu.edu.tw/~cjlin/libsvm/

The goodness of the different classification maps was assessed through their respective confusion matrices. The global overall accuracy (OA), average accuracy (AA) and Kappa coefficient, as well as the class-specific sensitivity and specificity were calculated. The overall accuracy is the percentage of correctly classified pixels, and the average accuracy is the mean of class-specific accuracies:

$$\mathrm{OA} = \frac{\sum_{i=1}^{C} n_i}{N} \tag{9.14}$$

$$\mathrm{AA} = \sum_{i=1}^{C} \frac{n_i}{N_i} \tag{9.15}$$

where n_i is the number of actual class i samples that have been correctly classified, N_i is the total number of actual class i samples, C is the total number of classes, and N is the total number of samples, so $\sum_i^C N_i = N$.

The kappa coefficient [148] is the percentage of corrected classified samples corrected by the number of samples correctly classified by chance:

$$\kappa = \frac{N \sum_k x_{kk} - \sum_k x_{k+} x_{+k}}{N^2 - \sum_k x_{k+} x_{+k}} \tag{9.16}$$

where x_{ij} denotes the element of the confusion matrix corresponding to the ith row and jth column, N is the total number of samples, x_{i+} denotes the sum over all columns for row i, and x_{+j} the sum over all rows for column j: $x_{i+} = \sum_j x_{ij}$ and $x_{+j} = \sum_i x_{ij}$.

9.4.2 Pavia University

9.4.2.1 Watershed segmentation

In order to apply the proposed LAAM-supervised orderings to the Pavia University scene we first induce the endmembers from the hyperspectral image using the ILSIA endmember induction algorithm. Figure 9.6 shows the induced endmembers. Following the proposed automatic methodology we defined the training sets used to obtain the LAAM-supervised orderings. For the basic LAAM-supervised ordering, endmember 4 was selected, $X = \{\mathbf{e}_4\}$. For the absolute and relative LAAM-supervised orderings endmember 1 and 2 were selected, $F = \{\mathbf{e}_1\}$ and $B = \{\mathbf{e}_2\}$. Figure 9.7 shows the h-function mappings of equation (9.1) obtained from Pavia University scene using each of the trained sets X, F and B. Pixels in blue are closer to the set of fixed points of the respective sets, $\mathcal{F}(X)$, $\mathcal{F}(F)$ and $\mathcal{F}(B)$; and, pixels in red far from them and thus being more lattice independent on the training sets.

Using the obtained LAAM h-mappings we performed erosions, $\varepsilon(G)$, and dilations, $\delta(G)$, with a disk structural element of increasing radius ($r = 2, 3, 4$) over the Pavia University image using the proposed LAAM-supervised orderings. We also eroded and dilated the scene using a component-wise ordering.

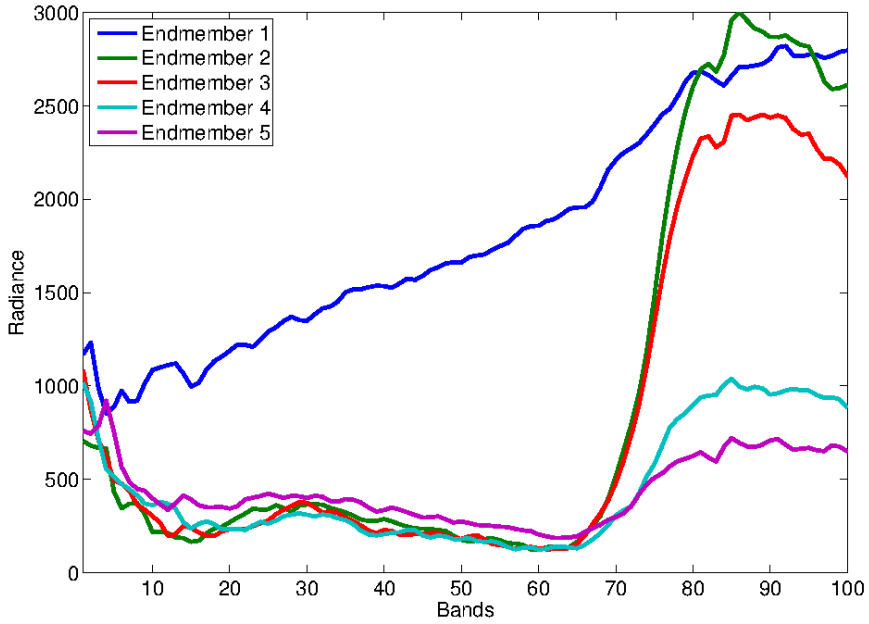
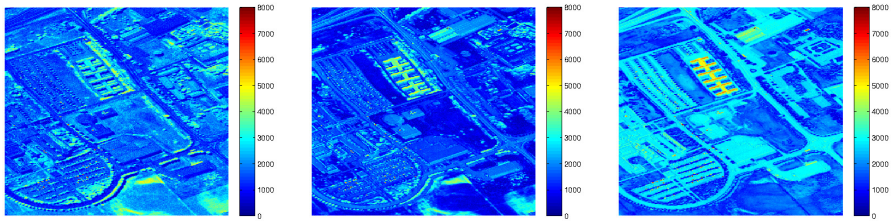


Figure 9.6: Endmembers induced by ILSIA algorithm from the Pavia University scene.



(a) (b) (c)

Figure 9.7: h-function mappings estimated by the training sets: (a) $X = \{\mathbf{e}_4\}$, (b) $F = \{\mathbf{e}_1\}$, (c) $B = \{\mathbf{e}_2\}$.

The Beucher gradients, $\triangle(G) = |\varepsilon(G) - \delta(G)|$, were calculated for the four orderings and three structural elements as shown in Figure 9.8.

Using the Beucher gradients we calculated their respective watershed segmentations. Figure 9.9 shows the watershed segmentations as black lines dividing the Pavia University scene in regions.

9.4.2.2 Experimental results

We trained a multi-class one-vs-all SVM classifier to perform pixel-wise classification of the Pavia University scene. We selected 200 instances per class by randomly sampling without replacement, so that the training data is well balanced.

Figure 9.10 shows the classification map obtained by the pixel-wise SVM. The overall accuracy is 88.97% and the average accuracy is 91.60% can be improved since there are isolated misclassified pixels all along the image. The NWHEDS and the WHEDS contextual methodology was applied to combine the pixel-wise SVM classification map with the watershed segmentations obtained by the different orderings. Figures 9.11 and 9.12 show the classification maps obtained by SVM-NWHEDS and SVM-WHEDS respectively. There is a performance increase from the pixel-wise SVM to the SVM-NWHEDS approach, and again a new boosting from SVM-NWHEDS to SVM-WHEDS as it can be deduced from the overall and average accuracies shown in tables 9.1-9.3 independently of the ordering or structural element used. The performance of the SVM-NWHEDS and SVM-WHEDS classification is similar for the component-wise ordering and the proposed LAAM-supervised orderings using the automatic and unsupervised training sets selection methodology, while the proposed methodologies does not provoke the apparition of spectral signatures in the eroded and dilated images not present in the original image. Figures 9.13 and 9.14 show respectively the class-specific sensitivity and specificity of the Pavia University classification for the pixel-wise SVM, the SVM-NWHEDS and the SVM-WHEDS confusion matrices, using the four orderings and a disk structural element with radius 3. The performance boosting appears in almost all the classes, except maybe those whose sensitivity or specificity values are close to 97-99%.

9.4.3 Indian Pines

9.4.3.1 Watershed segmentation

As with Pavia University scene we first induce the endmembers from the hyperspectral image using the ILSIA endmember induction algorithm and then, we defined the training sets used to obtain the LAAM-supervised orderings by means of the proposed automatic methodology. Figure 9.15 shows the induced endmembers. The training sets were defined as $X = \{\mathbf{e}_1\}$ and $F = \{\mathbf{e}_2\}$, $B = \{\mathbf{e}_3\}$. Figure 9.16 shows the h-function mappings of equation (9.1) obtained from Indian Pines scene using each of the trained sets X, F and B.

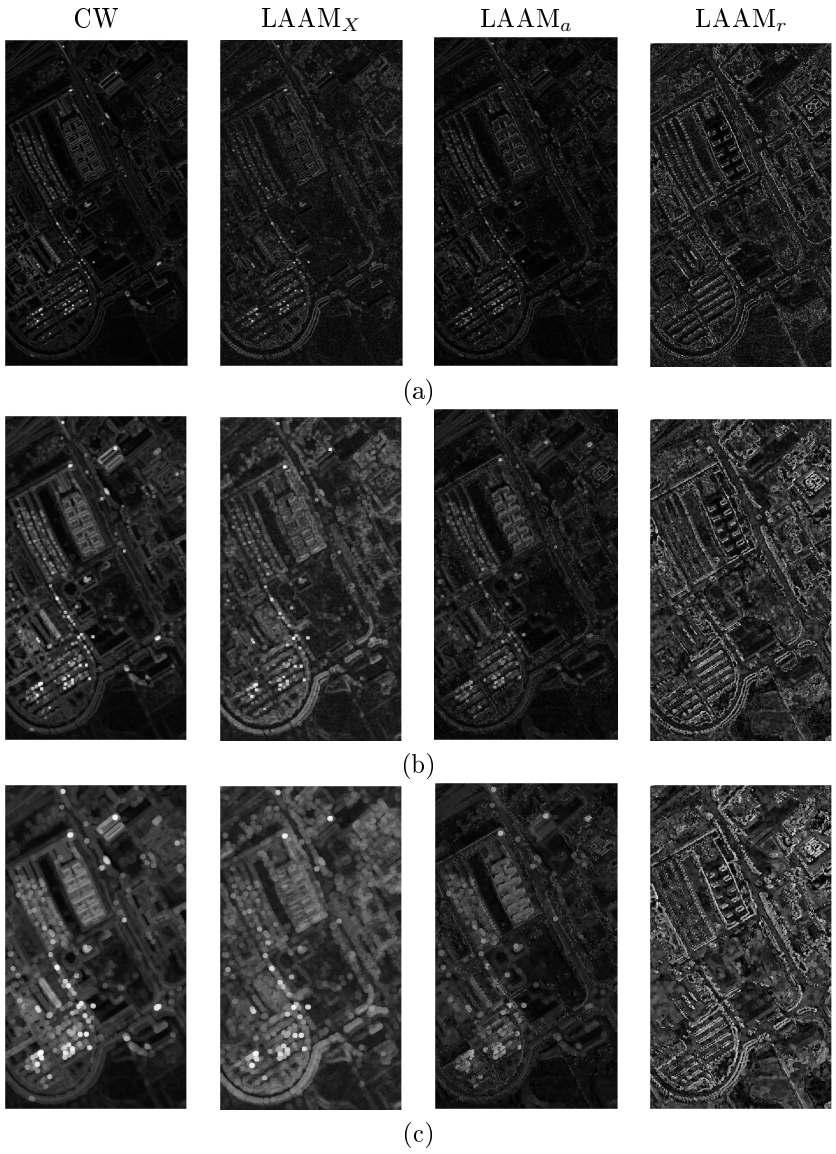


Figure 9.8: Beucher morphological gradients obtained from the Pavia University scene using the four different orderings and a disk structural element with increasing radius: (a) $r = 1$, (b) $r = 3$, (c) $r = 5$.

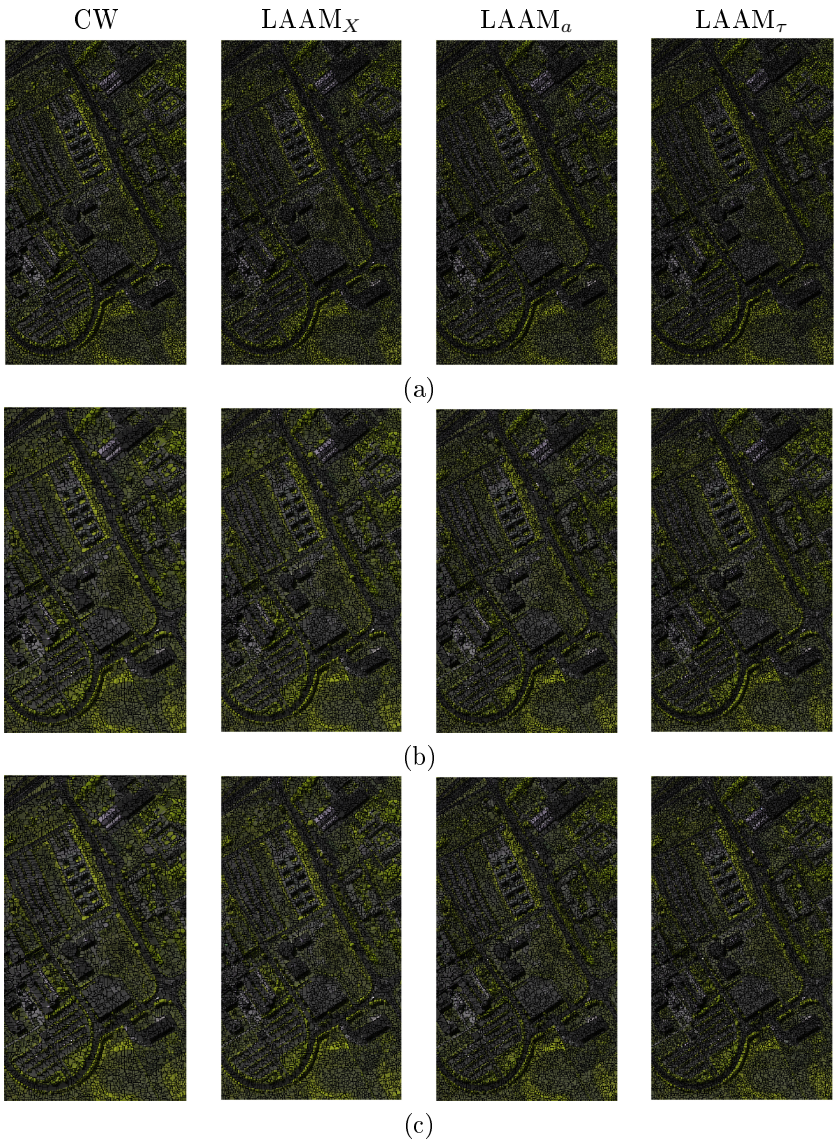


Figure 9.9: Watersheds obtained from the Pavia University scene using the Beucher morphological gradients obtained by the four different orderings and a disk structural element with increasing radius: (a) $r = 2$, (b) $r = 3$, (c) $r = 4$.

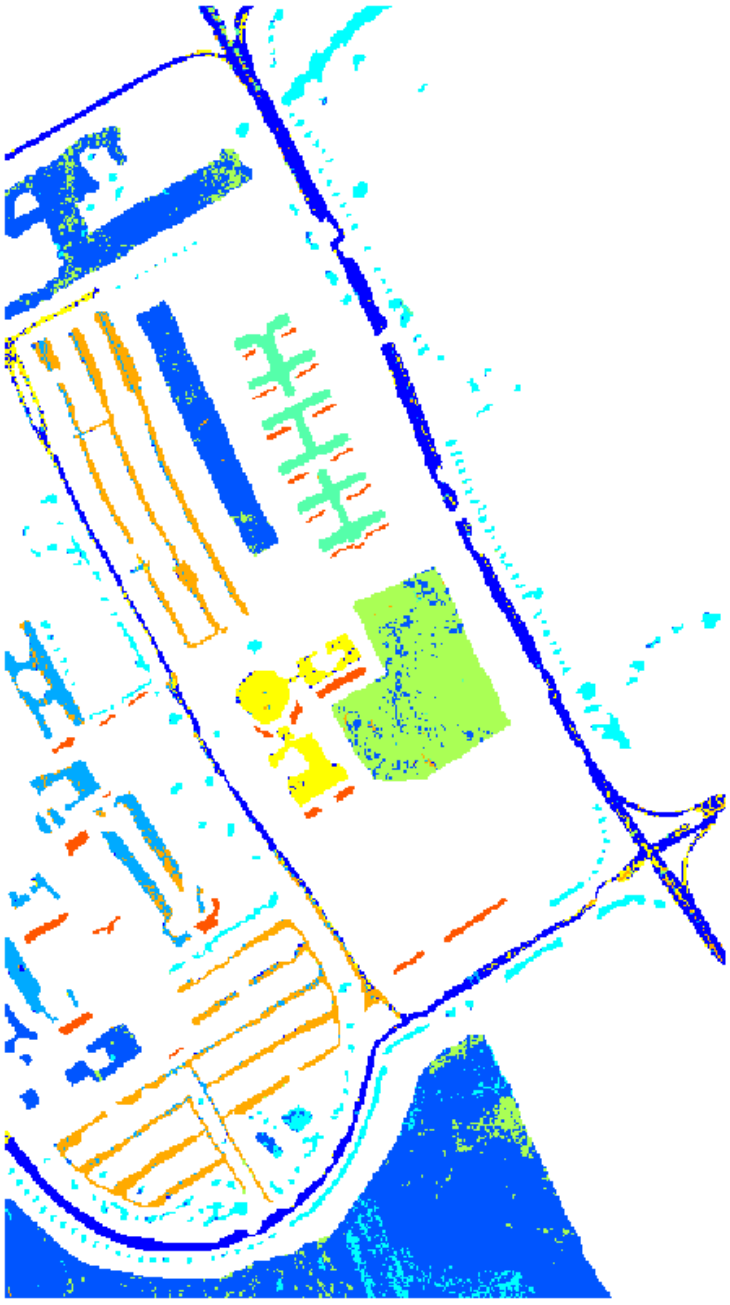


Figure 9.10: Classification map from the Pavia University scene obtained by the pixel-wise SVM.

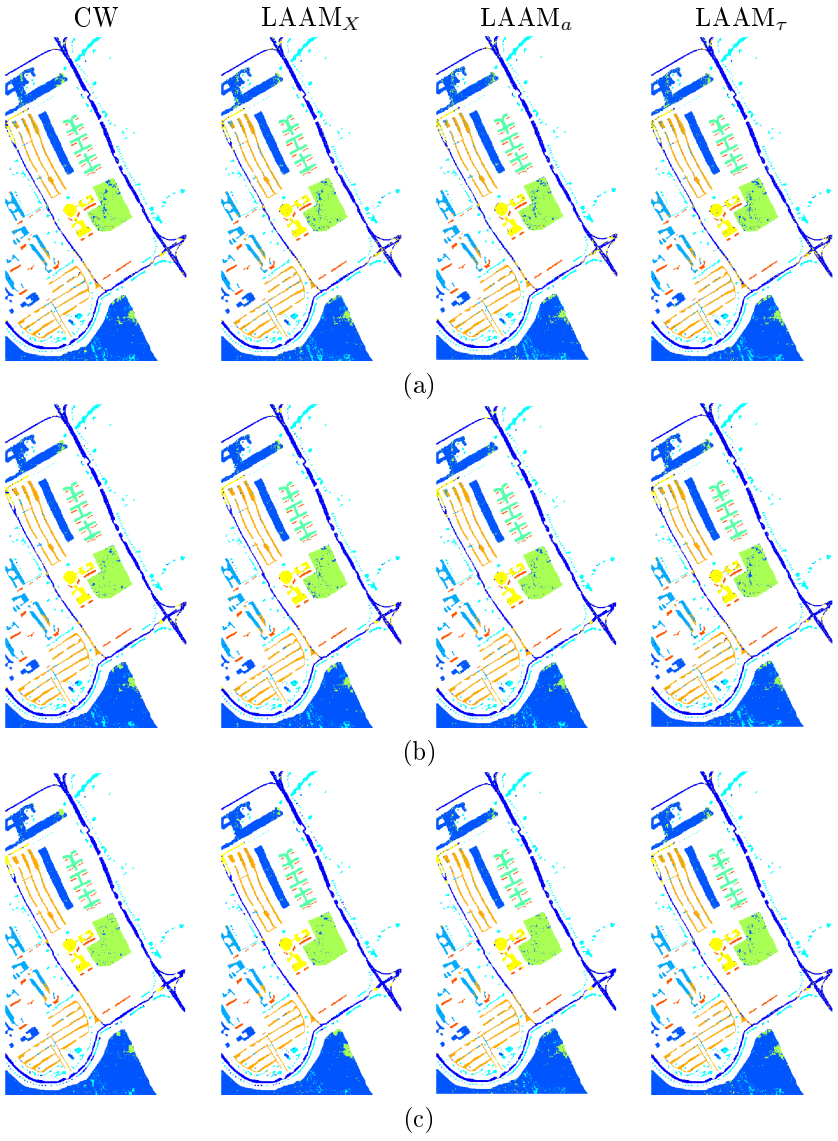


Figure 9.11: Classification maps from the Pavia University scene obtained by the SVM-NWHEDS using the watershed segmentations obtained by the four different orderings and a disk structural element with increasing radius: (a) $r = 2$, (b) $r = 3$, (c) $r = 4$.

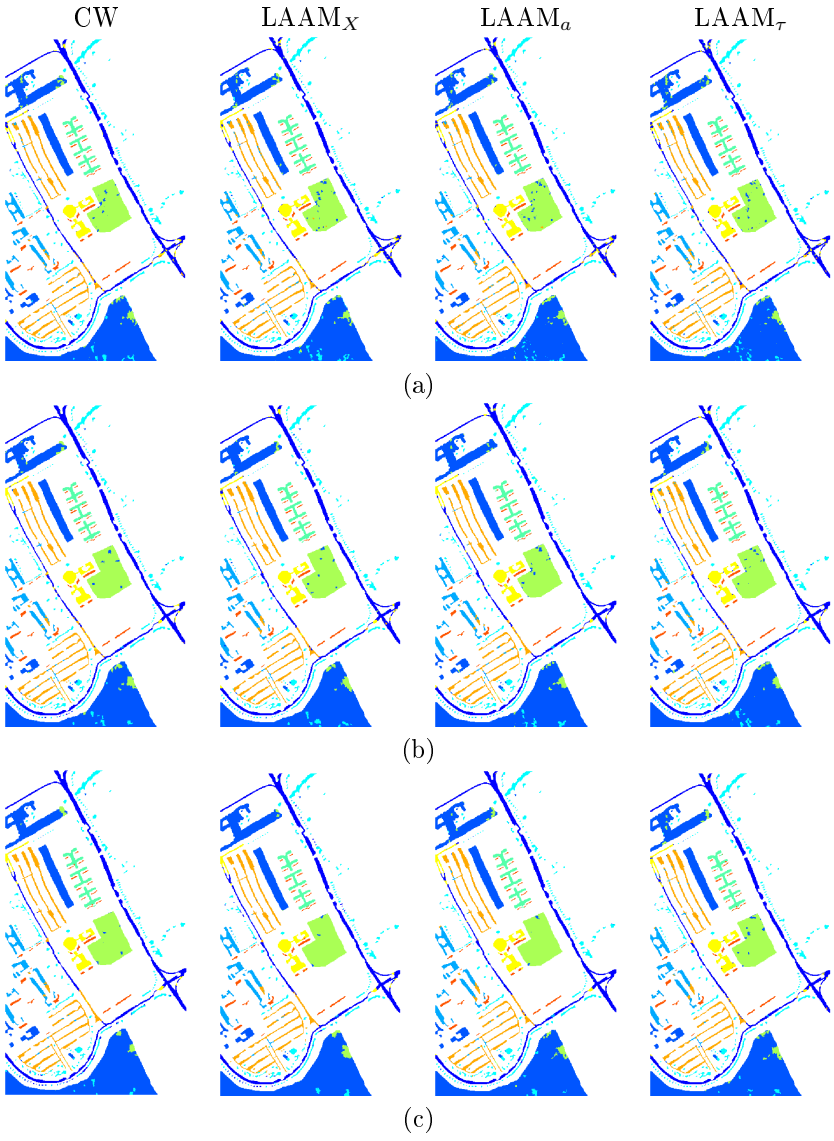


Figure 9.12: Classification maps from the Pavia University scene obtained by the SVM-WHEDS using the watershed segmentations obtained by the four different orderings and a disk structural element with increasing radius: (a) $r = 2$, (b) $r = 3$, (c) $r = 4$.

Method		OA	AA	κ
Pixel-wise SVM		88.97	91.60	0.8565
SVM + NWHED	CW	91.42	93.73	0.8880
	LAAM_X	90.91	93.16	0.8815
	LAAM_a	91.09	93.32	0.8838
	LAAM_r	90.81	92.90	0.8801
SVM+WHED	CW	94.46	96.33	0.9274
	LAAM_X	93.40	95.27	0.9136
	LAAM_a	93.99	95.78	0.9213
	LAAM_r	93.77	95.46	0.9184

Table 9.1: Global classification results of the Pavia University hyperspectral scene: overall accuracy (OA), average accuracy (AA) and Kappa (κ) values. Morphological results have been obtained using a disc of radius 1 as structural element.

Method		OA	AA	κ
Pixel-wise SVM		88.97	91.60	0.8565
SVM + NWHED	CW	92.87	94.83	0.9068
	LAAM_X	92.70	94.43	0.9045
	LAAM_a	92.81	94.46	0.9059
	LAAM_r	91.93	93.62	0.8944
SVM+WHED	CW	94.71	95.99	0.9306
	LAAM_X	94.90	96.27	0.9331
	LAAM_a	94.87	96.14	0.9326
	LAAM_r	94.69	95.83	0.9303

Table 9.2: Global classification results of the Pavia University hyperspectral scene: overall accuracy (OA), average accuracy (AA) and Kappa (κ) values. Morphological results have been obtained using a disc of radius 3 as structural element.

Method		OA	AA	κ
Pixel-wise SVM		88.97	91.60	0.8565
SVM + NWHED	CW	93.41	94.39	0.9135
	LAAM_X	93.65	94.72	0.9167
	LAAM_a	93.09	94.16	0.9096
	LAAM_r	92.61	93.84	0.9034
SVM+WHED	CW	95.46	95.86	0.9403
	LAAM_X	95.27	96.11	0.9378
	LAAM_a	95.15	95.62	0.9364
	LAAM_r	94.91	95.71	0.9332

Table 9.3: Global classification results of the Pavia University hyperspectral scene: overall accuracy (OA), average accuracy (AA) and Kappa (κ) values. Morphological results have been obtained using a disc of radius 5 as structural element.

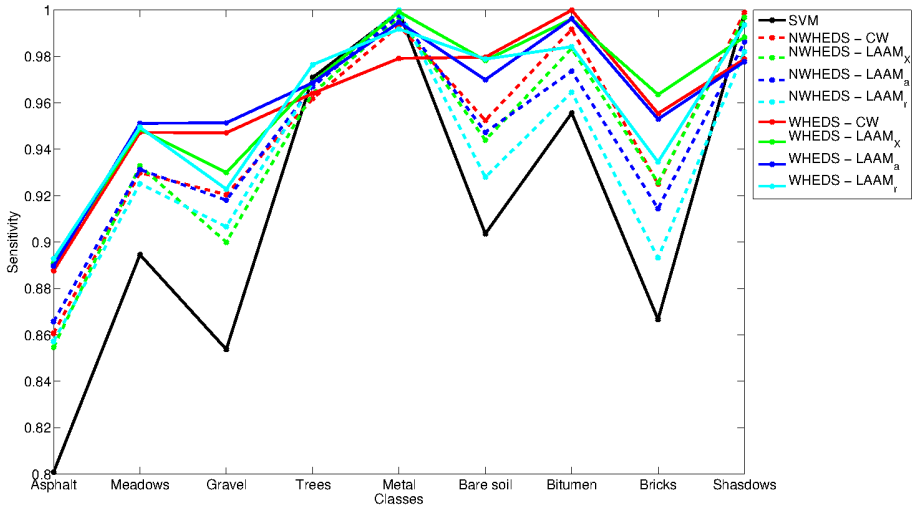


Figure 9.13: Class-specific sensitivity results for the classification of the Pavia University hyperspectral scene. Morphological results have been obtained using a disk of radius 3 as structural element.

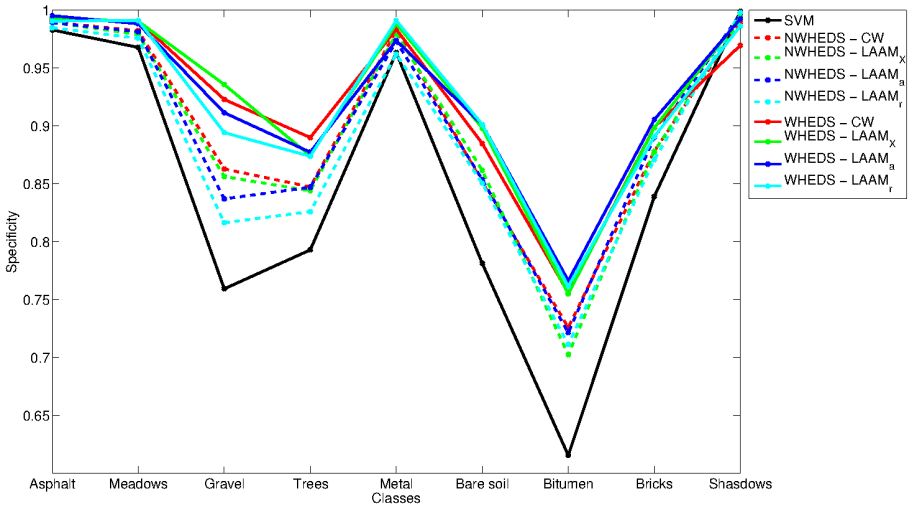


Figure 9.14: Class-specific specificity results for the classification results of the Pavia University hyperspectral scene. Morphological results have been obtained using a disc of radius 3 as structural element.

Pixels in blue are closer to the set of fixed points of the respective sets, $\mathcal{F}(X)$, $\mathcal{F}(F)$ and $\mathcal{F}(B)$; and, pixels in red far from them and thus being more lattice independent on the training sets.

Using the obtained LAAM h-mappings we performed erosions, $\varepsilon(G)$, and dilations, $\delta(G)$, with a disk structural element of increasing radius ($r = 1, 3, 5$) over the Indian Pines image using the proposed LAAM-supervised orderings. We also eroded and dilated the scene using a component-wise ordering. The Beucher gradients were calculated for the four orderings and three structural elements as shown in Figure 9.17.

Using the Beucher morphological gradients we calculated their respective watershed segmentations. Figure (9.18) shows the watershed segmentations as black lines dividing the Indian Pines scene in regions.

9.4.3.2 Experimental results

We trained a multi-class one-vs-all SVM classifier to perform pixel-wise classification of the Indian Pines scene. In this case there were not enough samples to get a balanced training set so we performed a 20-80 hold-out cross validation selecting 20% of instances per class to compose the training set. Figure 9.19 shows the classification map obtained by the pixel-wise SVM. The result is worse than in Pavia University scene, the overall accuracy is 86.59% and the average accuracy is 79.95%, since the dataset is more complicated with more classes and more spectral homogeneity among classes, and less samples per class with lot of unbalanced classes. Still there are well classified regions with iso-

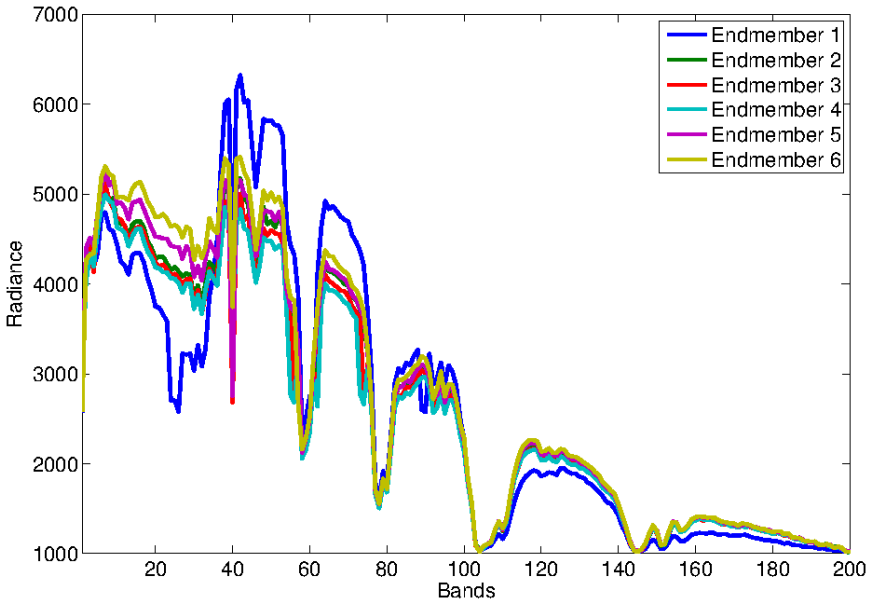
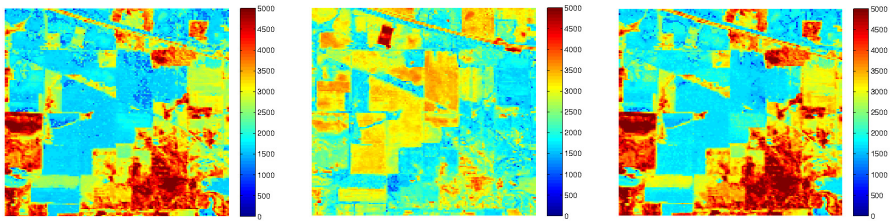


Figure 9.15: Endmembers induced by ILSIA algorithm from the Indian Pines scene.



(a) (b) (c)

Figure 9.16: h-function mappings estimated by the proposed LAAM supervised orderings for the given sets: (a) $X = \{\mathbf{e}_2\}$, (b) $F = \{\mathbf{e}_1\}$, (c) $B = \{\mathbf{e}_4\}$.

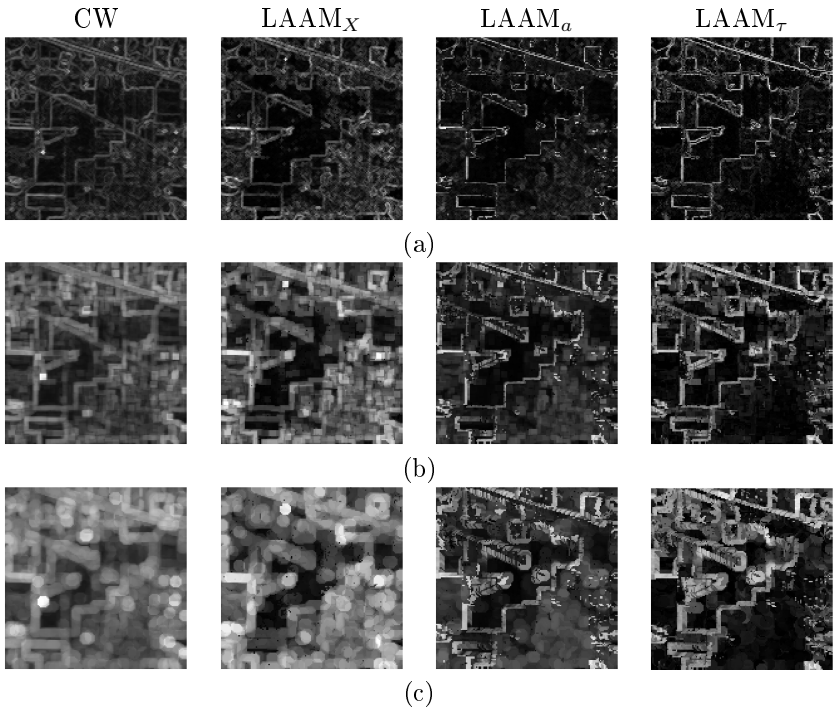


Figure 9.17: Beucher gradients obtained from the Indian Pines scene using the four different orderings and a disk structural element with increasing radius: (a) $r = 1$, (b) $r = 3$, (c) $r = 5$.

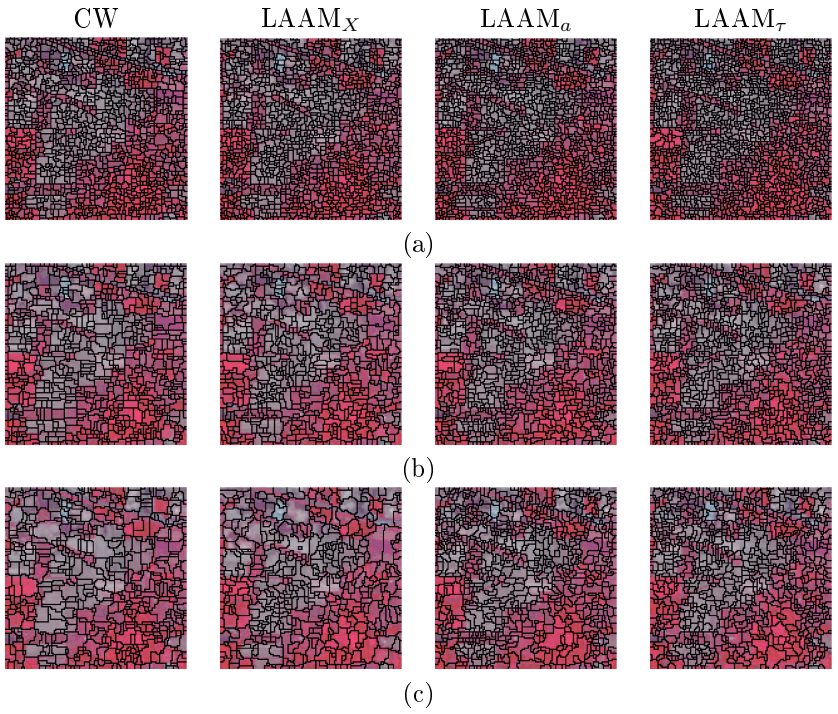


Figure 9.18: Watersheds obtained from the Indian Pines scene using the Beucher morphological gradients obtained by the four different orderings and a disk structural element with increasing radius: (a) $r = 2$, (b) $r = 3$, (c) $r = 4$.

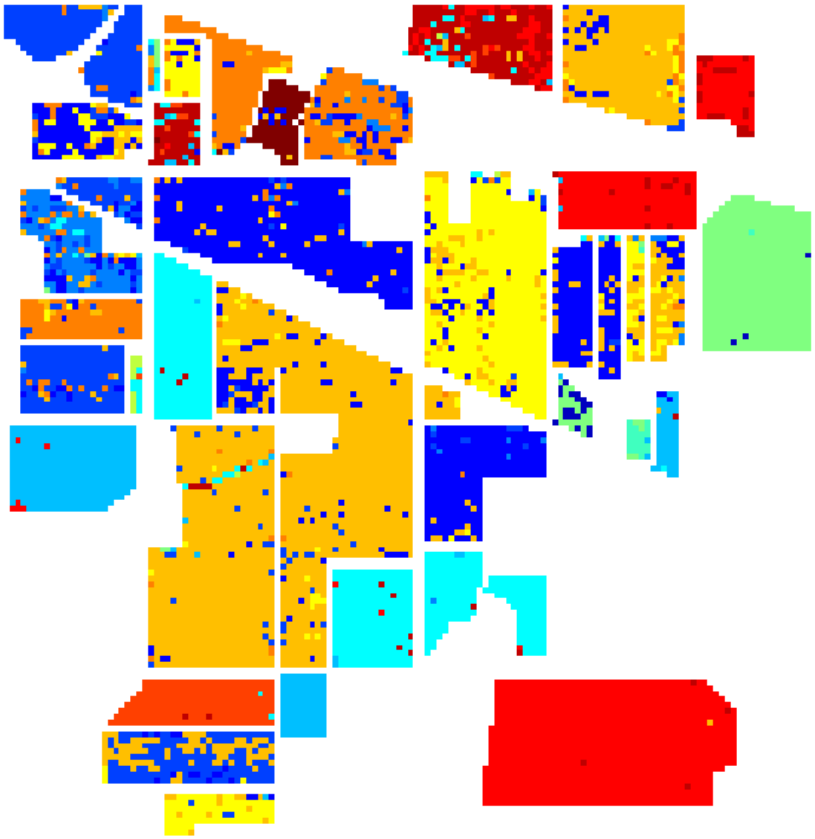


Figure 9.19: Classification map from the Indian Pines scene obtained by the pixel-wise SVM.

lated misclassified pixels within. The NWHEDS and the WHEDS contextual methodology improve the results as it can be visually checked in figures 9.20 and 9.21. The performance boosting from the pixel-wise SVM to the SVM-NWHEDS approach, and from SVM-NWHEDS to SVM-WHEDS is similar to the one in Pavia University scene. Overall and average accuracies for the pixel-wise SVM, the SVM-NWHEDS and the SVM-WHEDS are shown in tables 9.4-9.6. Again, the boosting in overall accuracy performance is similar for the different orderings. Figures 9.22 and 9.23 show the class-specific sensitivity and specificity. It can be shown that there are some problems with '*Alfalfa*' and '*Oats*' classes. Both classes have few samples, 46 and 28 respectively, and with a large structural element they can be included in regions labeled differently.

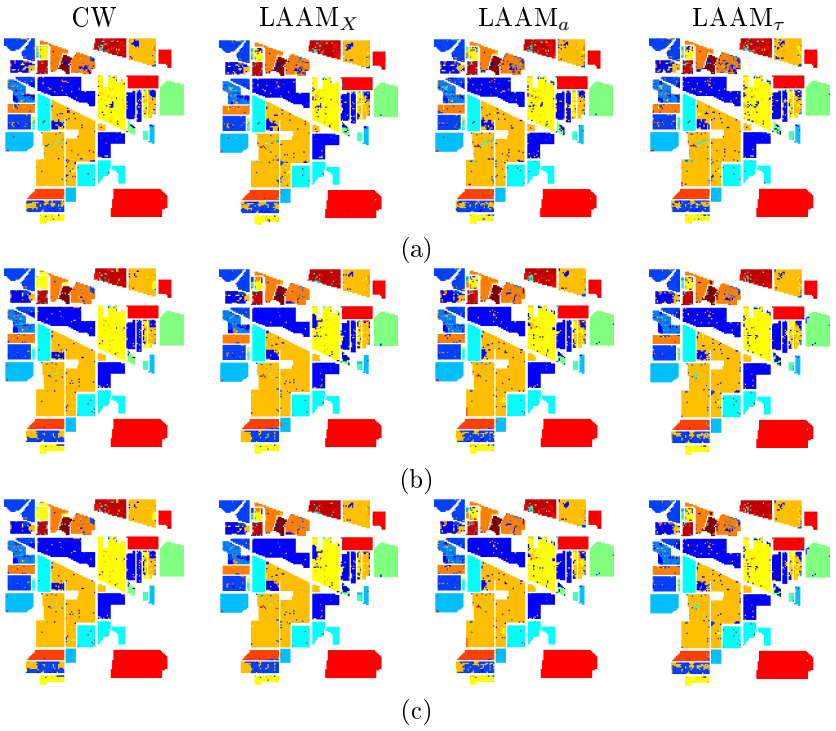


Figure 9.20: Classification maps from the Indian Pines scene obtained by the SVM-NWHEDS using the watershed segmentations obtained by the four different orderings and a disk structural element with increasing radius: (a) $r = 2$, (b) $r = 3$, (c) $r = 4$.

Method		OA	AA	κ
Pixel-wise SVM		86.59	79.95	0.8467
SVM + NWHED	CW	90.64	82.52	0.8930
	LAAM$_X$	90.02	82.63	0.8858
	LAAM$_a$	89.83	81.83	0.8838
	LAAM$_r$	88.80	79.25	0.8719
SVM+WHED	CW	95.33	85.69	0.9466
	LAAM$_X$	94.86	88.07	0.9412
	LAAM$_a$	93.92	86.00	0.9305
	LAAM$_r$	94.38	86.79	0.9358

Table 9.4: Global classification results of the Pavia University hyperspectral scene: overall accuracy (OA), average accuracy (AA) and Kappa (κ) values. Morphological results have been obtained using a disc of radius 1 as structural element.

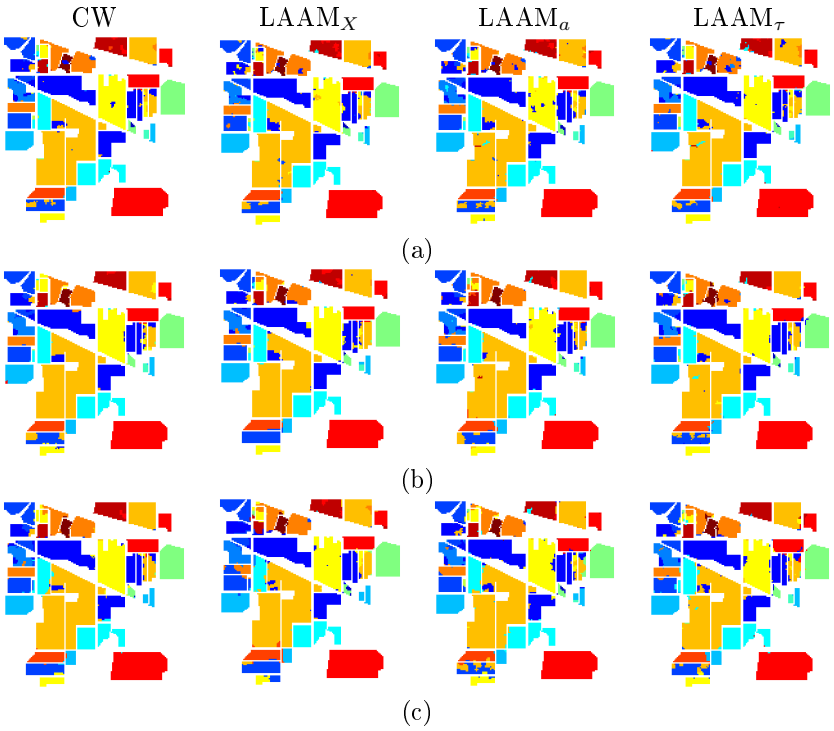


Figure 9.21: Classification maps from the Indian Pines scene obtained by the SVM-WHEDS using the watershed segmentations obtained by the four different orderings and a disk structural element with increasing radius: (a) $r = 2$, (b) $r = 3$, (c) $r = 4$.

Method		OA	AA	κ
Pixel-wise SVM		86.59	79.95	0.8467
SVM + NWHED	CW	92.17	82.28	0.9105
	LAAM$_X$	92.48	85.49	0.9140
	LAAM$_a$	90.42	82.01	0.8905
	LAAM$_\tau$	90.88	84.52	0.8958
SVM+WHED	CW	95.26	84.10	0.9458
	LAAM$_X$	95.68	84.47	0.9506
	LAAM$_a$	93.58	82.10	0.9266
	LAAM$_\tau$	94.57	85.78	0.9379

Table 9.5: Global classification results of the Pavia University hyperspectral scene: overall accuracy (OA), average accuracy (AA) and Kappa (κ) values. Morphological results have been obtained using a disc of radius 3 as structural element.

Method		OA	AA	κ
Pixel-wise SVM		86.59	79.95	0.8467
SVM + NWHED	CW	92.49	78.76	0.9141
	LAAM_X	92.01	82.73	0.9087
	LAAM_a	90.33	79.44	0.8894
	LAAM_r	90.21	80.95	0.8881
SVM+WHED	CW	94.65	76.88	0.9389
	LAAM_X	93.84	77.74	0.9297
	LAAM_a	93.47	82.15	0.9253
	LAAM_r	93.62	82.54	0.9271

Table 9.6: Global classification results of the Pavia University hyperspectral scene: overall accuracy (OA), average accuracy (AA) and Kappa (κ) values. Morphological results have been obtained using a disc of radius 5 as structural element.

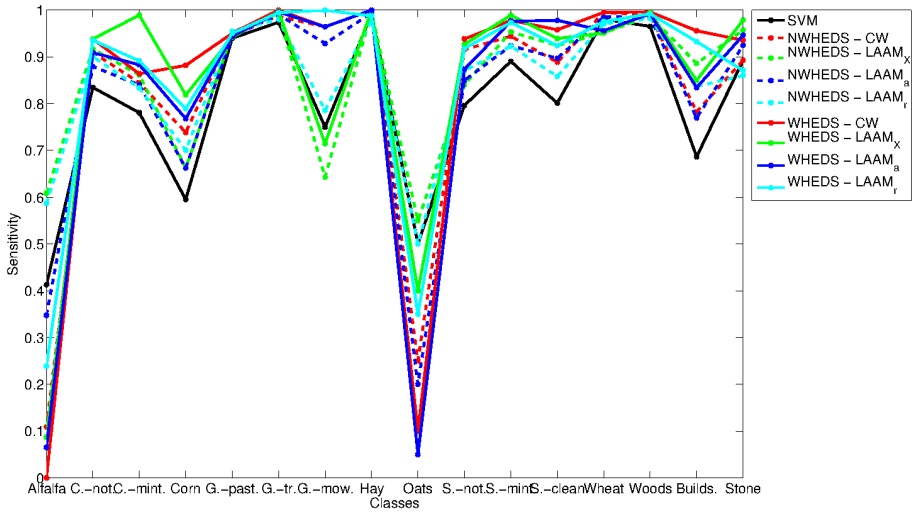


Figure 9.22: Class-specific sensitivity results for the classification of the Indian Pines hyperspectral scene. Morphological segmentation obtained using a disk of radius 3 as structural element.

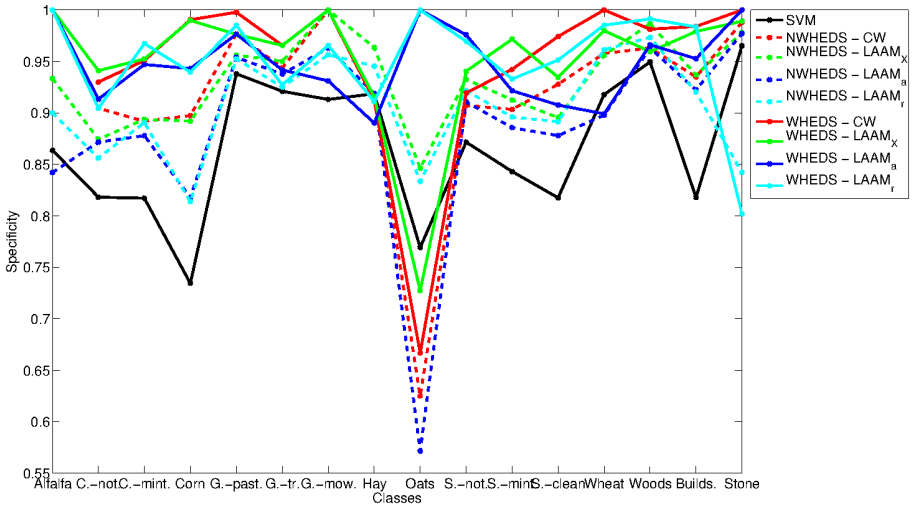


Figure 9.23: Class-specific specificity results for the classification results of the Indian Pines hyperspectral scene. Morphological segmentation obtained using a disc of radius 3 as structural element.

9.5 Conclusions

This chapter has proposed a way to perform multivariate mathematical morphology on multivariate image using lattice computing techniques. Specifically, using the recall distance of LAAMs as a reduced ordering mapping it is possible to perform mathematical morphology based on the induced supervised ordering. The chapter introduced a one-side and a background/foreground supervised orderings. The experiments with RGB color images show that the morphological operators are consistent from the intuitive point of view, so that erosions, dilations and subsequent filters defined on them do not introduce false color results. The extension to hypespectral images is immediate. It is possible to define morphological gradients, and the corresponding watershed segmentation algorithm. Validation of the morphological operators in a spatial-spectral classification approach combining the watershed segmentation with the pixel spectra classification using SVMs gives results comparable with state of the art approaches found in the literature for two well known hyperspectral benchmark images.

Chapter 10

Future work

The results and conclusions of a Thesis work unavoidably come together to ongoing works, open questions and new ideas. In this chapter we enumerate some of the research avenues that we have in mind for future works.

- One of the issues of inducing endmembers from an hyperspectral scene is that there are often objects or materials whose spectra differ along the image due to different luminosity conditions, sensor specifications, non-linear effects or to slightly changes in their own composition. This problem can cause that the induced endmembers are not enough representative of such objects spectral variability. Another issue is that some objects spectra can be masked due to they appear scarcely or highly mixed in the image, or since the endmember induction algorithm was not capable to retrieve a corresponding endmember. In both cases the result is a high residual error in the unmixing process. Some authors propose divide the scene in smaller windows such that luminosity conditions are homogeneous and analyze them separately. Others, propose an iterative process where a new spectra is selected from a library so the residual error can be reduced in those areas where is higher. We are working on the use of boosting and ensemble techniques, usually defined for classification or regression tasks, to address this issue. Furthermore, we want to take advantage of the rapid and light greedy LAAM-based endmember induction algorithms, such as EIA and ILSIA, to accomplish this task.

- We have followed up to now a computational and information theoretic point of view in the design of hyperspectral CBIR systems but, we are aware of the importance of the physical aspects of hyperspectral sensors. One of the most important aspects of hyperspectral data is the close relation between the sensed spectra and the physical properties of materials studied in lab. The high spectral resolution together to the contiguous non-overlapping measuring over different parts of the electromagnetic spectrum allows to observe absorption and emission patterns in specific wavelengths, that could uniquely identify some materials of interest. We

are interested on incorporate this relevant information to hyperspectral CBIR systems as some kind of a-priori information. This can be also of outstanding help to visualize and build up interfaces that improve the usability of relevance feedback processes.

- The spectral and spatial-spectral characterization of hyperspectral images can take advantage of the new developments in dimensionality reduction, spectral unmixing and endmember induction that we have commented in Chapter 1. We are also interested in improving the representation of spatial content. To that effect, we would like to try approaches based on the statistical model such as Markov Random Fields, as well as morphological approaches such as Extended Morphological Filters. We think that morphological approaches can be improved by the use of supervised orderings, as we have proposed in Chapter 9.

- The use of mathematical morphology can be of high interest to solve some issues related to the extraction of dictionaries from images. These techniques from Information Theory were defined with one-dimensional signals in mind, texts analysis for instance. When applied to images, there is not a single best way to convert an image to an array of symbols. Multiple paths can be follow to rearrange the images pixels and in all cases much of the spatial information is lost. We are exploring the possibility of using mathematical morphology to define morphological dictionaries in order to solve this issue.

- The semantic gap is the ultimate barrier on CBIR systems. We have proposed a relevance feedback methodology that makes use of dissimilarity spaces to face this issue in the proposed hyperspectral CBIR systems. We want to go deep into this research avenue and also explore other techniques such as image annotation or ontologies. Moreover, the dissimilarity spaces approach is a recent technique that have proof to be powerful and that gives rise to some revolutionary ideas in the machine learning theory. Bring similarity or closeness to the focus center of machine learning in the detriment of feature representation has a better sounding respect to the way humans learn. Furthermore, is a more flexible approach that links machine learning techniques to other artificial intelligence approaches such as structural or symbolic learning. We have in mind compare the use of dissimilarity spaces to well known machine learning techniques in hyperspectral analysis.

- An important issue that has been very present in this Thesis work is that of remote sensing CBIR systems validation. We have seen that there exist huge problems in collecting groundtruth data or experts assessment in the remote sensing field. This is probably one of the worst bottlenecks in the remote sensing community. This is even more severe in remote sensing CBIR, since there is the need to validate the proposed systems over huge databases. There are also big efforts in the remote sensing community to

provide automatic and robust image annotation algorithms that alleviate this problem. We have proposed an validation methodology that makes no use of any a-priori information by using automatic unsupervised clusterings. We are working on validate the proposed systems on well-known picture databases where traditional validation is easier to assess and so, we can give evidences of the validity of our proposal.

Appendix A

Lattice theory and lattice auto-associative memories

A.1 Lattice theory

Lattice Theory [18, 76, 77] is the study of sets of objects known as lattices, and provides a framework for unifying the study of classes or ordered sets in mathematics. Before defining the concept of lattice we need first to define the notion of ordering in a set. The order theoretic, and thus the topological properties of a set A, are expressed in terms of some ordering between its elements.

A binary relation ϱ on A can be defined as a subset of the Cartesian product $A^2 = A \times A$. The elements $a, b \in A$ are in relation with respect to ϱ iff $\langle a, b\rangle \in \varrho$. A more compact notation for $\langle a, b\rangle \in \varrho$ is $a \varrho b$. For all $a, b, c \in A$, the basic properties of a binary relation are as follows:

(Reflexivity) $a \varrho a$.

(Antisymmetry) $a \varrho b$ and $b \varrho a$ imply that $a = b$.

(Transitivity) $a \varrho b$ and $b \varrho c$ imply that $a \varrho c$.

(Linearity) $a \varrho b$ or $b \varrho a$.

A relation ϱ satisfying reflexivity, antisymmetry and transitivity properties is called an *ordering* and will usually be denoted by $\leq$. A non-empty set P equipped with such an ordering relation, $\langle P; \leq\rangle$, is called a *partially-ordered set* or *poset* and will be denoted by $\mathcal{P}$. A poset satisfying linearity property is called a *totally ordered set* or *chain*. Let $\mathcal{P} = \langle P; \leq\rangle$ be a poset, the elements $a, b \in P$ are *comparable* if $a \leq b$ or $b \leq a$; otherwise, a and b are *incomparable*, denoted as $a \parallel b$. In a chain all elements are comparable.

If $\langle P; \leq\rangle$ is an order then $\langle P; \geq\rangle$ is also an order, called the *dual* of $\langle P; \leq\rangle$ and denoted by $\mathcal{P}^\delta$. Now, if Φ is a statement about orders and we replace in Φ all the occurrences of $\leq$ by $\geq$, we get the dual of Φ, denoted by Φ^δ. This yields to the Duality Principle:

Lemma 22. *If a statement Φ is true in all orders, then its dual, Φ^δ, is also true in all orders.*

Let $X \subseteq P$ and $a \in P$. Then, a is an *upper-bound* of X if $a \geq x$ for all $x \in X$. An upper-bound a of X is the *least upper bound* of X, or *supremum* of X, if a is the least of all the upper bounds of H. The supremum is denoted by $a = \sup H$ or $a = \bigvee H$. The concepts of *lower bound* and *greatest lower bound* or *infimum* are dually defined; the latter is denoted by $a = \inf H$ or $a = \bigwedge H$. The supremum and infimum are unique if they exist.

A poset $\mathcal{L} = \langle L; \leq \rangle$ is a *lattice* if an infimum and a supremum exist for any pair of elements of L. The formal definition of a lattice is as follows:

Definition 23. A poset $\mathcal{L} = \langle L; \leq \rangle$ is a lattice if $\inf H$ and $\sup H$ exist for any finite non-empty subset $H \subseteq L$.

A lattice is a special type of poset so Duality Principle holds and, if $\mathcal{L} = \langle L; \leq \rangle$ is a lattice, so is its dual $\mathcal{L}^\delta = \langle L; \geq \rangle$. A lattice can also be described in an algebraic form by setting $a \wedge b = \inf\{a, b\}$ and $a \vee b = \sup\{a, b\}$. Thus, the algebra $\mathcal{L} = \langle L, \wedge, \vee \rangle$ is a lattice. Let the poset $\mathcal{L}^p = \langle L; \leq \rangle$ and the algebra $\mathcal{L}^a = \langle L, \wedge, \vee \rangle$ be lattices, then $(\mathcal{L}^a)^p = \mathcal{L}$ and $(\mathcal{L}^p)^a = \mathcal{L}$, stating that a lattice as an algebra or as a poset are equivalent concepts. An important kind of lattices is that of *complete lattices*.

Definition 24. A lattice $\mathcal{L} = \langle L; \leq \rangle$ is complete iff $\inf H$ and $\sup H$ exist for any subset $H \subseteq L$.

A complete lattice has both a smallest element called *bottom*, denoted as $\bot$, and a greatest element called top, denoted as $\top$. All finite lattices are complete lattices.

A.2 Lattice Auto-Associative Memories (LAAMs)

The work on Lattice Associative Memories (LAMs) stems from the consideration of the bounded lattice ordered group $(\mathbb{R}_{\pm\infty}, \vee, \wedge, +, +')$ as the alternative to the algebraic framework $(\mathbb{R}_{\pm\infty}, +, \cdot)$ for the definition of Neural Networks computation [151, 149], where $\mathbb{R}_{\pm\infty} = \mathbb{R} \cup \{-\infty, +\infty\}$ is the set of extended real numbers, the operators $\vee$ and $\wedge$ respectively denote the discrete max and min operators (sup and inf in a continuous setting), and $+, +'$ respectively denote addition and its dual operation such that $x +' y = y + x$, $\forall x \in \mathbb{R}$, $\forall y \in \mathbb{R}_{\pm\infty}$; $\infty +' (-\infty) = \infty = (-\infty) +' \infty$ and $\infty + (-\infty) = -\infty = (-\infty) + \infty$. Thus, the additive conjugate is given by $x^* = -x$.

Given a set of input/output pairs of patterns $(X, Y) = \left\{ \left(\mathbf{x}^\xi, \mathbf{y}^\xi \right); \xi = 1, .., k \right\}$, a linear heteroassociative neural network based on the pattern's cross correlation [88] is built up as $W = \sum_\xi \mathbf{y}^\xi \cdot \left(\mathbf{x}^\xi \right)'$. Mimicking this constructive procedure authors in [151, 149] proposed the following constructions of LAMs (denoted in those works as Morphological Associative Memories):

$$W_{XY} = \bigwedge_{\xi=1}^{k} \left[\mathbf{y}^{\xi} \times \left(-\mathbf{x}^{\xi} \right)' \right] \text{ and } M_{XY} = \bigvee_{\xi=1}^{k} \left[\mathbf{y}^{\xi} \times \left(-\mathbf{x}^{\xi} \right)' \right], \tag{A.1}$$

where $\times$ is any of the $\boxtimes$ or $\boxtimes$ operators, reducing the notational burden since $\mathbf{y}^{\xi} \boxtimes \left(-\mathbf{x}^{\xi} \right)' = \mathbf{y}^{\xi} \boxtimes \left(-\mathbf{x}^{\xi} \right)'$. Here $\boxtimes$ and $\boxtimes$ denote the max and min matrix product, respectively defined as follows:

$$C = A \boxtimes B = [c_{ij}] \Leftrightarrow c_{ij} = \bigvee_{k=1..n} \{a_{ik} + b_{kj}\}, \tag{A.2}$$

$$C = A \boxtimes B = [c_{ij}] \Leftrightarrow c_{ij} = \bigwedge_{k=1..n} \{a_{ik} + b_{kj}\}. \tag{A.3}$$

If $X = Y$ then W_{XX} and M_{XX} are called Lattice Auto-Associative Memories (LAAMs). LAAMs present some surprising properties: perfect recall for an unlimited number of stored patterns and convergence in one step for any input pattern. Following theorems [151, 149] formulate these properties:

Theorem 25. *$W_{XX} \boxtimes X = X = M_{XX} \boxtimes X$.*
$W_{XX} \boxtimes \mathbf{x} = \mathbf{x}$ iff $M_{XX} \boxtimes \mathbf{x} = \mathbf{x}$.
If $W_{XX} \boxtimes \mathbf{z} = \mathbf{v}$ and $M_{XX} \boxtimes \mathbf{z} = \mathbf{u}$, then $W_{XX} \boxtimes \mathbf{v} = \mathbf{v}$ and $M_{XX} \boxtimes \mathbf{u} = \mathbf{u}$.

Definition 26. A vector $\mathbf{x} \in \mathbb{R}^n_{\pm\infty}$ is called a fixed point or stable state of W_{XX} iff $W_{XX} \boxtimes \mathbf{x} = \mathbf{x}$. Similarly, $\mathbf{x}$ is a fixed point of M_{XX} iff $M_{XX} \boxtimes \mathbf{x} = \mathbf{x}$.

Theorem 25 establishes that W_{XX} and M_{XX} are perfect recall memories for any number of uncorrupted input vectors, theorem 25 implies one step convergence, and theorem 25 says that W_{XX} and M_{XX} share the same set of fixed points represented by $\mathcal{F}(X)$ [175, 150].

Theorem 27. *For every $\mathbf{x} \in \mathbb{R}^n_{\pm\infty}$, we have $W_{XX} \boxtimes \mathbf{x} = \hat{\mathbf{x}}$ and $M_{XX} \boxtimes \mathbf{x} = \check{\mathbf{x}}$, where $\hat{\mathbf{x}}$ denotes the supremum of $\mathbf{x}$ in the set of fixed points of W_{XX}, and $\check{\mathbf{x}}$ denotes the infimum of $\mathbf{x}$ in the set of fixed points of M_{XX}.*

LAAMs are extremely robust to erosive or dilative noise, but not to the presence of both. A distorted version $\tilde{\mathbf{x}}^{\gamma}$ of the pattern $\mathbf{x}^{\gamma}$ has undergone an erosive change whenever $\tilde{\mathbf{x}}^{\gamma} \leq \mathbf{x}^{\gamma}$, and a dilative change when $\tilde{\mathbf{x}}^{\gamma} \geq \mathbf{x}^{\gamma}$. Particularly, the erosive LAAM, W_{XX}, is extremely robust to erosive changes, while the dilative LAAM, M_{XX}, is so to dilative changes. Research on robust recall [149, 175, 144, 154, 176, 150] based on the so-called kernel patterns lead to the notion of Lattice Independence (LI) and the recall exact description in terms of the LAMs fixed points and their basis of attractions. The definition of Lattice Independence is closely tied to the study of the LAAM fixed points when they are interpreted as lattice transformations, as stated by the following theorems:

Definition 28. Given a set of vectors $\left\{\mathbf{x}^1, ..., \mathbf{x}^k\right\} \subset \mathbb{R}^n$ a *linear minimax combination* of vectors from this set is any vector $\mathbf{x} \in \mathbb{R}^n_{\pm\infty}$ which is a *linear minimax sum* of these vectors:

$$x = \mathcal{L}\left(\mathbf{x}^1, ..., \mathbf{x}^k\right) = \bigvee_{j \in J} \bigwedge_{\xi=1}^{k} \left(a_{\xi j} + \mathbf{x}^\xi\right),$$

where J is a finite set of indexes and $a_{\xi j} \in \mathbb{R}_{\pm\infty}$ $\forall j \in J$ and $\forall \xi = 1, ..., k$.

The *linear minimax span* of vectors $\left\{\mathbf{x}^1, ..., \mathbf{x}^k\right\} = X \subset \mathbb{R}^n$ is the set of all linear minimax sums of subsets of X, denoted $LMS\left(\mathbf{x}^1, ..., \mathbf{x}^k\right)$.

Given a set of vectors $X = \left\{\mathbf{x}^1, ..., \mathbf{x}^k\right\} \subset \mathbb{R}^n$, a vector $\mathbf{x} \in \mathbb{R}^n_{\pm\infty}$ is *lattice dependent* iff $x \in LMS\left(\mathbf{x}^1, ..., \mathbf{x}^k\right)$. The vector $\mathbf{x}$ is *lattice independent* iff it is not lattice dependent on X. The set X is said to be *lattice independent* iff $\forall \lambda \in \{1, ..., k\}$, $\mathbf{x}^\lambda$ is lattice independent of $X \backslash \left\{\mathbf{x}^\lambda\right\} = \left\{\mathbf{x}^\xi \in X : \xi \neq \lambda\right\}$.

Theorem 29. *Given a set of vectors* $X = \left\{\mathbf{x}^1, ..., \mathbf{x}^k\right\} \subset \mathbb{R}^n$, *a vector* $\mathbf{y} \in \mathbb{R}^n_{\pm\infty}$ *is a fixed point of* $\mathcal{F}(X)$, *that is* $W_{XX} \boxtimes \mathbf{y} = \mathbf{y} = M_{XX} \boxtimes \mathbf{y}$, *iff* $\mathbf{y}$ *is lattice dependent on* X.

Appendix B

Endmember induction algorithms

For the sake of completeness, in this appendix we present a brief review of the EIAs employed in the experiments, the geometric approach represented by N-FINDER, and the lattice computing approach by the EIHA. We will denote $\{\mathbf{f}(i) \in \mathbb{R}^q : i = 1, .., N\}$ the collection of spectra corresponding to the pixels of the image, and $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_p\}$ the collection of endmembers induced from the image. The vector $\bar{\mathbf{f}}$ denotes the vector of component-wise means over the image pixels. The vector $\boldsymbol{\sigma}$ denotes the vector of component-wise standard deviations over the image pixels.

B.1 N-FINDER

Algorithm B.1 presents the N-FINDER [193], which works by growing a simplex inside the data, beginning with a random set of pixels. The vertices of the simplex with higher volume are assumed to identify the endmembers. Previously, data dimensionality has to be reduced to $p-1$ dimensions, with p the number of endmembers searched for.

Let E be the matrix of endmembers augmented with a row of ones

$$E = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ \mathbf{e_1} & \mathbf{e_2} & \ldots & \mathbf{e}_p \end{bmatrix}, \tag{B.1}$$

where $\mathbf{e}_i$ is a column vector containing the spectra of the i-th endmember. The volume of the simplex defined by the endmembers is proportional to the determinant of E

$$V(E) = \frac{abs(\det(E))}{(p-1)!}. \tag{B.2}$$

The N-FINDER starts by selecting an initial random set of pixels as endmembers. Then, for each pixel and each stored endmember, the endmember is

Algorithm B.1 N-FINDER algorithm.

1. Compute Principal Component Analysis (PCA) retaining the first principal components accounting for 99% of the sum of the eigenvalues. Let $p-1$ be the number of eigenvectors retained.
2. Randomly select p vectors from the data to initialize the set of induced endmembers E.
3. Calculate the volume of the simplex $v = V(E)$ of eq. (B.2). $v_{actual} = v$.
4. For each endmember $\mathbf{e}_k; k = 1, \ldots p$:
   (a) For each data vector $\mathbf{f}(i) : i = 1, .., N$:
      i. Form a new matrix E' by substituting the endmember $\mathbf{e}_k$ by the data vector $\mathbf{f}(i)$.
      ii. Calculate the volume of the simplex $v' = V(E')$.
      iii. If $v' > v_{actual}$ then E' becomes E. $v_{actual} = v'$.
5. If $v_{actual} > v$ then $v = v_{actual}$. Go to step 4.

replaced with the spectrum of the pixel and the volume recalculated by equation B.2. If the volume of the new simplex increases, the endmember is replaced by the spectrum of the pixel. The procedure ends when no more replacements are done. The N-FINDER is a greedy algorithm, prone to fall in local maxima of the volume function defined by equation (B.2).

B.2 FIPPI

The Fast Iterative Pixel Purity Index (FIPPI) algorithm [31] is an improvement of the Pixel Purity Index (PPI) algorithm [20]. PPI has been widely used in hyperspectral image analysis community due to its availability in the ENVI software. However, its detailed implementation has never been made available in the literature. FIPPI improves PPi in several aspects: it estimates the number of endmembers required by virtual dimensionality and it is an iterative and usupervised algorithm as opposed to PPI, which requires human intervention.

The FIPPI algorithm B.2 works by projecting the sample data onto unit vectors named skewers. The idea is that those samples who projections are most extreme correspond to endmembers. The algorithm starts by calculating the number p of endmembers required to be induced, by the Harsanyi-Farrand-Chang (HFC) virtual dimensionality method [83]. Then the hyperspectral data dimensionality is reduced by applying a Minimum Noise Fraction (MNF) transform [78] and keeping the first p components. The initial set of skewers, skewer$^{(0)}$, is generated by selecting those pixels that correspond to target pixels selected by the Automatic Target Generation Process (ATGP) algorithm [138]. The itera-

Algorithm B.2 Fast Iterative Pixel Purity Index (FIPPI) algorithm.

1. Find the number of endmembers required p by HFC virtual dimensionality method.
2. Apply the MNF transform and reduce the data dimensinality by keeping the first p components.
3. Use the ATGP algorithm to initialize the set of p skewers, $\left\{\text{skewer}_j^{(0)}\right\}_{j=1}^{p}$.
4. At iteration $k \geq 0$:
   (a) Project all the sample vectors onto each $\text{skewer}_j^{(k)}$.
   (b) Find those samples whose projections are the most extreme and form the extrema set, $S_{\text{extrema}}\left(\text{skewer}^{(k)}\right)$.
   (c) Find the sample vectors with largest $N_{\text{PPI}}\left(\mathbf{r}^{(k)}\right)$.
   (d) Form the joint set $\left\{\text{skewer}^{(k+1)}\right\} = \left\{\text{skewer}^{(k)}\right\} \cup \left\{\mathbf{r}_j^{(k)}\right\}_{N_{\text{PPI}}\left(\mathbf{r}^{(k)}\right)>0}$.
5. Stop if $\left\{\text{skewer}^{(k+1)}\right\} = \left\{\text{skewer}^{(k)}\right\}$.
6. Return the induced endmembers $E = \{\mathbf{r}\}_{N_{\text{PPI}}\left(\mathbf{r}^{(k+1)}\right)>0}$.

tive process runs by projecting all the sample vectors $\mathbf{r}$ onto each skewer and then, finding those samples which lie in the extreme positions of the projection axis. Those samples form the extrema set, denoted as $S_{\text{extrema}}\left(\text{skewer}^{(k)}\right)$. Now, for each sample data, we calculate the PPI score, $N_{\text{PPI}}\left(\mathbf{r}^{(k)}\right)$. Let $I_S\left(\mathbf{r}\right)$ be an indicator function such that returns 1 if $\mathbf{r} \in S$, and otherwise returns 0. The PPI score is defined as:

$$N_{\text{PPI}}\left(\mathbf{r}^{(k)}\right) = \sum_j I_{S_{\text{extrema}}\left(\text{skewer}^{(k)}\right)}\left(\mathbf{r}\right). \tag{B.3}$$

Then, the new skewers set, $\text{skewer}^{(k+1)}$, is form by the union of the skewer set at iteration k with those samples whose PPI score is greater than zero. The algorithm continues until no new skewers are incorporated. Once the iteration process is over the PPI score is calculated again for all the sample data and, the set of induced endmembers is composed of those samples whose PPI score at iteration $k+1$ is greater than zero.

B.3 EIHA

The Endmember Induction Heuristic Algorithm (EIHA) proposed in [70, 73, 75] is based on the fact that sets of Strong Lattice Independent vectors are Affine Independent [155] and thus, they can be interpreted as a collection of endmembers for the analysis of hyperspectral data. The necessary conditions for Strong Lattice Independence are Lattice Independence and max/min dominance. Lattice Independence is detected based on results on fixed points for Lattice Autoassociative Memories (LAMs) [144, 150, 177, 155]. Max/min dominance can be tested directly [181], but in this algorithm it is a byproduct of the selection of the endmember candidate. Given a set of vector patterns $X = \left\{\mathbf{x}^{\xi}; \xi = 1, .., K\right\}$, and mimicking the constructive procedure for linear autoassociative memories, [151, 149] propose the constructions of dual LAMs as follows: $W_{XX} = \bigwedge_{\xi=1}^{k} \left[\mathbf{x}^{\xi} \times \left(-\mathbf{x}^{\xi}\right)'\right]$ and $M_{XX} = \bigvee_{\xi=1}^{k} \left[\mathbf{x}^{\xi} \times \left(-\mathbf{x}^{\xi}\right)'\right]$, where $\times$ is any of the $\boxtimes$ or $\boxtimes$ operators. Here $\boxtimes$ and $\boxtimes$ denote the max and min matrix products defined in [151, 149]. A lattice dependent vector will be a fixed point of anyone of the dual LAMs constructed with the current E.

The EIHA specified by Algorithm B.3 is a fast heuristic endmember induction method that passes only once over each pixel of the hyperspectral image. It starts with a randomly picked input vector as the single initial endmember, tests if each pixel spectrum in the input hyperspectral image data is lattice independent relative to the already discovered endmembers. If so, the pixel spectrum is added to the set of endmembers. The following notation is used in Algorithm B.3. The expression $(\mathbf{x} > \mathbf{0})$ denotes a vector of 0's and 1's, where the i-th component is 0 if $x_i \leq 0$, and 1 otherwise. Therefore, X is a collection of binary vectors corresponding to the sign of the endmembers' components. Also, $\mathbf{f}^{+}(i)$ and $\mathbf{f}^{-}(i)$ are binary vectors respectively corresponding to the sign of the components of the dilated and eroded pixel spectrum $\mathbf{f}(i)$ by the vector of standard deviations $\boldsymbol{\sigma}$ scaled by a gain factor α, whose conventional value is $\alpha = 2$. The lattice independence is tested by computing the vectors $\mathbf{y}^{+}(i)$ and $\mathbf{y}^{-}(i)$ recalled from the erosive and dilative LAM, respectively, built from X. If neither recall returns a vector already in X then we declare that the corresponding pixel spectrum is a new endmember. If we do not detect a new endmember but the pixel spectrum can be considered a dilation or an erosion of an already selected endmember, then it becomes the new endmember substituting the old one, ensuring the max/min dominance in the set of endmembers.

B.4 ILSIA

The Incremental Lattice Source Induction Algorithm (ILSIA) algorithm [68] stems in the EIHA algorithm detailed in Section B.3. ILSIA induce a set of endmembers by computing a set of Strongly Lattice Independent vectors using properties of lattice associative memories regarding Lattice Independence and Chebyshev best approximation. The dataset is denoted by $Y = \{\mathbf{y}_j;\ j = 1, \ldots, N\} \in \mathbb{R}^{n \times N}$ and the current set of the lattice independent sources at each step is de-

Algorithm B.3 Endmember Induction Heuristic Algorithm (EIHA) algorithm.

1. Shift the data sample to a global zero mean $\left\{\mathbf{f}^c(i) = \mathbf{f}(i) - \bar{\mathbf{f}}; i = 1, .., n\right\}$.
2. Initialize the set of endmembers $E = \{\mathbf{e}_1 = \mathbf{f}^c(i^*)\}$ where i^* is a randomly picked sample index. The initial set of endmember sample indices is $I = \{i^*\}$.
3. Initialize the set of lattice independent binary signatures $X = \{\mathbf{x}_1\} = \{\mathbf{e}_1 > \mathbf{0}\}$.
4. Construct the LAM's based on X: M_{XX} and W_{XX}.
5. For each input image feature vector $\mathbf{f}^c(i)$
   (a) Compute the variance induced dilations and erosions sign vectors $\mathbf{f}^+(i) = (\mathbf{f}^c(i) + \alpha\boldsymbol{\sigma} > \mathbf{0})$ and $\mathbf{f}^-(i) = (\mathbf{f}^c(i) - \alpha\boldsymbol{\sigma} > \mathbf{0})$.
   (b) Check Lattice Independence by computing $\mathbf{y}^+(i) = M_{XX} \boxtimes \mathbf{f}^+(i)$ and $\mathbf{y}^-(i) = W_{XX} \boxtimes \mathbf{f}^-(i)$.
      i. If $\mathbf{y}^+(i) \notin X$ and $\mathbf{y}^-(i) \notin X$ then $\mathbf{f}^c(i)$ is a new vertex to be added to E, and $\mathbf{f}^c(i) > \mathbf{0}$ is added to X. Execute step 4 with the new X and resume the exploration of the data samples.
      ii. if $\mathbf{y}^+(i) \in X$ and $\mathbf{f}^c(i) > \mathbf{e}_{\mathbf{y}^+}$ then the pixel spectral signature is a dilation of the endmember $\mathbf{e}_{\mathbf{y}^+}$, we substitute it by $\mathbf{f}^c(i)$.
      iii. if $\mathbf{y}^-(i) \in X$ and $\mathbf{f}^c(i) < \mathbf{e}_{\mathbf{y}^-}$ then the pixel spectral signature is an erosion of the endmember $\mathbf{e}_{\mathbf{y}^-}$, we substitute it by $\mathbf{f}^c(i)$.
6. The output set of endmembers is the set of original data vectors $\{\mathbf{f}(i) : i \in I\}$ corresponding to the data vectors selected as members of E.

noted by $X = \{\mathbf{x}_p;\, p = 1, \ldots, K\} \in \mathbb{R}^{n \times K}$. The variable K contains the number of lattice independent sources computed by the algorithm. The algorithm makes only one pass through the sample as in EIHA. The auxiliary variables $\mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^n$, respectively, count the times that a row stores the maximum, minimum of the component-wise differences (computed in $\mathbf{d} \in \mathbb{R}^n$) between pairs of lattice sources. Expression $(\mathbf{d} == m_1)$ in Algorithm B.4 denotes a vector of 0's or 1's, where 1 means that corresponding component of $\mathbf{d}$ equals the scalar value m_1.

The objective of the algorithm is to extract a set of Strong Lattice Independent vectors X from the input dataset. The algorithm's result is a set of Affine Independent vectors, furthermore they define a simplex that covers most of the data points in the dataset. To ensure that the resulting set of vectors is Strong Lattice Independent, we first ensure that we select vectors that are Lattice Independent in step 3(a) of Algorithm B.4. Each new input vector $\mathbf{y}_j$ is applied to the LAAM constructed with the already selected lattice independent sources W_{XX}. If recall is perfect, then $\mathbf{y}_j$ is Lattice Dependent on X, and can be discarded. If not, $\mathbf{y}_j$ is candidate to be added to X. We test in step 3(c) if X enlarged with the new input vector is max- or min-dominant. Note that we need to test the whole lattice source set because max- or min-dominance property is not guaranteed when we add a vector to a set of max- or min-dominant vectors. Furthermore, to test for Lattice Independence we only need to compute W_{XX} because the set of fixed points is the same for both types of LAAM. Nevertheless, we need to test both max- and min-dominance because Strong Lattice Independence requires either one of them or both to hold. If Strong Lattice Independence was the only criterion for including input vectors in X, then we would end up with a large set X, resulting in little significance of the abundance coefficients because many of the lattice independent sources will be closely placed in the data space. The strategy of Algorithm 1 discards Lattice Independent input vectors that are approximated by a fixed point of W_{XX} below a specified Chebyshev distance threshold. This test is performed in step 3(b) of the algorithm, before proceeding to test max- or min-dominance. We interpret this test in two ways: either the input vector is considered a noisy version of a vector which is Lattice Dependent on the current set X, or the input vector is too close to be lattice dependent on X.

Algorithm B.4 Incremental Lattice Source Induction Algorithm (ILSIA) algorithm.

1. Initialize the set of Lattice Independent Sources $X = \{\mathbf{x}_1\}$ with a randomly selected vector in the input dataset Y.
2. Construct the LAAM based on the current Lattice Independent Sources: W_{XX}.
3. For each data vector $\mathbf{y}_j$, $j = 1, \ldots, N$:
   (a) If $\mathbf{y}_j = W_{XX} \boxtimes \mathbf{y}_j$, then $\mathbf{y}_j$ is lattice dependent on the set X. Therefore, skip further processing.
   (b) If $\zeta\left(W_{XX} \boxtimes \left(\boldsymbol{\mu} + \mathbf{x}^{\#}\right), \mathbf{y}_j\right) < 0$, where $\mathbf{x}^{\#} = W_{XX}^{*} \boxtimes \mathbf{y}_j$ and $\boldsymbol{\mu} = \frac{1}{2}\left(W_{XX} \boxtimes \mathbf{x}^{\#}\right) \boxtimes \mathbf{y}_j$, then skip further processing.
   (c) Test max- and min-dominance to ensure Strong Lattice Independence, consider the enlarged set of lattice independent sources $X' = X \cup \{\mathbf{y}_j\}$:
      i. $\mu_1 = \mu_2 = 0$.
      ii. for $i = 1, \ldots, K+1$:
         A. $\mathbf{s}_1 = \mathbf{s}_2 = \mathbf{0}$.
         B. for $k = 1, \ldots, K+1$ and $k \neq i$: $\mathbf{d} = \mathbf{x}_i - \mathbf{x}_k$, $m_1 = \max(\mathbf{d})$, $m_2 = \min(\mathbf{d})$, $\mathbf{s}_1 = \mathbf{s}_1 + (\mathbf{d} == m_1)$, $\mathbf{s}_2 = \mathbf{s}_2 + (\mathbf{d} == m_2)$.
         C. $\mu_1 = \mu_1 + (\max(\mathbf{s}_1) == K)$ or $\mu_2 = \mu_2 + (\max(\mathbf{s}_2) == K)$.
      iii. If $\mu_1 = K+1$ or $\mu_2 = K+1$, then $X' = X \cup \{\mathbf{y}_j\}$ is Strong Lattice Independent. Go to step 2 with the enlarged set of lattice independent sources and resume exploration from $j+1$.
4. The set of lattice independent sources is X.

Appendix C

Synthetic hyperspectral images

The synthetic hyperspectral images are generated as linear mixtures of a set of spectra (the groundtruth endmembers) according to synthesized fractional abundance coefficients. Figure C.1 shows a flow diagram of the hyperspectral image synthesis process. First, we select a given number m of spectral signatures from a pool of spectral signatures. The same number m of synthetic groundtruth abundance images will be generated. The generation of the abundance coefficients is a spatial process performed independently for each desired endmember, which does not ensure for each pixel in the image the normalization properties required by the linear mixing model. To ensure that there are regions of almost pure endmembers, for each pixel we select the abundance coefficient with the greatest value, normalizing the remaining coefficients to ensure that the pixel's abundance coefficients sum up to one. The abundance images are used to produce linear combinations of endmembers at each pixel to obtain the resulting synthetic hyperspectral image.

Figure C.2 shows a flow diagram of the synthetic datasets creation process including the generation of noisy images, which has been used in Chapter 6.

The spectral signatures pool can be obtained from any of the available spec-

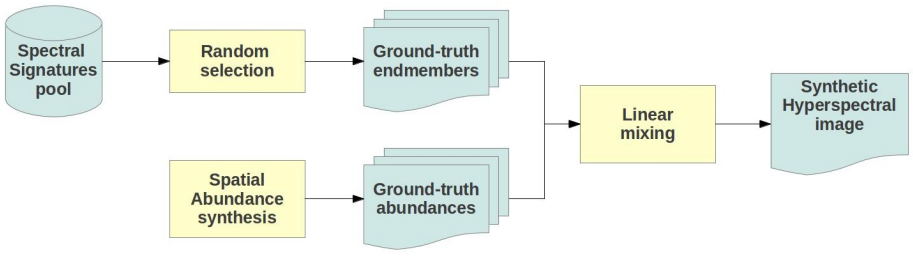


Figure C.1: Hyperspectral image synthesis process diagram.

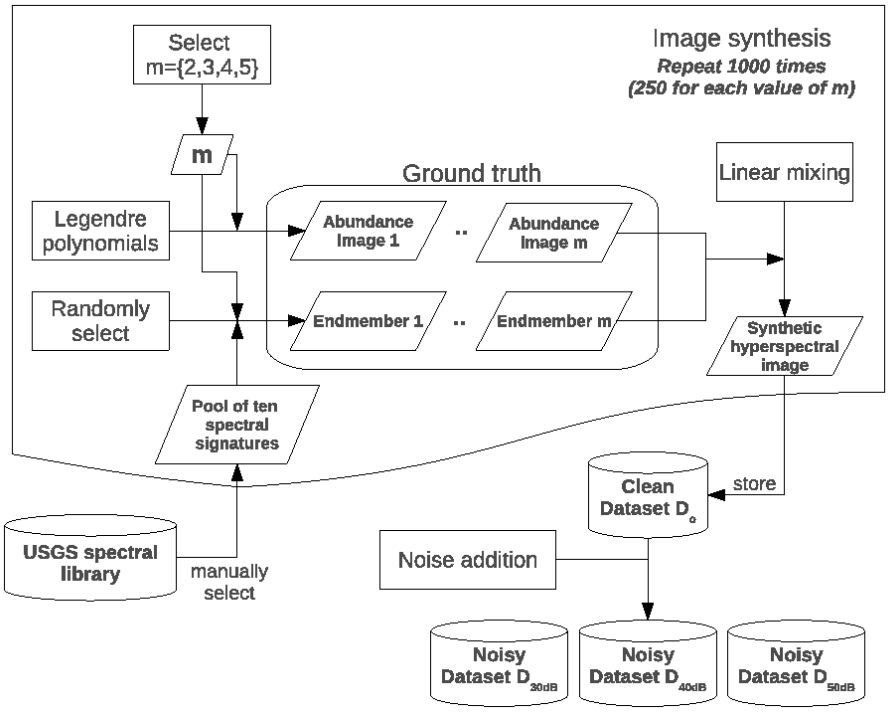


Figure C.2: Flow diagram of the synthetic hyperspectral datasets creation process when noisy images are also generated.

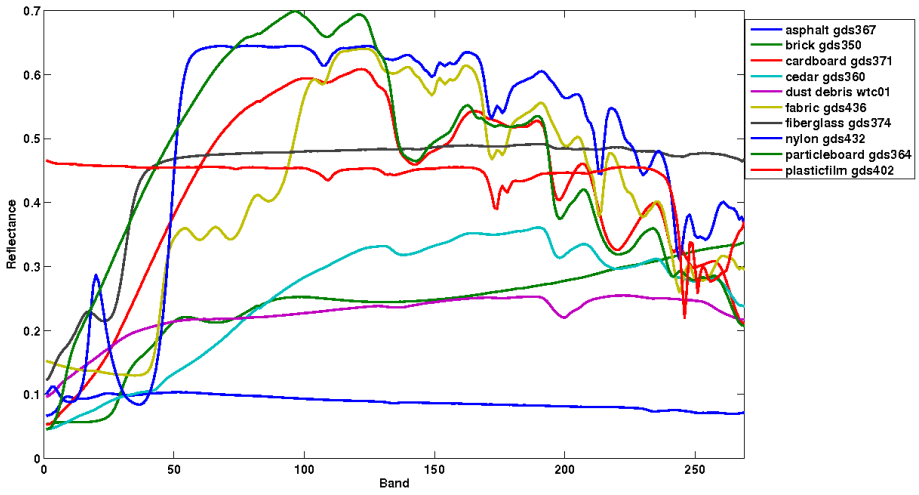


Figure C.3: Collection of endmembers selected from the USGS library to be the basis to synthesize the hyperspectral images in datasets from 10 endmember pools.

tral datasets, such as the USGS Digital Spectral Library[1], the ASTER library[2] or the SPECCHIO library[3]. The simulation of ground-truth abundance images is done as either a 2-D product of 1-D Legendre polynomials or as Gaussian random fields. For the later, we apply the procedure proposed by [102] for the efficient generation of Gaussian random fields with large domains.

Figure C.3 shows the endmembers pool for the generation of the 10-E dataset. Figure C.4 shows an example of randomly selected ground truth endmembers and corresponding generated abundance images used to synthesize an hyperspectral image from the 10E-256 $\times$ 256 pixels dataset. Simulated abundance image for each ground-truth endmember is generated sampling a Gaussian random field with Matern correlation function of parameters $\theta_1 = 10$ and $\theta_2 = 1$. Figure C.5 shows an example of abundances generated to synthesize an image with three endmembers.

We provide a toolbox for the synthetic hyperspectral images generation, freely available from the webpage of the Grupo de Inteligencia Computacional[4] from the Basque Country University (UPV/EHU). The toolbox contains a set of Matlab functions together to a collection of spectral signatures from the USGS spectral library. A graphical user interface is also provided as well as an user manual.

[1] http://speclab.cr.usgs.gov/spectral-lib.html
[2] http://speclib.jpl.nasa.gov/
[3] http://specchio.ch/
[4] http://www.ehu.es/ccwintco/index.php/GIC-source-code-free-libre

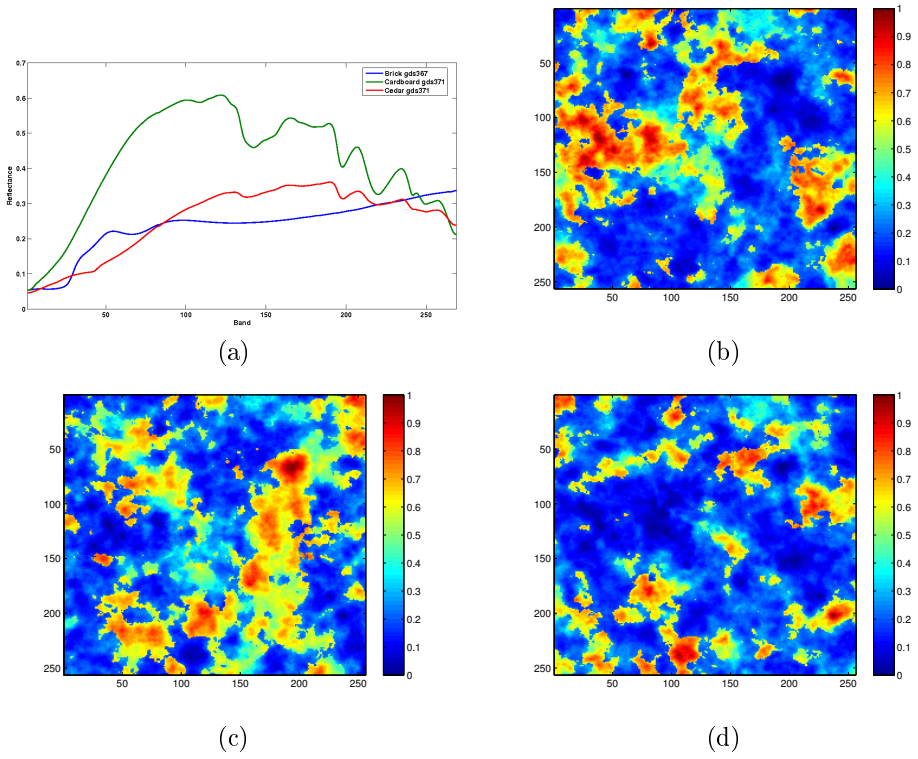


Figure C.4: Example of an image's ground truth endmembers and abundance images used to generate it. (a) The three ground-truth endmembers randomly selected from a pool of 10 endmembers from USGS spectral library. (b, c, d) 256×256 pixels synthetic abundance images corresponding to each of the endmembers in (a).

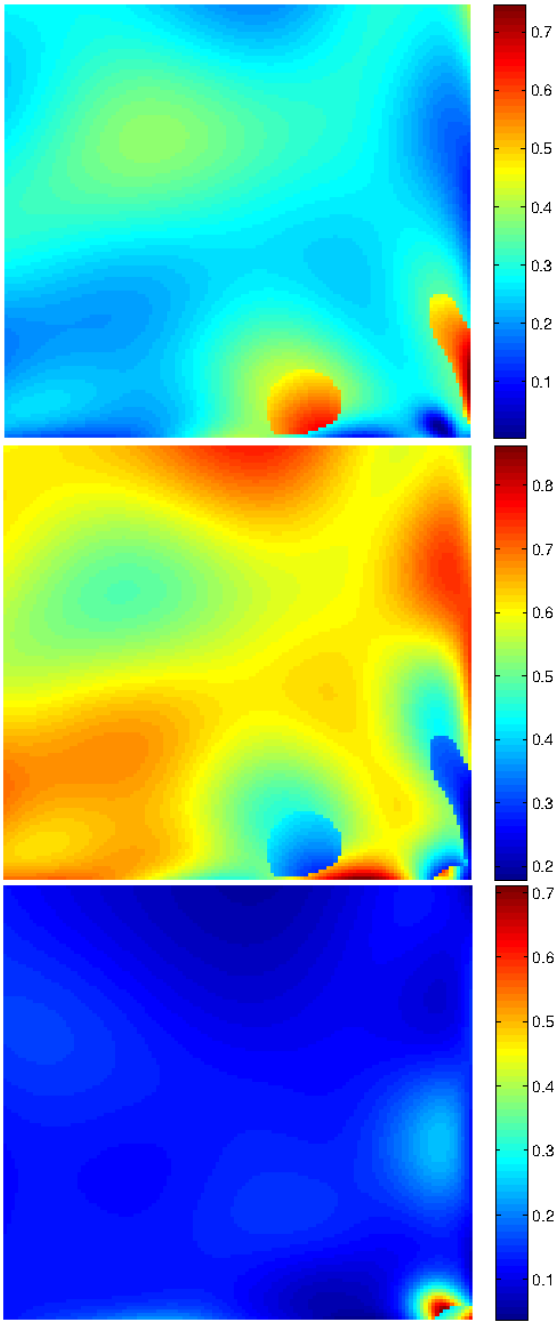


Figure C.5: An instance of synthetic fractional abundance images generated using random Legendre polynomials.

Appendix D

Real hyperspectral images

D.1 HyMAP dataset

The hyperspectral HyMAP data was made available from HyVista Corp. and German Aerospace Center's (DLR) optical Airborne Remote Sensing and Calibration Facility service[1]. The sensed scene corresponds to the radiance captured by the sensor in a flight line over the facilities of the DLR center in Oberpfaffenhofen (Germany) and its surroundings, mostly fields, forests and small towns. Figure D.1 shows the scene captured by the HyMAP sensor. The data cube has 2878 lines, 512 samples and 125 bands; and the pixel values are represented by 2-bytes signed integers.

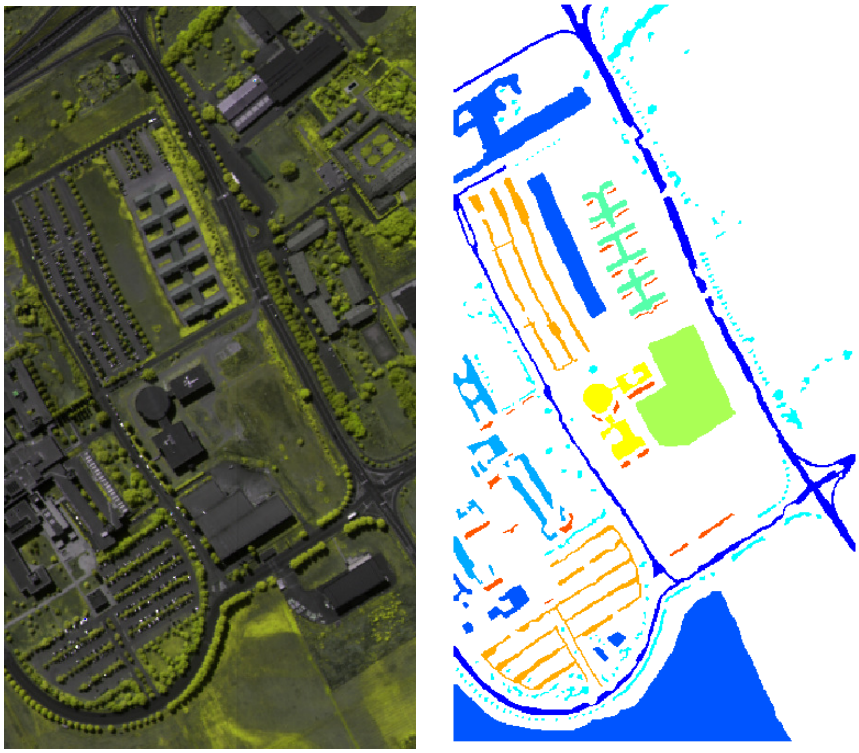
We cut the scene in patches of 64×64 pixels size for a total of 360 patches forming the hyperspectral database used in the CBIR experiments. We grouped the patches by visual inspection in five rough categories. The three main categories are 'Forests', 'Fields' and 'Urban Areas', representing patches that mostly belong to one of this categories. A 'Mixed' category was defined for those patches that presented more than one of the three main categories, being not any of them dominant. Finally, we defined a fifth category, 'Others', for those patches that didn't represent any of the above or that were not easily categorized by visual inspection. The number of patches per category are: (1) Forests: 39, (2) Fields: 160, (3) Urban Areas: 24, (4) Mixed: 102, and (5) Others: 35.

D.2 Pavia University

The Pavia University hyperspectral image was taken by the ROSIS-03 sensor over the facilities of the University of Pavia in Italy. The hyperspectral data has been provided by Prof. Paolo Gamba from the Telecommunications and Remote Sensing Laboratory, Pavia University (Italy). After discarding pixels with no information and noisy spectral bands, the image has a spatial size of

[1] http://www.OpAiRS.aero

Figure D.1: Hyperspectral scene by HyMAP sensor capturing the DLR facilities in Oberpfaffenhofen and its surroundings.



(a) (b)

Figure D.2: (a) Pavia University false color scene captured by ROSIS-03 sensor (bands 80, 90 and 70). (b) Available ground-truth for the Pavia University scene.

610×340 pixels with a spatial resolution of 1.3m per pixel, and 103 spectral bands comprised in the range of $430 - 860$ nm. Figures D.2a and D.2b show the hyperspectral scene in false color and its available ground-truth. In table D.1 the ground-truth classes are shown together to the number of labeled samples per class.

D.3 Indian Pines

The Indian Pines scene was gathered by airborne AVIRIS sensor over Northwestern Indiana and consists of 145×145 pixels and 224 spectral reflectance bands in the wavelength range 0.4–$2.5\mu m$. The spatial resolution is low compared to Pavia University image and the contains of the scene are more homogeneous, mainly two-thirds agriculture, and one-third forest or other natural perennial vegetation. There are two major dual lane highways, a railway line, as well as some low density housing, other built structures, and smaller roads.

#	Class	Samples
1	Asphalt	6631
2	Meadows	18649
3	Gravel	2099
4	Trees	3064
5	Metal sheets	1345
6	Bare Soil	5029
7	Bitumen	1330
8	Bricks	3682
9	Shadows	947

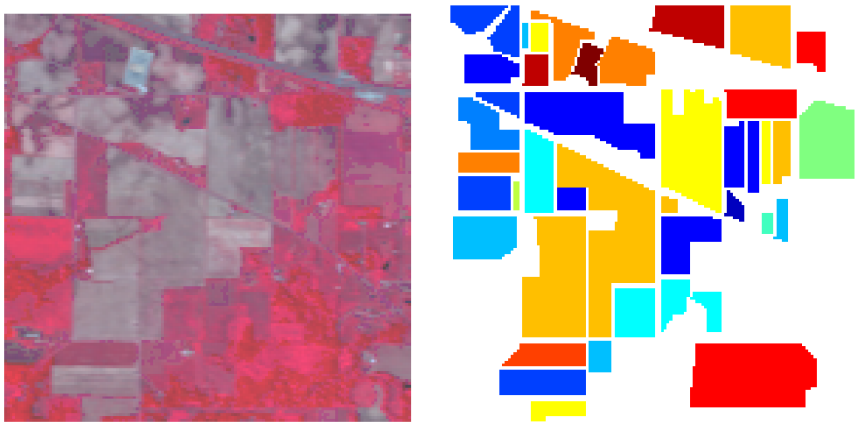
Table D.1: Ground-truth classes and number of samples per class for the Pavia University hyperspectral scene.

Since the scene is taken in June some of the crops present, corn, soybeans, are in early stages of growth with less than 5% coverage. The ground truth available is designated into sixteen classes with variable number of samples for each class (see table D.2). We have reduced the number of bands to 200 by removing bands covering the region of water absorption: $[104-108]$, $[150-163]$, 220. Indian Pines data are available through Purdue's university MultiSpec site[2]. Figure D.3 shows a false color representation and the ground truth of Indian Pines dataset.

D.4 Salinas

This scene was collected by the 224-band AVIRIS sensor over Salinas Valley, California, and is characterized by high spatial resolution (3.7-meter pixels). The area covered comprises 512 lines by 217 samples. As with Indian Pines scene, we discarded the 20 water absorption bands, in this case bands $[108-112]$, $[154-167]$ and 224. It includes vegetables, bare soils, and vineyard fields. Salinas ground truth contains 16 classes. Figure D.4 shows a sample band and the ground truth of Salinas dataset, and table D.3 shows the ground truth classes and their respective sample numbers.

[2] https://engineering.purdue.edu/~biehl/MultiSpec/

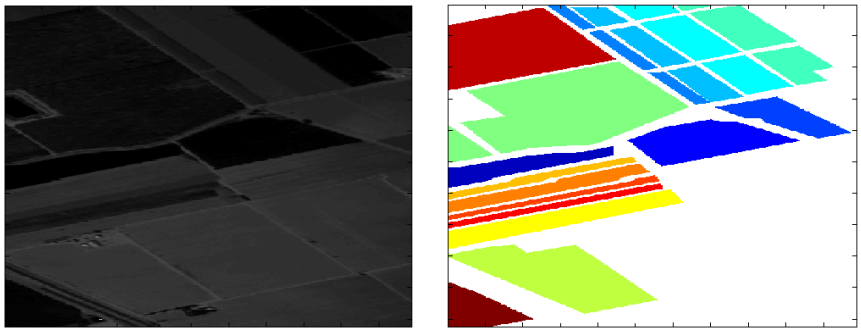


(a) (b)

Figure D.3: (a) Indian Pines false color image (bands 50, 27 and 17). (b) Indian Pines ground truth.

#	Class	Samples
1	Alfalfa	46
2	Corn-notill	1428
3	Corn-mintill	830
4	Corn	237
5	Grass-pasture	483
6	Grass-trees	730
7	Grass-pasture-mowed	28
8	Hay-windrowed	478
9	Oats	20
10	Soybean-notill	972
11	Soybean-mintill	2455
12	Soybean-clean	593
13	Wheat	205
14	Woods	1265
15	Buildings-Grass-Trees-Drives	386
16	Stone-Steel-Towers	93

Table D.2: Indian Pines ground truth classes and number of samples collected for each class.



(a) (b)

Figure D.4: (a) Salinas sample image, band 170. (b) Salinas ground truth.

#	Class	Number of samples
1	Broccoli_green_weeds_1	2009
2	Broccoli_green_weeds_2	3726
3	Fallow	1976
4	Fallow_rough_plow	1394
5	Fallow_smooth	2678
6	Stubble	3959
7	Celery	3579
8	Grapes_untrained	11271
9	Soil_vineyard_develop	6203
10	Corn_senesced_green_weeds	3278
11	Lettuce_romaine_4wk	1068
12	Lettuce_romaine_5wk	1927
13	Lettuce_romaine_6wk	916
14	Lettuce_romaine_7wk	1070
15	Vineyard_untrained	7268
16	Vineyard_vertical_trellis	1807

Table D.3: Salinas ground truth classes and number of samples for each class.

Bibliography

[1] Science for society: delivering earth system science knowledge for decision support in the year 2025. In *Geoscience and Remote Sensing Symposium, 2003. IGARSS '03. Proceedings. 2003 IEEE International*, volume 2.

[2] N. Acito, M. Diani, and G. Corsini. A new algorithm for robust estimation of the signal subspace in hyperspectral images in the presence of rare signal components. *Geoscience and Remote Sensing, IEEE Transactions on*, 47(11):3844–3856, 2009.

[3] N. Acito, M. Diani, and G. Corsini. Hyperspectral signal subspace identification in the presence of rare signal components. *Geoscience and Remote Sensing, IEEE Transactions on*, 48(4):1940–1954, 2010.

[4] 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*, 49(11):4194–4209, 2011.

[5] J. Angulo. Morphological colour operators in totally ordered lattices based on distances: Application to image filtering, enhancement and analysis. *Comput. Vis. Image Underst.*, 107(1-2):56–73, July 2007.

[6] C. M Bachmann, T. L Ainsworth, and R. A Fusina. Exploiting manifold geometry in hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3):441– 454, March 2005.

[7] C. M Bachmann, T. L Ainsworth, and R. A Fusina. Improved manifold coordinate representations of Large-Scale hyperspectral scenes. *IEEE Transactions on Geoscience and Remote Sensing*, 44(10):2786–2803, October 2006.

[8] P. Bajorski. Does virtual dimensionality work in hyperspectral images? In *Proceedings of SPIE*, pages 73341J–11, Orlando, FL, USA, 2009.

[9] P. Bajorski. On the reliability of PCA for complex hyperspectral data. In *Proceedings of WHISPERS*, pages 1–5, Grenoble, 2009.

[10] G.J.F. Banon and J. Barrera. Decomposition of mappings between complete lattices by mathematical morphology, part i. general lattices. *Signal Processing*, 30(3):299–327, February 1993.

[11] A. Baraldi, L. Bruzzone, and P. Blonda. Quality assessment of classification and cluster maps without ground truth knowledge. *Geoscience and Remote Sensing, IEEE Transactions on*, 43(4):857–873, 2005.

[12] A. Baraldi, L. Bruzzone, P. Blonda, and L. Carlin. Badly posed classification of remotely sensed images-an experimental comparison of existing data labeling systems. *Geoscience and Remote Sensing, IEEE Transactions on*, 44(1):214–235, 2006.

[13] C.A. Bateson, G.P. Asner, and C.A. Wessman. Endmember bundles: a new approach to incorporating endmember variability into spectral mixture analysis. *Geoscience and Remote Sensing, IEEE Transactions on*, 38:1083–1094, 2000.

[14] Z. Ben-Rabah, I.R. Farah, G. Mercier, and B. Solaiman. A new method to change illumination effect reduction based on spectral angle constraint for hyperspectral image unmixing. *IEEE Geoscience and Remote Sensing Letters*, 8(6):1110–1114, November 2011.

[15] C.H. Bennett, P. Gacs, M. Li, P.M.B. Vitanyi, and W.H. Zurek. Information distance. *Information Theory, IEEE Transactions on*, 44(4):1407–1423, 1998.

[16] M. Berman, H. Kiiveri, R. Lagerstrom, A. Ernst, R. Dunne, and J.F. Huntington. Ice: a statistical approach to identifying endmembers in hyperspectral images. *Geoscience and Remote Sensing, IEEE Transactions on*, 42:2085–2095, 2004.

[17] S. Beucher. *Segmentation d'images et morphologie mathematique*. PhD thesis, Ecole Nationale Superioure des Mines de Paris, Paris, 1990.

[18] G. Birkhoff. *Lattice theory*. AMS Bookstore, 1995.

[19] P. Blanchart and M. Datcu. A Semi-Supervised algorithm for Auto-Annotation and unknown structures discovery in satellite image databases. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, 3(4):698 –717, December 2010.

[20] J. Boardman, F. Kruse, and R. Green. Mapping target signatures via partial unmixing of aviris data. *Summaries of JPL Airborne Earth Science Workshop*, 1995.

[21] D. Bratasanu, I. Nedelcu, and M. Datcu. Bridging the semantic gap for satellite image annotation and automatic mapping applications. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 4(1):193–204, March 2011.

[22] C.F. Caiafa, E. Salerno, A.N. Proto, and L. Fiumi. Blind spectral unmixing by local maximization of non-Gaussianity. *Signal Processing*, 88(1):50–68, January 2008.

[23] G. Camps-Valls, J. Mooij, and B. Scholkopf. Remote sensing feature selection by kernel dependence measures. *Geoscience and Remote Sensing Letters, IEEE*, 7(3):587–591, 2010.

[24] K. Canham, A. Schlamm, A. Ziemann, B. Basener, and D. Messinger. Spatially adaptive hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 49(11):4248–4262, November 2011.

[25] D. Cerra and M. Datcu. Image retrieval using compression-based techniques. In *2010 International ITG Conference on Source and Channel Coding (SCC)*, pages 1–6. IEEE, January 2010.

[26] D. Cerra, A. Mallet, L. Gueguen, and M. Datcu. Algorithmic information Theory-Based analysis of earth observation images: An assessment. *IEEE Geoscience and Remote Sensing Letters*, 7(1):8–12, January 2010.

[27] G.J. Chaitin. *Algorithmic Information Theory*. Cambridge University Press, 2004.

[28] T.-H. Chan, C.-Y. Chi, Y.-M. Huang, and W.-K. Ma. A convex Analysis-Based Minimum-Volume enclosing simplex algorithm for hyperspectral unmixing. *IEEE Transactions on Signal Processing*, 57(11):4418–4432, November 2009.

[29] C.-I. Chang. *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*. 2003.

[30] C.-I. Chang and Q. Du. Estimation of number of spectrally distinct signal sources in hyperspectral imagery. *Geoscience and Remote Sensing, IEEE Transactions on*, 42(3):608–619, 2004.

[31] C.-I. Chang and A. Plaza. A fast iterative algorithm for implementation of pixel purity index. *Geoscience and Remote Sensing Letters, IEEE*, 3(1):63–67, 2006.

[32] C.-I. Chang, H. Ren, C.-C. Chang, F. D'Amico,, and J.O. Jensen. Estimation of subpixel target size for remotely sensed imagery. *Geoscience and Remote Sensing, IEEE Transactions on*, 42(6):1309–1320, 2004.

[33] C.-I. Chang and H. Safavi. Progressive dimensionality reduction by transform for hyperspectral imagery. *Pattern Recognition*, 44(10-11):2760–2773, 2011.

[34] C.-I. Chang and S. Wang. Constrained band selection for hyperspectral imagery. *Geoscience and Remote Sensing, IEEE Transactions on*, 44(6):1575–1585, 2006.

[35] C.-I. Chang, C.-C. Wu, and H.-M. Chen. Random pixel purity index. *IEEE Geoscience and Remote Sensing Letters*, 7(2):324–328, April 2010.

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

[37] C.-I. Chang, C.-C. Wu, C.-S. Lo, and M.-L. Chang. Real-Time simplex growing algorithms for hyperspectral endmember extraction. *IEEE Transactions on Geoscience and Remote Sensing*, 48(4):1834–1850, April 2010.

[38] C.-I. Chang, C.-C. Wu, and C.-T. Tsai. Random N-Finder (N-FINDR) endmember extraction algorithms for hyperspectral imagery. *Image Processing, IEEE Transactions on*, 20(3):641–656, 2011.

[39] C.-I. Chang, W. Xiong, W. Liu, M.-L. Chang, C.-C. Wu, and C.-C Chen. Linear spectral mixture analysis based approaches to estimation of virtual dimensionality in hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 48(11):3960–3979, November 2010.

[40] S. Chen and D. Zhang. Semisupervised dimensionality reduction with pairwise constraints for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 8(2):369–373, March 2011.

[41] X. Chen, J. Chen, X. Jia, B. Somers, J. Wu, and P. Coppin. A quantitative analysis of virtual endmembers' increased impact on the collinearity effect in spectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 49(8):2945–2956, August 2011.

[42] R. Cilibrasi and P.M.B. Vitanyi. Clustering by compression. *Information Theory, IEEE Transactions on*, 51(4):1523–1545, 2005.

[43] C.A. Coello, G.B. Lamont, and D.A. Van-Veldhuizen. *Evolutionary Algorithms for Solving Multi-Objective Problems (Genetic and Evolutionary Computation)*. Springer-Verlag New York, Inc., 2006.

[44] M.D. Craig. Minimum-volume transforms for remotely sensed data. *Geoscience and Remote Sensing, IEEE Transactions on*, 32:542–552, 1994.

[45] H. Daschiel and M. Datcu. Design and evaluation of human-machine communication for image information mining. *Multimedia, IEEE Transactions on*, 7(6):1036 – 1046, December 2005.

[46] H. Daschiel and M. Datcu. Information mining in remote sensing image archives: system evaluation. *Geoscience and Remote Sensing, IEEE Transactions on*, 43(1):188–199, 2005.

[47] M. Datcu, H. Daschiel, A. Pelizzari, M. Quartulli, A. Galoppo, A. Colapicchioni, M. Pastori, K. Seidel, P.G. Marchetti, and S. D'Elia. Information mining in remote sensing image archives: system concepts. *Geoscience*

and Remote Sensing, IEEE Transactions on, 41(12):2923 – 2936, December 2003.

[48] M. Datcu and K. Seidel. Human-centered concepts for exploration and understanding of earth observation images. *Geoscience and Remote Sensing, IEEE Transactions on*, 43(3):601 – 609, March 2005.

[49] M. Datcu, K. Seidel, and M. Walessa. Spatial information retrieval from remote-sensing images. i. information theoretical perspective. *Geoscience and Remote Sensing, IEEE Transactions on*, 36(5):1431 –1445, September 1998.

[50] R. Datta, D. Joshi, J. Li, and J.Z. Wang. Image retrieval: Ideas, influences, and trends of the new age. *ACM Comput. Surv.*, 40(2):1–60, 2008.

[51] K. Deb. *Multi-Objective Optimization Using Evolutionary Algorithms*. Wiley, 1 edition, June 2001.

[52] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, April 2002.

[53] N. Dobigeon, S. Moussaoui, M. Coulon, J.-Y. Tourneret, and A.O. Hero. Joint bayesian endmember extraction and linear unmixing for hyperspectral imagery. *Signal Processing, IEEE Transactions on*, 57(11):4355–4368, 2009.

[54] N. Dobigeon, J. -Y Tourneret, and Chein-I Chang. Semi-Supervised linear spectral unmixing using a hierarchical bayesian model for hyperspectral imagery. *IEEE Transactions on Signal Processing*, 56(7):2684–2695, July 2008.

[55] O. Duran and M. Petrou. Robust endmember extraction in the presence of anomalies. *IEEE Transactions on Geoscience and Remote Sensing*, 49(6):1986–1996, June 2011.

[56] S.S. Durbha and R.L. King. Semantics-enabled framework for knowledge discovery from earth observation data archives. *Geoscience and Remote Sensing, IEEE Transactions on*, 43(11):2563 – 2572, November 2005.

[57] S.S. Durbha, R.L. King, and N.H. Younan. An information semantics approach for knowledge management and interoperability for the global earth observation system of systems. *Systems Journal, IEEE*, 2(3):358–365, 2008.

[58] J. Eakins and M. Graham. Content-based image retrieval. Technical report, University of Northumbria at Newcastle, 1999.

[59] O. Eches, N. Dobigeon, C. Mailhes, and J.-Y. Tourneret. Bayesian estimation of linear mixtures using the normal compositional model. application to hyperspectral imagery. *Image Processing, IEEE Transactions on*, 19(6):1403–1413, 2010.

[60] O. Eches, N. Dobigeon, and J.-Y. Tourneret. Estimating the number of endmembers in hyperspectral images using the normal compositional model and a hierarchical bayesian algorithm. *IEEE Journal of Selected Topics in Signal Processing*, 4(3):582–591, June 2010.

[61] O. Eches, N. Dobigeon, and J.-Y. Tourneret. Enhancing hyperspectral image unmixing with spatial correlations. *IEEE Transactions on Geoscience and Remote Sensing*, 49(11):4239–4247, November 2011.

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

[63] M. Ferecatu and N. Boujemaa. Interactive Remote-Sensing image retrieval using active relevance feedback. *Geoscience and Remote Sensing, IEEE Transactions on*, 45(4):818 –826, April 2007.

[64] A.M. Filippi and R. Archibald. Support vector Machine-Based endmember extraction. *IEEE Transactions on Geoscience and Remote Sensing*, 47(3):771–791, March 2009.

[65] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker. Query by image and video content: the QBIC system. *Computer*, 28(9):23–32, September 1995.

[66] J. Goutsias and H.J.A.M. Heijmans. *Mathematical Morphology*. IOS Press, January 2000.

[67] M. Graña. A brief review of lattice computing. In *Fuzzy Systems, 2008. FUZZ-IEEE 2008. (IEEE World Congress on Computational Intelligence). IEEE International Conference on*, pages 1777–1781, 2008.

[68] M. Graña, D. Chyzhyk, M. Garcia-Sebastian, and C. Hernandez. Lattice independent component analysis for functional magnetic resonance imaging. *Information Sciences*, 181(10):1910–1928, May 2011.

[69] M. Graña and J. Gallego. Associative morphological memories for endmember induction. In *Geoscience and Remote Sensing Symposium, 2003. IGARSS '03. Proceedings. 2003 IEEE International*, volume 6, pages 3757–3759 vol.6, 2003.

[70] M. Graña and J. Gallego. Hyperspectral image analysis with associative morphological memories. In *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, volume 3, pages III–549–52 vol.2, 2003.

[71] M. Graña, J. Gallego, and C. Hernandez. Further results on amm for endmember induction. In *Advances in Techniques for Analysis of Remotely Sensed Data, 2003 IEEE Workshop on*, pages 237–243, 2003.

[72] M. Graña, A.M. Savio, M. Garcia-Sebastian, and E. Fernandez. A lattice computing approach for on-line fMRI analysis. *Image and Vision Computing*, 28(7):1155–1161, July 2010.

[73] M. Graña, P. Sussner, and G.X. Ritter. Associative morphological memories for endmember determination in spectral unmixing. In *Fuzzy Systems, 2003. FUZZ '03. The 12th IEEE International Conference on*, volume 2, pages 1285–1290 vol.2, 2003.

[74] M. Graña and M.A. Veganzones. Endmember induction by lattice associative memories and multi-objective genetic algorithms. *EURASIP Journal on Advances in Signal Processing*, In press.

[75] M. Graña, I. Villaverde, J.O. Maldonado, and C. Hernandez. Two lattice computing approaches for the unsupervised segmentation of hyperspectral images. *Neurocomput.*, 72(10-12):2111–2120, 2009.

[76] G. Gratzer. *General Lattice Theory*. Birkhauser Basel, 2 edition, January 2003.

[77] G. Gratzer. *Lattice Theory: Foundation*. Springer Basel, 1st edition. edition, February 2011.

[78] A.A. Green, M. Berman, P. Switzer, and M.D. Craig. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *Geoscience and Remote Sensing, IEEE Transactions on*, 26(1):65–74, 1988.

[79] L. Gueguen and M. Datcu. A similarity metric for retrieval of compressed objects: Application for mining satellite image time series. *Knowledge and Data Engineering, IEEE Transactions on*, 20(4):562 –575, April 2008.

[80] A. Gupta and R. Jain. Visual information retrieval. *Commun. ACM*, 40(5):70–79, May 1997.

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

[82] R.M. Haralick, S.R. Sternberg, and X. Zhuang. Image analysis using mathematical morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-9(4):532–550, July 1987.

[83] J.C. Harsanyi and C.-I. Chang. Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach. *Geoscience and Remote Sensing, IEEE Transactions on*, 32(4):779–785, 1994.

[84] P.W. Hawkes, H.J.A.M. Heijmans, and B. Kazan. *Morphological Image Operators*. Academic Press, December 1993.

[85] G. Healey and A. Jain. Retrieving multispectral satellite images using physics-based invariant representations. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 18(8):842 –848, August 1996.

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

[87] R. Heylen, D. Burazerovic, and P. Scheunders. Non-Linear spectral unmixing by geodesic simplex volume maximization. *IEEE Journal of Selected Topics in Signal Processing*, 5(3):534–542, June 2011.

[88] J.J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, 79(8):2554–2558, April 1982.

[89] H.-Y. Huang and B.-C. Kuo. Double nearest proportion feature extraction for Hyperspectral-Image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 48(11):4034–4046, November 2010.

[90] H.K. Huang. *PACS and Imaging Informatics: Basic Principles and Applications*. Wiley-Liss, 1 edition, April 2004.

[91] A. Huck, M. Guillaume, and J. Blanc-Talon. Minimum dispersion constrained nonnegative matrix factorization to unmix hyperspectral data. *Geoscience and Remote Sensing, IEEE Transactions on*, 48(6):2590–2602, 2010.

[92] D.P. Huijsmans and N. Sebe. How to complete performance graphs in content-based image retrieval: Add generality and normalize scope. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(2):245–251, 2005.

[93] A. Ifarraguerri and C.-I. Chang. Multispectral and hyperspectral image analysis with convex cones. *Geoscience and Remote Sensing, IEEE Transactions on*, 37:756–770, 1999.

[94] R. Green J. Boardman, F. Kruse. Mapping target signatures via partial unmixing of aviris data. 1995.

[95] S. Jia and Y. Qian. Constrained nonnegative matrix factorization for hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 47(1):161–173, January 2009.

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

[97] V. Karathanassi, D. Sykas, and K.N. Topouzelis. Development of a Network-Based method for unmixing of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 50(3):839–849, 2012.

[98] N. Keshava and J. F. Mustard. Spectral unmixing. *Signal Processing Magazine, IEEE*, 19(1):44–57, 2002.

[99] R.L. King. Putting information into the service of decision making: the role of remote sensing analysis. In *Proceedings of the 2003 IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, pages 25–29, 2003.

[100] B.C. Ko and H. Byun. Integrated region-based image retrieval using region's spatial relationships. In *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, volume 1, pages 196–199 vol.1, 2002.

[101] A. Konak, D.W. Coit, and A.E. Smith. Multi-objective optimization using genetic algorithms: A tutorial. *Reliability Engineering & System Safety*, 91(9):992–1007, September 2006.

[102] B. Kozintsev. *Computations with Gaussian Random Fields, PhD Thesis.* University of Maryland, 1999.

[103] O. Kuybeda, D. Malah, and M. Barzohar. Rank estimation and redundancy reduction of High-Dimensional noisy signals with preservation of rare vectors. *Signal Processing, IEEE Transactions on*, 55(12):5579–5592, 2007.

[104] J. Laaksonen, M. Koskela, S. Laakso, and E. Oja. PicSOM - content-based image retrieval with self-organizing maps. *Pattern Recognition Letters*, 21(13-14):1199–1207, December 2000.

[105] C.L. Lawson. *Solving Least Squares Problems.* Prentice Hall, June 1974.

[106] J. Li and R.M. Narayanan. Integrated spectral and spatial information mining in remote sensing imagery. *Geoscience and Remote Sensing, IEEE Transactions on*, 42(3):673 – 685, March 2004.

[107] J. Li, J.Z. Wang, and G. Wiederhold. IRM: integrated region matching for image retrieval. In *Proceedings of the eighth ACM international conference on Multimedia*, MULTIMEDIA '00, pages 147–156, Marina del Rey, California, United States, 2000. ACM.

[108] M. Li, X. Chen, X. Li, B. Ma, and P.M.B. Vitanyi. The similarity metric. *Information Theory, IEEE Transactions on*, 50(12):3250–3264, 2004.

[109] M. Li and P. Vitanyi. *An Introduction to Kolmogorov Complexity and Its Applications*. Springer, 2nd edition, February 1997.

[110] Y. Li and T.R. Bretschneider. Semantic-Sensitive satellite image retrieval. *Geoscience and Remote Sensing, IEEE Transactions on*, 45(4):853 –860, April 2007.

[111] M. Lienou, H. Maitre, and M. Datcu. Semantic annotation of satellite images using latent dirichlet allocation. *IEEE Geoscience and Remote Sensing Letters*, 7(1):28–32, January 2010.

[112] J. Liu and J. Zhang. A new maximum simplex volume method based on householder transformation for endmember extraction. *IEEE Transactions on Geoscience and Remote Sensing*, In press.

[113] K.-H Liu, E. Wong, E.Y. Du, C.-C. Chen, and C.-I. Chang. Kernel-Based linear spectral mixture analysis. *IEEE Geoscience and Remote Sensing Letters*, In press.

[114] L. Liu, B. Wang, and L. Zhang. Decomposition of mixed pixels based on bayesian self-organizing map and gaussian mixture model. *Pattern Recognition Letters*, 30(9):820–826, July 2009.

[115] L. Liu, B. Wang, and L. Zhang. An approach based on self-organizing map and fuzzy membership for decomposition of mixed pixels in hyperspectral imagery. *Pattern Recognition Letters*, 31(11):1388–1395, August 2010.

[116] Y. Liu, D. Zhang, G. Lu, and W.-Y. Ma. A survey of content-based image retrieval with high-level semantics. *Pattern Recognition*, 40(1):262–282, January 2007.

[117] W.Y. Ma and B.S. Manjunath. NeTra: a toolbox for navigating large image databases. In *Image Processing, 1997. Proceedings., International Conference on*, volume 1, pages 568 –571 vol.1, October 1997.

[118] A. Macedonas, D. Besiris, G. Economou, and S. Fotopoulos. Dictionary based color image retrieval. *Journal of Visual Communication and Image Representation*, 19(7):464–470, October 2008.

[119] M. Madugunki, D.S. Bormane, S. Bhadoria, and C.G. Dethe. Comparison of different CBIR techniques. In *Electronics Computer Technology (ICECT), 2011 3rd International Conference on*, volume 4, pages 372 –375, April 2011.

[120] B.S. Manjunath, J.-R. Ohm, V.V. Vasudevan, and A. Yamada. Color and texture descriptors. *IEEE Trans. on Circuits and Systems for Video Technology*, 11(6):703–715, 2001.

[121] G. Martin and A. Plaza. Region-Based spatial preprocessing for endmember extraction and spectral unmixing. *IEEE Geoscience and Remote Sensing Letters*, 8(4):745–749, July 2011.

[122] S. Mei, M. He, Z. Wang, and D. Feng. Spatial purity based endmember extraction for spectral mixture analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 48(9):3434–3445, September 2010.

[123] S. Mei, M. He, Y. Zhang, Z. Wang, and D. Feng. Improving Spatial-Spectral endmember extraction in the presence of anomalous ground objects. *IEEE Transactions on Geoscience and Remote Sensing*, 49(11):4210–4222, November 2011.

[124] M.J. Mendenhall and E. Merenyi. Relevance-Based feature extraction for hyperspectral images. *Neural Networks, IEEE Transactions on*, 19(4):658–672, 2008.

[125] F. Meyer. Topographic distance and watershed lines. *Signal Processing*, 38(1):113–125, July 1994.

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

[127] B. Mojaradi, H. Abrishami-Moghaddam, M.J.V Zoej, and R.P.W. Duin. Dimensionality reduction of hyperspectral data via spectral feature extraction. *IEEE Transactions on Geoscience and Remote Sensing*, 47(7):2091–2105, July 2009.

[128] M. Molinier, J. Laaksonen, and T. Hame. Detecting Man-Made structures and changes in satellite imagery with a Content-Based information retrieval system built on Self-Organizing maps. *Geoscience and Remote Sensing, IEEE Transactions on*, 45(4):861 –874, April 2007.

[129] H. Müller, W. Müller, D.McG. Squire, S. Marchand-Maillet, and T. Pun. Performance evaluation in content-based image retrieval: overview and proposals. *Pattern Recognition Letters*, 22(5):593–601, April 2001.

[130] J.M.P. Nascimento and J.M.B. Dias. Does independent component analysis play a role in unmixing hyperspectral data? *Geoscience and Remote Sensing, IEEE Transactions on*, 43:175–187, 2005.

[131] J.M.P. Nascimento and J.M.B. Dias. Vertex component analysis: a fast algorithm to unmix hyperspectral data. *Geoscience and Remote Sensing, IEEE Transactions on*, 43:898–910, 2005.

[132] B. Natarajan, K. Konstantinides, and C. Herley. Occam filters for stochastic sources with application to digital images. *Signal Processing, IEEE Transactions on*, 46(5):1434 –1438, may 1998.

[133] B. Paskaleva, M.M. Hayat, Z. Wang, J.S. Tyo, and S. Krishna. Canonical correlation feature selection for sensors with overlapping bands: Theory and application. *IEEE Transactions on Geoscience and Remote Sensing*, 46(10):3346–3358, October 2008.

[134] E. Pekalska and R.P.W. Duin. *The Dissimilarity Representation for Pattern Recognition: Foundations And Applications*. World Scientific Pub Co Inc, December 2005.

[135] E. Pekalska and B. Haasdonk. Kernel discriminant analysis for positive definite and indefinite kernels. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(6):1017–1032, June 2009.

[136] F.L. Pena, J.L. Crespo, and R.J. Duro. Unmixing Low-Ratio endmembers in hyperspectral images through gaussian synapse ANNs. *Instrumentation and Measurement, IEEE Transactions on*, 59(7):1834–1840, 2010.

[137] A. Pentland, R.W. Picard, and S. Sclaroff. Photobook: Content-based manipulation of image databases. *International Journal of Computer Vision*, 18(3):233–254, June 1996.

[138] A. Plaza and C.-I. Chang. Impact of initialization on design of endmember extraction algorithms. *Geoscience and Remote Sensing, IEEE Transactions on*, 44:3397–3407, 2006.

[139] A. Plaza, P. Martinez, R. Perez, and J. Plaza. Spatial/spectral endmember extraction by multidimensional morphological operations. *Geoscience and Remote Sensing, IEEE Transactions on*, 40:2025–2041, 2002.

[140] A. Plaza, P. Martinez, R. Perez, and J. Plaza. A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data. *Geoscience and Remote Sensing, IEEE Transactions on*, 42(3):650–663, 2004.

[141] A. Plaza, J. Plaza, A. Paz, and S. Blazquez. Parallel cbir system for efficient hyperspectral image retrieval from heterogeneous networks of workstations. In *Symbolic and Numeric Algorithms for Scientific Computing, 2007. SYNASC. International Symposium on*, pages 285–291, sept. 2007.

[142] A. Plaza, D. Valencia, J. Plaza, and C.-I. Chang. Parallel implementation of endmember extraction algorithms from hyperspectral data. *Geoscience and Remote Sensing Letters, IEEE*, 3:334–338, 2006.

[143] S. Prasad and L.M. Bruce. Limitations of principal components analysis for hyperspectral target recognition. *Geoscience and Remote Sensing Letters, IEEE*, 5(4):625–629, 2008.

[144] B. Raducanu, M. Graña, and F.X. Albizuri. Morphological scale spaces and associative morphological memories: Results on robustness and practical applications. *Journal of Mathematical Imaging and Vision*, 19(2):113–131, 2003.

[145] N. Raksuntorn and Q. Du. Nonlinear spectral mixture analysis for hyperspectral imagery in an unknown environment. *Geoscience and Remote Sensing Letters, IEEE*, 7(4):836–840, 2010.

[146] N. Renard, S. Bourennane, and J. Blanc-Talon. Denoising and dimensionality reduction using multilinear tools for hyperspectral images. *Geoscience and Remote Sensing Letters, IEEE*, 5(2):138–142, 2008.

[147] Y. Rezaei, M.R. Mobasheri, M.J.V. Zoej, and M.E. Schaepman. Endmember extraction using a combination of orthogonal projection and genetic algorithm. *IEEE Geoscience and Remote Sensing Letters*, 9(2):161–165, 2012.

[148] J.A. Richards and X. Jia. *Remote Sensing Digital Image Analysis: An Introduction*. Springer, 3rd edition, June 1999.

[149] G.X. Ritter, J.L. Diaz-de-Leon, and P. Sussner. Morphological bidirectional associative memories. *Neural Networks*, 12(6):851–867, July 1999.

[150] G.X. Ritter and P. Gader. Fixed points of lattice transforms and lattice associative memories. In Peter Hawkes, editor, *Advances in Imaging and Electron Physics,*, volume 144, pages 165–242. Academic Press, 2006.

[151] G.X Ritter, P. Sussner, and J.L. Diaz-de-Leon. Morphological associative memories. *Neural Networks, IEEE Transactions on*, 9(2):281–293, 1998.

[152] G.X. Ritter and G. Urcid. Lattice algebra approach to endmember determination in hyperspectral imagery. In Peter W. Hawkes, editor, *Advances in Imaging and Electron Physics,*, volume 160, pages 113–169. Academic Press, Burlington, 2010.

[153] G.X. Ritter and G. Urcid. A lattice matrix method for hyperspectral image unmixing. *Information Sciences*, 181(10):1787–1803, May 2011.

[154] G.X. Ritter, G. Urcid, and L. Iancu. Reconstruction of patterns from noisy inputs using morphological associative memories. *Journal of Mathematical Imaging and Vision*, 19(2):95–111, 2003.

[155] G.X. Ritter, G. Urcid, and M.S. Schmalz. Autonomous single-pass endmember approximation using lattice auto-associative memories. *Neurocomput.*, 72(10-12):2101–2110, 2009.

[156] D.M. 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*, 110:287–303, October 2007.

[157] Ronse C. Why mathematical morphology needs complete lattices. *Signal Processing*, 21(2):129–154, October 1990.

[158] Y. Rui, T.S. Huang, and S.-F. Chang. Image retrieval: Current techniques, promising directions, and open issues. *Journal of Visual Communication and Image Representation*, 10(1):39–62, March 1999.

[159] A. Samal, S. Bhatia, P. Vadlamani, and D. Marx. Searching satellite imagery with integrated measures. *Pattern Recognition*, 42(11):2502–2513, November 2009.

[160] M. Schroder, H. Rehrauer, K. Seidel, and M. Datcu. Spatial information retrieval from remote-sensing images. II. Gibbs-Markov random fields. *Geoscience and Remote Sensing, IEEE Transactions on*, 36(5):1446 –1455, September 1998.

[161] M. Schroder, H. Rehrauer, K. Seidel, and M. Datcu. Interactive learning and probabilistic retrieval in remote sensing image archives. *Geoscience and Remote Sensing, IEEE Transactions on*, 38(5):2288 –2298, September 2000.

[162] S.B. Serpico and G. Moser. Extraction of spectral channels from hyperspectral images for classification purposes. *IEEE Transactions on Geoscience and Remote Sensing*, 45(2):484–495, February 2007.

[163] J. Serra. *Image Analysis and Mathematical Morphology, Volume 1 (Image Analysis & Mathematical Morphology Series)*. Academic Press, February 1984.

[164] J. Serra. *Image Analysis and Mathematical Morphology, Vol. 2: Theoretical Advances*. Academic Press, 1st edition, February 1988.

[165] J. Setoain, M. Prieto, C. Tenllado, A. Plaza, and F. Tirado. Parallel morphological endmember extraction using commodity graphics hardware. *Geoscience and Remote Sensing Letters, IEEE*, 4:441–445, 2007.

[166] C.E. Shannon. A mathematical theory of communication. *SIGMOBILE Mob. Comput. Commun. Rev.*, 5(1):3–55, January 2001.

[167] J.L. Silvan-Cardenas and L. Wang. Fully constrained linear spectral unmixing: Analytic solution using fuzzy sets. *IEEE Transactions on Geoscience and Remote Sensing*, 48(11):3992–4002, November 2010.

[168] D.-G. Sim, O-K. Kwon, and R-H. Park. Object matching algorithms using robust hausdorff distance measures. *IEEE Transactions on Image Processing*, 8(3):425–429, 1999.

[169] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(12):1349 –1380, December 2000.

[170] J.R. Smith and S.F. Chang. VisualSEEk: a fully automated content-based image query system. *World Wide Web Internet And Web Information Systems*, pages 87–98, 1996.

[171] P. Soille. *Morphological Image Analysis: Principles and Applications*. Springer, 2nd edition, June 2004.

[172] R.J. Solomonoff. Algorithmic probability: Theory and applications. In *Information Theory and Statistical Learning*, pages 1–23. Springer US, Boston, MA, 2009.

[173] D.McG. Squire, W. Müller, H. Müller, and T. Pun. Content-based query of image databases: inspirations from text retrieval. *Pattern Recognition Letters*, 21(13-14):1193–1198, December 2000.

[174] H. Sun, S. Li, W. Li, Z. Ming, and S. Cai. Semantic-based retrieval of remote sensing images in a grid environment. *Geoscience and Remote Sensing Letters, IEEE*, 2(4):440 – 444, October 2005.

[175] P. Sussner. Fixed points of autoassociative morphological memories. In *IJCNN 2000, Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks, 2000*, volume 5, pages 611–616 vol.5. IEEE, 2000.

[176] P. Sussner. Generalizing operations of binary autoassociative morphological memories using fuzzy set theory. *Journal of Mathematical Imaging and Vision*, 19(2):81–93, September 2003.

[177] P. Sussner and M.E. Valle. Gray-scale morphological associative memories. *IEEE Transactions on Neural Networks*, 17(3):559–570, May 2006.

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

[179] Y. Tarabalka, J. Chanussot, and J.A. Benediktsson. Segmentation and classification of hyperspectral images using watershed transformation. *Pattern Recognition*, 43(7):2367–2379, July 2010.

[180] J.B. Tenenbaum, V. de Silva, and J.C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319 –2323, December 2000.

[181] G. Urcid and J.C. Valdiviezo. Generation of lattice independent vector sets for pattern recognition applications. In *Proceedings of SPIE*, pages 67000C–67000C–12, San Diego, CA, USA, 2007.

[182] G. Urcid, J.C. Valdiviezo, and G.X. Ritter. Endmember search techniques based on lattice auto-associative memories: a case of vegetation discrimination. In *SPIE, Proceedings*, volume 7477,74771D of *Image and Signal Processing for Remote Sensing XV*, pages 1–12. 2009.

[183] M.A. Veganzones and M. Graña. Lattice auto-associative memories induced supervised ordering defining a multivariate morphology on hyperspectral data. In *2012 4rd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*. IEEE, submitted.

[184] M.A. Veganzones and M. Graña. Endmember extraction methods: A short review. In *Proceedings of the 12th international conference on Knowledge-Based Intelligent Information and Engineering Systems, Part III*, KES '08, pages 400–407, Berlin, Heidelberg, 2008. Springer-Verlag.

[185] M.A. Veganzones and C. Hernandez. On the use of a hybrid approach to contrast endmember induction algorithms. In A. M. Savio E. S. Corchado-Rodriguez, M. Graña-Romay, editor, *Hybrid Artificial Intelligent Systems, Part II*, volume 6077 of *LNAI*, pages 69–76. Springer, 2010.

[186] M.A. Veganzones, J.O. Maldonado, and M. Grana. On Content-Based image retrieval systems for hyperspectral remote sensing images. In *Computational Intelligence for Remote Sensing*, volume 133 of *Studies in Computational Intelligence*, pages 125–144. Springer Berlin / Heidelberg, 2008.

[187] S. Velasco-Forero and J. Angulo. Supervised ordering in R^p: Application to morphological processing of hyperspectral images. *IEEE Transactions on Image Processing*, 20(11):3301–3308, November 2011.

[188] J. Wang and C.-I. Chang. Applications of independent component analysis in endmember extraction and abundance quantification for hyperspectral imagery. *Geoscience and Remote Sensing, IEEE Transactions on*, 44:2601–2616, 2006.

[189] J. Wang and C.-I. Chang. Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis. *Geoscience and Remote Sensing, IEEE Transactions on*, 44(6):1586–1600, 2006.

[190] J.Z. Wang, J. Li, and G. Wiederhold. SIMPLIcity: semantics-sensitive integrated matching for picture libraries. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(9):947 –963, September 2001.

[191] S. Wang and C.-I. Chang. Variable-Number Variable-Band selection for feature characterization in hyperspectral signatures. *Geoscience and Remote Sensing, IEEE Transactions on*, 45(9):2979–2992, 2007.

[192] T. Watanabe, K. Sugawara, and H. Sugihara. A new pattern representation scheme using data compression. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(5):579–590, 2002.

[193] M.E. Winter. N-FINDR: an algorithm for fast autonomous spectral endmember determination in hyperspectral data. In M. R. Descour and S. S.

Shen, editors, *Imaging Spectrometry V*, volume 3753 of *SPIE, Proceedings*, pages 266–275. October 1999.

[194] M.E. Winter, M.R. Descour, and S.S. Shen. N-FINDR: an algorithm for fast autonomous spectral end-member determination in hyperspectral data. volume 3753, pages 266–275, Denver, CO, USA, October 1999. SPIE.

[195] W. Xiong, C.-I. Chang, C.-C. Wu, K. Kalpakis, and H.-M. Chen. Fast algorithms to implement N-FINDR for hyperspectral endmember extraction. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 4(3):545–564, September 2011.

[196] Z. Yang, G. Zhou, S. Xie, S. Ding, J.-M. Yang, and J. Zhang. Blind spectral unmixing based on sparse nonnegative matrix factorization. *IEEE Transactions on Image Processing*, 20(4):1112–1125, April 2011.

[197] N. Yokoya, T. Yairi, and A. Iwasaki. Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 50(2):528–537, 2012.

[198] A. Zare and P. Gader. Sparsity promoting iterated constrained endmember detection in hyperspectral imagery. *Geoscience and Remote Sensing Letters, IEEE*, 4:446–450, 2007.

[199] A. Zare and P. Gader. Hyperspectral band selection and endmember detection using sparsity promoting priors. *IEEE Geoscience and Remote Sensing Letters*, 5(2):256–260, April 2008.

[200] A. Zare, P. Gader, and K.S. Gurumoorthy. Directly measuring material proportions using hyperspectral compressive sensing. *IEEE Geoscience and Remote Sensing Letters*, 9(3):323–327, 2012.

[201] H. Zeng and H.J. Trussell. Constrained dimensionality reduction using a Mixed-Norm penalty function with neural networks. *Knowledge and Data Engineering, IEEE Transactions on*, 22(3):365–380, 2010.

[202] B. Zhang, X. Sun, L. Gao, and L. Yang. Endmember extraction of hyperspectral remote sensing images based on the ant colony optimization (ACO) algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, 49(7):2635–2646, July 2011.

[203] B. Zhang, X. Sun, L. Gao, and L. Yang. Endmember extraction of hyperspectral remote sensing images based on the discrete particle swarm optimization algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, 49(11):4173–4176, November 2011.

[204] X.S. Zhou and T.S. Huang. Relevance feedback in image retrieval: A comprehensive review. *Multimedia Systems*, 8(6):536–544, April 2003.

[205] D. Ziou and S. Boutemedjet. An information filtering approach for the page zero problem. In *Multimedia Content Representation, Classification and Security*, volume 4105 of *Lecture Notes in Computer Science*, pages 619–626. Springer Berlin / Heidelberg, 2006.

[206] M. Zortea and A. Plaza. Spatial preprocessing for endmember extraction. *IEEE Transactions on Geoscience and Remote Sensing*, 47(8):2679–2693, August 2009.